

**Teil II: Eingehende Darstellung****PhysioAI: Bewertung und Übersetzung multimodaler Messinstrumente physischer Aktivität mittels Künstlicher Intelligenz zur Behandlung von Arthrose - PhysioAI; Teilvorhaben: Universitätsklinikum Hamburg Eppendorf**

Förderkennzeichen: 16DKWN115B

Laufzeit: 01.10.2022 bis 30.09.2025

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung, Technologie und Raumfahrt unter den Förderkennzeichen 16DKWN115 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei dem/der Autor/in bzw. den Autor/inn/en.

Finanziert durch die Europäische Union – NextGenerationEU. Die geäußerten Ansichten und Meinungen sind ausschließlich die des Autors/der Autoren und spiegeln nicht unbedingt die Ansichten der Europäischen Union oder der Europäischen Kommission wieder. Weder die Europäische Union noch die Europäische Kommission können für sie verantwortlich gemacht werden

Inhaltsverzeichnis

1	Aufgabenstellung und wissenschaftlicher und technischer Stand, an den angeknüpft wurde	2
2	Planung und Ablauf des Vorhabens	2
3	Inhaltliche Ergebnisse.....	7
4	Wichtigste Positionen des zahlenmäßigen Nachweises	12
5	Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten	12
6	Voraussichtlicher Nutzen, Verwertbarkeit der Ergebnisse und zukünftige Planungen im Sinne des Verwertungsplans.....	13
7	Fortschritt auf dem Gebiet des Vorhabens während der Durchführung des Vorhabens bei anderen Stellen	13
8	Erfolgte oder geplante Veröffentlichungen der Projektergebnisse	13
9	Literaturverzeichnis.....	14



1 Aufgabenstellung und wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Mit dem Projekt PhysioAI verfolgte das Konsortium das Ziel, Methoden der Data Science auf Basis von Künstlicher Intelligenz (KI) speziell für Anwendungen in der Physiotherapie zu entwickeln und zu validieren – einem Bereich, in dem solche Ansätze zuvor noch nicht etabliert waren. Der Schwerpunkt des Projekts lag auf der Untersuchung, wie KI-basierte Techniken dazu beitragen können, Behandlungsstrategien für Patienten mit Arthrose (Osteoarthritis, OA) zu optimieren.

Das übergreifende Ziel bestand darin, eine präzise und effiziente Messung körperlicher Aktivität sowohl im Forschungs- als auch im klinischen Kontext zu ermöglichen. Dadurch sollte erstmals eine kohortenübergreifende Kompatibilität und quantitative Vergleichbarkeit von Arthrose-Behandlungsprotokollen über verschiedene Studien hinweg sichergestellt werden. Durch diese Arbeit legte PhysioAI den Grundstein für ein neues Forschungsfeld, die „Computational Physiotherapy“ (computergestützte Physiotherapie), und verankerte diese nachhaltig in der wissenschaftlichen Gemeinschaft durch die Entwicklung und Implementierung KI-basierter Werkzeuge zur Messung, Bewertung und Optimierung physiotherapeutischer Interventionen.

Darüber hinaus befasste sich das Projekt mit der Herausforderung, verschiedene Metriken aus Gesundheitsumfragen zu übersetzen und zu harmonisieren, um die Vorhersage körperlicher Aktivität und die Vergleichbarkeit zwischen verschiedenen Messinstrumenten zu ermöglichen. Diese Entwicklungen schaffen die Basis für einen zuverlässigen Transfer und das Benchmarking von Studienergebnissen und Behandlungserfolgen über verschiedene Datensätze hinweg. Vor PhysioAI waren die Datenkompetenz (Data Literacy) und die Fähigkeiten zur KI-Anwendung in der physiotherapeutischen Forschung begrenzt; das Projekt adressierte diese Lücke direkt und trug durch die gezielte Entwicklung KI-gestützter Analysemethoden zur Modernisierung und digitalen Transformation des Fachbereichs bei.

2 Planung und Ablauf des Vorhabens

Arbeitspaket 1: Daten-Vorbereitung, Vorverarbeitung und Aufreinigung

Aufgabe 1.1 – MySQL-Datenbankerstellung (Leitung: UKE, Teilnahme: UHH)

Wir am UKE haben eine leere MySQL-Datenbank mit fünf leeren Tabellen (AP 1.1) erstellt, um die jeweiligen Datensätze darin zu integrieren. Wie später jedoch ersichtlich wird, haben wir uns aus praktischen Gründen für die Verwendung von Pandas-Dataframes anstelle von MySQL entschieden.

M1.1 Datenmanagementplan (UKE)

Wir haben einen Datenmanagementplan erstellt, der mit unseren Partnern an der UHH abgestimmt wurde.

M1.2 MySQL-Datenbankschema (UKE, UHH)

In Zusammenarbeit mit unserem Partner an der UHH haben wir ein MySQL-Datenschema entwickelt. Dieses Schema kann genutzt werden, um alle Variablen jedes Datensatzes nacheinander in eine MySQL-Datenbank oder eine andere Datenstruktur, wie z. B. einen Pandas-Dataframe, zu integrieren.

Aufgabe 1.2 – Integration von GLA:D-Kohortendaten (Leitung: UKE)

Wir am UKE haben zudem einen Antrag für den GLA:D®-Datensatz gestellt und den Zugang dazu erhalten (AP 1.2).

Aufgabe 1.3 – Integration von MOST-Kohortendaten (Leitung: UKE)

Anschließend haben wir am UKE einen Antrag für den MOST-Datensatz gestellt (AP 1.3). Aufgrund von Unterschieden in den Datenschutz- und Ethikbestimmungen zwischen den USA und der EU haben wir diese Daten jedoch nicht erhalten. Wir haben das Projekt daher mit den anderen vier Datensätzen fortgesetzt.

M1.3 Data-Warehouse-Pipeline (UKE)



Analog zu den anderen Verfahren haben wir die Datensätze OAI (AP 1.4), SHARE (AP 1.5) und DEAS (AP 1.6) beantragt und erhalten.

Aufgabe 1.4 – Integration von OAI-Kohortendaten (Leitung: UKE)

Am UKE haben wir die Rohdaten beantragt und erhalten. Diese wurden später zur Erstellung eines Pandas-Dataframes verwendet, der als Grundlage für alle nachfolgenden Aufgaben in Python diente.

Aufgabe 1.5 – Integration von SHARE-Kohortendaten (Leitung: UKE)

Am UKE haben wir die Rohdaten beantragt und erhalten. Diese wurden später zur Erstellung eines Pandas-Dataframes verwendet, der als Grundlage für alle nachfolgenden Aufgaben in Python diente.

Aufgabe 1.6 – Integration von DEAS-Kohortendaten (Leitung: UKE)

Am UKE haben wir die Rohdaten beantragt und erhalten. Diese wurden später zur Erstellung eines Pandas-Dataframes verwendet, der als Grundlage für alle nachfolgenden Aufgaben in Python diente.

M1.4 PhysioAI-Datenbank (UKE, UHH)

Wir am UKE haben die Daten von GLA:D, OAI, SHARE und DEAS in einem einzigen Pandas-Dataframe zusammengeführt. Die Größe dieses Dataframes war zu umfangreich für eine einzelne MySQL-Tabelle (in einer MySQL-Datenbank ist die Anzahl der Spalten pro Tabelle auf maximal 4.096 begrenzt). Daher beabsichtigen wir, die Daten für unsere praktischen Anwendungsfälle in der Python-basierten Downstream-Analyse als Pandas-Dataframe beizubehalten.

Zusammen mit den Partnern der UHH haben wir am UKE eine funktionsfähige Data-Warehouse-Pipeline entwickelt. Wir verfügen über Skripte für diese Pipeline, mit denen Nutzer, die über einen entsprechenden Zugang zu diesen Datenbanken (GLA:D, OAI, DEAS, SHARE) verfügen, ihr eigenes integriertes Data Warehouse generieren können.

Arbeitspaket 2: Entwicklung künstlicher Intelligenzen zur Kompatibilisierung von Messinstrumenten zur Erfassung physischer Aktivität

Aufgabe 2.1 – Implementierung einer Methoden-Pipeline (Leitung: UHH, Teilnahme: UKE)

Wir haben mit dem Partner der UHH zusammengearbeitet, um eine Evaluations-Pipeline zu entwickeln. Diese kann um beliebige Machine-Learning-Methoden erweitert werden, um Metriken wie R-Squared, RMSE, Genauigkeit (Accuracy) und AUROC zu berechnen.

M2.1 Pipeline-Evaluierung (UHH, UKE)

Dieselbe Pipeline wurde zunächst in GLA:D eingesetzt und anschließend in SHARE für die Vorhersage körperlicher Inaktivität sowie für weitere Aufgaben verwendet.

Aufgabe 2.2 – Transfer-Learning-Ansatz (Entwicklung und Implementierung in PyTorch) (Leitung: UHH, UKE)

In der Anfangsphase haben wir am UKE unseren Partner an der UHH dabei unterstützt, die Variablen zu identifizieren, die im Zusammenhang mit körperlicher Aktivität stehen. Zudem haben wir die Vergleichbarkeit (Verisimilitude) bestimmter Fragebögen über vier Datenbanken hinweg festgestellt, welche sich auf das Niveau der körperlichen Aktivität, Schmerzen, Lebensqualität etc. beziehen. Diese Informationen wurden später für Wissenstransfer-Aufgaben in SHARE, DEAS, GLA:D und OAI genutzt (AP 2.2).

Im Jahr 2024 erprobten UKE und UHH gemeinsam verschiedene Ansätze für den Wissenstransfer von einer Datenkohorte auf eine andere. Der Partner am UKE versuchte, ein generalisiertes Framework zu erstellen, um eine beliebige Datenbank auf eine andere abzubilden (**unidirektionaler DB-Konverter**). Hierfür begann das UKE zunächst mit synthetischen und halbsynthetischen Daten. Die UHH nutzte diese Methode, um sie auf reale Datensätze von SHARE und DEAS anzuwenden (AP 2.2).

Zur Entwicklung dieses generalisierten Frameworks führte das UKE zahlreiche Experimente bezüglich der Architektur, der Loss-Funktionen und des Hyperparameter-Tunings durch. Bei der Anpassung an die Realdaten musste die UHH geringfügige Änderungen an der Architektur (basierend auf der Input- und Output-Form) sowie ein minimales Hyperparameter-Tuning vornehmen, während die Loss-Komponenten unverändert blieben. Das UKE entwickelte später zudem eine verbesserte Technik für das Deep-Learning-basierte Mapping, welche die UHH später für Realdaten nutzen kann (AP 2.2). Die entsprechende Publikation befindet sich derzeit im Review-Prozess.

M2.2 Evaluierung und Implementierung des Transfer Learnings (UHH, UKE)

Das UKE entwickelte zudem verschiedene Evaluierungsmethoden, um die Qualität des Wissenstransfers (Knowledge Mapping) sowohl auf Populations- als auch auf Individualebene zu



überprüfen. Hierbei kamen der **L1- und L2-Fehler**, **PCA-Plots** (Hauptkomponentenanalyse) sowie **Korrelationsdiagramme** zum Einsatz. Diese Methoden wurden von der UHH genutzt, um die Ergebnisse für die Realdaten zu evaluieren (AP 2.3).

Aufgabe 2.3 Erweiterung der neuronalen Netzwerke und Fusion der Einzelmodelle (Leitung: UHH)

Meilenstein M2.3 Implementierung und Evaluierung der Fusions-Ansätze (UHH)

Aufgabe 2.4 – Registrierung der KI-Methodik aus Arbeitspaket 2 (Leitung: UKE)

Das generalisierte Transfer-Learning-Modell wurde in AIME registriert.

M2.4 AIME-Registrierung (UKE)

Dasselbe Modell aus AP 2.3 ist in AIME registriert. Die Registrierung ist unter folgendem Link einsehbar:

<https://aime.report/fbSPBM>

Arbeitspaket 3: Entwicklung künstlicher Intelligenzen zur Evaluation von individueller physischer Aktivität aus funktionalen Test-Daten und Patientendaten am Beispiel von Arthrose

Aufgabe 3.1 - Entwicklung, Implementierung und Validierung von maschinellen Lern-Methoden (Leitung: UKE, Teilnahme: UHH)

Wir am UKE haben in Zusammenarbeit mit den Partnern der UHH (UHAM) Pipelines für verschiedene Machine-Learning-Methoden entwickelt, darunter lineare Regression, logistische Regression und Random Forest unter Verwendung der Python-Bibliothek **scikit-learn**. Diese Pipelines waren hilfreich für die Vorhersage körperlicher Aktivität in SHARE und gemappten DEAS-Daten sowie für die Schmerzvorhersage in GLA:D und gemappten OAI-Daten. UKE und UHH haben gemeinsam Modelle auf Basis der SHARE-Daten entwickelt (zur Vorhersage körperlicher Inaktivität), die Teil einer Publikation sind, die sich derzeit im Review-Prozess befindet und bei der das UKE die Erstautorenschaft innehat. Diese Arbeit wurde bereits als Preprint auf arXiv veröffentlicht (DOI: 10.48550/arXiv.2409.07997) und in AIME registriert. Dasselbe Modell wurde später als vortrainiertes Modell für die kohortenübergreifende Evaluierung verwendet.

Es wurde eine vergleichende Analyse mit Daten aus dem GLA:D®-Programm (Interventionsgruppe) und der Osteoarthritis Initiative (OAI, Kontrollgruppe) durchgeführt, um Veränderungen der körperlichen Aktivität bei Personen mit Kniearthrose über einen Zeitraum von einem Jahr zu bewerten. Zur Angleichung der Baseline-Charakteristika und der Zielparameter (Outcome Measures) zwischen den Datensätzen wurde **Entropy Balancing** eingesetzt, was robuste statistische Vergleiche in Python ermöglichte. Während die Teilnehmer des GLA:D®-Programms eine strukturierte Schulung, überwachte Bewegungstherapie und individuelle Beratungen erhielten, erhielten die OAI-Teilnehmer keine spezifischen Interventionen.

Für den SHARE-Datensatz wurden am UKE 56 Variablen ausgewählt, die den Gesundheitszustand, Komorbiditäten und funktionelle Fähigkeiten erfassen, um körperliche Inaktivität (binäres Ergebnis) mittels eines **Random Forest Modells** vorherzusagen. Das Hyperparameter-Tuning führte dabei zu minimalen Schwankungen in der Genauigkeit. Lokale Modelle für einzelne SHARE-Länder wurden mit zentralisierten und föderierten Modellen verglichen, die auf aggregierten Daten aller teilnehmenden Länder trainiert wurden. Die Leistungsbewertung umfasste die Genauigkeit (Accuracy), AUROC sowie die Merkmalsgewichtung (**Feature Importance**), abgeleitet aus einer 5-fachen Kreuzvalidierung.

Im Bereich des Transfer Learnings erweiterten die Partner der UHH das Deep-Learning-Modell auf weitere Realdaten, indem sie Prädiktoren der körperlichen Aktivität auf äquivalente Variablen im DEAS-Datensatz mapp (Mapping).

M3.1 Implementierung und Evaluierung der maschinellen Lern-Modelle (UKE, UHH)

An der UHH wurde das vom UKE entwickelte Datenkonvertierungstool (**DB-Converter**) für SHARE- und DEAS-Daten implementiert. Später haben wir am UKE dasselbe für die OAI- und GLA:D-Daten umgesetzt. Für die Machine-Learning-Aufgaben wurde ein allgemeines Evaluierungsschema entwickelt, welches während des gesamten Projekts durchgängig Anwendung fand.

Aufgabe 3.2 - Entwicklung, Implementierung und Validierung von Deep-Learning-Methoden (Leitung: UHH, Teilnahme: UKE)

Für die kohortenübergreifende Validierung (Cross-Cohort Validation) wurde zunächst ein ML-Modell entwickelt, um körperliche Inaktivität innerhalb des SHARE-Frameworks vorherzusagen. Anschließend wendete die UHH das **DB-Converter**-Mapping-Tool an, um die DEAS-Daten in den Variablenraum von SHARE zu transformieren, und stellte den so gemappten DEAS-Datensatz bereit. Das vortrainierte



Modell wurde daraufhin verwendet, um die körperliche Inaktivität für die DEAS-Stichproben in diesem standardisierten Raum vorherzusagen. Abschließend wurden die gemappten Maße der körperlichen Inaktivität in DEAS gegen die Vorhersagen des Modells validiert, wobei die Metriken Genauigkeit (Accuracy) und AUROC zum Einsatz kamen.

M3.2 KI Modelle implementiert und evaluiert

Das KI-Modell wurde unter Verwendung derselben, zuvor entwickelten Pipeline anhand der Metriken Genauigkeit (Accuracy) und AUROC evaluiert.

Aufgabe 3.3 - Registrierung der KI-Methodik aus AP3 (Leitung: UKE)

Am UKE haben wir ausschließlich das oben genannte Modell basierend auf den SHARE-Daten bei AIME registriert. Für die OAI-Daten wählten wir Variablen zur körperlichen Aktivität und zu funktionellen Fähigkeiten zusammen mit Standard-Gesundheitsvariablen (ohne Komorbiditäten) aus, um die Komorbidität eines Teilnehmers vorherzusagen. Wir haben dieses Modell jedoch nicht weiterverfolgt, da die klinische Anwendbarkeit begrenzt war.

M3.3 AIME Registrierung AP3 (UKE)

Das Modell zur Vorhersage körperlicher Inaktivität in den SHARE-Daten ist registriert.

Arbeitspaket 4: Anwendung der entwickelten künstlichen Intelligenzen zur Evaluierung des Effekt des GLA:D® Programms auf die physische Aktivität

Aufgabe 4.1 - Übersetzung und Kalkulation eines Scores zur Bewertung physischer Aktivität (Leitung: UKE, Teilnahme: UHH)

Unsere Partner an der UHH führten das Mapping von SHARE auf DEAS basierend auf dem von uns am UKE entwickelten neuronalen Netzwerkmodell (DB-Converter) durch. Da wir die MOST-Daten nicht erhalten haben, schlossen wir unsere Aufgabe mit dem Mapping dieser vier Datensätze ab. In diesem Arbeitsschritt haben wir zwei verschiedene Methoden demonstriert: erstens einen auf neuronalen Netzen basierenden Ansatz (der Algorithmus wurde am UKE entwickelt und von der UHH auf SHARE- und DEAS-Daten angewendet) und zweitens ein auf Entropy-Balancing basierendes statistisches Verfahren (am UKE entwickelt und auf GLA:D- sowie OAI-Daten angewendet).

Aufgabe 4.2 - Anwendung des Einheits-Scoring-Systems (Leitung: UKE, Teilnahme: UHH)

In AP 4.2 haben wir diese Modelle evaluiert. Für GLA:D und OAI gab es keine einheitlichen Maße für körperliche Aktivität (PA), Lebensqualität (QOL) und Schmerz. Wir haben die statistische Methode des **Entropy Balancing** eingesetzt, wodurch wir nun PA, QOL und Schmerz vergleichen und die Wirkung des GLA:D-Programms in einem Ein-Jahres-Follow-up ab der Baseline (d. h. ein longitudinaler Vergleich) für das Management von Arthrose nachvollziehen können.

Diese Aufgabe hätte ohne das auf Entropy-Balancing basierende Mapping nicht realisiert werden können, da es in GLA:D keine Teilnehmer gibt, die nicht am Programm teilgenommen haben und somit als Kontrollgruppe fungieren könnten. Unsere Methode bietet somit ein Framework für das Testen von Hypothesen über nicht harmonisierte Datensätze hinweg (AP 4.2).

M4.1 GLA:D® Evaluation

Die Auswirkung des GLA:D-Programms auf die körperliche Aktivität (PA) wurde mithilfe harmonisierter Maße unter Anwendung von Entropy Balancing evaluiert.

Aufgabe 4.3 - Registrierung der KI-Methodik aus AP4 (Leitung: UKE)

Das zuvor genannte Modell ist in AIME registriert.

M4.2 AIME Registrierung AP4 (UKE)

Die Methoden aus AP4 sind in AIME registriert.

Arbeitspaket 5: Anwendbarkeit der entwickelten datenwissenschaftlichen Vorhersagewerkzeuge in der physiotherapeutischen Praxis und entsprechende Erweiterungsvorschläge für Guidelines zur Erhebung von physischer Aktivität

Aufgabe 5.1 - Nutzen des Einheits-Scoring-Systems (Leitung: UKE)

Unterschiedliche Bewertungssysteme verschiedener Variablen in den jeweiligen Datenbanken repräsentieren verschiedene Wege, um Ähnliches oder dasselbe zu messen – wie etwa bei der Erfassung der Lebensqualität.



Wir am UKE haben unser Tool zur Konvertierung einer Datenbank in eine andere, den sogenannten DB-Converter, implementiert, um die Bewertungssysteme zu vereinheitlichen, und haben dies anhand von zwei Anwendungsfällen (Use Cases) demonstriert. Bei der Konvertierung wird ein Bewertungssystem automatisch auf ein anderes abgebildet, indem das Modell aus den komplexen Beziehungen zu anderen Variablen lernt. Wir nutzten dabei einen auf Modelltransfer und Datenintegration basierenden Ansatz (jeweils für körperliche Inaktivität und Schmerzveränderung).

1. Unser Partner an der UHH konvertierte SHARE- und DEAS-Daten ineinander; anschließend haben wir DEAS in das Bewertungssystem von SHARE transformiert.
2. Zudem haben wir GLA:D und OAI ineinander konvertiert und konnten zeigen, dass die Vereinheitlichung von GLA:D und die Konvertierung von OAI in eine integrierte Datenbank die Genauigkeit der Vorhersagemodelle erhöht.

M5.1 Testen und Zwischenbilanz (UKE)

Während unser translationaler Ansatz für Variablen bei Umfragedaten und in der Forschung zur Vergrößerung der Datenbasis sowie zum Vergleich verschiedener Kohorten erfolgreich funktioniert, ist er nicht direkt auf die klinische Praxis übertragbar. Unser entwickeltes Modell zur Schmerzvorsage könnte jedoch für die klinische Praxis relevant sein. Dennoch ist derzeit noch ungewiss, ob die Modellleistung für eine tatsächliche Implementierung im klinischen Alltag ausreicht.

Um diese Frage zu adressieren und Forschende sowie Kliniker bei der Entscheidung zu unterstützen, ob ein Vorhersagemodell in ein klinisches Tool umgewandelt und implementiert werden sollte, haben wir ein Commentary-Paper mit entsprechenden Empfehlungen veröffentlicht [6].

Aufgabe 5.2 - Akzeptanz, Facilitatoren und Barrieren Evaluation (Leitung: UKE, Teilnahme: UHH)

Am UKE haben wir ein Tool entwickelt, das verdeutlicht, dass die Erstellung neuer Variablen als neues Bewertungssystem nicht erforderlich ist, sondern stattdessen das Mapping bestehender Variablen – also unterschiedlicher Bewertungssysteme – ermöglicht. Dieses Tool befindet sich derzeit im Review-Prozess für eine Publikation, in der wir aufzeigen, dass die Genauigkeit von Vorhersageaufgaben steigt, nachdem die Daten durch gegenseitige Konvertierung vereinheitlicht wurden. Die Entwicklung eines völlig neuen Bewertungssystems und dessen Akzeptanz im klinischen Bereich ist hürdenreich; daher ist die Integration durch das Mapping bereits klinisch etablierter Scores wesentlich effizienter.

Unser Partner an der UHH präsentierte zudem sein einsatzbereites Tool zur Schmerzvorsage in der **Science City Bahrenfeld (am 1. Juni 2024)**. Dort fand eine öffentliche Diskussion über die Anforderungen an die Modellleistung statt, die erfüllt sein müssen, damit eine Implementierung in die klinische Praxis relevant wird.

Aufgabe 5.3 - Guideline Empfehlung (Leitung: UKE, UHH)

Am UKE haben wir ein Tool entwickelt, das eine Konvertierung ermöglicht, um bestehende Variablen zu mappen, anstatt neue Variablen zu kreieren, die eine aufwendige klinische Zulassung erfordern würden. Folglich sind hierfür keine Änderungen an den klinischen Leitlinien notwendig. Im Gegensatz dazu haben wir für das Schmerzvorsage-Tool der UHH [6] die Anforderungen an die Modellleistung diskutiert, die erfüllt sein müssen, damit das Modell erfolgreich in die klinische Praxis implementiert werden kann.

M5.2 Publikation zur Guideline Empfehlung (UKE, UHH)

An der UHH haben wir unser einsatzbereites Tool zur Schmerzvorsage am 1. Juni 2024 in der **Science City Hamburg Bahrenfeld** vorgestellt, wobei eine öffentliche Diskussion über die Akzeptanz und die Barrieren für eine klinische Implementierung geführt wurde. Zudem haben wir ein Commentary-Paper [7] veröffentlicht, in dem die Anforderungen an die Modellleistung für die Implementierung [6] diskutiert werden, welche von unserem UHH-Partner entwickelt wurden.

Arbeitspaket 6: Erstellung eines datenwissenschaftlichen App Store für eine breite Anwendbarkeit in der physiotherapeutischen Forschung und Praxis zur Stärkung datenwissenschaftlicher Kompetenzen im Anwendungsfeld Physiotherapie

Aufgabe 6.1 - Implementierung der Prototypen-Methoden (Leitung: UHH, Teilnahme: UKE)

Wir am UKE haben unser Tool zur Datenbank-Konvertierung, den **DB-Converter**, in verschiedenen Versionen implementiert: als Docker-Container, als Python-Skript sowie als Python Notebook.

M6.1 Pipelines in GitHub veröffentlicht (UHH, UKE)

Sämtliche Skripte wurden in einem GitHub-Repository zur Verfügung gestellt.



Aufgabe 6.2 - Bereitstellung eines Web-Portals

(Leitung: UKE, Teilnahme: UHH)

Am UKE haben wir mit Unterstützung unserer Partner an der UHH ein Web-Interface für den Schmerzrechner erstellt. Dieses Tool wurde online veröffentlicht; der entsprechende Link ist im zugehörigen Ergebnisteil aufgeführt.

M6.2 PhysioAI App online (UKE, UHH)

Am UKE stellen wir das neuronale Netzwerk als Docker-Container, Python Notebook und als eigenständige Python-App bereit.

3 Inhaltliche Ergebnisse

Arbeitspaket 1: Daten-Vorbereitung, Vorverarbeitung und Aufreinigung

Aufgabe 1.1 - MySQL-Datenbankerstellung (Leitung: UKE, Teilnahme: UHH)

Ein Datenmanagementplan wurde erstellt und in Umlauf gebracht. Am UKE haben wir das MySQL-Datenbankschema entwickelt und vier Datensätze abgerufen. Der MOST-Datensatz konnte aufgrund von Unterschieden in den ethischen und datenschutzrechtlichen Bestimmungen zwischen den USA und der EU nicht bezogen werden, was zu einer bürokratischen Lücke führte.

M1.1 Data Management Plan (UKE)

Ein Datenmanagementplan wurde erstellt, mit der UHH geteilt und abgestimmt.

M1.2 MySQL-Datenbank-Schema (UKE, UHH)

Ein MySQL-Datenschema wurde erstellt und mit der UHH geteilt. Später wurde jedoch entschieden, **Pandas** anstelle von MySQL zu verwenden. Grund hierfür waren Datenmengenbeschränkungen in MySQL sowie der praktische Nutzen von Pandas für die nachfolgenden Python-basierten Analysen.

Aufgabe 1.2 - Integration von GLA:D Kohorten Daten (Leitung: UKE)

Die Daten wurden abgerufen und in ein Pandas-Dataframe konvertiert, welches an der UHH für die Integration mit den anderen drei Datensätzen genutzt wurde. Die folgende Tabelle beschreibt die Datensätze:



Tabelle 1: Übersicht der 4 Datensätze.

Name	DEAS	SHARE	GLAD	OAI
Eingeschlossene Datensätze	8 (with covid)	w1, w2, w3, w4, w5, w6, w7, w8, SCS1 (covid), SCS2(covid)	1 (baseline, 3m follow-up, 12m follow-up)	10
Genutzte Datensätze	2008-2011-2014-2017-2020	w8	baseline	1. Patienten Information 2. Klinischer Datensatz Medikation 3. Inventory (MIF) Datensatz
Anzahl der zusammengesetzten Datensätze (Variablen/Stichprobengröße)	(26626, 6308)	(571793, 6890)	(66215, 556)	(116386, 7197)
Summe aller (Variablen/Stichprobengröße)	(781020, 20951)			

Aufgabe 1.3 - Integration von MOST Kohorten Daten (Leitung: UKE)

Es gelang uns nicht, diese Daten zu beziehen, da Unterschiede in den ethischen und datenschutzrechtlichen Bestimmungen zwischen den USA und der EU zu einer bürokratischen Lücke führten. Dies beeinträchtigte jedoch nicht das Projektziel; wir verwendeten stattdessen die anderen vier Datensätze.

M1.3 Data Warehouse Pipeline (UKE)

Eine Data-Warehouse-Pipeline wurde generiert und in Pandas implementiert, um SHARE, OAI, DEAS und GLA:D zu integrieren.

Aufgabe 1.4 - Integration von OAI Kohorten Daten (Leitung: UKE)

Die Daten wurden abgerufen und in ein Pandas-Dataframe konvertiert, welches an der UHH zur Integration mit den anderen drei Datensätzen als Pandas-Dataframe genutzt wurde.

Aufgabe 1.5 - Integration von SHARE Kohorten Daten (Leitung: UKE)

Die Daten wurden abgerufen und in ein Pandas-Dataframe konvertiert, welches an der UHH zur Integration mit den anderen drei Datensätzen als Pandas-Dataframe genutzt wurde.

Aufgabe 1.6 - Integration von DEAS Kohorten Daten (Leitung: UKE)

Die Daten wurden abgerufen und in ein Pandas-Dataframe konvertiert, welches an der UHH zur Integration mit den anderen drei Datensätzen als Pandas-Dataframe genutzt wurde.

M1.4 PhysioAI-Datenbank (UKE, UHH)

Alle vier verfügbaren Datensätze (SHARE, DEAS, OAI und GLA:D) wurden in einem gemeinsamen Pandas-Data-Warehouse integriert. Die Gesamtübersicht aller Datensätze findet sich in der oben stehenden Tabelle 1.



Arbeitspaket 2: Entwicklung künstlicher Intelligenzen zur Kompatibilisierung von Messinstrumenten zur Erfassung physischer Aktivität

Aufgabe 2.1 - Implementierung einer Methoden-Pipeline (Leitung: UHH, Teilnahme: UKE)

Wir am UKE unterstützen unsere Partner an der UHH dabei, Machine-Learning-Module von **scikit-learn** zusammenzuführen, um **R-squared** (Bestimmtheitsmaß), **RMSE** (Wurzel der mittleren Fehlerquadratsumme), **Genauigkeit** (Accuracy) und **AUROC** für verschiedene ML-Aufgaben unter Verwendung der von der UHH entwickelten Pipeline zu berechnen. Die entsprechenden Ergebnisse für diese Metriken sind Teil einer Publikation, die sich derzeit im Review-Prozess befindet.

M2.1 Pipeline Evaluierung (UHH, UKE)

Dieselbe Pipeline wurde in GLA:D mit verschiedenen Variablensätzen und Modellen evaluiert.

Aufgabe 2.2 - Transfer-Learning-Ansatz (Entwicklung und Implementierung in PyTorch) (Leitung: UHH, UKE)

Es wurde sowohl ein unidirektionaler als auch ein bidirektionaler DB-Converter entwickelt. Die Ergebnisse hierzu befinden sich derzeit im Review-Prozess für eine wissenschaftliche Publikation.

M2.2 Evaluierung und Implementierung des Transfer Learnings (UHH, UKE)

Der unidirektionale und der bidirektionale DB-Converter wurden sowohl mit synthetischen als auch mit Realdaten (DEAS, SHARE, GLA:D, OAI) evaluiert. Bei der Konvertierung kontinuierlicher synthetischer Daten liegt der **L1-Fehler** im Bereich von 0,01 bis 0,03; bei Realdaten (GLA:D und OAI) bewegte er sich im Bereich von 0,38 bis 0,43.

Aufgabe 2.3 - Erweiterung der neuronalen Netzwerke und Fusion der Einzelmodelle (Leitung: UHH)

Meilenstein M2.3 Implementierung und Evaluierung der Fusions-Ansätze (UHH)

Aufgabe 2.4 - Registrierung der KI-Methodik aus Arbeitspaket 2 (Leitung: UKE)

Das oben genannte allgemeine Tool für den unidirektionalen DB-Converter wurde in **AIME** registriert.

M2.4 AIME Registrierung (UKE)

Dasselbe Modell aus AP 2.3 ist in AIME registriert (AP 2.4). Die Registrierung ist unter folgendem Link einsehbar: <https://aime.report/fbSPBM>.

Arbeitspaket 3: Entwicklung künstlicher Intelligenzen zur Evaluation von individueller physischer Aktivität aus funktionalen Test-Daten und Patientendaten am Beispiel von Arthrose

Aufgabe 3.1 - Entwicklung, Implementierung und Validierung von maschinellen Lern-Methoden (Leitung: UKE, Teilnahme: UHH)

Wir am UKE haben in Zusammenarbeit mit dem Partner an der UHAM Pipelines für verschiedene Machine-Learning-Methoden wie lineare Regression, logistische Regression und Random Forest unter Verwendung der Python-Bibliothek **scikit-learn** entwickelt. Wie oben für GLA:D erwähnt, ergaben die Aufgaben zur Schmerzvorschau ein **R^2 von 0,34** (mittels linearer Regression), und für die Vorhersage körperlicher Inaktivität in SHARE lag die **Genauigkeit bei 0,84** (mittels Random Forest).

M3.1 Implementierung und Evaluierung der maschinellen Lern-Modelle (UKE, UHH)

R-squared, RMSE, Genauigkeit (Accuracy) und AUROC wurden unter Verwendung der zuvor erstellten Pipeline für alle nachgelagerten ML-Aufgaben berechnet. Hier sind einige Beispiele für verschiedene Modelle in unterschiedlichen Datensätzen:

Tabelle 2: Performance von Maschine learning modellen auf GLA:D and SHARE daten .



data/ scenario	Model	Metric	Values
GLA:D®, National	Linear Regression	r-square	0.34
GLA:D®, National	Linear Regression	RMSE	18.2
GLA:D®, National	Random Forest Regression	r-square	0.32
GLA:D®, National	Random Forest Regression	RMSE	18.5
SHARE, International	Random Forest Classification	accuracy	0.84
SHARE, International	Random Forest Classification	AUROC	0.66

Aufgabe 3.2 - Entwicklung, Implementierung und Validierung von Deep-Learning-Methoden (Leitung: UHH, Teilnahme: UKE)

Die Variable für körperliche Inaktivität der DEAS-Stichproben wurde mittels des **DB-Converter**-Tools in den SHARE-Datenraum übertragen. Die Validierung erfolgte durch den Vergleich mit der vorhergesagten körperlichen Inaktivität (unter Verwendung eines in SHARE vortrainierten Modells) innerhalb der DEAS-Stichproben im SHARE-Raum, wobei die Metriken Genauigkeit (Accuracy) und AUROC herangezogen wurden. Diese Ergebnisse sind Bestandteil einer Publikation, die sich derzeit im Review-Prozess befindet.

M3.2 KI Modelle implementiert und evaluiert

Sowohl die interne Evaluierung (mit synthetischen und Realdaten) als auch die kohortenübergreifende Evaluierung (Cross-Cohort Evaluation mit Realdaten) belegen die Wirksamkeit des KI-Tools für das Mapping (DB-Converter). Die entsprechenden Ergebnisse wurden zur Publikation eingereicht.

Aufgabe 3.3 - Registrierung der KI-Methodik aus AP3 (Leitung: UKE)

Am UKE haben wir ausschließlich das oben genannte Modell zur Vorhersage körperlicher Inaktivität im SHARE-Raum bei **AIME** registriert.

M3.3 AIME Registrierung AP3 (UKE)

Das zuvor erwähnte Modell ist in AIME registriert und unter folgendem Link einsehbar: <https://aime.report/NirZqd>

Arbeitspaket 4: Anwendung der entwickelten künstlichen Intelligenzen zur Evaluierung des Effekt des GLA:D® Programms auf die physische Aktivität

Aufgabe 4.1 - Übersetzung und Kalkulation eines Scores zur Bewertung physischer Aktivität (Leitung: UKE, Teilnahme: UHH)

Unser Partner an der UHH begann mit dem Mapping von SHARE auf DEAS basierend auf dem von uns am UKE entwickelten neuronalen Netzwerkmodell. Die Vorhersage der körperlichen Inaktivität auf den übersetzten DEAS-Daten unter Verwendung des in SHARE trainierten Prädiktors stimmt mit der gemappten körperlichen Inaktivität mit einer **Genauigkeit von 0,97** und einem **AUROC von 0,86** überein. Dies belegt, dass der Harmonisierungsansatz für den Transfer des Vorhersagemodells der körperlichen Inaktivität erfolgreich funktioniert.

Aufgabe 4.2 - Anwendung des Einheits-Scoring-Systems (Leitung: UKE, Teilnahme: UHH)



Körperliche Aktivität, Lebensqualität und Schmerz – diese drei Variablen wurden zum Ausgangszeitpunkt (Baseline) und nach einem 3-monatigen Follow-up evaluiert, um die Wirkung von GLA:D zu verstehen. Dies ermöglichte es uns, den Effekt der GLA:D-Intervention zu erfassen: Beispielsweise erfuhren in der GLA:D®-Gruppe **12 % mehr Teilnehmer eine klinisch relevante Schmerzreduktion** im Vergleich zu den OAI-Teilnehmern.

M4.1 GLA:D® Evaluation

Der harmonisierte Datensatz (nach dem Entropy Balancing) ermöglichte einen Vergleich zwischen GLA:D-Teilnehmern und gematchten OAI-Patienten als Kontrollgruppe.

Aufgabe 4.3 - Registrierung der KI-Methodik aus AP4 (Leitung: UKE)

Das Modell für das Entropy Balancing und den Vergleich zwischen GLA:D und OAI ist in AIME dokumentiert.

M4.2 AIME Registrierung AP4 (UKE)

Das zuvor erwähnte Modell ist bei AIME registriert. Der Bericht ist unter folgendem Link einsehbar:

<https://aime.report/OS6tZ2>

Arbeitspaket 5: Anwendbarkeit der entwickelten datenwissenschaftlichen Vorhersagewerkzeuge in der physiotherapeutischen Praxis und entsprechende Erweiterungsvorschläge für Guidelines zur Erhebung von physischer Aktivität

Aufgabe 5.1 - Nutzen des Einheits-Scoring-Systems (Leitung: UKE)

Das vereinheitlichte Bewertungssystem wurde auf SHARE–DEAS und GLA:D–OAI angewendet. Die Harmonisierung verbesserte die Vorhersageleistung und ermöglichte studienübergreifende Vergleiche. Beispielsweise erhöhte die Datenintegration von gemappten OAI-Daten mit GLA:D® die Genauigkeit prognostischer Modelle auf **0,85**, verglichen mit einer Genauigkeit von **0,77**, wenn das Modell ausschließlich auf GLA:D®-Daten trainiert wurde.

M5.1 Testen und Zwischenbilanz (UKE)

Wir am UKE haben den **DB-Converter** bereitgestellt, der von der UHH für die Konvertierung der SHARE- und DEAS-Datensätze verwendet wurde. Anhand dieser konvertierten Datensätze wurde demonstriert, dass jedes bestehende Modell, das speziell für einen Datensatz entwickelt wurde, auch für andere Datensätze verwendet werden kann (gezeigt an SHARE- und DEAS-Daten). Wir am UKE konnten zeigen, dass die Vereinheitlichung von Metriken die Integration von Datensätzen ermöglicht und dadurch nachgelagerte ML-Aufgaben optimiert (gezeigt an OAI- und GLA:D-Daten). Diese Ergebnisse sind detailliert in einer Publikation beschrieben, die sich derzeit im Review-Prozess befindet.

Aufgabe 5.2 - Akzeptanz, Facilitatoren und Barrieren Evaluation (Leitung: UKE, Teilnahme: UHH)

Es gab sowohl interne als auch externe Diskussionen über die Validität des Tools und darüber, ob die Vorhersagen "gut" genug sind, um das Tool Patient:innen und/oder Kliniken zur Verfügung zu stellen. Insbesondere haben wir die Arbeit bei der Science City Bahrenfeld Veranstaltung vorgestellt und diskutiert.

Aufgabe 5.3 - Guideline Empfehlung (Leitung: UKE, UHH)

Die Ergebnisse der Diskussion von der Science City Bahrenfeld Veranstaltung wurden weiter mit Experten aus dem Bereich Ethik besprochen und strukturiert.

M5.2 Publikation zur Guideline Empfehlung (UKE, UHH)



Wir haben ein Commentary-Paper zur Akzeptanz der klinischen Implementierung des von unserem UHH-Partner entwickelten Schmerzvorhersage-Tools veröffentlicht [7]. Dieses empfiehlt, welche Faktoren berücksichtigt werden müssen, um zu entscheiden, ob ein personalisiertes Vorhersagemodell in ein klinisches Werkzeug umgewandelt und letztlich in die klinische Praxis implementiert werden sollte.

Arbeitspaket 6: Erstellung eines datenwissenschaftlichen App Store für eine breite Anwendbarkeit in der physiotherapeutischen Forschung und Praxis zur Stärkung datenwissenschaftlicher Kompetenzen im Anwendungsfeld Physiotherapie

Aufgabe 6.1 - Implementierung der Protoypen-Methoden (Leitung: UHH, Teilnahme: UKE)

Das Datenbank-Mapping-Tool **DB-Converter** wurde vollständig implementiert. Es steht ab sofort jedem Forschenden zum Download und zur Anwendung in eigenen Projekten zur Verfügung. Das Repository ist unter folgendem Link erreichbar: <https://github.com/Mycheaux/DB-conv>.

M6.1 Pipelines in GitHub veröffentlicht (UHH, UKE)

Die folgenden GitHub-Repositories werden vom UKE gepflegt:

- **DB-conv:** <https://github.com/Mycheaux/DB-conv> und <https://github.com/Mycheaux/KnowledgeTransfer.git>
- **Fed Survey (Föderiertes Lernen):** <https://github.com/Mycheaux/FedQuesPaper> und <https://github.com/Mycheaux/BetaSharingLinearRegression.git>

Aufgabe 6.2 - Bereitstellung eines Web-Portals (Leitung: UKE, Teilnahme: UHH)

Das Web-Interface für den Schmerzrechner ist unten in M6.2 aufgeführt; das Backend hierfür wird von der UHH verwaltet.

M6.2 PhysioAI App online (UKE, UHH)

Am UKE stellen wir unter folgendem Link ein Tool zur Datenbankkonvertierung bereit: <https://github.com/Mycheaux/DB-conv>.

Datenkompetenzsteigerung: Zur Förderung der interprofessionellen Zusammenarbeit und des gegenseitigen Fachverständnisses wurden im Projektverlauf zweiwöchentliche Jour-fixe-Termine etabliert. Diese Treffen dienten insbesondere der gezielten Vermittlung von Datenkompetenzen. Durch diesen Co-Creation-Ansatz wurden die Zielsetzungen der Förderrichtlinie – insbesondere die Stärkung der Data Literacy des wissenschaftlichen Nachwuchses – maßgeblich unterstützt.

4 Wichtigste Positionen des zahlenmäßigen Nachweises

Die größte Position im zahlenmäßigen Nachweis ist Kategorie 0812 (Beschäftigte E12-E15). Insgesamt hat ein Doktorand am UKE am Projekt gearbeitet. Außerdem sind Kosten in der Kategorie 0843 (sonst. allg. Verwaltungsausgaben) für Publikationsgebühren sowie in der Kategorie 0846 für Dienstreisen entstanden.

5 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Alle Stellen wurden so schnell wie möglich besetzt. Die Arbeitspakete wurden entsprechend der Projektbeschreibung abgearbeitet und die Personalkosten wurden somit zur Erfüllung der Aufgaben eingesetzt. Die für das Projekt eingeplanten Mittel wurden gemäß der Vorgaben eingesetzt.



6 Voraussichtlicher Nutzen, Verwertbarkeit der Ergebnisse und zukünftige Planungen im Sinne des Verwertungsplans

Das Projekt PhysioAI eröffnet neue Horizonte zur Lösung von Inkompatibilitätsproblemen bei Gesundheitsumfragedaten. Unsere Methoden können für die kohortenübergreifende Vereinheitlichung von Metriken, den Vergleich von Kennzahlen über verschiedene Kohorten hinweg, die studienübergreifende Datenintegration sowie die Übertragbarkeit von Modellen (Model Transferability) eingesetzt werden. In naher Zukunft sollte die Forschungsrichtung darauf abzielen, zu prüfen, ob unsere Methoden auf andere Datentypen in der Biologie und darüber hinaus generalisiert werden können, wo Dateninkompatibilität eine Hürde für die Data-Science-Gemeinschaft darstellt.

7 Fortschritt auf dem Gebiet des Vorhabens während der Durchführung des Vorhabens bei anderen Stellen

Seit dem Beginn des Projekts wurden zwar viele Techniken entwickelt, die auf **Self-Supervised Learning** basieren (wie z. B. Contrastive Learning [8]) und für ungelabelte, gepaarte Daten konzipiert sind. Diese bezogen sich jedoch hauptsächlich auf Bild- und Textdaten und nicht auf tabellarische Gesundheitsdaten [9]. Zudem erfordern diese Techniken die Generierung von positiven Paaren (aus derselben Stichprobe) und negativen Paaren (aus verschiedenen Stichproben) – eine Methode, die in unserem Fall nicht für das Mapping zweier tabellarischer Gesundheitsdatenbanken verwendet werden kann [10].

Obwohl bestimmte Transformer-basierte Modelle für tabellarische Daten entwickelt wurden, waren diese nicht für das Datenbank-Mapping ausgelegt [11]. Das Feld des **Manifold Mappings** (Abbildung von Mannigfaltigkeiten) zur Übertragung einer Datenbank in eine andere beschränkte sich weiterhin auf gepaarte Daten und nicht auf ungepaarte Daten. Daher sind wir unseres Wissens nach die Ersten, die eine **selbstüberwachte KI (Self-supervised AI)** entworfen haben, um Variablen einer Datenbank auf eine andere abzubilden, ohne dass dafür gepaarte Stichproben erforderlich sind.

8 Erfolgte oder geplante Veröffentlichungen der Projektergebnisse

a. Erfolgte Veröffentlichungen:

1. Rafiei M, Das S, Bakhtiari M, Roos E, Skou S, Grønne D, Baumbach J, Baumbach L. Personalized Predictions for Changes in Knee Pain Among Patients With Osteoarthritis Participating in Supervised Exercise and Education: Prognostic Model Study JMIR Rehabil Assist Technol 2025;12:e60162 URL: <https://rehab.jmir.org/2025/1/e60162> DOI: 10.2196/60162
2. Rafiei, M.; Das, S.; Roos, E.M.; Skou, S.T.; Baumbach, J.; Baumbach, L. Are Changes in Physical Activity, Pain, and Quality of Life in Patients with Knee Osteoarthritis After Exercise Therapy and Education Beyond Normal Fluctuations? A Comparative Study. J. Clin. Med. 2025, 14, 3406. <https://doi.org/10.3390/jcm14103406>
3. Ashry S, Das S, Rafiei M, Baumbach J, Baumbach L. Transfer learning of human activities based on IMU sensors: A review. IEEE Sensors Journal. 2024 Dec 9. DOI: [10.1109/JSEN.2024.3510097](https://doi.org/10.1109/JSEN.2024.3510097)
4. Baumbach L, Hötendorfer W, Baumbach J. To Implement or Not to Implement? A Commentary on the Pitfalls of Judging the Value and Risks of Personalized Prognostic Statistical Models. Journal of Medical Internet Research. 2025 May 19;27:e69341.

b. Geplante Veröffentlichungen:

1. S.Das, M.Rafiei, A.Maier, L.Baumbach, J.Baumbach. Self-supervised generative AI enables conversion of two non-overlapping cohorts. (Underview).
2. Das S, Rafie M, Kammer P, Skou ST, Grønne DT, Roos EM, et al. Privacy-preserving federated prediction of pain intensity change based on multi-center survey data 2024. <https://doi.org/10.48550/arXiv.2409.07997>.
3. S.Das, M.Rafiei, M.Burghardt, J.Baumbach, L.Baumbach. Origin of sexual dimorphism in osteoarthritis risk. (Underview).
4. M.Rafiei, N. Probul, J.Baumbach, L.Baumbach. Older adults with osteoarthritis show lower functional performance compared to those with diabetes or hypertension: Evidence from the SHARE dataset (Underview)



9 Literaturverzeichnis

- [1] Roos EM, Grønne DT, Ris I, et al. Patient characteristics in knee, hip and back pain patients participating in patient education and exercise therapy. a comparison of 46,273 patients from the glad registries in Denmark. *Osteoarthritis Cartilage* 2020; 28: S169–70.
- [2] Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D™): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. *BMC Musculoskelet Disord* 2017;18:72. <https://doi.org/10.1186/s12891-017-1439-y>.
- [3] Vogel C, Wettstein M, Daniela K, Svenja S, Lisa K, Lozano Alcántara A, et al. German Ageing Survey (DEAS): Instruments of the Survey 2020/2021 2021. <https://doi.org/10.5156/DEAS.2021.D.002>.
- [4] Michael C. Nevitt, David T. Felson, Gayle Lester. The osteoarthritis initiative n.d.
- [5] Börsch-Supan A, Brandt M, Hunkler C, Kneip T, Korbmacher J, Malter F, et al. Data Resource Profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *Int J Epidemiol* 2013;42:992–1001. <https://doi.org/10.1093/ije/dyt088>.
- [6] Baumbach L, Hötzendorfer W, Baumbach J. To Implement or Not to Implement? A Commentary on the Pitfalls of Judging the Value and Risks of Personalized Prognostic Statistical Models. *J Med Internet Res* 2025;27:e69341. <https://doi.org/10.2196/69341>.
- [7] Rafiei M, Das S, Bakhtiari M, Roos EM, Skou ST, Grønne DT, et al. Personalized Predictions for Changes in Knee Pain Among Patients With Osteoarthritis Participating in Supervised Exercise and Education: Prognostic Model Study. *JMIR Rehabil Assist Technol* 2025;12:e60162. <https://doi.org/10.2196/60162>.
- [8] Liu Z, Alavi A, Li M, Zhang X. Self-Supervised Contrastive Learning for Medical Time Series: A Systematic Review. *Sensors* 2023;23:4221. <https://doi.org/10.3390/s23094221>.
- [9] Huang S-C, Pareek A, Jensen M, Lungren MP, Yeung S, Chaudhari AS. Self-supervised learning for medical image classification: a systematic review and implementation guidelines. *Npj Digit Med* 2023;6:74. <https://doi.org/10.1038/s41746-023-00811-0>.
- [10] Ghesu FC, Georgescu B, Mansoor A, Yoo Y, Neumann D, Patel PK, et al. Contrastive self-supervised learning from 100 million medical images with optional supervision. *J Med Imaging* 2022;9:064503. <https://doi.org/10.1117/1.JMI.9.6.064503>.
- [11] Wang W, Kim B-H, Ganapathi V. RegCLR: A Self-Supervised Framework for Tabular Representation Learning in the Wild 2022. <https://doi.org/10.48550/arXiv.2211.01165>.
- [12] S.Das, M.Rafiei, A.Maier, L.Baumbach, J.Baumbach. Self-supervised generative AI enables conversion of two non-overlapping cohorts. (Underview) 2025.
- [13] Das S, Rafie M, Kammer P, Skou ST, Grønne DT, Roos EM, et al. Privacy-preserving federated prediction of pain intensity change based on multi-center survey data 2024. <https://doi.org/10.48550/arXiv.2409.07997>.
- [14] S.Das, M.Rafiei, M.Burghardt, J.Baumbach, L.Baumbach. Origin of sexual dimorphism in osteoarthritis risk. *Rev* n.d.



- [15] M.Rafiei, N. Probul, J.Baumbach, L.Baumbach. Older adults with osteoarthritis show lower functional performance compared to those with diabetes or hypertension: Evidence from the SHARE dataset n.d.
- [16] Ashry S, Das S, Rafiei M, Baumbach J, Baumbach L. Transfer Learning of Human Activities Based on IMU Sensors: A Review. IEEE Sens J 2025;25:4115–26. <https://doi.org/10.1109/JSEN.2024.3510097>.