

Anlage 8: Evaluationskonzept

Inhalt

1. Evaluation der klinischen Wirksamkeit.....	2
1.1 Rekrutierung und Intervention	2
1.1.1 Rekrutierungsstrategie und Einschlusskriterien	2
1.1.2 Randomisierung auf Landkreisebene	3
1.1.3 Multimodale Intervention.....	5
1.2 Datenfluss und technische Umsetzung	11
1.3 Evaluation der Wirksamkeit – methodisches Vorgehen.....	14
1.3.1 Zielgrößen	14
1.3.2 Fallzahlplanung	15
1.3.3 Statistische Analyse: Primäre Endpunkte	15
1.3.4 Analyse der sekundären Endpunkte und Verlaufsmessungen	17
1.3.5 Zusätzliche Analysen: Langzeiteffekte und Spätintervention	17
2. Begleitende Prozessevaluation	18
2.1 Perspektive der Ärzte.....	18
2.2 Fokusgruppen mit Mitarbeitenden des PoCo-Busses.....	18
2.3 Wahrnehmung von Terminen und Hausarztkonsultationen	19
3. Qualitative Analyse der Patientenperspektive	19
3.1 Charakterisierung der teilnehmenden Patienten	19
3.2 Rekrutierung	19
3.3 Interviewdurchführung	20
3.4 Auswertung.....	21
4. Gesundheitsökonomische Evaluation der Intervention	21
4.1 Zweck	21
4.2 Datengrundlage	21
4.3 Fragestellungen	22
4.4 Zentrale gesundheitsökonomische Zielindikatoren.....	24
4.4.1 Kostenindikatoren.....	24
4.4.2 Arbeitsunfähigkeitstage	25
4.5 Weitere für die gesundheitsökonomische Evaluation relevante Indikatoren	25
4.5.1 Interventionskosten	25
4.5.2 Arbeitsunfähigkeits-bedingter Bruttowertschöpfungsausfall	26
4.6 Beobachtungszeiträume	26
4.7 Statistische Analyse.....	27
4.8 Ergänzende Analysen: Langzeiteffekte und Effekt der Spätintervention.....	29
Literaturverzeichnis	30

1. Evaluation der klinischen Wirksamkeit

Das WATCH-Projekt ist als monozentrische randomisierte kontrollierte Studie konzipiert. Die Patienten werden in zwei Gruppen randomisiert: eine Frühinterventionsgruppe und eine Spätinterventionsgruppe (Kontrollgruppe). Die Frühinterventionsgruppe erhält eine sofortige 12-wöchige Intervention gefolgt von einem 12-wöchigen Nachbeobachtungszeitraum, während die Spätinterventionsgruppe zunächst die allgemein übliche Behandlung („treatment as usual“) inklusive Monitoring und nach 12 Wochen die Intervention erhält (*Abbildung 1*).

1.1 Rekrutierung und Intervention

1.1.1 Rekrutierungsstrategie und Einschlusskriterien

Die Rekrutierung geeigneter und interessierter Patienten erfolgt über niedergelassene Hausärzte in Thüringen. Über verschiedene Medieninstrumente werden diese auf die Möglichkeit zu einer Teilnahme am WATCH-Projekt durch die Kassenärztliche Vereinigung Thüringen (KVT) hingewiesen. Interessierte Ärzte werden im Rahmen einer Online-Fortbildung (vgl. § 3 (1) i. V. mit Anlage 7) durch die KVT über WATCH informiert. Hier wird unter anderem ein Video zur Vorstellung des Projektes gezeigt.

Teilnehmende Hausärzte haben zunächst die Aufgabe, potentiell geeignete Patienten auf WATCH hinzuweisen, über das Projekt aufzuklären, Ein- und Ausschlusskriterien zu prüfen, ein fragebogenbasiertes Screening durchzuführen und geeignete Patienten, die in eine Teilnahme einwilligen, der Studienzentrale im Post-COVID-Zentrum (PoCo-Zentrum) des Universitätsklinikums Jena (UKJ) zu melden.

Ab dem Startpunkt der Patientenrekrutierung werden alle Patienten mit anhaltenden Beschwerden infolge einer nachgewiesenen und mindestens vier Wochen zurückliegenden SARS-CoV-2-Infektion auf WATCH hingewiesen. Interessierten Patienten wird über das PoCo-Zentrum ein Fragebogenkatalog („Starterpaket“; Anhang 1 zur Anlage 1) zugesandt. Dieser enthält Fragen zu soziodemografischen Charakteristika, zu bestehenden Vorerkrankungen, zum Allgemeinbefinden während der akuten SARS-CoV-2-Infektion (Global Assessment of Functioning (GAF) scale; Hall, 1995), zur Zeit seit Infektion, zu aktuellen Beschwerden in Folge der SARS-CoV-2-Infektion (GAF scale) sowie den Short Form 36 (SF-36) Fragebogen (Bullinger et al., 1995; Morfeld et al., 2011).

Neben einer nachgewiesenen SARS-CoV-2-Infektion und persistierenden Beschwerden müssen Patienten die folgenden Einschlusskriterien erfüllen (siehe § 7 Versorgungsvertrag): Der Physical Component Score (PCS) des SF-36 muss unter 45 Punkten liegen und der Patient muss mindestens 18 Jahre alt sein, in Thüringen wohnen und ein mobiles Endgerät mit stabiler Internetverbindung besitzen. Von der Projektteilnahme ausgeschlossen sind schwangere oder stillende Personen, jene, die bereits außerhalb von WATCH an einer

klinischen Studie oder einem Versorgungsprojekt teilnehmen und Patienten, die sich in laufender psychiatrischer oder psychotherapeutischer Behandlung befinden.

Auf die Festlegung einer maximalen Zeitspanne zwischen dem SARS-CoV-2-Infektionsdatum und der Vorstellung bei dem Arzt wird verzichtet. Demnach können auch Patienten mit Post-COVID-Beschwerden, deren Infektion mehrere Monate zurückliegt oder die bereits eine Post-COVID-Diagnose erhalten haben, für WATCH rekrutiert werden. Zum Zeitpunkt des Einschlusses in WATCH müssen mindestens 12 Wochen seit der SARS-CoV-2-Infektion vergangen sein.

1.1.2 Randomisierung auf Landkreisebene

Die Randomisierung von Patienten erfolgt stratifiziert. Die Strata bilden die 22 thüringischen Landkreise beziehungsweise kreisfreien Städte. Für jedes Stratum wird die gleiche Anzahl an Patienten angestrebt. Die Randomisierungssequenz für das jeweilige Stratum wird im Vorfeld generiert. Nachdem ein Patient an das PoCo-Zentrum gemeldet und registriert wurde, erfolgt hier die auf der Randomisierungssequenz basierende Zuteilung in die Früh- oder Spätinterventionsgruppe. Durch diese Randomisierung auf Landkreisebene sollen strukturelle und sozioökonomische Unterschiede zwischen diesen adressiert, eine Balance zwischen den beiden Gruppen innerhalb der Strata angestrebt und letztlich eine Ergebnisverzerrung durch regionale Unterschiede vermieden werden.

Die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF) und Cochrane Deutschland sehen im „Manual zur Bewertung des Biasrisikos in Interventionsstudien“ eine stratifizierte Randomisierung eine Möglichkeit, eine potentielle Verzerrung des Behandlungseffekts aufgrund wichtiger prognostischer Faktoren auf der Ebene des Studiendesigns zu adressieren (Cochrane Deutschland et al., 2021). Gemäß der Leitlinie „statistical principles for clinical trials“ des International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH, 1998) sind diese Strata auch in der statistischen Analyse abzubilden (siehe Abschnitt 1.3.3.1).

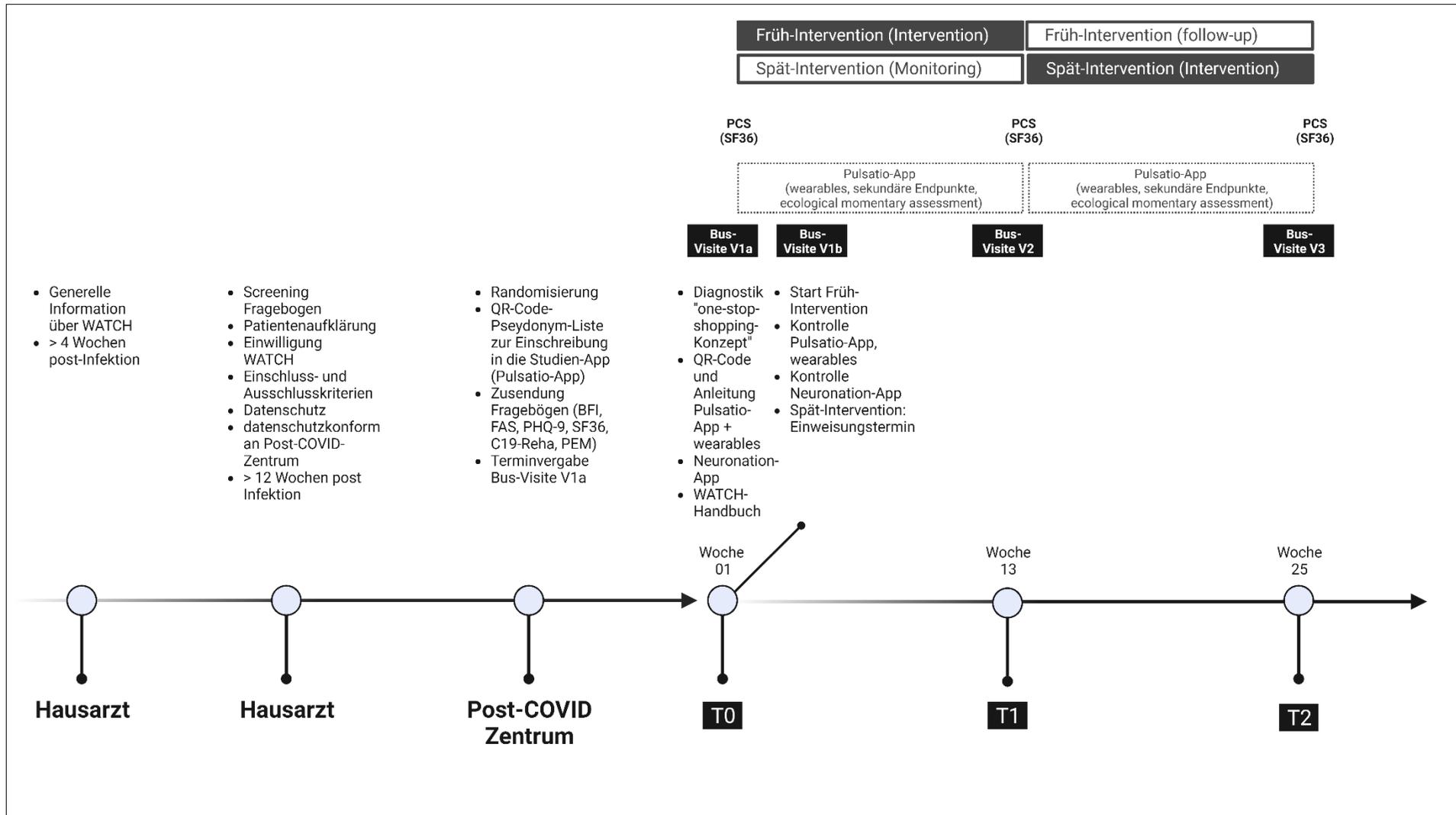


Abbildung 1. Flussdiagramm zum Ablauf des WATCH-Projekts

1.1.3 Multimodale Intervention

Nachdem ein Patient dem PoCo-Zentrum des UKJ gemeldet und hier randomisiert wurde, erfolgt über das PoCo-Zentrum die Zusendung eines Fragebogenkatalogs und die Terminvergabe für die erste Vorstellung (Visite 1a) im Post-COVID-Bus (PoCo-Bus; *Abbildung 1*). Im PoCo-Bus ist ein interdisziplinäres Team tätig, bestehend aus medizinischen Fachangestellten, Sportwissenschaftlern und Psychologen.

1.1.3.1 Visite 1a im PoCo-Bus

Bei Visite 1a (V1a) erfolgt eine umfassende interdisziplinäre Diagnostik („one-stop-shopping-Konzept“) zur Erfassung der Post-COVID-Beschwerden. Diese beinhaltet eine COVID-Anamnese, die Erhebung des primären Zielparameters (SF-36 PCS), einen Konzentrations-, Aufmerksamkeits- und Gedächtnistest (Oxford Cognitive Screen plus, [OCS-plus]; Demeyere et al., 2021), das Montreal Cognitive Assessment Tool (MoCA; Nasreddine et al., 2005), einen Sit-to-Stand-Test zur Erfassung der kardiorespiratorischen Leistungsfähigkeit (STST; Sevillano-Castaño et al., 2023), eine Handgriffkraftmessung zur Erfassung der muskulären Leistungsfähigkeit, eine Elektroimpedanztomografie (EIT) der Lunge zur Charakterisierung von Ventilationsstörungen sowie gegebenenfalls die Vervollständigung des vorab zugesandten Fragebogenkatalogs (*Tabellen 1 und 2*). Dieser Katalog beinhaltet neben dem SF-36 den Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001), die Fatigue Assessment Scale (FAS; Michielsen et al., 2004), den Essener Fragebogen zur Krankheitsverarbeitung (EFK; Franke & Jagla-Franke, 2021), das ICD-10-Symptom-Rating (ISR; Tritt et al., 2008), den Cognitive Failure Questionnaire (CFQ-G; Broadbent et al., 1982), den Pittsburgh Sleep Quality Index (PSQI; Hinz et al., 2017), das Immunologische Inventar für Erholungs- und Beanspruchungszustände (II-EBZ), den Fragebogen zur Post-Exertionellen-Malaise (PEM; Cotler et al., 2018) und den Work Ability Index (WAI; Hasselhorn & Freude, 2007).

Weiterhin erhalten die Patienten bei V1a in Vorbereitung für Visite 1b (V1b) eine schriftliche Anleitung sowie einen Link mit einem Video-Tutorial. Hier werden die Patienten in die für das Projekt relevanten Apps (Neuronation-App® und Pulsatio-App®) eingewiesen, über die drei Module der WATCH-Intervention (*Abbildung 2*) und die Installation und Nutzung des Fitness-Trackers („Wearables“) aufgeklärt. Die Teilnehmenden erhalten außerdem einen QR-Code, mit dem sie sich in die Apps einloggen können und der mit ihrem Pseudonym und damit mit der Projektdatenbank verknüpft ist (siehe Abschnitt 1.2).

Tabelle 1: Variablenliste und Parameter der primären und sekundären Endpunkte, Messzeitpunkte sowie Messinstrumente und Itemzahl

Variable (Intervention & Kontrolle)	Messzeitpunkt (T)	Instrument	Items (Anzahl)
Primärer Endpunkt			
Körperliche Gesundheit	T0p, TE4a, TE8a, T1p, T2p	SF-36 Physical component summary score (PCS)	36
Sekundäre Endpunkte/ Verlaufsmessungen - Apparative Diagnostik im Bus (TE)			
Mentale Gesundheit	T0p, TE4a, TE8a, T1p, T2p	SF-36 Mental component summary score (MCS)	36
Konzentrations-, Aufmerksamkeits- und Gedächtnistest	T0, T1, T2	Oxford screening plus (OCS-plus)	
Test zur Erfassung milder kognitiver Dysfunktion	T0, T1, T2	Montreal Cognitive Assessment (MoCA)	
Kardiorespiratorische körperliche Leistungsfähigkeit	T0, T1, T2	1-Minute Sit-to-Stand-Test inklusive Herzfrequenz, Sauerstoffsättigung	
Muskuläre körperliche Leistungsfähigkeit	T0, T1, T2	Handgriffkraftmesser	
Ventilationsstörungen	T0, T1, T2	Lungentonometrie	
Sekundäre Endpunkte/ Verlaufsmessungen			
Screening Depression	T0p, TE6a, T1p, TE18a, T2p	Patient Health Questionnaire (PHQ-9)	9
Aktuelle Bewältigungsanstrengungen zur Krankheitsverarbeitung (emotional, kognitiv)	T0p, TE6a, T1p, T2p	Essener Fragebogen zur Krankheitsverarbeitung (EFK)	45
Status und Schweregrad psychischer Störungen	T0p, TE6a, T12p, T24p	ICD-10-Symptom-Rating-Scale (ISR)	29
Körperliche und mentale Fatigue	T0p, TE2a, TE4a, TE6a, ..., T1a, ..., T2a	Fatigue assessment scale (FAS)	10
Tägliche kognitive Probleme	T0p, TE3a, TE6a, TE9a, ..., T1a, ..., T2a	Cognitive failure questionnaire (CFQ)	25
Schlafqualität	T0p, TE2a, TE4a, TE6a, ..., T1a, ..., T2a	Pittsburgh Sleep Quality Index (PSQI)	19
Immunologisches Inventar Erholung/Beanspruchung	T0p, TE2a, TE4a, TE6a, ..., T1a, ..., T2a	Immunologisches Inventar akute Erholung, Beanspruchung (II-EBZ)	18
Belastungsintoleranz	T0a-T2a (wöchentlich)	Post exertional malaise (PEM)	10
Aktuelle und zukünftige subjektive Arbeitsfähigkeit	T0p, T1p, T2p	Work Ability Index (WAI)	10
Interview: Nutzungsverhalten, Adhärenz, Akzeptanz	T3i, T9i, T18i	Befragung durch die DSck (Interviews): 20-24 Patient:innen	30-45 min
Wearable Daten			
Körperliche Belastung: Schrittzahl, metabolisches Äquivalent, zurückgelegte Strecke pro Tag	Kontinuierlich	Pulsatio-App	
Körperliche Beanspruchung: Ruheherzfrequenz Nacht und Tag, Schlafregulärität- und dauer, peak Herzfrequenz Tag	Kontinuierlich	Pulsatio-App	

Abkürzungen: p: Papier; a: Pulsatio-App; i: Interview, Telefon oder Video; DSck: Deutsche Stiftung für chronisch Kranke; T: Messzeitpunkt (entspricht Woche).

Tabelle 2: Erhobene Instrumente mit Itemzahl je Messzeitpunkt

Variable/Messzeitpunkt	Instrument	T0 Bus													T1 Bus													T2 Bus
		Visite V1b	W0	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18	W19	W20	W21	W22	W23	W24	W25
		Anzahl Items																										
Primärer Endpunkt																												
Körperliche Gesundheit	PCS	36				36				36					36													36
Sekundäre Endpunkte/ Verlaufsmessungen																												
Mentale Gesundheit	MCS	36				36				36					36													36
Screening Depression	PHQ-9	9							9						9													9
Aktuelle Bewältigungsanstrengungen zur Krankheitsverarbeitung (emotional, kognitiv)	EFK	45							45						45													45
Status und Schweregrad psychischer Störungen	ISR	29							29						29													29
Körperliche und mentale Fatigue	FAS	10		10		10		10		10		10		10		10		10		10		10		10		10		10
Tägliche kognitive Probleme	CFQ	25			25			25				25			25			25			25			25			25	
Schlafqualität	PSQI	19		19		19		19		19		19		19		19		19		19		19		19		19		19
Immunologische Erholung/Beanspruchung	II-EBZ	18		18		18		18		18		18		18		18		18		18		18		18		18		18
Belastungsintoleranz/neanspruchungsbedingte Symptomexazerbation	PEM	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Aktuelle und zukünftige subjektive Arbeitsfähigkeit	WAI	10													10													10
Summe		247	10	57	35	129	10	165	10	129	35	57	10	57	247	10	57	35	57	10	82	10	57	35	57	10	247	
Items pro Tag (entspricht Summe/7Tage)		35.3	1.4	8.1	5.0	18.4	1.4	23.6	1.4	18.4	5.0	8.1	1.4	8.1	35.3	1.4	8.1	5.0	8.1	1.4	11.7	1.4	8.1	5.0	8.1	1.4	35.3	
Sekundäre Endpunkte / Verlaufsmessungen - Apparative Diagnostik im Bus (TE)																												
Konzentrations-, Aufmerksamkeits- und Gedächtnistest	OCS-plus	X													X													X
Test zur Erfassung milder kognitiver Dysfunktion	MoCA	X													X													X
Kardiorespiratorische körperliche Leistungsfähigkeit	1-STS	X													X													X
Muskuläre körperliche Leistungsfähigkeit	HD	X													X													X
Ventilationsstörungen	LT	X													X													X
Interview: Nutzungsverhalten, Adhärenz, Akzeptanz	Interview				X						X										X							
Wearable Daten																												
Körperliche Belastung: Schrittzahl, metabolisches Äquivalent, zurückgelegte Strecke pro Tag	Pulsatio-App		kontinuierlich																									
Körperliche Beanspruchung: Ruheherzfrequenz Nacht und Tag, Schlafregulärität- und dauer, peak Herzfrequenz Tag	Pulsatio-App		kontinuierlich																									

Anmerkung: die Erfassung der Arbeitsunfähigkeit (AU-Tage) erfolgt zu denselben Zeitpunkten wie die Erhebung des WAI.

Abkürzungen: PCS, MCS: SF-36 Physical (PCS) und mental (MCS) component summary score, PHQ-9: Patient Health Questionnaire (PHQ-9), EFK: Essener Fragebogen zur Krankheitsverarbeitung (EFK); ISR: CD-10-Symptom-Rating-Scale, FAS: Fatigue assessment scale, CFQ: Cognitive failure questionnaire, PSQI: Pittsburgh Sleep Quality Index, II-EBZ: Immunologische Erholung/Beanspruchung (II-EBZ), PEM: Post exertional malaise, WAI: Work Ability Index (WAI), Interviews: Befragung durch die Deutsche Stiftung für chronisch Kranke (DSck).

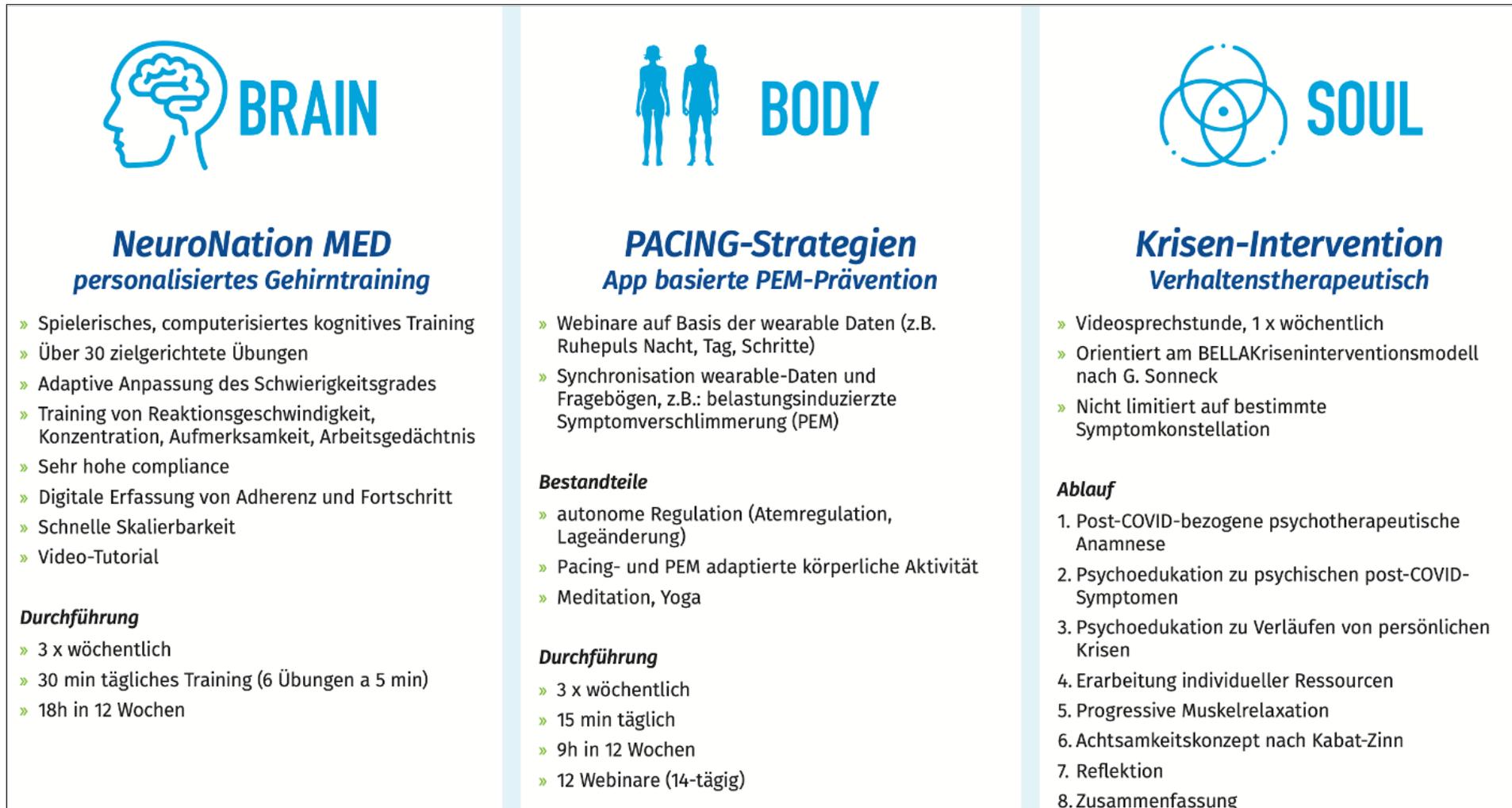


Abbildung 2. Schematische Darstellung der modularen symptomübergreifenden interdisziplinären telemedizinischen Intervention

1.1.3.2 *Digitale Bestandteile: Webinare, Neuronation-App®, Pulsatio-App® und Datenspende-Blog*

Die zentralen digitalen Bestandteile des WATCH-Projektes bilden die Projekthomepage (digitales Kursbuch), die Webinare, die **Pulsatio-App®** und der Datenspende-Blog sowie die **Neuronation-App®**. In der Pulsatio-App® werden zum einen WATCH-spezifische Fragebögen in regelmäßigen Abständen erfasst (*Tabellen 1 und 2*). Darüber hinaus werden die Wearables der Teilnehmenden mit der Pulsatio-App® verknüpft. Diese Daten zur körperlichen Beanspruchung und Belastung werden außerdem pseudonymisiert in aggregierten Analysen in Form von Gruppendaten gemeinsam mit der Technischen Universität Dresden (TU Dresden) analysiert und auf dem **Datenspende-Blog** (<https://corona-datenspende.de/science/>) zur Verfügung gestellt. Somit ist es den Teilnehmenden möglich, ihre individuellen Wearable-Daten im Vergleich zu den aggregierten Gruppendaten zu betrachten. Zudem werden 14-tägig bis wöchentlich auf der Projekthomepage Webinare und Podcasts zur Verfügung gestellt. Diese dienen der Vermittlung von Post-COVID- und modulspezifischem Wissen und umfassen Inhalte zum Umgang mit den Wearables und der Pulsatio-App®. Die Neuronation-App® ist zentraler Bestandteil des Moduls „BRAIN“ (siehe Abschnitt 1.1.3.4).

1.1.3.3 *Visite 1b im PoCo-Bus und Inhalte der drei Module „BRAIN“, „BODY“, „SOUL“*

V1b dient zur Kontrolle der Startbedingungen für die Frühintervention. Mit V1b beginnt die Frühintervention sowie das Monitoring der Spätinterventionsgruppe (*Abbildung 1*). Patienten der Frühinterventionsgruppe werden in die drei Module der telemedizinischen Intervention eingewiesen (*Abbildung 2*). Patienten der Spätinterventionsgruppe, die keine reguläre V1b erhalten, haben die Möglichkeit, bei aufkommenden Fragen in der entsprechenden Woche eine Vor-Ort-Beratung zu erhalten. Beiden Gruppen steht über den gesamten Projektverlauf hinweg das PoCo-Zentrum bei Fragen telefonisch zur Verfügung.

Teilnehmende Patienten erhalten über die 12 Wochen der Intervention 36 x BRAIN (3 x pro Woche, je 30 Minuten), 36 x BODY (3 x pro Woche je 15 min), 8 x SOUL (1 x pro Woche je 30 min) und somit insgesamt 80 Interventionseinheiten. Insgesamt ergeben sich 165 Minuten WATCH-Intervention pro Woche.

1.1.3.4 *Einweisung in das Modul „BRAIN“ – Verwendung der Neuronation-App®*

Die Einweisung der Patienten in das selbstständige Durchführen des App-basierten kognitiven Trainings in der Neuronation-App® erfolgt durch einen Psychologen des UKJ. Ein klinischer Psychologe sowie ein Arzt des UKJ sind für die Einweisung der Patienten, die Begleitung der Intervention und die Dokumentation des Behandlungsfortschritts mitverantwortlich. In der Neuronation-App® werden den Patienten spielartig konzipierte Aufgaben präsentiert, die auf

das Training von Aufmerksamkeit, Konzentration und Gedächtnis abzielen und sich dabei an das individuelle und zum Teil tagesformabhängige Leistungsniveau anpassen.

1.1.3.5 Einweisung in das Modul „BODY“ – Wearables, Pulsatio-App® und Datenspende-Blog

Die Einweisung in das Modul „BODY“ erfolgt durch die medizinischen Fachangestellten des UKJ. Es handelt sich bei dem Modul um ein digitales Rehabilitationsprogramm mit dem Ziel, die körperliche Fitness zu fördern und/oder wiederzuerlangen und den Umgang mit Belastungsintoleranz zu erlernen beziehungsweise zu verbessern. Verantwortung für das Modul (Eingangsdiagnostik bei V1a, Installation der telemedizinischen Komponenten, Umsetzung der Webinare und Podcasts) tragen Ärzte und Sportwissenschaftler des UKJ.

Grundlegende Modulinhalte sind die Antizipation erwarteter Belastung, die Anpassung und Dosierung von Belastung mit Energiemanagement-Strategien und „Pacing“, Atemregulationstraining, Yoga und Meditation sowie begleitende Post-COVID-spezifische Wissensvermittlung im Rahmen der Webinare und Podcasts auf der Projekthomepage. Darüber hinaus ist das Monitoring der eigenen körperlichen Aktivität und Erholung mittels Wearables sowie die Einsicht der aggregierten Gruppendaten über den Datenspende-Blog Bestandteil von „BODY“ (siehe Abschnitt 1.1.3.2). Das Training im Modul „BODY“ erfolgt drei Mal pro Woche (*Abbildung 2*).

1.1.3.6 Einweisung in das Modul „SOUL“

Bei „SOUL“ handelt es sich um eine digitale verhaltenstherapeutische Kurzzeit-Krisenintervention bestehend aus acht Sitzungen von je 30 Minuten (*Abbildung 2*). (*Abbildung 2*). Dabei zielt das Modul auf die Verbesserung der Krankheitsverarbeitung, die Ressourcenaktivierung zur Stärkung der Funktionsfähigkeit im Alltag sowie den Rückgang psychopathologischer Symptome – insbesondere Depressivität und Ängstlichkeit. Psychologen des UKJ sind verantwortlich für das Erstgespräch, die Terminkoordination, Vorbereitung, Durchführung und Nachbereitung der Sitzungen sowie für die Bearbeitung von Rückfragen und die Supervision der Patienten. Über die Projekthomepage werden begleitend zu jeder Sitzung Zusammenfassungen und Übungsmaterialien angeboten.

1.1.3.7 Visiten V2 und V3 im PoCo-Bus

Visite 2 (V2) findet 12 Wochen nach V1b statt. Für Patienten der Spätinterventionsgruppe entspricht V2 der V1b der Frühinterventionsgruppe: Patienten der Spätinterventionsgruppe werden hier in die Module eingewiesen und es erfolgt die Sicherstellung der Startbedingungen (Kontrolle der Wearables und App-Anwendungen). Für die Spätinterventionsgruppe beginnt nun die 12-wöchige Intervention (*Abbildung 1*). Patienten der Frühinterventionsgruppe erhalten zu V2 im Sinne eines Abschlussassessments alle Untersuchungen, die auch zu V1 erfolgen (*Tabelle 1*). Für die Patienten der Frühinterventionsgruppe erfolgt nun eine 12-

wöchige Nachbeobachtung. Visite 3 (V3) findet 12 Wochen nach V2 statt. Beide Gruppen – Früh- und Spätinterventionsgruppe – durchlaufen hier dieselben Untersuchungen wie bei V1 (*Tabelle 1*).

1.2 Datenfluss und technische Umsetzung

Der Datenfluss ist in *Abbildung 3* schematisch dargestellt. Die Studienzentrale registriert die teilnehmenden Patienten, vergibt eine eindeutige Projekt-ID und führt die Randomisierung durch. Diese Projekt-ID wird einem QR-Code und einem damit verknüpften Pseudonym zugeordnet. Während des WATCH-Projektes dient REDCap® als zentrales Datenerfassungstool. Die bei den Bus-Visiten erhobenen Daten werden über REDCap® erfasst und der Projekt-ID zugeordnet. Die zusätzlichen Fragebögen, die online in der Pulsatio-App® oder papierbasiert erhoben werden (*Tabellen 1 und 2*), sind über die Projekt-ID mit der REDCap®-Datenbank verknüpfbar. Patienten erhalten den mit ihrer Projekt-ID verknüpften QR-Code bei V1a. Dieser dient zum Login in die für WATCH-relevanten Apps. Damit ist es möglich, die in den Apps erhobenen Daten dem individuellen Pseudonym und somit mit der Projekt-ID zuzuordnen. Die Wearable-Daten werden durch die TU Dresden in Form aggregierter Gruppendaten ausgewertet und auf dem Datenspende-Blog öffentlich zugänglich gemacht (siehe Abschnitt 1.1.3.2).

Neben den Wearable-Daten und den in *Tabelle 1* dargestellten Endpunkten werden die folgenden Patientencharakteristika primär erhoben: Alter, Geschlecht, Gewicht, Größe, Datum der Infektion, Datum der Vorstellung im PoCo-Bus, Vorerkrankungen, Begleitmedikation, Symptomatik und ggf. Behandlung der akuten Infektion, Zeiten von Arbeitsunfähigkeit seit der SARS-CoV-2-Infektion, bereits erfolgte Post-COVID-Diagnostik inkl. der Befunde, sozioökonomische Charakteristika (Bildungsstatus, Berufsstatus, Einkommen) sowie der Impfstatus vor der Infektion.

Die Deutsche Stiftung für chronisch Kranke (DSck) erfasst im Rahmen des WATCH-Projektes die Perspektive der Teilnehmenden mittels qualitativer Interviews (siehe Abschnitt 3). Um geeignete Patienten auswählen zu können, stratifiziert das PoCo-Zentrum die an der Interviewstudie interessierten Patienten vorab hinsichtlich Geschlecht und Alter und führt eine entsprechende Meldung von Patienten an die DSck durch.

Für die vergleichenden Kostenanalysen (gesundheitsökonomische Evaluation, siehe Abschnitt 4) erfolgt für jene Projektteilnehmenden, die bei den datenliefernden Krankenkassen versichert sind, eine entsprechende Sekundärdatenabfrage. Diese Daten werden durch eine Datentreuhandstelle datenschutzkonform mit relevanten Primärdaten zusammengeführt und in pseudonymisierter Form an das Institut für Allgemeinmedizin am UKJ für die gesundheitsökonomische Evaluation übermittelt. Für die Abfrage und Verarbeitung der Sekundärdaten wird eine gesonderte Teilnahmeerklärung (Anlage 4a zum

Versorgungsvertrag) durch die Teilnehmenden des Projektes abgegeben. Die Übermittlung dieser Sozialdaten über die Datentreuhandstelle an das Institut für Allgemeinmedizin für die gesundheitsökonomische Evaluation bedarf darüber hinaus gemäß § 75 Abs. 4 SGB X der vorherigen Genehmigung durch die jeweils zuständige Aufsichtsbehörde.

Die zentralen Analysen – sowohl im Rahmen der *Evaluation der klinischen Wirksamkeit*, als auch der gesundheitsökonomischen Evaluation – werden von dem jeweils evaluierenden Institut verblindet durchgeführt. Eine Entblindung der Gruppenzugehörigkeit (Frühinterventions-/ Spätinterventionsgruppe) erfolgt nach Fertigstellung der entsprechenden Analysen.

1.3 Evaluation der Wirksamkeit – methodisches Vorgehen

1.3.1 Zielgrößen

In *Tabellen 1* und *2* sind die Messinstrumente der primären und sekundären Endpunkte und die entsprechenden Messzeitpunkte dargestellt.

1.3.1.1 Primäre Zielgröße - Hierarchische Testung

Die primäre Zielgröße ist die absolute Veränderung der „körperlichen Gesundheit“ gemessen mittels des **PCS des SF-36**. Angesichts der Erfahrungen aus dem PoCo-Zentrum am UKJ sowie der Testphase gehen wir davon aus, dass die Intervention den anvisierten Effekt gemessen am PCS erbringen kann und dies eine klinisch relevante Verbesserung für Patienten darstellt (Stallmach et al., 2023).

Vor dem Hintergrund des multimodalen Charakters der WATCH-Intervention und der von Post-COVID-Patienten berichteten typischen Beschwerden der körperlichen Belastungsinsuffizienz („Body“), der kognitiven Dysfunktionen („Brain“) sowie depressiver Verstimmungen („Soul“) wurde intensiv geprüft, ob zusätzliche Endpunkte zur (neuro-)psychologischen Gesundheit zu primären Endpunkten aufgewertet werden können (Han et al., 2022; Rochmawati et al., 2022; Stallmach et al., 2023). Hiernach würden sich drei gleichberechtigte primäre Endpunkte ergeben. Alternativ wäre die Kombination entsprechender dreier Endpunkte zu einem Composite-Endpunkt grundsätzlich möglich.

Ein entsprechendes Vorgehen wird in der Literatur jedoch sehr kritisch gesehen, sodass wir hiervon Abstand genommen haben (McCoy, 2018). Die Änderung der Anzahl von primären Endpunkten hätte zudem zur Folge, dass die Aussagekraft der Studienergebnisse bei gleicher Gruppengröße deutlich reduziert und entsprechend eine deutliche Anhebung der Stichprobengröße (siehe Abschnitt 1.3.2) notwendig gewesen wären. Dies schien vor dem Hintergrund der Komplexität der WATCH-Intervention sowie des zusätzlich anfallenden Personalaufwandes nicht umsetzbar. Jedoch werden wir die von Marcus et al. (1976) dargestellte Strategie zur Testung hierarchisch geordneter Hypothesen anwenden. Dem dort beschriebenen Prinzip folgend, wird zunächst der PCS des SF-36 als primärer Endpunkt betrachtet und eine zweiseitige Testung durchgeführt (siehe Abschnitt 1.3.3.1). Sollte der Unterschied zwischen den beiden Gruppen signifikant sein ($\alpha=0.05$), erfolgt die Betrachtung des zweiten Zielparameters, des Mental Component Scores (MCS) des SF-36, auf dieselbe Weise. Nach dem Prinzip des hierarchischen Testens ist hier keine Anpassung des α -Niveaus notwendig.

Formuliert als Nullhypothese gilt $H_{0PCS}: \mu_{\text{frühe_Intervention}} = \mu_{\text{späte_Intervention}}$ getestet gegen $H_{1PCS}: \mu_{\text{frühe_Intervention}} \neq \mu_{\text{späte_Intervention}}$. Führt dies zum Verwerfen von H_{0PCS} kann $H_{0MCS}: \mu_{\text{frühe_Intervention}} = \mu_{\text{späte_Intervention}}$ gegen $H_{1MCS}: \mu_{\text{frühe_Intervention}} \neq \mu_{\text{späte_Intervention}}$ zum vollen α -Niveau konfirmatorisch getestet werden. Im anderen Fall – dem Beibehalten von H_{0PCS} – ist ein Testen von H_{0MCS} vs. H_{1MCS} nicht möglich.

Die Erhebung des SF-36 erfolgt zu insgesamt fünf Zeitpunkten (*Tabellen 1 und 2*).

1.3.1.2 Sekundäre Endpunkte

Die sekundären Endpunkte werden im PoCo-Bus oder mittels Pulsatio-App® erfasst (*Tabelle 1, Tabelle 2*). Zu den Zeitpunkten V1a, V2 und V3 erfolgen die Erhebungen mittels Fragebögen, Tablet-basierten Untersuchungen oder apparativer Diagnostik im Bus. Die Erhebungen zu allen anderen Messzeitpunkten erfolgen mittels papierbasierter Fragebögen oder digitaler Fragebögen in der Pulsatio-App®.

1.3.2 Fallzahlplanung

Basierend auf den Erkenntnissen der Post-COVID-Ambulanz des UKJ (Stallmach et al., 2023) wird angenommen, dass bei einer Standardabweichung von $\pm 9,8$ Punkten vom PCS-Mittelwert eine Stichprobengröße von 486 Personen erforderlich ist. Um einen möglichen Drop-out von 22 % (138 Personen) zu berücksichtigen, sollen insgesamt 624 Personen rekrutiert werden. Mit der geplanten Stichprobengröße kann bei einer Power von 80 % und einem α -Niveau von 5 % ein Mittelwertunterschied im PCS von 2,5 Punkten aufgedeckt werden, der als klinisch relevant betrachtet wird. Diese Fallzahlschätzungen basieren vereinfacht auf einem Gruppenvergleich mittels t-Test. Zwar wird in der Analyse ein komplexeres Regressionsmodell verwendet, es kann jedoch davon ausgegangen werden, dass die Fallzahlberechnungen konservativ sind.

Der minimale klinisch relevante Unterschied im PCS ist erkrankungsabhängig. Verschiedene Interventionsstudien zeigen, dass für viele Erkrankungen eine klinisch relevante Verbesserung bei etwa 5 Punkten Unterschied im PCS liegt (Badhiwala et al., 2018; Carreon et al., 2010; Copay et al., 2008). Basierend auf den angeführten Annahmen besteht in diesem Projekt ausreichend statistischer "Puffer", um bereits einen kleineren Unterschied nachzuweisen.

1.3.3 Statistische Analyse: Primäre Endpunkte

1.3.3.1 Hauptanalyse

Die primären Endpunkte, der PCS und MCS des SF-36, werden mittels gemischtem Regressionsmodell analysiert. Die abhängige Variable stellt dabei die intraindividuelle Differenz in den PCS Werten zwischen V2 und V1a als metrische Variable dar. Die Gruppenzugehörigkeit (Früh-/Spätintervention) und der PCS-Wert zu V1a fließen als feste

Effekte, das Landkreisstratum als Zufallseffekt in das Modell ein. Es werden nur Patienten berücksichtigt, die in den benötigten Variablen keine fehlenden Werte aufweisen. Drop-outs und die entsprechenden Gründe für ein vorzeitiges Ausscheiden aus der Studie – wenn entsprechend freiwillige Angaben hierzu gemacht wurden – werden berichtet. Die Analyse erfolgt nach dem Intention-to-treat-Konzept. Entsprechend werden auch Personen, die die Intervention nicht vollständig wahrgenommen haben, in der Analyse berücksichtigt – sofern für sie die für die Analyse notwendigen Daten vorliegen. Eine Imputation wird nicht vorgenommen. Zeigt sich in dieser Analyse ein signifikanter Unterschied im PCS zwischen der Früh- und der Spätinterventionsgruppe, wird die Analyse im Sinne des hierarchischen Testens (siehe Abschnitt 1.3.1.1) für den zweiten primären Endpunkt, den MCS, wiederholt.

1.3.3.2 Adjustierung

Aufgrund des Randomisierungsprozesses wird nicht davon ausgegangen, dass sich die beiden Gruppen (Früh- und Spätinterventionsgruppe) in relevantem Ausmaß in prognostisch potentiell wichtigen Charakteristika voneinander unterscheiden. Jedoch garantiert eine randomisierte Gruppenzuweisung nicht, dass die beiden Gruppen hinsichtlich relevanter Baseline-Charakteristika dieselbe Verteilung aufweisen (Altman, 1985). Signifikanztests auf Unterschiede hinsichtlich Baseline-Charakteristika sind jedoch kein geeignetes Kriterium für die Entscheidung, für entsprechende Variablen zu kontrollieren, da sie lediglich ein Maß dafür sind, ob der Randomisierungsprozess erfolgreich war – entsprechend wären 5 % solcher Vergleiche signifikant auf einem alpha-Level von 5 % (Altman, 1985; Altman & Doré, 1990; Corbett et al., 2014; Roberts & Torgerson, 1999). Eine entsprechende Bewertung sollte auf der Basis der prognostischen Stärke des Faktors getroffen werden (Altman & Doré, 1990; Corbett et al., 2014). Denn während eine augenscheinlich große Dysbalance in einer jedoch unwichtigen Variablen potentiell ohne Konsequenzen bleibt, kann ein kleiner, aber relevanter Unterschied in einem wichtigen prognostischen Faktor ausreichen, Schlussfolgerungen aus der Studie zu beeinflussen (Corbett et al., 2014).

Basierend auf dem empfohlenen Vorgehen zum Umgang mit potentiell konfundierenden Größen werden im Rahmen der deskriptiven Statistik die Gruppen im Hinblick auf entsprechende Variablen gegenübergestellt. Vor dem Hintergrund des bisherigen Wissensstands zu Post-COVID und Post-COVID-Genesung erfolgt diese Gegenüberstellung für die folgenden Variablen: Alter, Geschlecht, Vorerkrankungen, Allgemeinbefinden während der akuten SARS-CoV-2-Infektion (GAF scale), Impfstatus und Allgemeinbefinden im Hinblick auf die Post-COVID-Erkrankung zur Baseline (GAF scale). Für metrische Variablen werden der Median und der Interquartilsabstand (IQR) sowie Mittelwert und Standardabweichung (SD) berichtet, für kategoriale Variablen absolute und relative Häufigkeiten.

Zunächst wird die Regressionsanalyse ohne Berücksichtigung dieser Größen durchgeführt und dann unter Berücksichtigung jener Variablen dieser Liste wiederholt, für die eine Dysbalance vorliegt. Als Orientierung für eine Dysbalance wird a priori ein Unterschied von mehr als zwei Prozentpunkten in den kategorialen Variablen, ein Unterschied im medianen Alter von mehr als zwei Jahren und ein Unterschied in der mittleren GAF scale von mehr als zwei Punkten festgesetzt. Diese Annahmen werden als konservativ betrachtet mit dem Ziel, dem Einfluss selbst kleiner Dysbalancen zu begegnen.

1.3.3.3 Sensitivitätsanalysen

Zusätzlich zur Hauptanalyse, die als Intention-to-treat Analyse erfolgt, wird für jeden der zwei primären Endpunkte eine Per-Protocol-Auswertung, bei der die Anzahl der wahrgenommenen Interventionstermine berücksichtigt wird, durchgeführt. Nur Personen der Frühinterventionsgruppe, die die Intervention vollständig wahrgenommen haben, werden in der Analyse berücksichtigt.

Unabhängig von der angeführten Überlegung zu potentiellen Dysbalancen (Abschnitt 1.3.3.2) werden Subgruppenanalysen durchgeführt. So werden die Hauptanalysen unter anderem stratifiziert nach Geschlecht oder nach dem Vorliegen bestimmter selbstberichteter Erkrankungen wiederholt.

1.3.4 Analyse der sekundären Endpunkte und Verlaufsmessungen

Die sekundären Endpunkte (siehe auch *Tabellen 1* und *2*) werden ebenfalls in je einem gemischten Modell analysiert. Des Weiteren erfolgen sowohl für die primären (PCS, MCS), als auch die sekundären Endpunkte die Auswertung der Verlaufsmessungen. Im Rahmen von Mehrebenenmodellen werden eben diese Zielparameter im zeitlichen Verlauf der Intervention im Vergleich zum zeitlichen Verlauf ohne Therapie analysiert. Neben dem Prädiktor der Gruppenzugehörigkeit und dem Outcome (dem jeweiligen Zielparameter, z. B. dem PCS), werden mehrere Ebenen berücksichtigt. Auf der ersten Ebene befindet sich der Messzeitpunkt, auf der zweiten Ebene der Patient beziehungsweise die Patientin, auf der dritten Ebene der Landkreis. Diese Analysen der sekundären Endpunkte sowie der Verlaufsmessungen erfolgen explorativ, es werden Konfidenzintervalle der Schätzer, aber nicht die Signifikanz berichtet.

1.3.5 Zusätzliche Analysen: Langzeiteffekte und Spätintervention

Zur Analyse der Nachhaltigkeit der Intervention erfolgt zusätzlich eine Auswertung der primären und sekundären Endpunkte zu V3, 12 Wochen nach Abschluss der Intervention in der Frühinterventionsgruppe. Hier ist kein Vergleich mit der Kontrollgruppe möglich, da diese zu dieser Zeit bereits ihre Intervention erhält. Ebenso erfolgt eine Auswertung des Interventionseffektes nach der Intervention in der Spätinterventionsgruppe. Auch hier ist keine

Vergleichsgruppe vorhanden, es soll lediglich die Frage beantwortet werden, ob im Rahmen der WATCH-Intervention eine Verbesserung der Post-COVID-Symptomatik auch nach der Phase der „standard care“ erfolgt.

2. Begleitende Prozessevaluation

Die Prozessevaluation zielt darauf ab, auf der Basis von qualitativen und quantitativen Erhebungen die Perspektive hinsichtlich Machbarkeit, Nutzen und Akzeptanz der an der Intervention beteiligten Mitarbeitenden im PoCo-Bus sowie der beteiligten Ärzte zu erfassen. Die Sicht der Patienten wird unter Abschnitt 3 beleuchtet. Zusätzlich erfolgt eine Auswertung hinsichtlich der Auslastung des Busses sowie der Häufigkeit von Hausarztconsultationen im Projektverlauf.

2.1 Perspektive der Ärzte

Es erfolgt eine qualitative Befragung von teilnehmenden Ärzten mittels Leitfadenterviews zu Beginn der Interventionsphase bis Woche 12. Die Rekrutierung erfolgt über direkte Ansprache per Telefon und E-Mail. Dabei wird auf eine gleichmäßige Verteilung hinsichtlich der Region, Alter und Geschlecht geachtet. Die Interviews werden in Präsenz oder auf Wunsch auch online durchgeführt. Die Aufzeichnung erfolgt über ein digitales Audiogerät mit anschließender wörtlicher Transkription. Die Auswertung erfolgt mit der Software MaxQDA nach den Prinzipien der qualitativen Inhaltsanalyse nach Kuckartz. Über offenes Kodieren werden Kategorien generiert, welche im Anschluss auf die restlichen Interviews angewendet werden. Ziel ist es, erste Erfahrungen der Ärzte mit den Projektprozessen zu erfassen, sowie Barrieren und Facilitatoren für den Einschluss von Patienten zu identifizieren. Die Ergebnisse der Interviews werden zudem genutzt, um die Kommunikation sowie Prozesse im weiteren Projektverlauf anzupassen und einen Fragebogen zur quantitativen Erhebung der Zufriedenheit, Machbarkeit und Akzeptanz zu entwickeln. Dieser Fragebogen wird am Ende des Interventionszeitraumes an alle, am WATCH-Projekt teilnehmenden Ärzte ausgehändigt und dient zur Vorbereitung der Verstetigung der neuen Versorgungsform.

2.2 Fokusgruppen mit Mitarbeitenden des PoCo-Busses

Zur Erfassung der Abläufe im Bus und möglichen Optimierungspotentialen erfolgt die Befragung von Mitarbeitenden des UKJ, die an der Arbeit im Bus beteiligt sind. Diese Befragung erfolgt in Form von Fokusgruppeninterviews unter Verwendung eines semi-strukturierten Interviewleitfadens. Die Durchführung und Auswertung orientiert sich an den Ausführungen von Tausch und Menold (2015). Die Fokusgruppensprechungen werden mit einem digitalen Audiogerät aufgezeichnet, transkribiert und letztlich mittels qualitativer Inhaltsanalyse ausgewertet. Diese Erhebung soll primär der Identifikation von Verbesserungsmöglichkeiten im gesamten Prozess der Intervention, in den die

Mitarbeitenden eingebunden sind (z. B. Abläufe, Kommunikation) dienen. Zusätzlich sollen im Projektverlauf wiederholte kurze Befragungen der Mitarbeitenden des PoCo-Busses erfolgen. Hier können im Sinne der Qualitätssicherung eventuelle Prozessanpassungen bewertet werden, die ggf. nach der Fokusgruppendifkussion erfolgt sind, und neu aufkommende Verbesserungsmöglichkeiten identifiziert werden.

2.3 Wahrnehmung von Terminen und Hausarztkonsultationen

Um die Auslastung des Busses über den Interventionsverlauf zu quantifizieren, werden geplante Bustermine („Slots“) mit wahrgenommenen Terminen abgeglichen. Darüber hinaus werden die dokumentierten Gründe für etwaige Ausfälle ausgewertet.

Des Weiteren wird im Bus erfasst, ob und wenn ja wie oft, die Teilnehmenden seit der letzten Visite Ihren Hausarzt konsultiert haben.

3. Qualitative Analyse der Patientenperspektive

Um die Perspektive der Patienten innerhalb von WATCH, inklusive der telemedizinischen Elemente, abzubilden, sind über den gesamten Projektverlauf explorative, qualitative Interviews zum Nutzungsverhalten, zur Adhärenz und zur Akzeptanz sowie den Bedarfen vorgesehen. In der Interviewstudie werden Veränderungen des Nutzungs- und Gesundheitsverhaltens über die Zeit sowie Anforderungen an eine langfristige Adhärenz betrachtet. Hieraus sollen u. a. Maßnahmen zur Steigerung der Nutzungsadhärenz für digitale Lösungen unter Beachtung patientenbezogener, therapiebezogener, sozioökonomischer und Gesundheitssystem-bezogener Faktoren abgeleitet werden.

3.1 Charakterisierung der teilnehmenden Patienten

Die Patienten, die an der Interviewstudie teilnehmen, setzen sich aus einer Stichprobe der Patienten der Früh- und Spätinterventionsgruppe zusammen. Für die Interviews wird ein ausgeglichenes Geschlechterverhältnis innerhalb der Gruppen angestrebt. Da Post-COVID in allen Altersgruppen auftritt und sich die Meinungen und Bedarfe je nach Altersgruppe unterscheiden können, sollten pro Gruppe mindestens 3 Patienten aus jeder der drei Altersklassen 18 bis 39, 40 bis 59 und 60 oder älter vertreten sein.

Es gelten die gleichen Ein- und Ausschlusskriterien wie für das Hauptprojekt. Es sind keine ergänzenden Ein- und Ausschlusskriterien für die Teilnahme an der Interviewstudie zu definieren.

3.2 Rekrutierung

Um verschiedene Wahrnehmungen und Bedarfe im Hinblick auf die neue Versorgungsform ermitteln zu können, sollen bis zur inhaltlichen Sättigung 10-12 Patienten pro

Interventionsgruppe an der Interviewstudie teilnehmen. Es ergibt sich eine Gesamtzahl von 20-24 teilnehmenden Patienten. Die Teilnahme an der Interviewstudie ist freiwillig.

Im Zuge der ersten Vorstellung im Bus (Woche 0; V1a) werden die Patienten über die ergänzende Interviewstudie informiert und aufgeklärt. Besteht Interesse an einer Teilnahme, wird eine Einwilligungserklärung eingeholt. Die Mitarbeitenden im Bus stellen anschließend den Kontakt zu den Interview-Verantwortlichen der DSck) her (Mail und/oder Telefon). Die Absprache der konkreten Interviewtermine erfolgt telefonisch oder per E-Mail direkt durch die, für die Interviews verantwortliche Person. Die Durchführung der Interviews erfolgt, je nach Präferenz des Teilnehmenden, telefonisch oder online per Videokonferenz.

3.3 Interviewdurchführung

Alle teilnehmenden Patienten werden mittels semi-strukturierter Interviews zu drei Zeitpunkten befragt: Woche 3 (+2/-1), Woche 9 (+2/-1), Woche 18 (+2/-1). Die Interviewleitfäden werden unter Berücksichtigung bestehender Theorien erstellt (siehe Abschnitt 3.4). Ziel des ersten Interviews ist es, retrospektiv das Nutzungsverhalten der Patienten in den ersten 1-3 Wochen der Intervention sowie bereits wahrgenommene Veränderungen des eigenen Gesundheitsverhaltens zu erfassen. Veränderungen des Nutzungsverhaltens, des Gesundheitsverhaltens und der Adhärenz über den Interventionsverlauf sowie das geplante Gesundheitsverhalten nach Abschluss der Intervention werden im zweiten Interview erfasst. Das dritte Interview wird während der Nachbeobachtungszeit durchgeführt und zielt auf das tatsächliche Gesundheitsverhalten nach Interventionsende im Vergleich zum geplanten Gesundheitsverhalten und die Nachhaltigkeit der Intervention ab. Dabei werden u. a. Fragen zur Weiternutzung der Neuronation-App® sowie zur Aufzeichnung von Vitaldaten gestellt. Der Messplan ist *Tabelle 3* zu entnehmen.

Tabelle 3: Messplan und Erhebungszeitpunkte der qualitativen Interviewstudie

Messzeitpunkt	Woche 3 (+2/-1)	Woche 9 (+2/-1)	Woche 18 (+2/-1)
Endpunkte			
Nutzungsverhalten	X	X	
Gesundheitsverhalten	X	X	X
Nutzungsadhärenz	X	X	
Akzeptanz		X	
Bedarf		X	
Nachhaltigkeit			X

Die Wahl der Messzeitpunkte weicht bewusst von den Messzeitpunkten des Hauptprojektes (siehe Abschnitt 1.1.3) im PoCo-Bus ab, um eine Überlastung der Patienten zu vermeiden. Die Dauer der Interviews ist auf je 30-45 Minuten angesetzt. Die Interviews werden mit einem Audiogerät aufgezeichnet und anschließend transkribiert.

Die Festlegung der Messzeitpunkte erfolgte im Sinne der partizipativen Studienplanung in enger Abstimmung mit einer Vertreterin der Betroffeneninitiative „Long COVID Deutschland“, die selbst seit mehr als 2 Jahren von Post-COVID betroffen ist. Im Rahmen dieses Dialoges wurde aufgrund der zu erwartenden Erschöpfungszustände und kognitiven Einschränkungen (insbesondere Konzentrationsschwäche) des einzuschließenden Kollektivs beschlossen, dass die Interviewteilnehmenden die Kernfragen der Interviews jeweils etwa eine Woche vor Interviewdurchführung per E-Mail erhalten.

3.4 Auswertung

Im Rahmen der Interviewvorbereitungen werden Hauptkategorien gebildet, welche die Basis der Interviewleitfäden für die semi-strukturierten Interviews darstellen. Die transkribierten Interviews werden entsprechend einer qualitativen Inhaltsanalyse nach Mayring ausgewertet. Hierzu wird das bereits deduktiv gebildete Kategoriensystem um induktiv gebildete Klassen erweitert. Die Kodierung der Interviews erfolgt dabei durch zwei unabhängige Mitarbeitende der DSck. Unterschiede in der Kodierung werden gemeinsam diskutiert. Bei Uneinigkeiten wird ein dritter Mitarbeiter mit Erfahrung in der Auswertung qualitativer Daten hinzugezogen.

4. Gesundheitsökonomische Evaluation der Intervention

4.1 Zweck

Die gesundheitsökonomische Evaluation der neuen Versorgungsform WATCH schließt an den Leitgedanken an, dass mit dieser Intervention die körperliche, kognitive und psychische Gesundheit von Patienten verbessert, die Heilung beschleunigt und die Chronifizierungswahrscheinlichkeit vermindert werden kann. Sie soll die Grundlage für Verhandlungen mit den Krankenkassen zur Überführung der Versorgungsform in die Regelversorgung bilden.

4.2 Datengrundlage

Die gesundheitsökonomische Evaluation beruht maßgeblich auf Routinedaten von drei am Projekt beteiligten, datenliefernden Krankenkassen. Die gesundheitsökonomische Evaluation basiert entsprechend auf jenen Projektteilnehmenden, die bei einer dieser drei Kassen versichert sind. Da diese drei Krankenkassen in Thüringen etwa 70 % des GKV-Marktes abdecken, gehen wir für die gesundheitsökonomische Evaluation von einer Fallzahl von 436 Teilnehmenden aus, die sich durch den eingeplanten Drop-out von 22 % auf etwa

340 Personen verringern wird. Fragestellungen, die sich auf die Arbeitsunfähigkeit beziehen (siehe Abschnitt 5.3), können nur in der Subgruppe der Erwerbstätigen untersucht werden. Für jene Fragestellungen reduziert sich die Fallzahl entsprechend.

Datengrundlage der gesundheitsökonomischen Auswertung bilden die Stamm- und Leistungsdaten der Krankenkassen der folgenden Bereiche:

- Stammdaten der Versicherten (§ 288 SGB V inkl. DMP-Teilnahme chronisch Kranker gemäß § 137f SGB V),
- Daten zur Arbeitsunfähigkeit (§ 295 SGB V),
- Ambulante Abrechnungsdaten (§ 295 SGB V, auch in Verbindung mit selektivvertraglicher Versorgung §§ 73b und 140a SGB V),
- Ambulante Arzneimittelverordnungsdaten (§ 300 SGB V),
- Stationäre Abrechnungsdaten inklusive Krankenhausambulanzen sowie stationäre Rehabilitation (§§ 301, 295 i. V. m. 120, 116 SGB V),
- Heilmittelversorgung bei Ergo-, Physio- und Sprachtherapeuten (§ 302 SGB V),
- Daten der häuslichen Krankenpflege (§ 302 SGB V),
- Rettungsdiensteinsätze und Krankentransporte (§ 302 SGB V),
- Krankengeld (§ 107 SGB IV).

Die in die gesundheitsökonomische Analyse einfließenden Interventionskosten werden geschätzt (siehe Abschnitt 4.5.1).

Des Weiteren geht eine Auswahl primär erhobener Daten (siehe auch Abschnitt 1.) in die Analyse ein. Darunter der PCS und MCS des SF-36 sowie die Anzahl der selbst berichteten Arbeitsunfähigkeitstage in den vergangenen 3 Monaten (sowie die selbst berichtete Erwerbstätigkeit). Diese Daten werden jeweils während der Bus-Visiten V1a, V2 und V3 erhoben.

Ergänzend werden folgende einmalig gemessene Primärdaten in die Analyse einbezogen: Alter, Geschlecht, Landkreis, selbstberichtete Vorerkrankungen, GAF scale, Zeit seit Infektion und Impfstatus. Diese Informationen werden einmalig im Rahmen des Starterpakets vor Beginn der Intervention erhoben.

4.3 Fragestellungen

Im Rahmen der gesundheitsökonomischen Evaluation sollen die folgenden Fragen beantwortet werden:

- (1) Werden die Gesamtversorgungskosten (Leistungsausgaben) durch die Intervention reduziert?
- (2) Wird die Anzahl der Arbeitsunfähigkeitstage durch die Intervention reduziert?

Einem hierarchischen Vorgehen folgend wird nach positiver Beantwortung jeweils eine Anschlussfrage untersucht:

- (1.1) Werden die Interventionskosten durch die eingesparten Gesamtversorgungskosten (Leistungsausgaben, vgl. Frage 1) im Zeitraum von 3 Monaten nach Interventionsbeginn kompensiert (aus der Perspektive der Krankenkassen als Kostenträger)?
- (2.1) Welche inkrementellen Kosten (inklusive der durchschnittlichen Interventionskosten je Frühinterventionsteilnehmer) sind je verhindertem Arbeitsunfähigkeitstag aufzuwenden (aus der Perspektive der Krankenkassen als Kostenträger)?

Sowie, wenn entweder (1) oder (2) positiv beantwortet wurde:

- (1.2) Werden die Interventionskosten durch die eingesparten Gesamtversorgungskosten (Leistungsausgaben) aus der Perspektive der Krankenkassen inklusive des eingesparten arbeitsunfähigkeitsbedingten Bruttowertschöpfungsausfalls im Zeitraum von 3 Monaten nach Interventionsbeginn kompensiert (aus volkswirtschaftlicher Perspektive)?

Die gesundheitsökonomische Evaluation soll darüber hinaus Informationen darüber liefern, welche Kosten für eine Verbesserung der körperlichen beziehungsweise mentalen Gesundheit aufzuwenden sind. Voraussetzung für die Prüfung dieser Frage ist, dass die Intervention eine entsprechende Verbesserung erzielt. Dies wird für das Projekt WATCH im Rahmen der regulären Evaluation, basierend auf den Daten aller Projektteilnehmenden, geprüft. Da die gesundheitsökonomische Evaluation nicht alle Projektteilnehmer betrachten kann (siehe Abschnitt 4.2), muss in der für die gesundheitsökonomische Evaluation vorliegenden Stichprobe zunächst geprüft werden:

- (3) Wird die körperliche Gesundheit (PCS) durch die Intervention verbessert?
- (4) Wird die mentale Gesundheit (MCS) durch die Intervention verbessert?

Im Sinne eines hierarchischen Vorgehens werden nach positiver Beantwortung im Anschluss die bereits angesprochenen Fragen untersucht:

Welche inkrementellen Kosten (inklusive der durchschnittlichen Interventionskosten je Frühinterventionsteilnehmer) sind je hinzugewonnenem Punkt auf der PCS-Skala aufzuwenden (aus der Perspektive der Krankenkassen als Kostenträger)?

- (4.1) Welche inkrementellen Kosten (inklusive der durchschnittlichen Interventionskosten je Frühinterventionsteilnehmer) sind je hinzugewonnenem

Punkt auf der MCS-Skala aufzuwenden (aus der Perspektive der Krankenkassen als Kostenträger)?

4.4 Zentrale gesundheitsökonomische Zielindikatoren

4.4.1 Kostenindikatoren

Aufgrund der weitreichenden Ausstrahlungseffekte der Post-COVID-Erkrankung auf alle Lebensbereiche der Betroffenen (bis hin zur potenziell eingeschränkten medizinischen Leistungsanspruchnahme aufgrund von anderen Erkrankungen) steht bewusst die Betrachtung der Gesamtversorgungskosten (Leistungsausgaben) im Vordergrund. Erst in zweiter Linie wird die Abgrenzung einzelner, potenziell Post-COVID-assoziiertes Leistungen und damit verbundener Versorgungskosten (Leistungsausgaben) vorgenommen und zwischen den beiden Gruppen verglichen.

Zentrale Zielgröße sind die Gesamtversorgungskosten (Leistungsausgaben) je Projektteilnehmer (aus der Perspektive der Krankenkassen als Kostenträger), operationalisiert als Summe der Kosten aller pro Person in den folgenden Leistungsbereichen in Anspruch genommenen Leistungen (jeweils in Euro):

1. ambulante ärztliche (und psychologisch-psychotherapeutische) Behandlung,
2. stationäre Krankenhausbehandlung,
3. ambulante Krankenhausbehandlung,
4. stationäre (Anschluss-)Reha-Behandlung¹,
5. Rettungsdienst- und Krankentransporte,
6. Arzneimittel,
7. Heilmittel,
8. Hilfsmittel,
9. häusliche Krankenpflege.

Das (arbeitsunfähigkeitsveranlasste) Krankengeld wird separat zu den Gesamtversorgungskosten (Leistungsausgaben) berichtet, weil es sich dabei um eine Transferleistung handelt, die nicht zu den direkten Versorgungskosten aus der Perspektive der Krankenkassen zählt.

Eine Abgrenzung der Daten erfolgt anhand des Verordnungs- beziehungsweise Leistungstags, -monats beziehungsweise -quartals der abgerechneten Leistung. Bei nicht tagesgenau dokumentierten Leistungen wird eine Zuordnung nach dem Datum des Leistungsbeginns beziehungsweise -endes beziehungsweise eine anteilige Zuordnung zum Bezugszeitraum vorgenommen. Ebenso erfolgt die jeweilige Kostenabgrenzung. Nach

¹ Eingeschlossen werden können ausschließlich Leistungen, für die die Krankenkasse als Kostenträger fungiert. Berufliche Rehabilitationsmaßnahmen können nicht abgebildet werden.

Möglichkeit werden Nettokosten, d. h. Kosten nach Abzug versichertenseitiger Zuzahlungen betrachtet. Die Daten enthalten Abrechnungsbeträge in Euro mit Ausnahme der ambulanten ärztlichen Leistungen, die der Budgetierung unterliegen. Zur Ermittlung der Kosten werden die hier angegebenen Punktwerte je Leistung mit den jeweils geltenden (bundesdurchschnittlichen) Orientierungswerten multipliziert.

Für ein tieferes Verständnis der Kosten erfolgt darüber hinaus eine explorative Bezifferung der Inanspruchnahme in den genannten Leistungsbereichen im Sinne von Leistungsmengen beziehungsweise der Inanspruchnahme potenziell Post-COVID-assoziiertes Leistungen, darunter:

- Anzahl Arztkontakte,
- Anzahl Krankenhausaufenthalte, Krankenhaustage (ggf. bei spezifischen dokumentierten Diagnosen), Notaufnahmen,
- Krankengeldtage,
- Inanspruchnahme spezifischer Fachspezialisten (insbesondere: Hausärzte, Fachärzte für innere Medizin, Kardiologen, Pneumologen, HNO-Ärzte, Neurologen, Psychiater),
- Inanspruchnahme psychotherapeutischer Behandlung,
- Inanspruchnahme Kassenärztlicher Notdienst,
- Inanspruchnahme bestimmter Arzneimittelwirkstoffe (z. B. Inhalativa),
- Inanspruchnahme von Psychiatrischen Institutsambulanzen,
- Einsatz bildgebender Verfahren,
- Einsatz von Labordiagnostik,
- Einsatz spezifischer Heilmittel (u. a. Atemtherapie).

4.4.2 Arbeitsunfähigkeitstage

Arbeitsunfähigkeitstage werden sekundär erhoben. Gemessen werden die pro Person in den Kassendaten dokumentierten Arbeitsunfähigkeitstage.

Ergänzend, mit dem Ziel der Validierung, werden auch die per Selbstbericht primär erhobenen Arbeitsunfähigkeitstage zu den drei Zeitpunkten V1a, V2 und V3 jeweils mit Zeitbezug auf die zurückliegenden 3 Monate betrachtet (*Tabelle 2*). Bei augenscheinlichen Abweichungen im Vergleich zu den in den Kassendaten dokumentierten Arbeitsunfähigkeitstagen werden die Analysen unter Berücksichtigung der selbstberichteten Arbeitsunfähigkeitstage wiederholt.

4.5 Weitere für die gesundheitsökonomische Evaluation relevante Indikatoren

4.5.1 Interventionskosten

Die Schätzung der Kosten pro durchschnittlichem Patient beruht zum einen auf den Kosten, welche durch die Intervention selbst entstehen („eigentliche Interventionskosten“, z. B.

Personalkosten, Materialkosten, Betriebskosten) sowie den Kosten, die durch die Einbindung der Ärzte entstehen.

Für die im Rahmen der WATCH-Intervention vorgesehenen Leistungen durch Haus- und Fachärzte wurden Leistungspauschalen festgelegt. Bei Überführung der Versorgungsform in die Regelversorgung würde die Bewertung der haus- und fachärztlichen Leistungen durch den Gemeinsamen Bundesausschuss (G-BA) erfolgen. Für die Projektdauer rechnen die Haus- und Fachärzte die erbrachten Leistungen über die KVT ab, die wiederum die entsprechenden Kosten aus den Projektmitteln zurückerstattet bekommen. Entsprechend liegen dem Evaluator Übersichten über die bei der KVT abgerechneten Leistungen vor (siehe auch Versorgungsvertrag, Anlage 3). Hierbei handelt sich um drei von den Hausärzten abrechenbare Leistungskomplexe (Leistungskomplex 1: Einschreibung/Basis-Assessment, Leistungskomplex 2: Überleitung an das PoCo-Zentrum und Leistungskomplex 3: Patientenbegleitung/Motivation) und einen von Fachärzten abrechenbaren Leistungskomplex (Leistungskomplex Differentialdiagnostik).

Im Rahmen der eigentlichen Interventionskosten erfolgt die Kostenschätzung für zwei Szenarien: einmal konkret für das WATCH-Projekt (Szenario 1) und zum anderen hypothetisch bei Überführung der Intervention in die Regelversorgung (Szenario 2). Für beide Szenarien werden die Kosten pro durchschnittlichem Patient geschätzt. Für Szenario zwei werden die eigentlichen Interventionskosten plus abgerechnete Leistungen – basierend auf den für WATCH festgelegten Leistungspauschalen – pro durchschnittlichem Patient pro Jahr geschätzt. Die erwartete Patientenzahl soll hierbei auf den aktuellsten Prävalenzschätzungen von Post-COVID beruhen.

4.5.2 Arbeitsunfähigkeits-bedingter Bruttowertschöpfungsausfall

Die durchschnittliche Bruttowertschöpfung betrug gemäß Berechnungen der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) für das Jahr 2021 219 Euro pro Tag (BAuA, 2022). Die Zahlen werden jährlich fortgeschrieben. Der in die Analyse eingehende arbeitsunfähigkeitsbedingte Bruttowertschöpfungsausfall wird mit dem zum Auswertungszeitpunkt veröffentlichten aktuellen Betrag in Euro je Arbeitsunfähigkeitstag beziffert.

4.6 Beobachtungszeiträume

Die Zielindikatoren (siehe Abschnitt 4.4) werden betrachtet für drei Zeiträume:

- Vorbeobachtungszeitraum (= 3-Monats-Zeitraum vor Beginn der Frühintervention)
- Evaluationszeitraum (= 3-Monats-Zeitraum nach Beginn der Frühintervention)
- Nachbeobachtungszeitraum (= 3-Monats-Zeitraum nach Ende der Frühintervention)

Die drei Zeiträume werden dabei durch folgende Zeitpunkte abgegrenzt, wobei T0 V1a entspricht:

- Vorbeobachtung: T0 minus 12 Wochen bis T0
- Evaluation: T0 bis T0 plus 12 Wochen
- Nachbeobachtung: T0 plus 12 Wochen bis T0 plus 24 Wochen

Die Bezugszeiträume werden pro Person berechnet, basierend auf dem individuellen T0-Termin. Mit dem Vorgehen, die Zeiträume ausschließlich basierend auf T0 abzugrenzen, soll sichergestellt werden, dass durchweg gleich große Zeiträume betrachtet werden. Wenn die tatsächlichen Bus-Visiten V2 und V3, etwa bedingt durch individuelle Terminvergabe, deutlich von den genannten Zeitpunkten abweichen, erfolgt in explorativen Zusatzanalysen die Abgrenzung der Zeiträume basierend auf den tatsächlichen Zeitpunkten der Bus-Visiten.

In Routinedatenanalysen können kleine Beobachtungszeiträume problematisch sein. Um die Inanspruchnahme vor der Intervention zuverlässig abschätzen zu können, wird der Vorbeobachtungszeitraum von 3 Monaten möglicherweise nicht ausreichen. Darum wird für explorative Anschlussanalysen ein erweiterter Vorbeobachtungszeitraum von 12 Monaten definiert (T0 minus 364 Tage).

4.7 Statistische Analyse

Zur Beantwortung von Fragen 1 und 2 werden die im Abschnitt 4.4 beschriebenen Zielindikatoren für den Evaluationszeitraum sowie für den Vorbeobachtungszeitraum betrachtet. Dem Vorgehen des Gesamtprojektes folgend werden intra-individuelle Differenzen bestimmt, also die Veränderung der Zielindikatoren zwischen diesen beiden Zeiträumen. Diese Differenzen werden als metrische Outcome-Variablen in gemischten Regressionsmodellen mit Gruppenzugehörigkeit (Intervention/Kontrolle) und dem Wert des Zielindikators im Vorbeobachtungszeitraum als feste Effekte sowie dem Landkreisstratum als Zufallseffekt analysiert. Sollte sich zwischen Interventions- und Kontrollgruppe hinsichtlich relevanter Variablen – Alter, Geschlecht, selbstberichteter Vorerkrankungen, Allgemeinbefinden während der akuten SARS-CoV-2-Infektion (GAF scale), Impfstatus und Allgemeinbefinden im Hinblick auf die Post-COVID-Erkrankung zur Baseline (GAF scale) – eine Dysbalance zeigen, werden die entsprechenden Variablen in der Analyse als Adjustierungsgrößen berücksichtigt. Explorativ wird eine Auswahl dieser Variablen auch für Subgruppenanalysen in Erwägung gezogen. Die gesundheitsökonomische Evaluation folgt hier dem Vorgehen der Evaluation des Gesamtprojektes (siehe Abschnitt 1.3.3.3).

Um die Validität der per Selbstbericht gemessenen Vorerkrankungen zu prüfen, werden diese zusätzlich auch über die Krankenkassendaten sekundär erhoben. Bei augenscheinlichen Abweichungen wird die Analyse mittels der in den Kassendaten dokumentierten Vorerkrankungen wiederholt, um Schlussfolgerungen abzusichern.

Eine Betrachtung intra-individueller Differenzen kann in Kostenanalysen aufgrund der kleinen Beobachtungszeiträume problematisch sein. In explorativen Anschlussanalysen erfolgt darum ein (querschnittlicher) Vergleich von Interventions- und Kontrollgruppe hinsichtlich der Zielindikatoren im Evaluationszeitraum adjustiert für die Inanspruchnahme im erweiterten Vorbeobachtungszeitraum (12 Monate vor der Intervention).

Zur Beantwortung der Fragestellungen 3 und 4 wird analog zum Gesamtprojekt vorgegangen. Entsprechend werden intra-individuelle Unterschiede in den PCS-/MCS-Werten am Ende (V2) und am Anfang (V1a) der Intervention gebildet und als metrische Outcome-Variablen in einem gemischten Regressionsmodell mit Gruppenzugehörigkeit (Intervention/Kontrolle) und dem PCS-/MCS-Wert am Anfang der Intervention (V1a) als feste Effekte und dem Landkreisstratum als Zufallseffekt analysiert. Hinsichtlich Adjustierung und Subgruppenanalyse wird vorgegangen wie zuvor beschrieben.

Zur Beantwortung von Fragestellung 1.1 werden, getrennt für Interventions- und Kontrollgruppe, die mittleren Gesamtversorgungskosten (Leistungsausgaben) im Evaluationszeitraum und im Vorbeobachtungszeitraum je Patienten ermittelt. Sollte eine Adjustierung nötig sein, wird hier mit adjustierten Vorhersagen des entsprechenden Regressionsmodells gearbeitet. Für beide Gruppen wird nun jeweils die Differenz berechnet und schließlich die Differenz der Kontrollgruppe von der der Interventionsgruppe subtrahiert. Das Ergebnis, die eingesparten Gesamtversorgungskosten (Leistungsausgaben), wird (rein deskriptiv) mit den Interventionskosten pro durchschnittlichem Patienten verglichen.

Zur Beantwortung der Fragestellungen 2.1, 3.1 und 4.1 wird jeweils eine inkrementelle Kosten-Effektivitäts-Relation (ICER) errechnet. Hierzu wird jeweils ein Quotient gebildet, mit der Differenz der Gesamtversorgungskosten (Leistungsausgaben) inklusive Interventionskosten von Interventions- vs. Kontrollgruppe im Zähler, geteilt durch folgenden Effekt der Intervention:

- Fragestellung 2.1: Differenz bzgl. Veränderung der AU-Tage,
- Fragestellung 3.1: Differenz bzgl. Veränderung der körperlichen Gesundheit (PCS),
- Fragestellung 4.1: Differenz bzgl. Veränderung der mentalen Gesundheit (MCS).

Zur Beantwortung der Fragestellung 1.2 wird geprüft, ob die inkrementellen Kosten je verhindertem Arbeitsunfähigkeitstag (für Fragestellung 2.1 berechnete ICER) kleiner sind als der Bruttowertschöpfungsausfall je Arbeitsunfähigkeitstag (siehe Abschnitt 4.5.2).

Für alle im Rahmen der gesundheitsökonomischen Evaluation untersuchten Fragestellungen gilt, analog zum Gesamtprojekt, dass nur diejenigen Personen in die Analyse eingehen, für die die jeweils benötigten Daten vollständig vorliegen.

4.8 Ergänzende Analysen: Langzeiteffekte und Effekt der Spätintervention

Analog zum Gesamtprojekt (siehe Abschnitt 1.3.5) erfolgt zusätzlich eine Auswertung der zentralen Zielindikatoren im Nachbeobachtungszeitraum und zwar getrennt für beide Gruppen:

- Frühinterventionsgruppe: Leitend ist hier die Frage, ob die Effekte der Intervention anhalten. Dabei ist kein Vergleich mit der Kontrollgruppe möglich, da diese zu dieser Zeit bereits ihre Intervention erhält.
- Spätinterventionsgruppe: Der Effekt der Spätintervention soll beziffert werden. Auch hier ist kein Vergleich mit einer Kontrollgruppe möglich.

Zusätzlich sind explorative Analysen geplant über den Verlauf der Inanspruchnahme seit der selbst berichteten COVID-Infektion bzw. seit der Dokumentation der Post-COVID-Erkrankung bis 24 Wochen nach Interventionsbeginn (Frühinterventionsgruppe) bzw. 12 Wochen nach Interventionsbeginn (Spätinterventionsgruppe).

Literaturverzeichnis

- Altman, D. G. (1985). Comparability of Randomised Groups. *The Statistician*, 34(1), 125. Doi: 10.2307/2987510
- Altman, D. G. & Doré, C. J. (1990). Randomisation and baseline comparisons in clinical trials. *Lancet*, 335(8682), 149–153. Doi: 10.1016/0140-6736(90)90014-v
- Badhiwala, J. H., Witiw, C. D., Nassiri, F., Akbar, M. A., Jaja, B., Wilson, J. R. & Fehlings, M. G. (2018). Minimum Clinically Important Difference in SF-36 Scores for Use in Degenerative Cervical Myelopathy. *Spine*, 43(21), E1260-E1266. Doi: 10.1097/BRS.0000000000002684
- Broadbent, D. E., Cooper, P. F., FitzGerald, P. & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *The British Journal of Clinical Psychology*, 21(1), 1–16. Doi: 10.1111/j.2044-8260.1982.tb01421.x
- Bullinger, M., Kirchberger, I. & Ware, J. (1995). Der deutsche SF-36 Health Survey Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität. *Journal of Public Health*, 3(1), 21–36. Doi: 10.1007/BF02959944
- Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (2022). *Volkswirtschaftliche Kosten durch Arbeitsunfähigkeit 2021*. Verfügbar unter: <https://www.baua.de/DE/Themen/Monitoring-Evaluation/Zahlen-Daten-Fakten/Kosten-der-Arbeitsunfaehigkeit.html>
- Carreon, L. Y., Glassman, S. D., Campbell, M. J. & Anderson, P. A. (2010). Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. *The Spine Journal*, 10(6), 469–474. Doi: 10.1016/j.spinee.2010.02.007
- Cochrane Deutschland, Institut für Medizinische Biometrie und Statistik Freiburg, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften - Institut für Medizinisches Wissensmanagement & Ärztliches Zentrum für Qualität in der Medizin (2021). *Manual Bewertung des Biasrisikos in Interventionsstudien (2. Aufl.)*. Doi: 10.6094/UNIFR/194900
- Copay, A. G., Glassman, S. D., Subach, B. R., Berven, S., Schuler, T. C. & Carreon, L. Y. (2008). Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *The Spine Journal*, 8(6), 968–974. Doi: 10.1016/j.spinee.2007.11.006
- Corbett, M. S., Higgins, J. P. T. & Woolacott, N. F. (2014). Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Research Synthesis Methods*, 5(1), 79–85. Doi: 10.1002/jrsm.1090
- Cotler, J., Holtzman, C., Dudun, C. & Jason, L. A. (2018). A Brief Questionnaire to Assess Post-Exertional Malaise. *Diagnostics*, 8(3), 66. Doi: 10.3390/diagnostics8030066
- Demeyere, N., Haupt, M., Webb, S. S., Strobel, L., Milosevich, E. T., Moore, M. J., Wright, H., Finke, K. & Duta, M. D. (2021). Introducing the tablet-based Oxford Cognitive Screen-Plus (OCS-Plus) as an assessment tool for subtle cognitive impairments. *Scientific Reports*, 11(1), 8000. Doi: 10.1038/s41598-021-87287-8
- Franke, G. H. & Jagla-Franke, M. (2021). *EFK - Essener Fragebogen zur Krankheitsverarbeitung*. Doi: 10.23668/psycharchives.4893

- Hall, R. C. (1995). Global assessment of functioning. A modified scale. *Psychosomatics*, 36(3), 267–275. Doi: 10.1016/S0033-3182(95)71666-8
- Han, Q., Zheng, B., Daines, L. & Sheikh, A. (2022). Long-Term Sequelae of COVID-19: A Systematic Review and Meta-Analysis of One-Year Follow-Up Studies on Post-COVID Symptoms. *Pathogens*, 11(2), 269. Doi: 10.3390/pathogens11020269
- Hasselhorn, H.-M. & Freude, G. (2007). *Der Work Ability Index: Ein Leitfaden*. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin / Sonderschrift: S 87. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAUA). Verfügbar unter: <https://www.baua.de/DE/Angebote/Publikationen/Schriftenreihe/Sonderschriften/S87.html>
- Hinz, A., Glaesmer, H., Brähler, E., Löffler, M., Engel, C., Enzenbach, C., Hegerl, U. & Sander, C. (2017). Sleep quality in the general population: psychometric properties of the Pittsburgh Sleep Quality Index, derived from a German community sample of 9284 people. *Sleep Medicine*, 30, 57–63. Doi: 10.1016/j.sleep.2016.03.008
- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (1998). *ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9* (4. Aufl.). Verfügbar unter: <https://www.ich.org/page/efficacy-guidelines>
- Kroenke, K., Spitzer, R. L. & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. Doi: 10.1046/j.1525-1497.2001.016009606.x
- Marcus, R., Peritz, E. & Gabriel, K. R. (1976). On Closed Testing Procedures with Special Reference to Ordered Analysis of Variance. *Biometrika*, 63(3), 655. Doi: 10.2307/2335748
- McCoy, C. E. (2018). Understanding the Use of Composite Endpoints in Clinical Trials. *The Western Journal of Emergency Medicine*, 19(4), 631–634. Doi: 10.5811/westjem.2018.4.38383
- Michielsen, H. J., Vries, J. de, Van Heck, G. L., Van de Vijver, F. J. & Sijtsma, K. (2004). Examination of the Dimensionality of Fatigue. *European Journal of Psychological Assessment*, 20(1), 39–48. Doi: 10.1027/1015-5759.20.1.39
- Morfeld, M., Kirchberger, I. & Bullinger, M. (2011). SF-36 Fragebogen zum Gesundheitszustand: Deutsche Version des Short Form-36 Health Survey. In: *SF-36 Fragebogen zum Gesundheitszustand. Deutsche Version des Short Form-36 Health Survey* (2. Aufl., S. 1-221). Hogrefe-Verlag. Verfügbar unter: [https://fis-uke.de/portal/de/publications/sf36-fragebogen-zum-gesundheitszustand-deutsche-version-des-short-form36-health-survey\(4931d936-6b7c-4b49-b5ab-600b27e65e99\).html](https://fis-uke.de/portal/de/publications/sf36-fragebogen-zum-gesundheitszustand-deutsche-version-des-short-form36-health-survey(4931d936-6b7c-4b49-b5ab-600b27e65e99).html)
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L. & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. Doi: 10.1111/j.1532-5415.2005.53221.x
- Roberts, C. & Torgerson, D. J. (1999). Understanding controlled trials: baseline imbalance in randomised controlled trials. *BMJ*, 319, 185. Doi: 10.1136/bmj.319.7203.185
- Rochmawati, E., Iskandar, A. C. & Kamilah, F. (2022). Persistent symptoms among post-COVID-19 survivors: A systematic review and meta-analysis. *Journal of Clinical Nursing*, 00, 1-11. Doi: 10.1111/jocn.16471

- Sevillano-Castaño, A. I., Peroy-Badal, R., Torres-Castro, R., Cañuelo-Márquez, A. M., Rozalén-Bustín, M., Modrego-Navarro, Á., Sousa-De Sousa, L. de, Ramos-Álvarez, J. J., Maté-Muñoz, J. L. & García-Fernández, P. (2023). Test–Retest Reliability and Minimal Detectable Change in Chester Step Test and 1-Minute Sit-to-Stand Test in Long COVID Patients. *Applied Sciences*, *13*(14), 8464. Doi: 10.3390/app13148464
- Stallmach, A., Katzer, K., Besteher, B., Finke, K., Gizzas, B., Gremme, Y., Abou Hamdan, R., Lehmann-Pohl, K., Legen, M., Lewejohann, J. C., Machnik, M., Moshmoh Alsabbagh, M., Nardini, L., Puta, C., Stallmach, Z. & Reuken, P. A. (2023). Mobile primary healthcare for post-COVID patients in rural areas: a proof-of-concept study. *Infection*, *51*(2), 337–345. Doi: 10.1007/s15010-022-01881-0
- Tausch, A. P. & Menold, N. (2015). Methodische Aspekte der Durchführung von Fokusgruppen in der Gesundheitsforschung: welche Anforderungen ergeben sich aufgrund der besonderen Zielgruppen und Fragestellungen? *GESIS Papers 2015/12*. Köln: GESIS – Leibnitz-Institut für Sozialwissenschaften. Doi: 10.21241/ssoar.44016
- Tritt, K., Heymann, F. von, Zaudig, M., Zacharias, I., Söllner, W. & Loew, T. (2008). Entwicklung des Fragebogens "ICD-10-Symptom-Rating" (ISR). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, *54*(4), 409–418. Doi: 10.13109/zptm.2008.54.4.409

Halle, den 12.12.2023

Unterschrift des Evaluators

A handwritten signature in blue ink, appearing to read 'Mikolajczyk', written in a cursive style.

Prof. Dr. med. Rafael Mikolajczyk, MSc