Evidenzbericht des Forschungsprojekts „Arbeitsmedizin digital – eine Informations- und Evidenzplattform für die betriebsärztliche Praxis“

Präventionsdimension: Tertiärprävention, digitale Anwendungen bei „Return to Work“

1. Hintergrund

Langzeiterkrankungen prägen das Arbeitsunfähigkeitsgeschehen in Deutschland in besonderem Maße. Rund 40 % aller Fehltage entfallen auf Arbeitsunfähigkeitsfälle mit einer Dauer von mehr als sechs Wochen (AOK/Fehlzeiten-Report 2024). Diese Fälle sind zahlenmäßig eine Minderheit, bestimmen aber den betrieblichen und volkswirtschaftlichen Folgeaufwand weit überwiegend. Die durch Arbeitsunfähigkeit insgesamt verursachten Produktionsausfälle beziffert die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin für Deutschland auf rund 134 Milliarden Euro jährlich (BAuA 2024).

Digitale Gesundheitsanwendungen haben in der Regelversorgung deutlich an Bedeutung gewonnen. Seit ihrer Einführung im Jahr 2020 wurden etwa eine Million Nutzungen von DiGA verzeichnet, mit einem Verordnungsanstieg von rund 80 % innerhalb eines Jahres (DiGA-Report 2024, SVDGV). Damit sind digitale Interventionen im Alltag der Versorgung angekommen, noch bevor ihre Wirksamkeit für arbeitsbezogene Endpunkte systematisch getestet wurde.

Genau hier liegt die Forschungslücke. Bisherige Übersichtsarbeiten zu digitalen Interventionen haben vor allem klinische Endpunkte betrachtet (Depressivität, Angstsymptomatik, Schmerzintensität, Funktionsfähigkeit). Der Return to Work (RTW) als primärer Endpunkt wurde kaum systematisch untersucht. Für die arbeitsmedizinische Praxis ist diese Unterscheidung relevant. Eine Besserung der Symptomatik bedeutet nicht zwangsläufig, dass die Arbeitsfähigkeit wiederhergestellt ist. Für die Rückkehr an den Arbeitsplatz spielen darüber hinaus betriebliche und psychologische Faktoren eine Rolle, die sich mit klinischen Skalen allein unzureichend abbilden lassen.

2. Methodik

Die vorliegende Übersichtsarbeit geht daher der Frage nach, wie wirksam digitale Interventionen zur Unterstützung des Return to Work nach einer Langzeiterkrankung sind. Die Übersichtsarbeit wurde gemäß den PRISMA-2020-Leitlinien durchgeführt und vorab im internationalen Register PROSPERO registriert (CRD42024618831). Die Forschungsfrage wurde gemäß dem PICO-Schema formuliert. Als **Population (P)** wurden erwerbstätige Personen mit klinisch diagnostizierter Langzeiterkrankung berücksichtigt, die sich im Prozess der beruflichen Wiedereingliederung befanden. Die **Intervention (I)** umfasste digitale und telemedizinische Interventionen, darunter eHealth-, Telehealth- und Remote-Care-Angebote sowie digitale Gesundheitsanwendungen und Online-Programme. Als **Komparator (C)** wurden Usual Care, konventionelle Rehabilitationsmaßnahmen oder keine Intervention zugelassen. Als **Outcomes (O)** wurden arbeitsbezogene Endpunkte definiert, insbesondere die Rückkehr zur Arbeit (Return to Work, RTW), die Zeit bis zum RTW, die Anzahl der Arbeitsunfähigkeitstage sowie die Arbeitsfähigkeit.

Am 25. November 2024 wurde die Datenbanksuche in PubMed (1.144 Treffer), Scopus (1.052 Treffer) und Web of Science (555 Treffer) durchgeführt, insgesamt fanden wir 2.751 Ergebnisse. Es wurden Publikationen aus dem Zeitraum von Februar 1995 bis November 2024 in deutscher und englischer Sprache berücksichtigt.

Nach der Entfernung von 825 Duplikaten verblieben 1.926 Treffer für das Titelscreening, bei dem 1.686 Arbeiten ausgeschlossen wurden. Im anschließenden Abstract-Screening entfielen weitere 214 Treffer, sodass insgesamt 26 Volltexte vollständig geprüft wurden. Eingeschlossen wurden schließlich 9 Publikationen, die 6 Primärstudien repräsentieren. Screening, Datenextraktion und Bias-Assessment erfolgten durch zwei unabhängige Reviewer.

Das Bias-Assessment wurde instrumentenspezifisch bewertet: randomisierte Studien mit dem Cochrane-Instrument RoB 2, nicht-randomisierte Studien mit ROBINS-I.

Aufgrund der ausgeprägten Heterogenität hinsichtlich Indikation, Studiendesign, Interventionsform und Endpunktdefinition wurde auf eine Metaanalyse verzichtet.

3. Ergebnisse

3.1 Charakteristika der eingeschlossenen Studien

Die sechs Primärstudien umfassen 1.107 Teilnehmer aus fünf Ländern (Australien, Schweden, Spanien, Niederlande, USA) und decken drei Studientypen ab: eine cluster-randomisierte kontrollierte Studie, Kohortenstudien sowie Pilot- beziehungsweise Machbarkeitsstudien. Sie befassen sich inhaltlich mit drei gesundheitlichen Bereichen: psychische Erkrankungen, chronische Schmerzen und muskuloskelettale Erkrankungen.

Studie	Land	Design	n	Indikation	Intervention	Bias-Risiko
ECO-Projekt	NL	Cluster-RCT	220	Psychische Erkrankungen	Blended Care (Online-KVT/PST + ärztliches Decision-Aid)	Some concerns (RoB 2)
Workable	AU	Pilotstudie	13	Psychische Erkrankungen	KVT + wöchentliches Video-Coaching	Serious (ROBINS-I)
LeapForward	AU	Retrospektive Kohorte	78	Gemischt (körperlich)	KI-gestützter Lifestyle-Coach + Telehealth	Serious (ROBINS-I)
SWEPPE	SE	Machbarkeits-studie	20	Chronischer Schmerz	Smartphone-App + Arbeitgeber-Web-App	Serious (ROBINS-I)
Whiplash -Studie	ES	Kohortenstudie + Prä-Post-Vergleich	387 + 200	HWS-Schleudertrauma	Hybrid-Telerehabilitation	Serious (ROBINS-I)
SWORD Health	USA	Prospektive Kohorte	189	Hand-/Handgelenk-schmerz	Volldigitales Programm mit Biofeedback	Serious (ROBINS-I)

3.2 Digitale Interventionsformen

Die eingeschlossenen Interventionen lassen sich sechs Kategorien zuordnen, die sich im Grad der menschlichen Begleitung deutlich unterscheiden.

- **Blended eHealth:** Online-Module auf Basis kognitiver Verhaltenstherapie und Problemlösetherapie, kombiniert mit ärztlichem Decision-Aid und psychiatrischer Konsultation (ECO/Return@Work, Niederlande)
- **Therapeutengestütztes Programm:** wöchentliche Online-Module ergänzt um 45-minütige Video-Coaching-Calls mit Psychologen (Workable, Australien)
- **Smartphone-basierte App:** Selbstmonitoring, Zielplanung, Selbstbewertung und strukturierte Arbeitgeber-Kommunikation (SWEPPE, Schweden)

- **Hybrid-Telerehabilitation:** fünf Präsenzsitzungen kombiniert mit zehn Heimsitzungen über eine audiovisuelle Online-Plattform (Whiplash-Studie, Spanien)
- **Volldigitales Programm:** IMU-Biofeedback, asynchrones physiotherapeutisches Monitoring und digitale kognitive Verhaltenstherapie (KVT)-/ Akzeptanz- sowie Commitment-Therapie (ACT)-Module ohne Präsenztermine (SWORD Health, USA)
- **KI-gestütztes Lifestyle-Programm:** regelbasierter virtueller Coach mit täglichen Check-ins, ergänzt durch Telehealth-Calls (LeapForward, Australien)

Die Vielfalt der identifizierten Ansätze verdeutlicht, dass es kein einheitliches Standardmodell für digitale RTW-Interventionen gibt. Stattdessen hängt die Auswahl des spezifischen Ansatzes vom jeweiligen Krankheitsbild, den verfügbaren Ressourcen und den individuellen Merkmalen der Patienten ab.

3.3 Psychische Erkrankungen

Das ECO-Projekt aus den Niederlanden ist die methodisch stärkste Studie dieser Übersichtsarbeit, eine cluster-randomisierte kontrollierte Studie mit 220 Teilnehmenden. Die Intervention Return@Work kombiniert KVT- und Problem-Solving Therapy (PST)-basierte Online-Module für Beschäftigte mit einem automatisierten Decision-Aid für Betriebsärzte und bedarfsweiser psychiatrischer Konsultation.

Der erste Return to Work erfolgte median nach 50 Tagen, verglichen mit 77 Tagen in der Kontrollgruppe (HR 1,39; 95 %-KI 1,034–1,870; p = 0,03). Die Remissionsrate nach 9 Monaten lag bei 56 % gegenüber 37 % (OR 2,23; p = 0,02). Gesundheitsökonomisch zeigte sich ein deutlicher Return on Investment sowohl aus Arbeitgeber- als auch aus gesellschaftlicher Perspektive, mit einem Break-even bereits nach 6 Monaten. Die Prozessevaluation belegt eine hohe Implementierungsreichweite: 100 % der Betriebsärzte und 76,3 % der Arbeitnehmer wurden erreicht.

Die Einschränkungen sind ebenso deutlich. Beim dauerhaften vollständigen RTW zeigte sich kein signifikanter Unterschied (p = 0,15), ebenso wenig bei den gesamten Krankheitstagen (174 vs. 228 Tage; p = 0,10). Der Remissionsvorteil war nach 12 Monaten nicht mehr signifikant. Nur 29 % der Teilnehmenden besprachen das Tool tatsächlich mit dem Betriebsarzt, die Interventionsgruppe war mit der betriebsärztlichen Betreuung signifikant weniger zufrieden als die Kontrollgruppe. Der letztgenannte Befund verdient arbeitsmedizinische Aufmerksamkeit, Blended-Care-Ansätze können die persönliche Arzt-Patient-Beziehung offenbar belasten. Nach RoB 2 wurde die Studie mit „Some concerns" bewertet.

Die australische Pilotstudie Workable untersuchte über 6 Wochen KVT-Module in Kombination mit wöchentlichen Video-Coaching-Calls an 13 Teilnehmern. Die Adhärenz war mit 92,3 % sehr hoch, 60 % hatten bei Programmende den RTW vollzogen, die Fehltage sanken von 25,5 auf 11,2 Tage (p = 0,001; g = 1,69). Für Depressivität (PHQ-9: g = 1,16), Angstsymptomatik (GAD-7: g = 1,10) und Arbeitsfähigkeit (g = 0,73) ergaben sich große Effektstärken. Angesichts des Einzelgruppendesigns, der sehr kleinen Stichprobe und ausschließlich selbstberichteter Endpunkte wurde die Studie mit ROBINS-I als „Serious" eingestuft. Effektstärken dieser Größenordnung sind ohne Kontrollgruppe nicht kausal interpretierbar.

LeapForward, eine retrospektive Kohortenstudie mit 78 Teilnehmenden, setzte über 6 Wochen hinweg einen KI-gestützten Coach („Lucy") mit täglichen Check-ins und ergänzenden Telehealth-Calls ein. 72 % schlossen das Programm ab. Der Anteil erwerbstätiger Teilnehmender stieg von 2,6 % auf 25,6 %. Die psychische Belastung (K10) war signifikant reduziert (p ≤ 0,001; r = 0,47), ebenso verbesserten sich Depressivität und Wohlbefinden signifikant. Ein bemerkenswerter Negativbefund ist, dass die spezifische Angst vor der Rückkehr an den Arbeitsplatz unverändert blieb (p = 0,631). Dies deutet darauf hin, dass 6 Wochen möglicherweise nicht ausreichen, um tiefgreifende arbeitsbezogene Barrieren zu überwinden. Aufgrund des retrospektiven Designs ohne Kontrollarm wurde auch diese

Studie als „Serious" bewertet. Die Veränderung der Erwerbsquote lässt sich ohne Vergleichsgruppe nicht von natürlicher Erholung und Selektionseffekten trennen.

3.4 Chronischer Schmerz

Das schwedische SWEPPE-Projekt mit 20 Teilnehmenden untersuchte eine Smartphone-App mit 6 Modulen für Patientinnen und Patienten nach abgeschlossener Schmerzrehabilitation, ergänzt um eine Web-Anwendung für Arbeitgeber.

Alle Teilnehmenden nutzten das Selbstmonitoring mindestens einmalig, im Median an 26 Tagen mit täglichem Selbst-Rating. 69 % formulierten konkrete arbeitsbezogene Ziele, 56 % teilten Ziele und Barrieren aktiv mit ihren Arbeitgebern. Insgesamt wurden 51 konkrete Unterstützungsmaßnahmen am Arbeitsplatz identifiziert: Aufgabenanpassung (60 %), Haltungskorrektur (60 %), Pausen und Kurzruhen (53 %) sowie ergonomische Anpassungen (33 %).

Der Hauptvorteil liegt in der klaren Kommunikation zwischen Beschäftigten und Arbeitgebern. Unsichtbare Barrieren am Arbeitsplatz werden identifizierbar, ohne dass die betriebsärztliche Vermittlung jeden einzelnen Fall begleiten muss. Als Machbarkeitsstudie ohne Kontrollgruppe und mit n=20 bietet SWEPPE jedoch keine Aussagekraft zur Wirksamkeit; die ROBINS-I-Bewertung ist „Serious". Bei hoher Schmerzintensität zeigt sich zudem eine praktische Limitation: Die Adhärenz sinkt, da die täglichen Eingaben als zusätzliche Belastung wahrgenommen werden.

3.5 Muskuloskelettale Erkrankungen

Die spanische Whiplash-Studie besteht aus einer Kohortenstudie mit 387 Teilnehmenden und einem Prä-post-Vergleich mit 200 Teilnehmenden. Untersucht wurden Beschäftigte mit HWS-Schleudertrauma, die ein Hybridprotokoll mit 5 Präsenz- und 10 Heimübungen über eine Online-Plattform durchführten. Im vorgeschalteten Prä-post-Vergleich sank die Zahl der Präsenztermine von 13 auf 5 Sitzungen, in der Kohortenstudie von 9 auf 7 Sitzungen, während die Dauer der Arbeitsunfähigkeit dort mit 31 gegenüber 32 Tagen nahezu gleich blieb. Die Autoren werten dies als Nichtunterlegenheit (Non-Inferiority), also vergleichbare Wirksamkeit bei reduziertem Aufwand. Bei einer Rücklaufquote von 79 % berichteten alle Befragten von Verbesserungen, als Hauptvorteil wurde der Wegfall der Anfahrtszeiten zur Klinik genannt. ROBINS-I bewertet die Studie als „Serious", hauptsächlich aufgrund möglicher Confounding-Effekte. Da die Zuteilung zur Telerehabilitation nicht randomisiert war, ist nicht auszuschließen, dass die Gruppen sich systematisch unterscheiden. Daher sind kausale Schlussfolgerungen zur Wirksamkeit nur eingeschränkt möglich.

SWORD Health aus den USA ist ein achtwöchiges, vollständig remote durchgeführtes Programm, das sensorgestützte Bewegungsrückmeldung (IMU-Biofeedback), asynchrones physiotherapeutisches Monitoring sowie digitale Module zu Edukation, kognitiver Verhaltenstherapie und Akzeptanz- sowie Commitment-Therapie umfasst. In einer einarmigen, prospektiven Kohorte mit 189 Teilnehmern, die unter Hand- und Handgelenkschmerzen litten, sank die Schmerzintensität (NPRS) um 51,3 % (mittlere Differenz 2,26 Punkte; 95 %-KI 1,73-2,78), die Funktionseinschränkung (QuickDASH) um 52,1 % und die Produktivitätsverluste (WPAI) um 68,2 %. Bei Teilnehmern mit anfänglicher Operationsabsicht reduzierte sich diese um 76,1 %, während der Anteil mit Analgetikaeinnahme von 22,5 % auf 7,1 % sank. Die Abschlussquote lag bei 78,8 %, die durchschnittliche Zufriedenheit betrug 8,5 von 10 Punkten. Der Studie fehlt ein Kontrollarm, sämtliche Endpunkte beruhen auf Selbstauskunft. ROBINS-I bewertet die Studie als „Serious".

3.6 Gesamtschau und Evidenzqualität

Über alle 6 Primärstudien hinweg weist die Effektrichtung in dieselbe Richtung. Im ECO-Projekt erfolgte der erste Return to Work nach 50 statt nach 77 Tagen, in der Whiplash-Studie blieb die Arbeitsunfähigkeitsdauer bei deutlich reduziertem Präsenzaufwand mit 31 gegenüber 32 Tagen vergleichbar. Workable berichtet eine RTW-Rate von 60 % bei Programmende, LeapForward einen Anstieg der Erwerbstätigkeit von 2,6 % auf 25,6 % innerhalb von 6 Wochen. SWORD Health verzeichnete eine Reduktion der Produktivitätsverluste um 68,2 % über 8 Wochen, wobei diese Verbesserung fast ausschließlich auf Präsentismus beruht, da nur wenige Teilnehmende zu Studienbeginn Fehlzeiten aufwiesen. Wo psychische Endpunkte erhoben wurden, verbesserten sich Depressivität und Angst durchgängig signifikant, teils mit großen Effektstärken. Gesundheitsökonomische Daten liegen bislang ausschließlich aus dem ECO-Projekt vor, das einen deutlichen Return on Investment mit einem Break-even nach 6 Monaten aufweist.

Diese Übereinstimmung darf nicht als gegenseitige Bestätigung gelesen werden. 5 der 6 Primärstudien wurden nach ROBINS-I mit einem schwerwiegenden Verzerrungsrisiko bewertet, allein das ECO-Projekt erreichte nach RoB 2 die Einstufung „Some concerns". Die Gründe wiederholen sich über die Studien hinweg: fehlende Kontrollarme, kleine Stichproben, ausschließlich selbstberichtete Endpunkte und teilweise retrospektive Datenerhebung. Diese Limitationen sind damit nicht zufällig verteilt, sondern strukturell gleichgerichtet. Studien ohne Kontrollgruppe können natürliche Verbesserungen, Regression zur Mitte und Selektionseffekte nicht vom tatsächlichen Interventionseffekt unterscheiden. Zudem neigen unverblindete Selbstberichte dazu, systematisch positive Ergebnisse zu berichten. Dass ähnlich verzerrte Designs ähnliche Resultate zeigen, ist vorhersehbar. Die Übereinstimmung in der Richtung der Effekte deutet eher auf ein plausibles Wirkprinzip hin, anstatt auf eine gesicherte Wirksamkeit. Kausale Schlüsse erlaubt derzeit nur das ECO-Projekt.

Belastbarer sind die Befunde zu Akzeptanz und Umsetzbarkeit, da sie weniger vom fehlenden Kontrollarm abhängen. Die Abschluss- und Adhärenzraten liegen zwischen 78,8 % und 92,3 %, die Zufriedenheit ist durchweg hoch und die Programme wurden erfolgreich in 5 sehr unterschiedlichen Gesundheitssystemen eingeführt. Dies belegt die gute Machbarkeit und Akzeptanz digitaler RTW-Interventionen, während ihre Wirksamkeit auf arbeitsbezogene Endpunkte noch vorläufig ist.

4. Ausblick

4.1 Einordnung für die arbeitsmedizinische Praxis

Die Befunde legen 4 Schlüsse für die betriebsärztliche Praxis nahe. Zum Ersten zeigen Blended-Care-Ansätze durchweg stärkere Effekte im Vergleich zu reinen digitalen Selbsthilfeprogrammen: Die Interventionen mit den eindeutigsten RTW-Ergebnissen – ECO und Workable – verbinden digitale Inhalte mit menschlicher Betreuung, während bei SWEPPE nur 13-18 % der Teilnehmer die Coach-Funktion nutzen. Menschliche Begleitung bleibt anhand der aktuellen Datenlage ein entscheidender Wirkfaktor.

Zweitens basieren die effektivsten Ergebnisse auf KVT und Problemlösetherapie – bewährte Ansätze, die sich gut digitalisieren lassen. Das bedeutet für die Praxis, dass bekannte Methoden in digitale RTW-Programme integriert werden können, anstatt komplett neue Konzepte zu entwickeln.

Drittens zeigt die Whiplash-Studie, dass Telerehabilitation die Anzahl der erforderlichen Präsenztermine deutlich reduzieren kann, ohne die Dauer der Arbeitsunfähigkeit zu verlängern. Angesichts begrenzter Versorgungskapazitäten und des Fachkräftemangels ist das ein praktisches Argument.

Viertens ist die Kosteneffektivität das belastbarste Argument gegenüber Kostenträgern und Arbeitgebern. Das ECO-Projekt weist einen deutlichen Return on Investment sowohl aus Arbeitgeber- als auch aus gesellschaftlicher Perspektive auf, mit einem Break-Even nach 6 Monaten.

4.2 Grenzen dieser Übersichtsarbeit

Die Aussagekraft dieser Übersichtsarbeit ist in mehreren Punkten eingeschränkt. Am schwerwiegendsten wiegt die Qualität der verfügbaren Evidenz. 5 der 6 Primärstudien weisen ein erhebliches Verzerrungsrisiko auf, was die Ziehung kausaler Schlüsse erschwert. Eine natürliche Erholung könnte die beobachteten Verbesserungen ebenfalls erklären. Außerdem ist die statistische Power zweier Studien mit nur 13 (Workable) und 20 Teilnehmern (SWEPPE) gering, was die Übertragbarkeit der Ergebnisse einschränkt. Die Nachbeobachtungszeiten von 6 Wochen bis 12 Monaten sind zudem teilweise zu kurz, um die Nachhaltigkeit der Effekte zu beurteilen.

2 weitere Einschränkungen betreffen die Übersichtsarbeit selbst. Berücksichtigt wurden ausschließlich deutsch- und englischsprachige Publikationen, womit Studien aus anderen Sprachräumen unberücksichtigt blieben. Und da positive Ergebnisse häufiger publiziert werden als Nullbefunde, ist in einem so jungen Forschungsfeld mit einem Publikationsbias zu rechnen, sodass Studien ohne nachweisbaren Effekt unterrepräsentiert sein dürften.

4.3 Forschungsbedarf und Fazit

Der Befund einer breiten, aber methodisch eingeschränkten Evidenzbasis ist typisch für ein junges Forschungsfeld. Die Untersuchung von Return to Work als primärem Endpunkt digitaler Interventionen ist erst seit wenigen Jahren im Fokus, die vorliegenden Pilot- sowie Machbarkeitsstudien erfüllen ihre Rolle. Sie demonstrieren Akzeptanz, Adhärenz und praktische Umsetzbarkeit, was die Grundlage für größere, bestätigende Studien bildet. Eine randomisiert-kontrollierte Studie zum SWEPPE-Programm ist bereits im Gange.

Künftige Studien sollten 4 Anforderungen erfüllen. Sie brauchen einen randomisierten Kontrollarm gegenüber Usual Care sowie den Return to Work als präspezifizierten primären Endpunkt. Die Arbeitsunfähigkeitsdaten sollten registerbasiert statt über Selbstauskunft erhoben werden, um die Abhängigkeit von unverblindeten Angaben zu verringern. Der Nachbeobachtungszeitraum sollte mindestens 12 Monate umfassen, damit sich eine dauerhafte Wiedereingliederung von einem frühen ersten Arbeitsversuch unterscheiden lässt. Diese Übersichtsarbeit liefert die bislang fehlende systematische Grundlage für solche Studien.

Digitale Anwendungen ersetzen keine gute arbeitsmedizinische Betreuung, bieten jedoch eine vielversprechende und potenziell kostengünstige Ergänzung bei der beruflichen Wiedereingliederung. Richtig eingesetzt, können sie die Versorgungskapazitäten entlasten und Beschäftigte früher wieder in den Beruf integrieren.