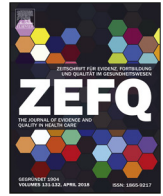




Contents lists available at ScienceDirect

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ)

journal homepage: <http://www.elsevier.com/locate/zefq>

Evidenz in der Gesundheitsversorgung / Evidence in Health Care

## Mehr Versorgung für Patient\*innen dank nutzerfreundlicher Praxissoftware? Eine explorative Studie

*More care for patients thanks to user-friendly practice software? An exploratory study*

Doreen Müller<sup>a,b,\*</sup>, Michael Erhart<sup>c,d</sup>, Jakob Holstiege<sup>a</sup>, Dominik Graf von Stillfried<sup>a</sup>

<sup>a</sup>Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland, Berlin, Deutschland

<sup>b</sup>Charité – Universitätsmedizin Berlin, Centrum für Prävention, Human- und Gesundheitswissenschaften, Institut für Medizinische Soziologie und Rehabilitationswissenschaft, Berlin, Deutschland

<sup>c</sup>Alice Salomon Hochschule Berlin, Fachbereich Gesundheits-/Rehabilitationswissenschaften, Berlin, Deutschland

<sup>d</sup>Apollon Hochschule der Gesundheitswirtschaft, Fachbereich Psychologie, Bremen, Deutschland

### ARTIKEL INFO

#### Artikel-Historie:

Eingegangen: 21. August 2025  
Revision eingegangen: 11. Dezember 2025  
Akzeptiert: 11. Februar 2026  
Online gestellt: xxxx

#### Schlüsselwörter:

Praxisverwaltungssysteme  
System Usability Scale  
Nutzerfreundlichkeit  
Arbeitsabläufe  
Versorgung  
Leistung

### ZUSAMMENFASSUNG

**Hintergrund:** Angesichts eines steigenden Versorgungsbedarfs und begrenzter Arztzeit untersucht diese Studie, ob die Usability (Nutzer\*innenfreundlichkeit) von Praxisverwaltungssystemen (PVS) mit der Leistungsfähigkeit von Arztpraxen zusammenhängt. Analysiert wird, ob nutzerfreundlichere PVS mit mehr Behandlungsfällen oder einem höheren Leistungsvolumen einhergehen (primäre Hypothesen), insbesondere abhängig von der jeweiligen Fallzahl bzw. dem Leistungsvolumen der Praxis (sekundäre Hypothesen).

**Methoden:** Untersucht wurden 24.140 (Anzahl Behandlungsfälle) bzw. 24.434 (Leistungsbedarf in Euro) Einzelpraxen niedergelassener Ärzt\*innen im Jahr 2024. Mithilfe linearer Mehrebenenmodelle wurde sowohl für die Gesamtstichprobe als auch getrennt nach Praxen ober- und unterhalb des Modalwerts der abhängigen Variable geprüft, ob die Anzahl der Behandlungsfälle bzw. das Leistungsvolumen mit der Usability des PVS zusammenhängt. In einer Sensitivitätsanalyse wurden Patient\*innenstruktur und Krankheitslast sowie regionale sozioökonomische Deprivation und Arztstruktur berücksichtigt.

**Ergebnisse:** In der Gesamtstichprobe zeigte sich in Bezug auf die primären Hypothesen kein signifikanter Zusammenhang zwischen Usability und Behandlungsfallzahl ( $\beta = 7,9$ ; 90%-Konfidenzintervall (KI)  $[-4,6; 20,4]$ ) bzw. Leistungsvolumen ( $\beta = 505,4$ ; 90%-KI  $[-226,5; 1.237,7]$ ). Bei Praxen im Segment mit einem höheren Leistungsvolumen assoziiert ( $\beta = 530,1$ ; 90%-KI  $[13,9; 1.046,8]$ ), auch nach Sensitivitätsanalyse adjustiert für Patient\*innenfaktoren der Praxis und regionale Strukturen ( $\beta = 518,1$ ; 90%-KI  $[9,6; 1.027,3]$ ).

**Schlussfolgerungen:** Die Studie legt ein Potenzial von Praxissoftware zur Effizienzsteigerung in der Arztpraxis nahe, besonders bei noch nicht ausgeschöpften Kapazitäten. Diese Ergebnisse können Hersteller von PVS dazu anregen, die Usability ihrer Systeme gezielt weiterzuentwickeln und stärker an den praktischen Anforderungen des Versorgungsalltags auszurichten.

### ARTICLE INFO

#### Article History:

Received: 21 August 2025  
Received in revised form: 11 December 2025  
Accepted: 11 February 2026  
Available online: xxxx

### ABSTRACT

**Background:** Given the increasing demand for care and limited physician time, this study investigates whether the usability of practice management systems (PMS) is associated with the performance of medical practices. It analyzes whether more user-friendly PMS enable a higher number of treatment cases and greater service volumes (primary hypotheses), depending, in particular, on case load or service volume (secondary hypotheses).

**Methods:** A total of 24,140 individual practices (number of treatment cases) and 24,434 individual practices (service volume in euros) of office-based physicians were analyzed in 2024. Using linear multilevel

\* Korrespondenzadresse: Doreen Müller, Zentralinstitut für die kassenärztliche Versorgung, Salzufer 8, 10587 Berlin, Deutschland.  
E-Mail: [dmueller@zi.de](mailto:dmueller@zi.de) (D. Müller).

**Keywords:**

Practice management systems  
System Usability Scale (SUS)  
Usability  
Workflows  
Care  
Performance

models, it was tested whether the number of treatment cases or the service volume was associated with the usability of the PMS, both for the total sample and separately for practices above and below the mode of the dependent variable. In a sensitivity analysis, patient characteristics as well as regional socioeconomic deprivation and physician structure were taken into account.

**Results:** In the multilevel models of the total sample (concerning the primary hypotheses), no significant association was found between usability ( $\beta = 7.9$ ; 90% confidence interval (CI)  $[-4.6; 20.4]$ ) or the number of treatment cases or service volume ( $\beta = 505.4$ ; 90% CI  $[-226.5; 1,237.7]$ ). However, in practices below the mode, higher usability was significantly associated with higher service volumes ( $\beta = 530.1$ ; 90% CI  $[13.9; 1,046.8]$ ), even after sensitivity analysis, adjusted for patient-level factors of the practice and regional structures ( $\beta = 518.1$ ; 90% CI  $[9.6; 1,027.3]$ ).

**Conclusions:** The study highlights the potential of PMS to increase efficiency, especially where capacities are not fully utilized. The findings may encourage developers to further improve PMS usability and better tailor systems to everyday practice needs.

## Hintergrund

In Deutschland sind derzeit rund 190.000 niedergelassene Ärzt\*innen in circa 100.000 Praxen tätig [1]. Bis zum Jahr 2040 werden laut aktuellen Schätzungen zwischen 30.000 und 50.000 ambulant tätige Mediziner\*innen fehlen [2]. Doch auch schon heute berichten Patient\*innen zunehmend von Schwierigkeiten, einen Termin zu bekommen, wie die aktuellste GKV-Versichertenbefragung aus dem Jahr 2025 zeigt [3]. Eine mögliche Ursache liegt darin, dass jüngere Ärzt\*innen im Schnitt weniger Arbeitszeit leisten als die Generation, die in den Ruhestand geht. Gleichzeitig trägt der demographische Wandel unserer älter werdenden Gesellschaft dazu bei, dass der Versorgungsbedarf weiter steigt. Vor diesem Hintergrund gewinnt die Frage an Bedeutung, wie sich die verfügbare Arztzeit effizienter für die medizinische Behandlung nutzen lässt.

Häufig wird darauf hingewiesen, dass vor allem bürokratische Erfordernisse die Zeit begrenzen, die ein\*e Ärzt\*in für die direkte Patient\*innenversorgung aufwenden kann [4]. Doch auch andere Faktoren können dazu führen, dass weniger Zeit für die eigentliche medizinische Tätigkeit bleibt. So dürfte die Effizienz der Arbeitsabläufe in einer Arztpraxis maßgeblich beeinflussen, wie viele zeitliche Kapazitäten den ärztlichen Leistungserbringenden für die Versorgung ihrer Patient\*innen zur Verfügung steht. Im Mittelpunkt der Betrachtung dieser Studie stehen hierbei die Abläufe im Zusammenhang mit Praxisverwaltungssystemen (PVS).

Jede Praxis nutzt für die Dokumentation und Abrechnung ihrer Leistungen ein PVS nach den Vorgaben des SGB V (§ 294) und der Kassenärztlichen Bundesvereinigung (KBV). Die Nutzer\*innenfreundlichkeit, auch als Usability bezeichnet, beschreibt, wie effektiv, auch als Usability bezeichnet, beschreibt, wie effektiv, effizient und zufriedenstellend ein Produkt, wie zum Beispiel eine Software, angewendet werden kann [5]. Sie ist ein Maß dafür, wie reibungslos und zügig ein System bedient werden kann und in welcher Weise es die positiven Erwartungen der Nutzer\*innen erfüllt. Sie ist damit nicht nur ein zentraler Indikator für die Akzeptanz und den Erfolg eines Systems, sondern kann auch Hinweise auf mögliche Störfaktoren und Reibungsverluste im Arbeitsablauf geben. In Bezug auf PVS bedeutet eine hohe Usability, dass sie ein übersichtliches Layout enthalten, keine überflüssigen Informationen dargestellt werden, die Arbeitsabläufe nicht redundant sind, keine irrelevanten Warnmeldungen erscheinen und dass das System insgesamt fehlerfrei läuft [6].

Studien zeigen, dass eine geringere Usability von Software die Leistungsfähigkeit der Anwender\*innen beeinträchtigen könnte [7]. Für PVS konnte darüber hinaus ein signifikanter Zusammenhang zwischen der Usability und der Zahl an Klicks sowie der Bearbeitungsdauer pro Vorgang nachgewiesen werden: Je mehr Zeit und Interaktionen für die Durchführung eines Vorgangs erforderlich sind, umso niedriger ist die subjektive Usability [8]. Auch tragen

häufigere Fehlersituationen zu einer geringeren Usability bei [9] und dürften die zeitliche Effizienz der Praxisabläufe verringern. Eine Metaanalyse zeigte ebenfalls auf, dass elektronische Patientenakten, die eine geringe Usability aufweisen, den Workflow unterbrechen und die Arbeitszeit erhöhen, die für die Dokumentation aufgewendet werden muss [10].

Nach Kenntnis der Autor\*innen wurde bislang nicht untersucht, ob und in welchem Ausmaß die Usability eines PVS mit Leistungskennzahlen einer Arztpraxis zusammenhängt. Es ist zu vermuten, dass ein nutzerfreundlicheres PVS effizientere Arbeitsabläufe ermöglicht und die Praxis dadurch in der Lage ist, ein höheres Leistungsvolumen zu erbringen. Dies kann sowohl über die Anzahl der Behandlungsfälle als auch über den erbrachten Leistungsbedarf in Euro erfasst werden.

Gleichzeitig ist es denkbar, dass dieser Zusammenhang nicht für alle Arztpraxen in gleichem Maße gilt. So könnten Praxen mit einem hohen Patient\*innenaufkommen besonders profitieren, da sich kleine Zeitersparnisse bei vielen Fällen stärker aufsummieren. Andererseits könnte ein besser bedienbares Praxisverwaltungssystem bei kleineren Praxen im Verhältnis zur Gesamtleistungsmenge einen größeren Effekt erzielen. Diese Studie untersucht daher zusätzlich, ob Praxen aus dem oberen im Vergleich zum unteren Segment der Fallzahlen und des Leistungsvolumens unterschiedlich stark von einer höheren Usability profitieren. Folgende Forschungsfragen wurden bearbeitet:

**Primäre Forschungsfragen:** Erbringt eine Arztpraxis, je nutzerfreundlicher ihr PVS ist,

- I a) umso mehr Behandlungsfälle?
- I b) umso mehr Leistungsvolumen?

**Sekundäre Forschungsfragen:** Unterscheidet sich dies zwischen Arztpraxen

- II a) im Segment mit geringeren versus mit höheren Fallzahlen?
- II b) im Segment mit geringerem versus mit höherem Leistungsvolumen?

## Methode

### Stichprobe

In die Stichprobe wurden Arztpraxen eingeschlossen, welche im Jahr 2024 in Deutschland kassenärztliche Leistungen an gesetzlich Versicherten erbrachten. Um eine Vergleichbarkeit zwischen den Leistungsindikatoren der Arztpraxen zu gewährleisten, wurden nur Einzelpraxen ausgewählt, also jene Arztpraxen, welche nicht mehr als eine\*n Leistungserbringer\*in aufwiesen. Da sich die Fall- und Leistungsstruktur zwischen psychotherapeutisch versus

ärztlich tätigen Leistungserbringer\*innen stark unterscheidet, wurden zudem ärztliche und nicht-ärztliche Psychotherapeut\*innen aus der Analyse ausgeschlossen. Zusätzlich wurden nur Praxen berücksichtigt, welche im Jahr 2024 ausschließlich ein einziges PVS verwendeten; Praxen ohne Information zum verwendeten PVS wurden ebenfalls von den Analysen ausgeschlossen. Insgesamt wurden somit 45.522 Praxen in die Studienpopulation aufgenommen.

#### Datenbasis und Maße

Für die Identifikation der Arztpraxen und alle weiteren Informationen auf Praxis-Ebene wurden die kassenärztlichen Abrechnungsdaten nach §295 SGB V des Abrechnungsjahres 2024 verwendet. Aus den Abrechnungsdaten wurden als abhängige Variablen die Gesamtzahl der Behandlungsfälle mit gesetzlich Versicherten sowie das insgesamt erbrachte Leistungsvolumen nach Euro-Gebührenordnung der gesetzlichen Krankenversicherung des Jahres 2024 ermittelt. Da in den Daten auch Leistungserbringer\*innen mit außergewöhnlich geringen oder hohen Fall- und Leistungsbedarfszahlen zu finden sind und die vorliegende Studie ihr Erkenntnisinteresse auf die durchschnittliche Arztpraxis richtet, wurden für die Analyse jene Leistungserbringer\*innen eingeschlossen, welche sich jeweils zwischen dem 20. und 80. Perzentil des entsprechenden Leistungsindikators befanden. Es wurde zudem der Versorgungsbereich der Praxis (hausärztlich, fachärztlich, sonstiger Bereich) extrahiert, da dieser signifikant mit der Bewertung der Usability zusammenhängt [9].

Pro Praxis wurde das dort installierte PVS anhand einer Prüfnummer aus den Abrechnungsdatentransfer-(ADT)-Abrechnungen identifiziert. Als erklärende Variable wurde auf Ebene der PVS die Bewertung der Usability im Rahmen des System Usability Scores (SUS) [11] aus einer Untersuchung anhand von über 10.000 Bewertungen für 39 PVS hinzugespielt. Der SUS wird anhand eines Fragebogens aus zehn Fragen zur Verwendbarkeit der Software gebildet. Er kann Werte zwischen 0 (geringste Nutzerfreundlichkeit) und 100 (höchste Nutzerfreundlichkeit) annehmen. Die Reliabilität und Validität der deutschen Fassung des SUS für die Bewertung von PVS wurde nachgewiesen; zur weiteren Methodik der Erhebung siehe Müller et al. [9].

Für eine Sensitivitätsanalyse wurden jeder Praxis Faktoren ihrer Patient\*innenstruktur hinzugefügt: Der Anteil weiblicher Patientinnen, das durchschnittliche Patient\*innenalter sowie der Charlson-Comorbidity-Index [12] zur Adjustierung der Krankheitslast pro Praxis. Für das Hinzuspielen weiterer Kontrollvariablen auf Kreisebene wurde den Praxen aus den Abrechnungsdaten der Kreis zugeordnet. Hierfür wurde jener Kreis ausgewählt, in welchem die meisten Patient\*innen der jeweiligen Praxis wohnten. Anschließend wurde darauf basierend der German Index of Socioeconomic Deprivation (GISD) vom Robert-Koch-Institut in der aktuellsten verfügbaren Version (Jahr 2021) als Kontrollvariable hinzugespielt [13]. Der GISD kondensiert regionale Informationen zur Bildungs-, Beschäftigungs- und Einkommenssituation anhand von Werten zwischen 0 (geringste Deprivation) und 1 (höchste Deprivation). Er ist ein anerkanntes und häufig verwendetes Maß für Analysen in Bezug auf das deutsche Gesundheitswesen, da er stark mit Krankheitsindikatoren und Inanspruchnahmefaktoren zusammenhängt. Als zusätzliche Kontrollvariablen auf Kreisebene wurden aus dem Bundesarztregister der Kassenärztlichen Bundesvereinigung die Indikatoren Anteil weiblicher Ärztinnen in Prozent, Durchschnittsalter der Ärzt\*innen in Jahren, Anteil Vertragsärzt\*innen in Prozent, Teilnahmumfang pro Kopf sowie die Arztdichte (Ärzt\*innen pro 10.000 Einwohner) angespielt [14].

#### Analysen

Die Information über die Usability eines PVS lag nicht pro Praxis, sondern auf Ebene des PVS vor. Zudem unterscheiden sich Abläufe, Fallstrukturen und Leistungsvolumina zwischen Praxen

aus dem hausärztlichen, fachärztlichen und sonstigen Bereich [15]. Daher wurde mit Hilfe einer Intraklassenkorrelation (ICC) zunächst geprüft, ob die Varianz der Beobachtungen zwischen den PVS- und Versorgungsbereichs-Gruppen größer ist als die Varianz innerhalb der Gruppen, und damit eine Mehrebenen-Analyse angezeigt ist. Die ICC mit einem Wert von 0,86 zeigte, dass ein großer Teil der Gesamtvarianz auf Unterschiede zwischen den Gruppen PVS  $\times$  Versorgungsbereich zurückzuführen ist. Daher war die Anwendung von linearen Mehrebenen-Modellen sinnvoll.

In den anschließenden linearen Mehrebenen-Analysen waren die Praxen auf Level I angeordnet und in Gruppen von PVS  $\times$  Versorgungsbereich auf Level II-a geschachtelt. Für die Beantwortung der primären Forschungsfragen wurden die Analysen für die beiden abhängigen Variablen Behandlungsfallzahl und Leistungsbedarf durchgeführt, welche auf Level I vorlagen. Auf Level II-a wurde der SUS-Wert als erklärende Variable hinzugefügt. Für beide abhängigen Variablen (Zahl der Behandlungsfälle und Leistungsbedarf in Euro) wurden drei Mehrebenen-Modelle gerechnet: Zunächst wurde das Modell für den gesamten Bereich der abhängigen Variable errechnet (primäre Forschungsfragen I a und I b). Anschließend wurde die Verteilung der abhängigen Variablen an ihrem am häufigsten vorkommenden Wert (Modalwert) in zwei Bereiche geteilt, und das Modell wurde für die Prüfung der sekundären Forschungsfragen II a und II b getrennt für das unterhalb beziehungsweise oberhalb des Modalwertes liegende Verteilungssegment geschätzt. Ziel dieser Analysen war es, die Zusammenhänge innerhalb jedes Segments jeweils separat in Bezug auf die abhängigen Variablen zu untersuchen.

Der Modalwert wurde gewählt, da er den am häufigsten vorkommenden und damit den empirisch „typischsten“ Wert abbildet. Dadurch lassen sich Praxen ober- und unterhalb des real häufigsten Leistungsniveaus differenzieren. Diese Herangehensweise erscheint für die Fragestellung geeigneter als eine Aufteilung nach Median, da hier nicht die Position innerhalb einer theoretischen Ordnung, sondern das tatsächlich vorherrschende Leistungsniveau als Referenzpunkt dient. Für die Ermittlung des Modalwertes wurde die Anzahl der Behandlungsfälle in 100er und der Leistungsbedarf in Euro in 10.000er Kategorien eingeordnet, um die Daten in sinnvoll gerundete Klassen zu überführen und so die Modalwerte besser identifizieren zu können. Diese Klasseneinteilung war erforderlich, da bei stetigen und sehr fein aufgelösten Variablen wie Abrechnungsvolumina in Euro in der Regel keine identischen Werte zwischen Praxen vorkommen. Durch die Bildung von Kategorien wird es erst möglich, Häufigkeitsschwerpunkte und damit empirisch typische Werte sinnvoll zu bestimmen, ohne den inhaltlichen Aussagekern der Verteilung zu verändern. Bei der Anzahl der Fälle ergaben sich somit 51 Klassen, während dieses Vorgehen beim Leistungsbedarf in Euro zu 35 Klassen führte.

Die erklärende sowie die Kontroll-Variablen wurden für die Mehrebenen-Analysen am Gesamtmittelwert zentriert. Da die Forschungsfragen a priori einen positiven Zusammenhang zwischen dem SUS-Wert und der Anzahl der Fälle beziehungsweise dem Leistungsvolumen in Euro annehmen, erfolgte der Signifikanztest einseitig. Dieses Vorgehen entspricht der Logik der Forschungshypothese und reduziert das Risiko falsch-negativer Befunde bei gerichtetem Erwartungswert [16].

Wenn der Zusammenhang der erklärenden Variable signifikant war, so wurden dem Modell in einer Sensitivitätsanalyse die Kontrollvariablen auf Praxis- und Kreisebene hinzugefügt, um den Haupteffekt auf seine Robustheit zu prüfen. Eine Korrektur für multiples Testen ist nicht erforderlich, wenn nur eine erklärende Variable untersucht wird und die übrigen Variablen ausschließlich der Kontrolle von Störfaktoren dienen, da anderenfalls die Wahrscheinlichkeit des Typ-II-Fehlers erhöht wird [17]. Es wurde mit „Hilfe“ der Varianzinflationsfaktoren (VIF) geprüft, ob eine Multi-

kollinearität vorlag. Da die Kreise ebenfalls eine Level-II-Ebene darstellen (Level II-b), sich die PVS × Versorgungsbereich-Gruppen (Level II-a) jedoch nicht innerhalb der Kreise einordnen lassen, sondern dazu über Kreuz liegen, wurde hierfür der Cross-Classified Mehrebenen-Ansatz gewählt [18]. Eine schematische Übersicht über die Mehrebenen-Modellierung und der Variablen für jedes Level findet sich in [Abbildung 1](#).

Praxen mit fehlenden Werten wurden zeilenweise aus den Analysen ausgeschlossen. Für die Datenaufbereitung wurden die Programme Oracle SQL Developer (Version 21.2.1.204) und R (Version 4.1.2) verwendet, die Mehrebenen-Analysen wurden in R mit den Paketen lme4 (Version 1.1.32), lmerTest (Version 3.1.3) und performance (Version 0.15.0) durchgeführt. Das vorliegende Manuskript wurde entsprechend der STROBE-Richtlinien für querschnittliche Studien erstellt [19].

**Ergebnisse**

*Deskription*

Insgesamt wurden 45.522 Praxen als grundsätzlich für die Analysen in Frage kommend identifiziert. Pro Praxis stammten im Durchschnitt 77,3 Prozent (Standardabweichung = 19,4 Prozent) der Versorgten aus dem Kreis, welchem die Praxis zugeordnet wurde. Die Anzahl der Behandlungsfälle pro in die Analyse einfließender Arztpraxis lag zwischen 602 (20. Perzentil) und 5.557 (80. Perzentil) mit einem Modalwert von 4.100. Das Leistungsvolumen in Euro lag zwischen 45.499 (20. Perzentil) und 394.460 (80. Perzentil) mit einem Modalwert von 280.000 ([Tabelle 1](#)). Für 78,7 Prozent der Praxen (N = 35.995) lagen Informationen zum mittleren SUS-Wert des von ihnen verwendeten PVS vor. Der SUS-Wert hatte über die 39 PVS einen Mittelwert von 58,8 (Standardabweichung = 11,5).

*Forschungsfragen 1: Mehr Leistung bei nutzerfreundlicherem PVS?*

In das Analysemodell zur Anzahl der Behandlungsfälle konnten 24.140 Praxen aufgenommen werden, für welche der mittlere SUS-

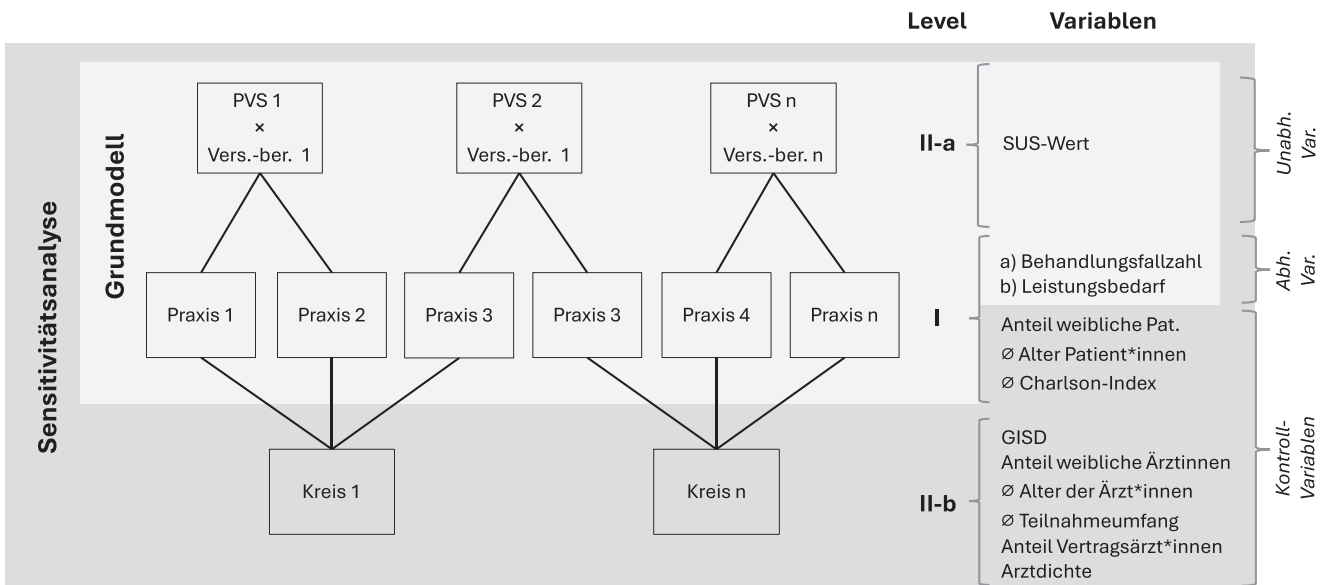
Wert des verwendeten PVS vorlag, sowie welche im Segment zwischen dem 20. und 80. Perzentil des Leistungsindikators lagen. Insgesamt 24.434 Praxen konnten analog in das Modell zum Leistungsbedarf in Euro eingeschlossen werden. In [Tabelle 2](#) und [Tabelle 3](#) werden die Ergebnisse der Mehrebenen-Modelle dargestellt. Es zeigte sich, dass im Modell über alle eingeschlossenen Arztpraxen ein höherer SUS-Wert nicht statistisch signifikant mit einer höheren Behandlungsfallzahl zusammenhing ( $\beta = 7,9$ ; 90%-KI [-4,6; 20,4]). Auch ein höheres Leistungsvolumen konnte im Gesamtmodell nicht durch einen höheren SUS-Wert erklärt werden ( $\beta = 505,4$ ; 90%-KI [-226,5; 1.237,7]).

*Forschungsfragen 2: Unterschiede nach Ausmaß der Leistung?*

Die Anzahl der Behandlungsfälle war weder im unteren ( $\beta = 4,8$ ; 90%-KI [-3,6; 13,1]) noch im oberen ( $\beta = 0,3$ ; 90%-KI [-0,5; 1,0]) Verteilungssegment signifikant mit der Usability assoziiert ([Tabelle 2](#)). Beim Leistungsvolumen gab es in Arztpraxen, die oberhalb des Modalwertes lagen, ebenfalls keinen statistisch signifikanten Zusammenhang zum SUS-Wert ( $\beta = 28,6$ ; 90%-KI [-25,2; 81,9], [Tabelle 3](#)). Bei Arztpraxen im unteren Verteilungssegment zeigte sich jedoch, dass eine höhere Usability mit einem höheren Leistungsvolumen in Euro statistisch signifikant einherging ( $\beta = 530,1$ ; 90%-KI [13,9; 1.046,8], [Tabelle 3](#)).

In der Sensitivitätsanalyse ([Tabelle 4](#)) blieb der Effekt bestehen ( $\beta = 518,1$ ; 90%-KI [9,6; 1.027,3]). Bezüglich der Varianzinflationsfaktoren gab es keine Auffälligkeiten. In der Sensitivitätsanalyse zeigten sich zudem auf Seiten der Kontrollvariablen weitere signifikante Zusammenhänge. So waren ein höherer Anteil weiblicher Patientinnen pro Praxis ( $\beta = 157,9$ ; 95%-KI [92,9; 222,8]) sowie ein höherer Wert des Kreis-bezogenen GISD ( $\beta = 26.024,6$ ; 95%-KI [16.702,7; 35.351,2]) mit einem höheren Leistungsbedarf in Euro assoziiert.

Im Grundmodell wurden durch den fixen Effekt der Usability 0,7 Prozent (Sensitivitätsanalyse: 1,7 Prozent) der Varianz im Leistungsvolumen erklärt. Die erklärte Gesamtvarianz des Leistungsbedarfs in Euro durch das Modell für das untere Leistungssegment inklusive der Unterschiede auf Level II-a mit der Gruppenzugehör-



**Abbildung 1.** Schematische Darstellung der Mehrebenen-Modelle (Grundmodell und Sensitivitätsanalyse) mit den verwendeten Variablen.

**Tabelle 1**

Deskriptive Beschreibung der Stichprobe und Variablen.

		Anzahl Fälle			Leistungsbedarf in Euro		
		Gesamt	Unteres Segment	Oberes Segment	Gesamt	Unteres Segment	Oberes Segment
<b>Anzahl Praxen</b>							
	N	24.140	15.020	9.120	24.434	15.156	9.278
<b>Abhängige Variablen auf Level I (Praxen)</b>							
Anzahl der Behandlungsfälle	Min	603,0	603,0	4.101,0	19,0	19,0	243,0
	MW	3.486,9	2.703,7	4.776,8	3.665,8	2.978,1	4.789,0
	SD	1.302,8	1.000,7	411,3	1.700,8	1.492,7	1.394,2
Leistungsbedarf in Euro	Max	5.557,0	4.100,0	5.557,0	13.164,0	12.134,0	13.164,0
	Min	5.414,2	5.414,2	36.347,1	45.499,3	45.499,3	280.001,0
	MW	268.994,4	227.732,7	336.949,5	241.437,3	185.475,8	332.852,8
	SD	155.388,0	159.339,1	121.150,1	89.984,6	64.476,5	32.566,4
	Max	4.417.801,6	4.417.801,6	3.827.756,9	394.460,3	279.998,9	394.460,3
<b>Unabhängige Variable Level II-a (PVS × Versorgungsbereich)</b>							
System Usability Scale (SUS-Wert)	Min	42,8	42,8	42,8	42,8	42,8	42,8
	MW	58,8	58,7	58,9	58,7	58,7	58,7
	SD	11,6	11,4	11,8	11,5	11,4	11,6
	Max	83,5	83,5	83,5	83,5	83,5	83,5
<b>Kontrollvariablen auf Level I (Praxen)</b>							
Prozentualer Anteil weiblicher Patientinnen	Min	3,7	3,7	9,0	3,7	5,2	3,7
	MW	60,4	60,5	60,4	60,7	61,5	59,3
	SD	17,3	17,2	17,6	17,6	18,2	16,5
	Max	100,0	100,0	100,0	100,0	100,0	100,0
Alter der Patient*innen	Min	3,8	3,8	4,1	3,8	3,8	4,1
	MW	51,6	51,5	51,6	51,5	51,4	51,7
	SD	13,5	13,5	13,5	12,9	12,3	13,8
	Max	86,0	86,0	84,0	86,0	86,0	83,3
Charlson-Comorbidity-Index	Min	0,0	0,0	0,0	0,0	0,0	0,0
	MW	1,1	1,1	1,1	1,1	1,0	1,1
	SD	0,7	0,7	0,6	0,6	0,6	0,6
	Max	6,4	6,4	5,9	6,2	6,2	6,0
<b>Kontrollvariablen auf Level II-b (Kreise)</b>							
German Index of Socioeconomic Deprivation (GISD)	Min	0,0	0,0	0,0	0,0	0,0	0,0
	MW	0,5	0,5	0,5	0,5	0,5	0,5
	SD	0,2	0,2	0,2	0,2	0,2	0,2
	Max	1,0	1,0	1,0	1,0	1,0	1,0
Prozentualer Anteil weiblicher Ärztinnen	Min	34,4	34,4	34,4	34,4	34,4	34,4
	MW	56,0	56,1	55,9	55,9	55,9	55,9
	SD	6,4	6,4	6,5	6,4	6,4	6,5
	Max	69,9	69,9	69,9	69,9	69,9	69,9
Durchschnittliches Alter der Ärzt*innen	Min	49,8	49,8	49,8	49,8	49,8	49,8
	MW	54,0	54,0	53,9	54,0	54,0	53,9
	SD	1,1	1,1	1,2	1,1	1,1	1,2
	Max	58,3	58,3	58,3	58,3	58,3	58,3
Teilnahmeumfang	Min	0,7	0,7	0,7	0,7	0,7	0,7
	MW	0,8	0,8	0,8	0,8	0,8	0,8
	SD	0,1	0,1	0,1	0,1	0,1	0,1
	Max	1,0	1,0	1,0	1,0	1,0	1,0
Vertragsärzt*innen in Prozent	Min	44,4	44,4	44,4	44,4	44,4	44,4
	MW	70,3	70,4	70,3	70,3	70,4	70,2
	SD	5,0	4,9	5,1	4,9	4,9	5,1
	Max	89,9	89,9	89,9	89,9	89,9	89,9
Arztdichte	Min	88,4	88,4	88,4	88,4	88,4	88,4
	MW	195,8	200,2	188,5	195,9	199,0	190,8
	SD	57,4	59,5	52,9	57,7	59,9	53,4
	Max	413,5	413,5	413,5	413,5	413,5	413,5

Anmerkungen: N = Anzahl, Min = Minimum, MW = Mittelwert, SD = Standardabweichung, Max = Maximum.

rigkeit PVS × Versorgungsbereich betrug 23,6 Prozent (Tabelle 3 – „Unteres Segment“; Sensitivitätsanalyse: 25,1 Prozent; Tabelle 4).

## Diskussion

In der vorliegenden Studie wurde überprüft, inwieweit die Usability der verwendeten Praxissoftware mit Leistungsindikatoren einer Arztpraxis statistisch zusammenhängt.

Hierfür wurde die Assoziation zwischen der durchschnittlichen Usability eines PVS mit der Anzahl der Behandlungsfälle sowie mit dem erbrachten Leistungsvolumen pro Arztpraxis ausgewertet. Zudem wurde untersucht, ob sich der Zusammenhang je nach Leistungsstärke der Arztpraxis unterscheidet, also zwischen Praxen im oberen und solchen im unteren Segment der Verteilung der Behandlungsfallzahl beziehungsweise des Leistungsvolumens.

**Tabelle 2**  
Multilevel-Analysen zur Anzahl der Behandlungsfälle.

<b>Gesamtstichprobe</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	2.866,6	2.683,0; 3.047,2	0,000
SUS-Wert <sup>b</sup>	7,9	–4,6; 20,4	0,151
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	695.945,3		
$\sigma^2$	1.508.742,7		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 24.140	<i>N</i> Level II-a 104	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,318	Marginales R <sup>2</sup> 0,004	
<b>Unteres Segment (bis Modalwert)</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	2.333,4	2.209,7; 2.455,2	0,000
SUS-Wert <sup>b</sup>	4,8	–3,6; 13,1	0,178
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	302.190,3		
$\sigma^2$	875.273,6		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 15.020	<i>N</i> Level II-a 104	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,258	Marginales R <sup>2</sup> 0,002	
<b>Oberes Segment (ab Modalwert)</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	4.777,8	4.765,8; 4.788,7	0,000
SUS-Wert <sup>b</sup>	0,3	–0,5; 1,0	0,272
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	452,1		
$\sigma^2$	168.812,7		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 9.120	<i>N</i> Level II-a 81	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,003	Marginales R <sup>2</sup> 0,000	

Anmerkungen: \* p &lt; 0,05; \*\* p &lt; 0,01; \*\*\* p &lt; 0,001.

<sup>a</sup> 95%-Konfidenzintervall.<sup>b</sup> 90%-Konfidenzintervall aufgrund einseitiger Testung.

Über die Gesamtverteilung der beiden Leistungsindikatoren hinweg konnte kein statistisch signifikanter Zusammenhang zur Usability eines PVS nachgewiesen werden. Beim Blick auf die einzelnen Verteilungsgruppen jedoch wurde evident, dass das Leistungsvolumen jener Praxen, welche sich im unteren Segment befanden, signifikant höher war, je höher der durchschnittliche SUS-Wert des verwendeten PVS ausfiel. Bisherige Studien zeigten, dass eine höhere Usability die Aufgabenbearbeitung positiv beeinflusst [20,21] und dass eine geringere Usability mit einer erhöhten Bearbeitungszeit einhergeht [10,22]. Je mehr Zeit für die Dokumentation aufgewendet werden muss, umso mehr empfinden ärztliche Leistungserbringer\*innen Autonomieverlust, Einschränkungen ihrer Work-Life-Balance sowie kognitive Erschöpfung [23]. Zusätzlich zum Zeitverlust selbst könnten PVS mit geringer Usability also zu erhöhten arbeitsbezogenen

**Tabelle 3**  
Multilevel-Analysen zum Leistungsbedarf in Euro

<b>Gesamtstichprobe</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	209.730,9	198.931,0; 220.377,2	0,000
SUS-Wert <sup>b</sup>	505,4	–226,5; 1.237,7	0,131
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	2.348.989.127,1		
$\sigma^2$	7.482.389.405,0		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 24.434	<i>N</i> Level II-a 104	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,242	Marginales R <sup>2</sup> 0,003	
<b>Unteres Segment (bis Modalwert)</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	163.473,7	155.869,2; 170.959,1	0,000
SUS-Wert <sup>b</sup>	530,1	13,9; 1.046,8	0,048
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	1.119.643.398,1		
$\sigma^2$	3.750.197.207,6		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 15.156	<i>N</i> Level II-a 103	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,236	Marginales R <sup>2</sup> 0,007	
<b>Oberes Segment (ab Modalwert)</b>			
FIXE EFFEKTE	Schätzer	90/95%- Konfidenzintervall	p- Wert
Interzept <sup>a</sup>	332.944,7	332.207,2; 333.708,3	0,000
SUS-Wert <sup>b</sup>	28,6	–25,2; 81,9	0,188
<b>ZUFÄLLIGE EFFEKTE</b>			
	Varianz		
$\tau_{00}$ Level II-a	818.586,8		
$\sigma^2$	1.059.710.979,0		
<b>MODELLSTATISTIKEN</b>			
N	<i>N</i> Level I 9.278	<i>N</i> Level II-a 87	
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,001	Marginales R <sup>2</sup> 0,000	

Anmerkungen: \* p &lt; 0,05; \*\* p &lt; 0,01; \*\*\* p &lt; 0,001.

<sup>a</sup> 95%-Konfidenzintervall.<sup>b</sup> 90%-Konfidenzintervall aufgrund einseitiger Testung.

nen Beanspruchungen beitragen und damit potenziell die Leistungsfähigkeit beeinträchtigen.

Dass sich im vorliegenden Studiendesign ein positiver Zusammenhang der Usability mit dem Leistungsumfang zunächst nur bei Praxen mit geringerem Leistungsumfang zeigte, könnte auf einen Deckeneffekt hinweisen. Insbesondere in Ein-Personen-Praxen mit bereits hohen Leistungszahlen ist es denkbar, dass diese ihre Kapazitätsgrenzen weitgehend ausgeschöpft haben. Es zeigt sich, dass in Praxen mit höherem Durchlauf weniger Zeit pro Patient\_in aufgewendet werden kann [24], auch wenn die Behandlungsziele dennoch erfüllt werden [25]. Eine weitere Effizienzsteigerung durch ein nutzerfreundliches PVS wäre dort möglicherweise weniger realisierbar. In Praxen mit moderatem Leistungsvolumen hingegen besteht vermutlich noch Optimierungsspielraum, der durch leistungsfähige und intuitiv bedienbare Praxissoftware besser genutzt werden kann.

**Tabelle 4**

Sensitivitäts-Analyse zum Leistungsbedarf in Euro im unteren Segment (bis Modalwert).

FIXE EFFEKTE	Schätzer	90/95%-Konfidenzintervall	p-Wert
Interzept <sup>a</sup>	165.558,8		0,000
SUS-Wert <sup>b</sup>	518,1	9,6; 1.027,3	0,049
KONTROLLVARIABLEN LEVEL I (PRAXIS)			
Anteil weiblicher Patientinnen <sup>a</sup>	157,9	92,9222,8	0,000
Ø Alter der Patient*innen <sup>a</sup>	-29,5	-154,196,9	0,644
Ø Charlson-Comorbidity-Index der Pat. <sup>a</sup>	2.354,3	-497,65.175,3	0,103
KONTROLLVARIABLEN LEVEL II-B (KREISE)			
GISD <sup>a</sup>	26.024,6	16.702,7; 35.351,2	0,000
Anteil weiblicher Ärztinnen <sup>a</sup>	796,8	526,2; 1.067,9	0,822
Ø Alter der Ärzt*innen <sup>a</sup>	166,1	-1.274,6; 1.601,8	0,197
Ø Teilnahmeumfang <sup>a</sup>	20.368,2	-10.279,3; 51.113,4	0,146
Anteil Vertragsärzt*innen <sup>a</sup>	195,7	-65,6; 457,3	0,110
Arztdichte <sup>a</sup>	-26,4	-58,6; 5,7	0,112
ZUFÄLLIGE EFFEKTE			
	Varianz		
$\tau_{00}$ Level II-a	1.083.289.294,2		
$\tau_{00}$ Level II-b	60.001.307,8		
$\sigma^2$	3.657.232.339,7		
MODELLSTATISTIKEN			
N	$N_{\text{Level I}}$ 15.156	$N_{\text{Level II-a}}$ 103	$N_{\text{Level II-b}}$ 400
R <sup>2</sup>	Konditionales R <sup>2</sup> 0,251	Marginales R <sup>2</sup> 0,017	

Anmerkungen: \* p &lt; 0,05; \*\* p &lt; 0,01; \*\*\* p &lt; 0,001.

<sup>a</sup> 95%-Konfidenzintervall.<sup>b</sup> 90%-Konfidenzintervall aufgrund einseitiger Testung.

Kein signifikanter Zusammenhang zeigte sich bei der Anzahl der Behandlungsfälle, während das Leistungsvolumen mit Unterschieden in der PVS- Usability statistisch zusammenhing. Dies deutet darauf hin, dass der Leistungsbedarf in Euro ein differenzierterer Indikator für die tatsächliche ärztliche Arbeitslast sein könnte, da im Gegensatz zum Leistungsvolumen die Fallzahl sowohl kurze als auch zeitintensive Behandlungen umfassen kann. In der internationalen Forschung werden deshalb häufig „work relative value units“ (wRVU) verwendet, die Zeitaufwand, Fähigkeiten und Ressourcen einer Leistung messen und in den USA auch zur Ermittlung der Vergütungshöhe dienen. Es wurde gezeigt, dass wRVU ein angemessenerer Indikator für die Erfassung der Leistung sind als die bloße Zahl der Patient\*innen oder Kontakte [26]. Damit könnte der Leistungsbedarf einem valideren Maß für den tatsächlichen Arbeitsaufwand im Vergleich zur reinen Behandlungsfallzahl entsprechen.

In der Sensitivitätsanalyse zeigten sich weitere signifikante Assoziationen: Eine höhere Deprivation war mit einer größeren Morbiditätslast und einer höheren Inanspruchnahme von Gesundheitsleistungen verknüpft, was konsistent zur Literatur ist [27]. Für das höhere Leistungsvolumen bei einem größeren Anteil weiblicher Patientinnen in der Praxis finden sich hingegen keine eindeutigen Befunde in der bisherigen Literatur. Zwar gibt es Hinweise auf tendenziell höhere Gesundheitskosten bei Frauen, dieser Effekt ist jedoch variabel und nicht robust [28].

Das vergleichsweise geringe marginale R-Quadrat des fixen Effekts weist darauf hin, dass die Usability des PVS zwar einen statistisch nachweisbaren, jedoch insgesamt geringen Anteil der Varianz im Leistungsindikator statistisch erklärt. Dies legt nahe, dass die Leistungsfähigkeit einer Arztpraxis maßgeblich auch mit anderen Faktoren assoziiert ist, etwa mit der Fachrichtung, der Praxisorganisation, mit personellen Ressourcen oder regionalen Versorgungsbedingungen [29]. In komplexen Verhaltens- und Sozialdaten sind niedrige R-Quadrate jedoch nicht unüblich; ihre Ergebnisse können dennoch valide und relevante Erkenntnisse liefern [30].

Zusammengefasst zeigen die vorliegenden Ergebnisse, dass eine höhere Usability von Praxissoftware grundsätzlich das Potenzial haben könnte, die Effizienz ärztlicher Leistungserbringung zu steigern, insbesondere in Praxen mit noch nicht ausgeschöpften Kapazitäten. Ein kausaler Zusammenhang ist aufgrund des querschnittlichen Designs nicht ableitbar, da Ursache- und Wirkungsrichtung nicht eindeutig bestimmt werden können.

### Limitationen

Die Analysen unterliegen verschiedenen Limitationen. So war die Studienpopulation eingeschränkt durch die Konzentration auf Praxen mit nur einem\*r Leistungserbringer\*in. Zudem lag die Information zur Usability nur auf Ebene der PVS vor. Unterschiedliche Leistungserbringer\*innen bewerten das selbe PVS jedoch durchaus verschieden; weiterhin hängt die Bewertung möglicherweise auch von genutzten oder nicht genutzten Zusatzfunktionen ab. Auch wenn der durchschnittliche SUS-Wert pro PVS reliabel und valide gemessen wurde, konnten interindividuelle Unterschiede hier nicht berücksichtigt werden. Zudem können zusätzliche Faktoren auf regionaler Ebene das Geschehen beeinflussen, wie beispielsweise regional unterschiedliche Zulassungsregularien, die die Fluktuation und damit den Neuerwerb von PVS-Lizenzen betreffen können.

### Schlussfolgerungen

Die vorliegende explorative Studie legt nahe, dass das eingesetzte PVS die Leistungsfähigkeit einer Arztpraxis beeinflussen könnte. Um die Zusammenhänge sowie relevante Kontextfaktoren differenzierter verstehen zu können, sind weiterführende Untersuchungen mit größeren und repräsentativeren Stichproben sowie längsschnittlichen Designs erforderlich. Dennoch kann diese Studie als Impuls für Entwickler\*innen und Hersteller\*innen von PVS dienen, die Usability ihrer Systeme gezielt weiterzuentwickeln und sie

noch besser an die praktischen Anforderungen im Versorgungsalltag anzupassen.

### Interessenkonflikt

Die Autor\*innen erklären, dass kein Interessenkonflikt besteht.

### Autor\*innenschaft

Doreen Müller: Konzeption, Methodik, formale Analyse, Datenaufbereitung, Schreiben – ursprünglicher Entwurf. Michael Erhart: Methodik. Jakob Holstiege: Datenaufbereitung, Schreiben – Überarbeitung und Korrektur. Dominik Graf von Stillfried: Konzeption, Betreuung, Schreiben – Überarbeitung und Korrektur.

### Literatur

- [1] KBV - Zahlen und Fakten n.d. <https://www.kbv.de/infothek/zahlen-und-fakten>. [Accessed 31 July 2025].
- [2] Lipovsek J, Schulz M, Hering R, Czihal T, Kroll LE. Bedarfsprojektion für Medizinstudienplätze in Deutschland – Aktualisierung 2024 2024. <https://doi.org/10.20364/FB3-24.01>.
- [3] Licht und Schatten: So sehen die Versicherten das deutsche Gesundheitswesen n. d. [https://www.gkv-90prozent.de/ausgabe/40/autorenbeitrag/40\\_versichertenbefragung/40\\_versichertenbefragung.html](https://www.gkv-90prozent.de/ausgabe/40/autorenbeitrag/40_versichertenbefragung/40_versichertenbefragung.html). [Accessed 31 July 2025].
- [4] Zi-Praxis-Panel - Jahresbericht 2023. Wirtschaftliche Situation und Rahmenbedingungen in der vertragsärztlichen Versorgung der Jahre 2019 bis 2022. Berlin: Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland (Zi); 2024.
- [5] Bevan N, Carter J, Harker S. ISO 9241-11 Revised: what Have We Learnt About Usability Since 1998? *Hum-Comput Interact Des Eval* 2015.
- [6] Ratwani RM, Savage E, Will A, Arnold R, Khairat S, Miller K, et al. A usability and safety analysis of electronic health records: a multi-center study. *J Am Med Inform Assoc JAMIA* 2018;25:1197–201. <https://doi.org/10.1093/jamia/ocy088>.
- [7] Longo L. Experienced mental workload, perception of usability, their interaction and impact on task performance. *PLoS ONE* 2018;13:e0199661. <https://doi.org/10.1371/journal.pone.0199661>.
- [8] Müller D, Nieporte T, Graf Von Stillfried D. Praxisverwaltungssysteme: Usability, Fehlersituationen und Perspektiven von PVS-Wechslern 2025. <https://doi.org/10.20364/zipaper-2025.30>.
- [9] Müller D, Nieporte T, Graf von Stillfried D. Praxisverwaltungssysteme: Deutschlandweite Ergebnisse zu Usability, Nutzerzufriedenheit und Wechselbereitschaft aus 10.245 Bewertungen. *GMS Med Inform Biom Epidemiol* 2024;20:Doc13. <https://doi.org/10.3205/mibe000269>.
- [10] Olakotan O, Samuriwo R, Ismaila H, Atiku S. Usability challenges in electronic health records: impact on documentation burden and clinical workflow: a scoping review. *J Eval Clin Pract* 2025;31:e70189. <https://doi.org/10.1111/jep.70189>.
- [11] Brooke JSUS. A quick and dirty usability scale. *Usability Eval Ind* 1995;189.
- [12] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8).
- [13] Michalski N, Reis M, Tetzlaff F, Herber M, Kroll LE, Hövener C, et al. In: German Index of Socioeconomic Deprivation (GISD): Revision Aktualisierung und Anwendungsbeispiele. <https://doi.org/10.25646/10640>.
- [14] KBV - Gesundheitsdaten - Zahlen, Trends und Analysen n.d. <https://www.kbv.de/infothek/zahlen-und-fakten/gesundheitsdaten>. [Accessed 31 July 2025].
- [15] Kopetsch T, Schmitz H. Regional variation in the utilisation of ambulatory services in Germany. *Health Econ* 2014;23:1481–92. <https://doi.org/10.1002/hec.3001>.
- [16] Ludbrook J. Should we use one-sided or two-sided P values in tests of significance? *Clin Exp Pharmacol Physiol* 2013;40:357–61. <https://doi.org/10.1111/1440-1681.12086>.
- [17] Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43.
- [18] Fielding A, Goldstein H. *Cross-classified and Multiple Membership Structures in Multilevel Models*, 2006.
- [19] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344–9. <https://doi.org/10.1016/j.jclinepi.2007.11.008>.
- [20] Kortum P, Peres SC. The relationship between system effectiveness and subjective usability scores using the system usability scale. *Int J Human-Computer Interact* 2014;30:575–84. <https://doi.org/10.1080/10447318.2014.904177>.
- [21] Ferreira JM, Rodríguez F, Santos A, Acuña ST, Juristo N. Impact of usability mechanisms: a family of experiments on efficiency, effectiveness and user satisfaction. *IEEE Trans Softw Eng* 2023;49:251–67. <https://doi.org/10.1109/TSE.2022.3149586>.
- [22] Dunn Lopez K, Chin C-L, Leitão Azevedo RF, Kaushik V, Roy B, Schuh W, et al. Electronic health record usability and workload changes over time for provider and nursing staff following transition to new EHR. *Appl Ergon* 2021;93:103359. <https://doi.org/10.1016/j.apergo.2021.103359>.
- [23] Kruse CS, Mileski M, Dray G, Johnson Z, Shaw C, Shirodkar H. Physician burnout and the electronic health record leading up to and during the first year of COVID-19: systematic review. *J Med Internet Res* 2022;24:e36200. <https://doi.org/10.2196/36200>.
- [24] Campbell JL, Ramsay J, Green J. Practice size: impact on consultation length, workload, and patient assessment of care. *Br J Gen Pract J R Coll Gen Pract* 2001;51:644–50.
- [25] Vamos EP, Pape UJ, Bottle A, Hamilton FL, Curcin V, Ng A, et al. Association of practice size and pay-for-performance incentives with the quality of diabetes management in primary care. *CMAJ Can Med Assoc J J Assoc Medicales Can* 2011;183:E809–16. <https://doi.org/10.1503/cmaj.101187>.
- [26] Glass KP, Anderson JR. Relative value units and productivity: Part 2 of 4. *J Med Pract Manag MPM* 2002;17:285–90.
- [27] Jansen L, Schwettmann L, Behr C, Eberle A, Holleczeck B, Justenhoven C, et al. Trends in cancer incidence by socioeconomic deprivation in Germany in 2007 to 2018: An ecological registry-based study. *Int J Cancer* 2023;153:1784–96. <https://doi.org/10.1002/ijc.34662>.
- [28] Colombo D, Simoni L, Zagni E. Gender medicine and pharmacoeconomics: a narrative review of the international literature of the last 5 years. a revision of evidences about the relationship between gender and economic consumption in health. *Curr Pharm Des* 2017;23:3713–22. <https://doi.org/10.2174/1381612823666170510142933>.
- [29] de Graaf-Ruizendaal WA, van der Hoek L, de Bakker DH. Are low and high utilization related to the way GPs manage their practices? An observational study. *BMC Fam Pract* 2018;19:46. <https://doi.org/10.1186/s12875-018-0732-7>.
- [30] Funder DC, Ozer DJ. Evaluating effect size in psychological research: sense and nonsense. *Adv Methods Pract Psychol Sci* 2019;2:156–68. <https://doi.org/10.1177/2515245919847202>.