Aus der

Klinik für Orthopädie und Unfallchirurgie

Klinikum der Ludwig-Maximilians-Universität München

Predictive Modeling of Fall Risk in Orthogeriatric Patients using Machine Learning Techniques

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von
Moritz Kraus

aus
Starnberg

Jahr
2025

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Wolfgang Böcker
Zweites Gutachten:	Prof. Dr. Henning Wackerhage
Drittes Gutachten:	Prof. Dr. Michael Ingrisch

Dekan:	Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung:	24.03.2025

Affidavit

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Promotionsbüro
Medizinische Fakultät

Eidesstattliche Versicherung

Kraus, Moritz

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Predictive Modeling of Fall Risk in Orthogeriatric Patients using Machine Learning Techniques

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 29.06.2025 Kraus, Moritz

Ort, Datum Unterschrift Doktorandin bzw. Doktorand

Inhaltsverzeichnis

Abbildungsverzeichnis

Publikationsliste

Veröffentlichte Originalarbeiten

- **Kraus M**, Stumpf UC, Keppler AM, Neuerburg C, Böcker W, Wackerhage H, Baumbach SF, Saller MM. Development of a Machine Learning-Based Model to Predict Timed-Up-and-Go Test in Older Adults. Geriatrics (Basel). 2023 Oct 7;8(5):99. doi: 10.3390/geriatrics8050099. PMID: 37887972; PMCID: PMC10606325.
- Obermeier KT, Liokatis P, Haidari S, Kakoschke TK, **Kraus M**, Smolka W. Lymph nodal recurrence in levels IV and V in oral squamous cell carcinoma after neck dissection. ANZ J Surg. 2023 Jun;93(6):1688-1693.
- **Kraus, M**.; Neuerburg, C.; Thomasser, N.; Stumpf, U.C.; Blaschke, M.; Plötz, W.; Saller, M.M.; Böcker, W.; Keppler, A.M. Reduced Awareness for Osteoporosis in Hip Fracture Patients Compared to Elderly Patients Undergoing Elective Hip Replacement. Medicina2022, 58, 1564.
- Haidari S, Obermeier KT, **Kraus M**, Otto S, Probst FA, Liokatis P. Nodal Disease and Survival in Oral Cancer: Is Occult Metastasis a Burden Factor Compared to Preoperatively Nodal Positive Neck? Cancers (Basel). 2022 Aug 31;14(17):4241.
- Stumpf U, **Kraus M**, Ladurner R, Neuerburg C, Böcker W. (2022). Osteoporose: Diagnostik und Behandlung Osteoporosis: diagnostics and treatment. Der Chirurg. 93. 1107–1120 (2022).
- Obermeier, K.T.; **Kraus, M**.; Henkel, J.; Saller, T.; Otto, S.; Liokatis, P. Postoperative Delirium in Patients with Oral Cancer: Is Intraoperative Fluid Administration a Neglected Risk Factor? Cancers 2022, 14.10.3390/cancers14133176
- **Kraus M**, Saller MM, Baumbach S, Neuerburg C, Stumpf U, Böcker W, Keppler A. Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor InsoleBased Gait Analysis and Machine Learning Algorithms: Cross-sectional Study, JMIR Med Inform 2022;10(1):e32724
- Stumpf, U., **Kraus, M.** & Hadji, P. Influence of denosumab on bone mineral density in a severe case of pregnancy-associated osteoporosis. Osteoporos Int (2021). Nov;32(11):2383-2387.
- Keppler AM, **Kraus M**, Blaschke M, Thomasser N, Kammerlander C, Böcker W, Neuerburg C, Stumpf UC. Reduced Awareness for Osteoporosis in Distal Radius Fracture Patients Compared to Patients with Proximal Femur Fractures. J Clin Med. 2021 Feb 19;10(4):848.
- Stumpf U, **Kraus M**, Gosch M, Hesse E, Kammerlander C. Fracture liaison service (FLS): A structured care concept for optimized secondary prevention of osteoporotic fractures, Osteologie 2021, 2;
- Stumpf UC, Kraus M, Baumbach SF, et al. Bone marrow edema syndrome: Interdisciplinary diagnostic strategy and treatment options. Osteologie 2021,1;
- Keppler AM, **Kraus M**, Neuerburg C, Fürmetz J, Böcker W, Kammerlander C. Mobilitätsmessungen bei Patienten mit Schenkelhalsfrakturen Jatros 03/2021

Präsentierte Kongressbeiträge

2023

- **Kraus M,** Saller MM, Neuerburg C, Keppler AM, Böcker W, Stumpf UC. Prospektive Einschätzung des Sturzrisikos anhand Machine-Learning-basierter Ganganalyse während des TUG-Tests mit 2 Jahres Follow-Up. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- **Kraus M,** Saller MM, Neuerburg C, Keppler AM, Böcker W, Stumpf UC. Prospektiver Vergleich zur Einschätzung des Sturzrisikos anhand etablierter Scores in der Alterstraumatologie mit 2 Jahres "real-world" Evaluation. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- **Kraus M,** Rieken JN, Ditto J, Zeckey C. Einschränkung der Sprunggelenksbeweglichkeit durch Orthesen führt zu einer erhöhten Kniegelenksbelastung. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- **Kraus M,** Illner J, Warnhoff M, Audigé L, Moro F, Brunner M, Góralczyk A, Scheibel M. Inakzeptabel hohe Lockerungs- und Revisionsrate nach anatomischer Schultertotalendoprothese mit zementierter pegged Polyethylen-Glenoidkomponente. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- **Kraus M,** Illner J, Warnhoff M, Brunner M, Scheibel M. Komplikationen und Revisionen nach anatomischer TEP mit metal-back Glenoiden - ein klinischer und radiologischer Vergleich von Kalottenersatz vs. Schaftendoprothese. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- Sanchez Carbonel JF; Brunner M; Hinz M; **Kraus M**; Bertoni G; Mehl J; Siebenlist S; Winkler P. Tibial slope and lateral femoral condyle ratio: Indicators for early ACL reconstruction to prevent concomitant injuries. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2023
- **Kraus M**, Klöpfer-Krämer I, Mühling M, Gabel J, Augat P, Brand A. Foot function after calcaneus fracture assessed by application of unsupervised machine learning on pedobarographic gait data. Annual meeting of the European society for movement analysis in adults and children, Athen, Griechenland 2023
- **Kraus M,** Illner J, Warnhoff M, Audigé L, Moro F, Brunner M, Góralczyk A, Scheibel M. Inakzeptabel hohe Lockerungs- und Revisionsrate nach anatomischer Schultertotalendoprothese mit zementierter pegged Polyethylen-Glenoidkomponente. Jahreskongres der D-A-CH-Vereinigung für Schulter- und Ellenbogenchirurgie, Hamburg, Deutschland. 2023
- **Kraus M,** Illner J, Warnhoff M, Brunner M, Scheibel M. Komplikationen und Revisionen nach anatomischer TEP mit metal-back Glenoiden - ein klinischer und radiologischer Vergleich von Kalottenersatz vs. Schaftendoprothese. 83. Jahreskongress swiss orthopeadics, St. Gallen, Schweiz. 2023
- **Kraus M.** Aktuelle Strategien: KI in der Optimierung der Diagnostik der „frailty", Übersichtsvortrag. Jahreskongress DVO OSTEOLOGIE, Salzburg, Österreich. 2023
- **Kraus M,** Saller MM, Neuerburg C, Keppler AM, Böcker W, Stumpf UC. Prospektive Einschätzung des Sturzrisikos anhand mobiler Ganganalysen während des TUG-Tests mit 2 Jahres Follow-up. Jahreskongress der Vereinigung der Süddeutschen Orthopäden und Unfallchirurgen, Baden-Baden, Deutschland. 2023
- **Kraus M,** Rieken JN, Ditto J, Zeckey C. Significant influence of ankle orthoses on knee joint kinematics. Gamma Kongress, Heidelberg, Deutschland. 2023

2022

- **Kraus M,** Rieken JN, Ditto J, Zeckey C, Saller MM, Keppler AM, et. al. „Unsupervised clustering of gait analysis sensor data to determine outcome after tibial plateau fractures. Jahreskongress der deutschen Kniegesellschaft. München, Deutschland. 2022
- **Kraus M**, Saller MM, Keppler AM, et. al. KI basierte Erkennung des Sturzrisikos mittels einer sohlenbasierten Ganganalyse bei orthogeriatrischen Patienten, Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2022
- **Kraus M**, Saller MM, Keppler AM, et. al. „Clusteranalyse der funktionellen Kniekinematik nach Tibiaplateau-Frakturen zur Bestimmung der Rehabilitationsergebnisse basierend auf Algorithmen des maschinellen Lernens. Jahreskongress der Deutschen Gesellschaft für Biomechanik, Köln, Deutschland. 2022
- **Kraus M,** Saller MM, Keppler AM, et. al. Entwicklung eines KI-gestützten Entscheidungstools zur Bewertung der Physical Frailty bei immobilisierten orthogeriatrischen Patienten, Jahreskongress DVO OSTEOLOGIE, Baden-Baden, Deutschland. 2022
- **Kraus M,** Saller MM, Keppler AM, et. al. Einschätzung der physical frailty bei orthogeriatrischen Patienten aus Nicht-Mobilitätsdaten mit Hilfe eines AI-powered Clinical Decision Support Systems. Alterstraumatologie Kongress. München, Deutschland. 2022
- **Kraus M,** Keppler AM, Saller MM, et. al. Reduziertes Bewusstsein für Osteoporose bei Patienten mit Hüftfrakturen im Vergleich zu Patienten, mit elektivem Hüftgelenkersatz. Jahreskongress der Vereinigung der Süddeutschen Orthopäden und Unfallchirurgen, Baden-Baden, Deutschland. 2022

2021

- Keppler AM, **Kraus M**, Saller MM, et. al. Identifikation von alterstraumatologischen "patients at risk" für ein postoperatives Mobilisationsdefizit mittels Handkraftmessung. Deutscher Kongress für Orthopädie und Unfallchirurgie, Berlin, Deutschland. 2021

1.1 Beitrag zur 1. Publikation

Kraus M, Saller MM, Baumbach SF et al (2022) **Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor Insole--Based Gait Analysis and Machine Learning Algorithms: Cross-sectional Study.** JMIR Med Inf 10:e32724. https://doi.org/10.2196/32724

Am Beginn des Projekts stand eine ausführliche Literaturrecherche, um die Konzeptionalisierung der Arbeit gemeinsam mit meinem wissenschaftlichen Betreuer Dr. rer. nat. Maximilian Saller durchführen zu können. Die so festgelegte Projektbasis wurde von mir selbst weiterentwickelt und die Umsetzung im Detail mit meinen ärztlichen und wissenschaftlichen Betreuern und Co-Autoren geplant. Alle in die vorliegende Arbeit eingeflossenen Daten wurden eigenständig in der Osteoporosesprechstunde des osteologischen Schwerpunktzentrums der LMU durch mich selbst erhoben und in einer extra hierfür durch mich aufgesetzten REDCap Studiendatenbank gespeichert und verwaltet, um die Datenqualität und -validität so hoch wie möglich zu halten. Die Auswertung der erhobenen Daten erfolgte mittels der Open Source Software R durch meine Person. Nach gemeinsamer Interpretation der Daten mit den Co-Autoren habe ich die Erstfassung des Manuskripts erstellt und dieses in Rücksprache mit den Co-Autoren eingereicht und anschließend mit Unterstützung insbesondere von Dr. rer. nat. Maximilian

Saller die Reviewer-Kommentare adressiert und das Manuskript erneut eingereicht. Nach Publikation des Manuskripts im Januar 2022 haben Dr. med. Alexander und ich im April 2022 gemeinsam die Bewerbung für den Digitalisierungspreis der Deutschen Gesellschaft für Orthopädie und Unfallchirurgie für die beste Publikation formuliert, welcher uns im Oktober 2022 zuerkannt wurde.

1.2 Beitrag zur 2. Publikation

Kraus, M., Stumpf, U. C., Keppler, A. M., Neuerburg, C., Böcker, W., Wackerhage, H., Baumbach, S. F., & Saller, M. M. (2023**). Development of a Machine Learning-Based Model to Predict Timed-Up-and-Go Test in Older Adults**. *Geriatrics (Basel, Switzerland)*, *8*(5), 99. https://doi.org/10.3390/geriatrics8050099

Zu Beginn des Projekts führte ich eine umfassende Literaturrecherche durch, um die Planung der Arbeit zusammen mit meinem wissenschaftlichen Betreuer Dr. rer. nat. Maximilian Saller zu gestalten. Basierend auf dieser Projektgrundlage habe ich das Projekt weiterentwickelt und im Detail mit den beteiligten Ärzten und wissenschaftlichen Betreuern geplant. Sämtliche Daten, die in dieser Arbeit verwendet wurden, wurden durch mich selbst, während der Osteoporosesprechstunde im osteologischen Schwerpunktzentrum erhoben und in einer speziell dafür durch mich eingerichteten REDCap-Studiendatenbank gespeichert und verwaltet, um die Qualität und Gültigkeit der Daten bestmöglich zu gewährleisten. Die Auswertung der gesammelten Daten wurde von mir unter Verwendung der Open-Source-Software R durchgeführt. Nach der gemeinsamen Interpretation der Daten mit den Co-Autoren habe ich die erste Version des Manuskripts erstellt und dieses in Absprache mit meinen Co-Autoren eingereicht. Anschließend habe ich eigenständig die Kommentare der Gutachter bearbeitet und das Manuskript erneut eingereicht, wobei ich insbesondere auf die Unterstützung von Dr. rer. nat. Maximilian Saller zählen konnte. Nach Annahme des Abstracts wurde unsere Publikation vom Journal als Cover-Paper der entsprechenden Ausgabe ausgewählt. Hierzu habe ich ein „Graphical-Abstract“ über unser Projekt erstellt, das in dieser Dissertation am Beginn der 2. Publikation abgebildet ist.

2. Einleitung

2.1.1 Introduction

Falls are a serious public health concern, especially for high-risk groups such as orthogeriatric patients and individuals undergoing trauma surgery. Falls can result in severe injuries, fractures, and other adverse health outcomes, leading to increased morbidity, mortality, and healthcare costs. In 2007, the World Health Organization report revealed that 32-42% of all people over 70 years of age fall once a year [1]. The economic burden of non-fatal falls in elderly patients in the United States of America was approximately $50 billion in 2017. [2] Therefore, accurately predicting fall risk and implementing appropriate preventive strategies is crucial to improve patient outcomes and to reduce the burden on healthcare systems as well as implementing efficient fall prevention initiatives. [3] About 5% of falls in orthogeriatric patients result in fractures, with an even higher incidence rate in patients suffering osteoporosis. [4] Osteoporotic fractures present a significant public health concern, particularly in an aging society where the incidence of fractures rises continuously. These fractures not only lead to reduced quality of life, but also place a substantial burden on healthcare resources. To address this issue, it is crucial to implement effective preventive measures that can mitigate the occurrence of fractures and optimize the utilization of healthcare resources. In this context, the risk assessment by gait analysis and non-mobility data among orthogeriatric patients holds great potential. [5]

2.1.2 Definition of Osteosarcopenia

Osteosarcopenia is a medical condition that poses a significant threat to orthogeriatric patients, particularly in relation to falls, fractures, and other adverse health events. [6] Osteosarcopenia is characterized by the simultaneous presence of two age-related conditions: osteoporosis, which is the loss of bone mass and deterioration of bone tissue, and sarcopenia, which is the progressive loss of muscle mass, strength, and function. [7] Both, osteoporosis and sarcopenia, individually contribute to an increased risk of falls and fractures, but when combined, they create a synergistic effect that significantly amplifies the vulnerability of orthogeriatric patients. [8]

The interaction between osteoporosis and sarcopenia extends beyond falls and fractures. The presence of these conditions in orthogeriatric patients is associated with an increased likelihood of experiencing adverse health events. [9] Individuals with osteosarcopenia may have reduced functional capacity, leading to difficulties in performing activities of daily living, compromised independence, and a higher risk of institutionalization. Additionally, the coexistence of osteoporosis and sarcopenia can result in prolonged recovery periods following fractures, higher rates of postoperative complications, and increased mortality rates. [10]

A multifaceted strategy is required to manage the hazard of osteosarcopenia in orthogeriatric patients. [8] This may involve implementing preventive measures, such as exercise programs to improve muscle strength and balance, optimizing nutrition to support bone health and muscle function, and ensuring appropriate pharmacological interventions to manage osteoporosis. [11] Furthermore, interdisciplinary collaborations among orthopedic surgeons, geriatricians, physical therapists, and nutritionists are essential to develop individual tailored treatment plans and interventions that address the specific needs of these patients. [12], [13]

2.1.3 Current Tools for Risk Stratification for Osteosarcopenia

It is crucial to assess the fall risk in orthogeriatric patients to identify those at risk early and enable timely preventive measures before a fall occurs (Figure 1). This approach can prevent a cascade of falls and ensuing fractures that may affect the active elderly, as seen in the top part of figure 1. It is critical to reduce the number of falls. This technique has the potential to reduce hospitalization rates while allowing patients to retain high levels of mobility and quality of life for as long as feasible. These facts suggest that fall prevention is better than any post-fall rehabilitation approach. As a result, improved risk assessment systems that leverage various, multivariate data synthesis, including artificial intelligence technologies in the future, are urgent.

Figure 1: Graphical abstract presenting the need for fall risk assessment in orthogeriatric patients

Various risk stratification tools have been developed to assess the risk of osteosarcopenia, considering both bone health (osteoporosis) and muscle function (sarcopenia) [14]. Two commonly used tools are the European Working Group on Sarcopenia in Older People (EWGSOP) [15] algorithm and the International Osteoporosis Foundation (IOF) Fracture Risk Assessment Tool ® (FRAX) [16].

The EWGSOP algorithm evaluates muscle mass, strength, and physical performance, along with bone mineral density (BMD) measured by dual x-ray absorptiometry (DXA), providing a comprehensive assessment of osteoporosis and sarcopenia [17]. In contrast, FRAX primarily assesses fracture risk due to osteoporosis using clinical risk factors [18].

While the EWGSOP algorithm provides a more accurate diagnosis of osteosarcopenia, further testing, such as DXA or BIA, is necessary. Conversely, FRAX is more convenient and pragmatic, encompassing some muscle-related risk factors indirectly through factors like previous falls. Nevertheless, it may not comprehensively assess sarcopenia.

The decision to use one of these tools depends on the specific clinical context, availability of resources, institutional infrastructure such as body impedance analyzers, and the primary outcome of interest - whether it's a comprehensive assessment of osteosarcopenia or an assessment of fracture risk.

2.2 Guidelines for the Assessment and Treatment of Osteosarcopenia

2.2.1 Guidelines for the Diagnosis and Management of Osteoporosis

Osteoporosis, a common skeletal disorder characterized by decreased bone density and increased fracture risk, has attracted significant attention from renowned organizations, including the Dachverband Osteolologie [19], the National Osteoporosis Guidelines Group [20], and the International Osteoporosis Foundation [21]. These organizations formulated protocols to assist the diagnosis and treatment of osteoporosis, furnishing invaluable perspectives on optimal methodologies and evidence-supported suggestions [22].

While aiming to improve patient outcomes, each set of guidelines exhibits subtle variations in their approach. The "Dachverband Osteolologie" guidelines prioritize a multifaceted osteoporosis management strategy, emphasizing bone mineral density analysis, clinical risk assessment tools, and detailed evaluation of fracture risk factors. This comprehensive approach facilitates tailored interventions based on thorough assessment of individual risk profiles.

Conversely, the National Osteoporosis Guidelines Group emphasizes the role of fracture risk assessment tools like FRAX® in guiding treatment decisions. They emphasize the assessment of each person's fracture risk taking into account measures of bone mineral density and clinical risk factors. These guidelines emphasize the importance of evaluating individual fracture risk, based on 20 million patient years [23] by taking clinical risk factors and bone mineral density measurements into account, to determine the most appropriate course of action. [14]

Similarly, the International Osteoporosis Foundation guidelines stress assessing fracture risk with FRAX® and advocate integrating pharmacological interventions, recommending specific medications such as bisphosphonates, Denosumab, and teriparatide-based on fracture risk profiles and patient characteristics.

Despite nuanced differences, all three guideline sets converge on the importance of lifestyle modifications, including weight-bearing exercise, smoking cessation, and sufficient calcium and vitamin D intake, as foundational components of osteoporosis management.

I have referred to the references in the preparation of the sub-projects of this dissertation and I was already able to present the collected baseline data and follow-up data at the annual congress of the DVO and “Deutscher Kongress für Orthopädie und Unfallchirurgie (DKOU). [24]–[27]

2.2.2 Guidelines for the Assessment of Sarcopenia

Sarcopenia is a disorder that is characterized by a progressive loss of muscle mass and function because of aging and has received significant attention from medical experts as a serious public health problem. To tackle this issue, key organizations such as the EWGSOP [17] and the Asian Working Group for Sarcopenia (AWGS) [28] have developed guidelines to enhance the diagnosis and treatment of sarcopenia. While both guidelines have a shared objective, a comparative analysis reveals important distinctions in their approaches.

The EWGSOP recommendations place a strong emphasis on using physical performance tests, muscle strength evaluations, and measurements of muscle mass to diagnose sarcopenia. They stress the significance of integrating these criteria into a comprehensive diagnostic algorithm, which facilitates a more precise assessment of muscle health in older adults. The EWGSOP guidelines offer cut-off points and reference values for each diagnostic criterion, assisting clinicians in efficiently interpreting and integrating the guidelines.

Conversely, the AWGS guidelines deem grip strength as a primary diagnostic criterion for sarcopenia and recommend lower grip strength thresholds for Asian populations based on age-related muscle loss in local cohorts. Additionally, the AWGS guidelines highlight the significance of additional markers such as gait speed and body composition in diagnosing sarcopenia in Asian populations. [28]

Both guidelines agree on the importance of regular physical activity, adequate protein intake, and resistance training as fundamental components of sarcopenia management. To optimize patient outcomes, a multidisciplinary approach involving health care professionals from different specialties is required and recognized as necessary by health care experts.

2.3 Gait Analysis in Orthogeriatric Patients

Gait analysis holds significant value to assess orthogeriatric patients, due to its ability to provide objective and quantitative data on gait parameters and patterns. [29] Gait analysis encompasses the methodical assessment of multiple facets of ambulation, including stride length, cadence, velocity, step width and time-and-space factors. [30] It contributes to a better understanding of the functional constraints and biomechanical changes that occur in orthogeriatric patients, giving useful insights for professional treatment and research.

Gait analysis is a highly effective technique in clinical practice for the diagnosis, monitoring, and management of orthogeriatric diseases. [31] Through the examination of gait patterns, healthcare professionals are able to detect variations from typical patterns and ascertain the root causes. For example, gait analysis can reveal asymmetries or abnormalities in weight-bearing distribution, indicating potential limb or joint pathologies. [32] It can provide information on balance impairments, muscle weakness, or joint stiffness that may contribute to increased fall risk in orthogeriatric patients. [33] This objective data assists to develop personalized treatment plans, to monitor the regeneration progress, and evaluating the effectiveness of interventions. [34]

Gait analysis is an essential component of orthogeriatric patient research. The investigators have the ability to examine the impact of various interventions, such as fitness programs, assistive devices, or surgical operations, on gait performance and functional outcomes through the analysis of gait patterns. [35] Gait analysis provides objective measures that can be compared across different patient groups or treatment modalities, to facilitate evidence-based decision-making and to advance our understanding of the impact of interventions on gait function in orthogeriatric patients. [36]

Gait analysis has the potential to become a tool for patient outcome control [35], [37]. By establishing standardized gait parameters and normative data for orthogeriatric populations, gait analysis can be used as an objective outcome measure. [38] This would enable doctors to analyze changes more precisely and quantitatively in gait performance over time, assess the effectiveness of therapies, and monitor functional recovery. Objective gait measurements might help to construct prediction models for identifying those who are at a higher risk of falling or having an adverse event, allowing for early interventions to reduce such risks. [39]

To realize the full potential of gait analysis as an outcome control tool, further advancements are needed. [35] Portable and cost-effective gait analysis systems, such as the Insole3 (Moticon, Munich, Germany), offer a widely accessible solution for comprehensive gait analysis in clinical environments, eliminating the need for a dedicated gait laboratory. [40] These systems enable detailed assessment of gait parameters without significant logistical or financial constraints.

In real-world situations, these insoles are effortlessly incorporated into patients' footwear and offer continuous and objective gait data. Mobile sensor insoles obviously have numerous benefits compared to standard gait labs. This innovative technique allows for the analysis of gait patterns during everyday motions and under changing environmental circumstances. Consequently, healthcare providers can obtain a better understanding of their patients' gait patterns and mobility restrictions, and eventually offer individualized intervention options.

Additionally, the establishment of standardized protocols and reference databases specific to orthogeriatric patients would enhance the interpretation and comparability of gait analysis results across different centers and studies. This will help to improve objective measurement of gait performance and functional recovery in orthogeriatric patients as well as predictive information on potential rehabilitation capacity and occurrence of adverse health events. [5]

2.4 Machine Learning

Machine learning (ML) is a branch of artificial intelligence focused on creating algorithms and models that learn from data to make predictions or decisions. [41] In orthogeriatric care, ML can help to advance to risk assessment and predicting rehabilitation outcomes through supervised and unsupervised learning approaches. [42], [43] There are two primary branches of machine learning: supervised and unsupervised learning. Supervised learning involves training a model with labeled data to predict outcomes based on patient characteristics and preoperative data. [44] This can help develop risk assessment tools for adverse events in orthopedic surgery. [45] Unsupervised learning, on the other hand, trains models without labels to discover hidden patterns, such as patient subgroups with similar characteristics or treatment outcomes. This can aid in identifying profiles with successful rehabilitation outcomes [46] or predict patient-reported outcomes. [47] ML models can provide more accurate risk assessment tools by incorporating a wide range of patient factors, from a large reference cohort. They should be used as decision support tools rather than replacing clinical judgment, as previous results could not achieve a significant improvement of the therapy results, mainly the treatment safety can be improved. [48] Surgeons need to critically evaluate and interpret the model outputs considering the specific context and individual patient characteristics.

2.4.1 Machine Learning for Evaluation of Gait-Analysis and Multidimensional Data

The combination of mobile sensor insoles, machine learning, and gait analysis has the potential to revolutionize orthogeriatric care, concept shown in figure 2 and is often referred to the term “smart gait”. [49] This multidimensional approach allows for the assessment of real-world gait data, accurate fall risk prediction, and personalized interventions for fracture prevention. The evaluation of gait patterns in orthogeriatric patients plays a crucial role in identifying potential risks and implementing preventive measures to mitigate the occurrence of fractures. Traditional gait analysis methods, predominantly conducted in specialized gait labs, have provided valuable insights into biomechanical parameters. [50] However, these methods are often limited to controlled environments, making it challenging to capture real-world gait patterns and predict fall risks accurately. In recent years, the integration of ML techniques and mobile sensor insoles has emerged as a promising approach to make it more accessible, enhance gait analysis, risk stratification, and fracture prevention in orthogeriatric patients. [50]

Figure 2: Graphical abstract for reporting vs. reality in fall risk

When evaluating fall risk, a considerable gap exists between standard assessments and the actual risk, particularly among orthogeriatric patients. This discrepancy is primarily due to the use of patient-reported questionnaires, which are prone to a high reporting bias. When physical tests such as the Timed-Up-and-Go-Test or a balance test are conducted, there is often examiner bias, as well as an increased patient motivation leading to false positive test results, when compared to real world performance. Therefore, it is critical to monitor patients with objective tools such as mobile wearables like wristbands or sensor soles that collect data around-the-clock for a specific period and accurately record real-world mobility. When analyzing data, it is preferable to use algorithms to ensure maximum objectivity and validity.

ML techniques have proven valuable in analyzing vast amounts of multidimensional gait data obtained by mobile sensor insoles. [49] These algorithms can detect tiny patterns and correlations in data that human viewers may miss. Researchers have built prediction models for fall risk assessment and patient categorization using ML algorithms. Such models consider a wide range of gait metrics, demographic variables, and comorbidities, allowing for tailored risk assessment and targeted treatments. Healthcare workers may get useful insights into patients' mobility patterns, identify deviations from normal gait, and more accurately forecast fall risk by merging mobile sensor insoles, medical record data, and machine learning approaches. [51] This information can aid in the development of targeted interventions, including exercise programs, environmental changes, and customized assistive devices, to enhance patient outcomes and decrease the occurrence of fractures. The integration of these advancements in clinical practice has potential to notably boost patient outcomes, enrich quality of life, and decrease healthcare expenses linked to orthogeriatric fractures. [52] Additional research and validation studies should be conducted to fully establish the effectiveness and clinical usefulness of this innovative approach.

2.5 Current State of Physical Frailty Assessment in Bedridden Orthogeriatric Patients

Assessing physical frailty in bedridden orthogeriatric patients remains a critical challenge in clinical practice. Currently, clinicians heavily rely on subjective evaluations such as the Clinical Frailty Scale (CFS) [53] or patient-reported questionnaires, which are prone to reporting bias and can result in an overestimation of individual abilities. [54] This leads to discrepancies between subjective assessments and objective measures [54] It is crucial to develop more objective and accurate methods for evaluating frailty in this population. The discrepancies highlight the necessity for improved assessment tools that precisely capture the multifaceted nature of physical frailty in this demographic. [55]

2.5.1 Discrepancies Between Clinical and Objective Assessments

The apparent discrepancies between clinical and objective assessments in immobilized orthogeriatric patients raise significant concerns about the accuracy and reliability of conventional assessment techniques. [56] Research reveals the inherent limitations of relying solely on subjective clinical appraisals, which reveal substantial disparities in comparison to objective assessments. This discrepancy emphasizes the necessity of supplementing traditional assessment approaches with objective and measurable standards to achieve complete understanding of physical frailty in immobilized patients. Additionally, exploration of ML methods presents a promising path for advancing assessment precision through integration of multifaceted clinical-, demographic- and mobility data and activities of daily living. [52] Utilizing cutting-edge ML methods may enable a more comprehensive assessment of physical frailty and fall risk, providing a more nuanced understanding that transcends the limitations of traditional assessments, particularly for individuals facing immobilization challenges.

2.5.2 Challenges in Conducting Physical Tests in Immobilized Patients

Conducting physical tests on immobilized orthogeriatric patients presents numerous challenges that hinder accurate and comprehensive assessment. These patients' limited mobility and functional ability considerably impedes the feasibility and reliability of conventional physical tests. Furthermore, the inability to conduct standard physical assessments due to immobilization exacerbates the difficulties in obtaining precise measurements of physical frailty. [57] Alternative assessment methodologies that surpass mobility-related constraints are crucial for a comprehensive evaluation of physical frailty in this susceptible patient population. Therefore, it is imperative to develop new approaches to overcome these challenges.

Traditional approaches on assessing physical frailty in bedridden orthogeriatric patients are often limited to subjective scales. Innovative methodologies are being sought to bridge the gap between subjective and objective assessments, and to overcome the difficulties involved in conducting physical tests on immobilized individuals.

2.5.3 Need for Assessment of Physical Status in Bedridden Patients

The evaluation of physical condition in bedridden patients, especially in cases of acute trauma resulting in immobilization, is crucial in clinical decision-making. [58] When a person suffers an acute hip fracture, the injury makes them immobile. Nothing was known about the patients' baseline physical fragility before to the incident, making it unable to assess their pre-injury mobility capabilities. It is vital to identify the individual's previous physical competence since it serves as the foundation for designing specific therapy techniques. In these cases, it is necessary to accurately examine the patients' physical health before to the accident in order to choose the most effective treatment procedures. The clinicians' capacity to make educated decisions about the most suited treatment method is severely impaired without this information of their pre-injury health.

Accurately assessing pre-injury physical condition is essential in clinical scenarios involving mobility and frailty assessment, highlighting the crucial role of predictive modeling. It is imperative to consider various clinical and demographic factors to understand the intricate interplay of variables that impact an individuals' physical status. ML algorithms provide a strong method for managing complex datasets and estimating the impact of individual variables relative to each other. [59] Utilizing these algorithms, clinicians can develop predictive models that are crucial in multiple clinical settings where assessing mobility or frailty is essential. These models not only assist in estimating the risk of falls but also provide insight for patient rehabilitation planning, resource allocation, and the development of personalized care strategies. [60] The implementation of predictive modeling through machine learning techniques is seen as a revolutionary approach to navigating the complexities in assessing physical status in immobilized patients, especially in acute trauma scenarios where pre-injury status is uncertain. [61]

2.5.4 Aim of the Dissertation

The dissertation endeavors to tackle the challenges of evaluating physical frailty in elderly patients, specifically those with orthopedic impairments, by employing novel methodologies based on ML and multivariate non-mobility data. The present research, which is based on two complimentary studies, attempts to rethink, and revolutionize the evaluation of physical frailty in this specific patient group. My main goal is to improve the precision, objectivity, and inclusiveness of physical frailty evaluation approaches.

The fundamental goal of this research is to provide and test alternative paradigms that go beyond the limitations of existing evaluation procedures, which usually suffer from difficulties of subjectivity, inaccuracy, and limited applicability among immobile persons. This dissertation aims to demonstrate, through extensive investigation and analysis, that ML-based gait analysis outperforms traditional questionnaires and physical exams in diagnosing physical frailty in orthogeriatric patients.

Furthermore, it aims to address a gap in existing approaches by developing objective ML models that may integrate a variety of non-mobility elements. These models aim to reliably anticipate the time required to complete the Timed-Up-and-Go test by reducing reliance on mobility-related information, supporting an impartial and automated diagnosis approach for physically weak individuals. The objective is to offer full and precise evaluations, allowing for better clinical decision-making and tailored treatments in orthogeriatric care settings.

The dissertations' comprehensive approach examines the complexity of diagnosing physical frailty using empirical methodologies, ML techniques, and sophisticated statistical analysis. This emphasizes the significance of powerful prediction models and deep understanding of multifactorial characteristics in modernizing the process of measuring physical frailty in aged orthogeriatric patients.

3. Zusammenfassung:

Die Beurteilung der körperliche Gebrechlichkeit (physical frailty) bei älteren Patienten, insbesondere bei solchen mit orthopädischen Beeinträchtigungen, stellt in der klinischen Praxis eine große Herausforderung dar. Dies liegt in erster Linie an der Subjektivität, Unzuverlässigkeit und Zeitintensität etablierter Bewertungsmethoden. Diese stützen sich in der Regel auf mobilitätsbezogene Daten und sind bei immobilen Personen größtenteils nicht anwendbar. Angesichts dieser Einschränkungen wurden zwei sich ergänzende Studien durchgeführt, um die Einschätzung der körperlichen Gebrechlichkeit in dieser demografischen Gruppe zu verbessern.

Ziel der ersten Studie war es, die Vorhersagegenauigkeit von Gangdaten, die während des Timed-Up-and-Go-Tests erhoben wurden, mit etablierten Fragebögen zur Einschätzung der „physical Frailty“ zu vergleichen. In dieser Studie wurden Algorithmen des maschinellen Lernens eingesetzt, um körperliche Gebrechlichkeit, definiert durch die Short Physical Performance Battery (SPPB), bei Patienten im Alter von über 60 Jahren zu identifizieren, die selbstständig gehfähig sind und keine geistigen oder neurologischen Beeinträchtigungen aufweisen. Diese Querschnittsuntersuchung erfasste verschiedene Parameter, die mit körperlicher Gebrechlichkeit assoziiert sind, und zeigte signifikante Unterschiede in den Gangparametern zwischen Gruppen mit und ohne körperliche Gebrechlichkeit. Darüber hinaus wies der Timed-Up-and-Go-Test im Vergleich zum SARC-F-Fragebogen (Strength, Assistance with walking, Rise from a chair, Climb stairs and Falls) einen höheren prädiktiven Wert auf, was durch eine "Area under the curve“ der „reciever operator characteristics“ (AUROC) von 0,862 gegenüber 0,639 belegt wird. Mithilfe rekursiver Variablen Auswahl identifizierten Algorithmen des maschinellen Lernens neun entscheidende Parameter, die aus digitalen Gangmessungen stammten. Mit diesen Parametern ließ sich eine robuste Vorhersagegenauigkeit erreichen, die zu AUROCs zwischen 0,801 und 0,919 führte. Die vorliegende Studie unterstreicht die Überlegenheit der auf maschinellem Lernen basierenden Ganganalyse bei der effizienten Identifizierung der körperlichen Gebrechlichkeit bei orthogeriatrischen Patienten im Vergleich zu herkömmlichen Methoden.

Die zweite Studie zielte darauf ab, die Grenzen der Bewertung körperlicher Gebrechlichkeit zu überwinden, indem objektive Modelle entwickelt wurden, die multifaktorielle Parameter nutzen, die nicht auf Mobilitätsmessungen beruhen. Mit diesem Ansatz wird die Abhängigkeit von mobilitätsbezogenen Daten überwunden und die Timed-Up-and-Go-Testzeit dennoch möglichst genau abgeschätzt. Unter Verwendung von sechs verschiedenen Algorithmen zur Feature-Selektion und 67 multifaktoriellen Parametern wurden in der Studie vier maschinelle Lernalgorithmen trainiert, darunter ein Generalized Linear Model, eine Support Vector Machine, ein Random Forest Algorithmus und ein Extreme Gradient Boost Algorithmus. Der Random Forest Algorithmus zeigte die höchste Genauigkeit bei der Vorhersage der Timed-Up-and-Go-Testzeit, mit einem mittleren absoluten Fehler von 2,7 Sekunden. Die Methodik der Variablenauswahl hatte nur minimalen Einfluss auf die Gesamtleistung des Modells. Allerdings neigten alle Algorithmen dazu, die Zeit für schnellere Patienten zu überschätzen und für langsamere Patienten zu unterschätzen. Diese Ergebnisse zeigen, dass es möglich ist, die Timed-Up-and-Go-Testzeit ohne Mobilitätsdaten vorherzusagen, was eine objektive Bewertung und automatische Identifizierung von körperlich gebrechlichen Patienten ermöglicht. Die Fortschritte haben das Potenzial, die Patientenversorgung und die Behandlungsplanung in der Orthogeriatrie zu verbessern und stellen einen revolutionären Ansatz für die klinische Entscheidungsfindung und personalisierte Interventionen dar.

4. Abstract (English)

Clinical practice has a barrier when assessing physical frailty in older patients, especially those with orthopedic limitations. This is mostly because standard assessment techniques are subjective, unreliable, and time-consuming. They also frequently depend on data relating to mobility, which may not be applicable to people who are immobile. Considering these limitations, two complementary studies were conducted to redefine the evaluation of physical frailty in this demographic group. The aim is to improve the evaluation and assessment of physical frailty.

The primary objective of the initial study was to examine and compare the efficacy of utilizing insole data obtained from the Timed-Up-and-Go test in comparison to known benchmark questionnaires and physical tests. The present study employed machine learning algorithms to detect physical frailty, as determined by the Short Physical Performance Battery (SPPB), in a cohort of individuals aged 60 years and above who possessed independent ambulation and did not exhibit any cognitive or neurological disorders. This study conducted a cross-sectional analysis to examine several factors related to physical frailty. The results showed notable disparities in gait metrics between individuals with and without physical frailty. Furthermore, the Timed-Up-and-Go test exhibited superior predictive value, when compared to the SARC-F (Strength, Assistance with walking, Rise from a chair, Climb stairs and Falls) questionnaire, as evidenced by an AUROC of 0.862 versus 0.639. Machine learning algorithms discovered nine critical characteristics, mostly from digital insole gait data, using recursive feature elimination. Robust predictive accuracy was achieved using these settings, with AUROCs ranging from 0.801 to 0.919. In summary, this research shows that machine learning-based gait analysis is superior to conventional evaluations when it comes to accurately detecting physical fragility in elderly individuals. The second study aimed to address the limitations of assessing physical frailty by developing objective machine models that utilize multifactorial non-mobility parameters. This approach dissociates reliance on mobility-related data and predicts the Timed-Up-and-Go test time accurately. Four machine learning methods—a generalized linear model, a support vector machine, a random forest algorithm, and an extreme gradient boost technique—were compared using six distinct feature selection approaches and 67 multifactorial variables. The random forest algorithm demonstrated the highest accuracy in predicting Timed-up-and-Gotest time, with a mean absolute error of 2.7 seconds. The variable selection methodology had minimal influence on the overall model performance. For slower patients, all algorithms tended to underestimate time, whereas for faster individuals, they tended to overestimate it. These results highlight the potential for Timed-Up-and-Go test time prediction in the absence of mobility data, enabling the automated identification and objective evaluation of patients who are physically frail. With this approach to clinical decision-making and tailored interventions, these developments might have the potential to significantly improve patient care and treatment planning in orthogeriatric settings.

Publication I

Veröffentlicht am: 05.01.2022

DOI: 10.2196/32724

4.1 Aim of Publication I

In the Paper “Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor Insole-Based Gait Analysis and Machine Learning Algorithms: Cross-sectional Study” we aimed at assessing the physical frailty of older patients, researchers sought to leverage modern insole wearables and ML algorithms to enhance the accuracy of evaluation methods. By contrasting the insole data obtained from the Timed-Up-and-Go-Test with traditional evaluations like the SARC-F (Strength, Assistance with walking, Rise from a chair, Climb stairs and Falls) questionnaire, the research sought to determine the most efficient method for assessing physical frailty, as defined by the Short Physical Performance Battery (SPPB). Through comprehensive analysis of multiple parameters, including body composition and gait patterns captured by digital sensor insoles, the study revealed that ML algorithms outperformed traditional methods in identifying physical frailty. This innovative gait analysis approach using sensor soles showcased its potential to revolutionize physical frailty assessments for orthogeriatric patients through the innovative use of machine learning algorithms and sensor soles, leading to more accurate and effective evaluation methods that can inform individualized therapies and improve the quality of care for patients at fall and fracture risk.

JMIR MEDICAL INFORMATICS Kraus et al

Original Paper

Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor Insole–Based Gait Analysis and Machine Learning Algorithms: Cross-sectional Study

Moritz Kraus; Maximilian Michael Saller, Dr rer nat; Sebastian Felix Baumbach, Dr med; Carl Neuerburg, Dr med, PD; Ulla Cordula Stumpf, Dr med; Wolfgang Böcker, Prof Dr med; Alexander Martin Keppler, Dr med

Department of Orthopaedics and Trauma Surgery, Musculoskeletal University Center Munich, Ludwig-Maximilians Universität Munich, Munich, Germany

Corresponding Author:
Alexander Martin Keppler, Dr med
Department of Orthopaedics and Trauma Surgery
Musculoskeletal University Center Munich
Ludwig-Maximilians Universität Munich
Marchioninistr. 15
Munich, 81377
Germany
Phone: 49 89 4400 0
Email: alexander.keppler@med.uni-muenchen.de

Abstract

Background: Assessment of the physical frailty of older patients is of great importance in many medical disciplines to be able to implement individualized therapies. For physical tests, time is usually used as the only objective measure. To record other objective factors, modern wearables offer great potential for generating valid data and integrating the data into medical decision-making.

Objective: The aim of this study was to compare the predictive value of insole data, which were collected during the Timed-Up-and-Go (TUG) test, to the benchmark standard questionnaire for sarcopenia (SARC-F: strength, assistance with walking, rising from a chair, climbing stairs, and falls) and physical assessment (TUG test) for evaluating physical frailty, defined by the Short Physical Performance Battery (SPPB), using machine learning algorithms.

Methods: This cross-sectional study included patients aged >60 years with independent ambulation and no mental or neurological impairment. A comprehensive set of parameters associated with physical frailty were assessed, including body composition, questionnaires (European Quality of Life 5-dimension [EQ 5D 5L], SARC-F), and physical performance tests (SPPB, TUG), along with digital sensor insole gait parameters collected during the TUG test. Physical frailty was defined as an SPPB score≤8. Advanced statistics, including random forest (RF) feature selection and machine learning algorithms (K-nearest neighbor [KNN] and RF) were used to compare the diagnostic value of these parameters to identify patients with physical frailty.

Results: Classified by the SPPB, 23 of the 57 eligible patients were defined as having physical frailty. Several gait parameters were significantly different between the two groups (with and without physical frailty). The area under the receiver operating characteristic curve (AUROC) of the TUG test was superior to that of the SARC-F (0.862 vs 0.639). The recursive feature elimination algorithm identified 9 parameters, 8 of which were digital insole gait parameters. Both the KNN and RF algorithms trained with these parameters resulted in excellent results (AUROC of 0.801 and 0.919, respectively).

Conclusions: A gait analysis based on machine learning algorithms using sensor soles is superior to the SARC-F and the TUG test to identify physical frailty in orthogeriatric patients.

(JMIR Med Inform 2022;10(1):e32724) doi: 10.2196/32724

KEYWORDS

wearables; insole sensors; orthogeriatric; artificial intelligence; prediction models; machine learning; gait analysis; digital sensors; digital health; aging; prediction algorithms; geriatric; mobile health; mobile insoles

Introduction

The physiological process of aging is inevitably connected to a decrease in physical performance [1]. It has been estimated that approximately 30% of the US population above the age of 55 years suffer from moderate to severe physical limitations [2]. In an orthogeriatric patient population, the assessment of physical frailty is of particular importance, as it is not only strongly associated with falls but also to an inferior outcome following surgery [3]. Consequently, it is of upmost importance to test for and thereby objectify physical impairment (ie, frailty).

Various individual parameters have been proposed to assess physical performance, including handgrip strength, daily step count, and gait speed. However, all of these have considerable interindividual variation [4]. Along with individual physiologic parameters, a variety of questionnaires such as the Barthel index [5], De-Morton Mobility index [6], or FRAIL scale [7] have been developed to quantify frailty. However, these questionnaires have proven to be inferior to the more complex physical assessments [8]. The Short Physical Performance Battery (SPPB) [9] is often considered one of the benchmark tests to assess frailty [8]. The SPPB combines multiple physical assessments, including gait, balance, and strength [10]. There is a consensus that screening for physical frailty is not only the prerequisite for successful individual patient care but also for cost-effectiveness [11]. Nonetheless, an international consensus on the most appropriate screening method is still missing [12].

As outlined above, comprehensive physical stance and gait assessments might be the most effective approach to quantify frailty. A new approach to assess physical activity and gait parameters includes the use of wearables and physical activity monitors [13]. These devices enable physicians and researchers to assess physical activity comprehensively under real-life conditions, and they have already been successfully applied to assist in the diagnosis of musculoskeletal diseases and to monitor rehabilitation [14-17]. A more recent development is sensor insoles with pressure and gyroscope sensors. These insoles can be easily inserted into any shoe and allow for the assessment of several gait parameters in an outpatient setting and also during various daily activities. This might provide a more feasible alternative to time-consuming assessments in specialized gait laboratories.

Although sensor insoles might help in the assessment of frailty, the large number of data points generated necessitates advanced statistical analysis. The random forest (RF) based on decision trees or the K-nearest neighbor (KNN) based on the Euclidean distance between points in high-dimensional space are two suitable strategies to develop clinical decision algorithms [18].

The aim of this study was to compare the classification capability of insole data collected during the Timed-Up-and-Go (TUG) test—a clinical gait test to assess a patient's mobility and risk of falling—to SARC-F (a five-item questionnaire for the quick assessment of the risk of sarcopenia, assessing strength, assistance with walking, rising from a chair, climbing stairs, and falls) and the TUG test to assess physical frailty, defined by the SPPB, using machine learning algorithms.

Methods

Patient Selection

Patients presenting to our orthogeriatric outpatient clinic for an osteoporosis diagnosis or therapy between December 2020 and March 2021 were invited to participate in this study. Inclusion criteria were aged >60 years, independent ambulation without any walking aids, and no mental or neurological impairment. Patients were informed of the study details, including the anonymized evaluation of the collected data, and then provided written consent. This cross-sectional study was approved by the local ethics committee (#19-177).

General Data Assessment

All data were collected in a standardized fashion by a unique, specially trained investigator. Demographic data included age, weight, height, BMI, body composition, general health-related quality of life assessed by the European Quality of Life 5-dimension (EQ-5D-5L) questionnaire [19], and the sarcopenia and physical frailty screening questionnaire SARC-F [20]. All questionnaires were completed together with the patients to obtain the highest possible data quality. Body composition (ie, body fat and muscle percentages) was measured using a clinically validated body composition monitor (BF511, Omron-Healthcare, Kyoto, Japan).

Assessment of Physical Frailty

Physical frailty was assessed by three different means: the SPPB, the TUG test, and digital insole gait parameters assessed during the TUG test using sensor insoles (Science3, Moticon, Munich, Germany).

The SPPB [9] is considered the benchmark test to assess physical frailty and was therefore used as the primary outcome parameter [8]. The SPPB is comprised of multiple tests for gait and stance safety, as well as lower-extremity strength and performance [10]. This tool scores the ability to stand in three different positions for 10 seconds, the time required to walk 3 meters, and the time it takes to rise from and sit down on a chair 5 times. Points are awarded for each subtest according to the time achieved, with a maximum score of 12 and a minimum score of 0. Patients with SPPB scores≤8 are considered to be physically frail [21,22]. The binary SPPB score (not physically frail vs physically frail) was used as the classification label for the machine learning models applied in this study.

The TUG test measures the time a patient takes to rise from a chair (height 46 centimeters), walk 3 meters, turn 180 degrees, and return to their initial seating position [23]. A duration of 12 seconds or longer has been associated with a higher probability of physical frailty [24]. Therefore, a cut-off value of 12 seconds was chosen to classify patients into physically frail and not physically frail groups.

The gait parameters were assessed by Science3 digital sensor insoles during the TUG test. Each of these insoles has 19 pressure sensors and a 3D gyroscope sensor to measure a variety of temporal, spatial, and local gait parameters, including gait speed and pressure distribution [25,26]. The parameters assessed are outlined in detail in Table 1.

Table 1. Overview of all insole gait parameters assessed.

Parameter	Unit
TUG[a] test time	seconds
Steps	number
Mean length of gait line	millimeters
Standard deviation x/y of gait line	meters
Mean total force during stance	Newtons
Mean gait cycle time	seconds
Mean gait cadence	strides/minute
Mean double support time	seconds
Mean acceleration over gait cycle (x/y/z)	*g*
Mean stride length	meters
Mean fraction of stance phase	%
Mean fraction of swing phase	%
Walking distance	meters
Mean walking speed	meters/second
COP[b] variability (left/right)	meters
COP trace length (left/right)	meters

[a]TUG: Timed-Up-and-Go.

[b]COP: center of pressure.

General Statistical Analysis

Unpaired *t* tests were used with α adjustment according to the Benjamini and Hochberg method [27] to compare interval-scaled, normally distributed variables (demographics, questionnaires, and gait parameters) between patients with and without physical frailty. Data are expressed as mean (SD). The effect size is expressed as the standardized mean difference.

Prediction Algorithms

To train the prediction algorithms, all collected performance- and nonperformance-related variables were used to train a recursive feature elimination algorithm that can identify the most relevant parameters for distinguishing patients with (SPPB score≤8) and without (SPPB score>8) physical frailty. For this purpose, the feature elimination algorithm was used to choose the best suitable variables based on an RF algorithm from the ranger package [28]. Gini impurity was used to rank the variables in order of their importance, as this measure is particularly suited to assess how well certain variables divide up a data set [29]. Based on this ordering of the variables, the variables were gradually removed until the lowest possible classification error was achieved. The classification error was chosen as the performance measure for the recursive feature selection, since the main focus was on maximizing the accuracy of the models developed later.

Two supervised machine learning algorithms, KNN [30] and RF, were used for further analysis using the previously selected variables. Both algorithms rely on being trained with labeled training data with a subsequent performance evaluation using test data. Prior to the training and tuning processes, the data were split into a training and a testing data set at a 70:30 ratio. The training process included an internal 3-fold cross-validation step. As hyperparameter tuning is essential for supervised machine learning algorithms to increase the accuracy of the classification [31], both algorithms were subjected to a tuning process that optimizes all variables to be tuned simultaneously, exclusively using the training data set. For the KNN, the tuning range for the number of neighbors was set from 1 to 22. For the type of kernels, the four variants rectangular, Gaussian, rank, and optimal were tested. For the unit of measurement of the distance, the options Euclidean distance, absolute distance, and Minkowski distance were available. For the RF, the number of variables considered as split candidates within a tree was tuned in the range of 1 to 7, the maximum number of branches in a tree was in the range of 2 to 10, and the number of trees in the RF was set from 100 to 1000. The nested resampling technique was used to enable better estimation of the true model performance on unseen data [32]. The 30% of the data not seen by the model were used to compare the performance of the different models subsequently.

To compare the generated algorithms to the classification properties of the TUG and SARC-F, confusion matrices and receiver operating characteristic (ROC) curves were created based on a logistic regression for the SARC-F using solely the score achieved and for the TUG using only the time taken to complete the test so as to compare the different prediction strategies. All data were collected in a REDCap study database [33] and analyzed in a standardized manner with RStudio software (version 1.3.1093), R (version 4.0.3), using the packages dplyr (version 1.0.2), Hmisc (version 4.6-0), ggplot2 (version 3.3.2), caret (version 6.0-86), and mlr3 (version 6.0-86)

Table 2. Comparison of demographics, body composition, physical activity, physical performance, and health questionnaire scores between patients with and without physical frailty.

Variable	No physical frailty (n=34)	Physical frailty (n=23)	*P* value	SMD[a]
Age (years), mean (SD)	74.76 (5.92)	80.00 (5.82)	.002	0.892
BMI (kg/m^2), mean (SD)	24.42 (4.81)	24.66 (3.79)	.84	0.055
Height (cm), mean (SD)	160.94 (6.37)	160.56 (7.84)	.85	0.053
Weight (kg), mean (SD)	62.77 (9.72)	63.45 (9.61)	.80	0.070
Body fat (%), mean (SD)	30.15 (8.55)	32.14 (7.86)	.37	0.243
Visceral fat (%), mean (SD)	7.95 (3.21)	8.71 (2.72)	.34	0.254
Muscle mass (%), mean (SD)	30.26 (4.20)	28.52 (3.29)	.09	0.460
Resting metabolism (kcal), mean (SD)	1345.32 (110.40)	1341.29 (123.22)	.90	0.034
Calf circumference, mean (SD)	35.04 (3.12)	34.31 (3.30)	.41	0.228
EQ-5D-5L[b] index, mean (SD)	0.84 (0.16)	0.65 (0.27)	.007	0.818
SPPB[c] score (points), mean (SD)	11.30 (0.79)	6.44 (2.06)	<.001	–3.106
SPPB score≤8, n (%)	0 (0)	23 (40)	<.001	
SARC-F[d] score, n (%)			.01	1.002
0	22 (65)	6 (26)		
1	8 (24)	7 (30)		
2	2 (6)	3 (13)		
3	0 (0)	4 (17)		
4	2 (6)	3 (13)		
Number of falls in past year, n (%)			.31	0.422
0	24 (71)	12 (52)		
1-3	7 (21)	9 (39)		
>3	3 (9)	2 (9)		
BMD[e] femoral neck (g/cm^3), mean (SD)	0.61 (0.06)	0.59 (0.06)	.27	0.303
BMD lumbar spine (g/cm^3), mean (SD)	0.85 (0.12)	0.91 (0.16)	.17	0.391
Smoking, n (%)			>.99	0.005
No	31 (91)	21 (91)		
Yes	3 (9)	2 (9)		
Self-sustaining, n (%)			.74	0.103
No	6 (18)	5 (22)		
Yes	28 (82)	18 (78)		
Daily leaving apartment, n (%)			.05	0.566
No	4 (12)	8 (35)		
Yes	30 (88)	15 (65)		
Weekly sports activity (>3 h), n (%)			.06	0.569
No	10 (29)	13 (57)		
Yes	24 (71)	10 (43)		

[a]SMD: standardized mean difference.

[b]EQ-5D-5L: European Quality of Life 5-dimension questionnaire.

[c]SPPB: Short Physical Performance Battery.

[d]SARC-F: sarcopenia test (strength, assistance with walking, rising from a chair, climbing stairs, and fall).

[34]. The code used to create and compare the models to the established tests has been made publicly available on GitHub [35].

Results

All of the 57 eligible consecutive orthogeriatric patients were included in the final analysis. The patients' mean age was 77 (SD 6) years and 93% were women. Classified by the SPPB, 23 patients (40%) had physical frailty. Table 2 shows the comparison of all assessed general parameters between the patients with and without physical frailty. Only age, EQ-5D-5L index, and SARC-F score differed significantly between the two groups. It should be emphasized that the average age of the patients with physical frailty was more than 5 years above the average age of the patients without physical frailty. In parallel, the mean health index of the patients with physical frailty determined by the EQ-5D-5L was almost 0.2 points below that of the patients without physical frailty. All other collected demographic data such as weight, height, BMI, body fat, and muscle mass did not differ significantly between the two groups.

The between-groups comparison of the digital gait analysis is presented in Table 3. The two groups differed significantly for all insole-generated gait parameters (all $P<.05$).

The classification errors of the TUG test and SARC-F to identify patients with physical frailty were 0.333 and 0.316, respectively. However, the area under the ROC curve (AUROC) for the TUG test was higher when compared with that of the SARC-F (0.862 vs 0.639; Figure 1A, Figure 1B).

The RF-based recursive feature elimination algorithm was trained to extract the most important features for classifying physical frailty using all parameters collected, except the SPPB, TUG test, and SARC-F, as they either define the result or represent the classification methods to be compared.

Based on the defined criteria, the 9 parameters outlined in Figure 2 were included. Notably, 8 out of the 9 parameters selected were gait parameters collected by the insoles (Figure 2). The number of steps and the step length were the most decisive factors for the identification of physical frailty by the algorithm. The gait speed followed in third place. Of the variables selected, double support seemed to have the least effect on classification.

These variables were then used to train the two classification algorithms KNN and RF. The tuning process resulted in an optimal combination of hyperparameters for the KNN as follows: k=15, a "rank" kernel, and the Minkowski distance. The optimal combination for the RF was 7 split variables, 6 branches, and 550 trees.

To compare the classification abilities of the TUG and the SARC-F with the algorithms created, a logistic regression was carried out on the SARC-F score and the TUG time on the dependent variable physical frailty and the ROC curve was drawn (Figure 1A-D). Table 4 summarizes the prediction accuracy of the four classifiers. Both classical approaches were outperformed by the machine learning–based models in terms of classification error (KNN=0.246, Figure 1D; RF=0.281, Figure 1C). The AUROC for the RF was slightly superior to that of the KNN (Table 4). Overall, the KNN showed the lowest error rate in classification at 24.6% (Figure 1). RF showed the largest AUROC value and thus appears to be the most suitable for classification. In the conventional tests, the TUG test was far superior to the SARC-F in terms of area under the ROC curve and classification error. The KNN showed the lowest classification error rate, but had a slightly smaller AUROC value than those of the RF and the TUG test.

Table 2. Comparison of demographics, body composition, physical activity, physical performance, and health questionnaire scores between patients with and without physical frailty.

Variable	No physical frailty (n=34)	Physical frailty (n=23)	*P* value	SMD[a]
Age (years), mean (SD)	74.76 (5.92)	80.00 (5.82)	.002	0.892
BMI (kg/m^2), mean (SD)	24.42 (4.81)	24.66 (3.79)	.84	0.055
Height (cm), mean (SD)	160.94 (6.37)	160.56 (7.84)	.85	0.053
Weight (kg), mean (SD)	62.77 (9.72)	63.45 (9.61)	.80	0.070
Body fat (%), mean (SD)	30.15 (8.55)	32.14 (7.86)	.37	0.243
Visceral fat (%), mean (SD)	7.95 (3.21)	8.71 (2.72)	.34	0.254
Muscle mass (%), mean (SD)	30.26 (4.20)	28.52 (3.29)	.09	0.460
Resting metabolism (kcal), mean (SD)	1345.32 (110.40)	1341.29 (123.22)	.90	0.034
Calf circumference, mean (SD)	35.04 (3.12)	34.31 (3.30)	.41	0.228
EQ-5D-5L[b] index, mean (SD)	0.84 (0.16)	0.65 (0.27)	.007	0.818
SPPB[c] score (points), mean (SD)	11.30 (0.79)	6.44 (2.06)	<.001	–3.106
SPPB score≤8, n (%)	0 (0)	23 (40)	<.001	
SARC-F[d] score, n (%)			.01	1.002
0	22 (65)	6 (26)		
1	8 (24)	7 (30)		
2	2 (6)	3 (13)		
3	0 (0)	4 (17)		
4	2 (6)	3 (13)		
Number of falls in past year, n (%)			.31	0.422
0	24 (71)	12 (52)		
1-3	7 (21)	9 (39)		
>3	3 (9)	2 (9)		
BMD[e] femoral neck (g/cm^3), mean (SD)	0.61 (0.06)	0.59 (0.06)	.27	0.303
BMD lumbar spine (g/cm^3), mean (SD)	0.85 (0.12)	0.91 (0.16)	.17	0.391
Smoking, n (%)			>.99	0.005
No	31 (91)	21 (91)		
Yes	3 (9)	2 (9)		
Self-sustaining, n (%)			.74	0.103
No	6 (18)	5 (22)		
Yes	28 (82)	18 (78)		
Daily leaving apartment, n (%)			.05	0.566
No	4 (12)	8 (35)		
Yes	30 (88)	15 (65)		
Weekly sports activity (>3 h), n (%)			.06	0.569
No	10 (29)	13 (57)		
Yes	24 (71)	10 (43)		

[a]SMD: standardized mean difference.

[b]EQ-5D-5L: European Quality of Life 5-dimension questionnaire.

[c]SPPB: Short Physical Performance Battery.

[d]SARC-F: sarcopenia test (strength, assistance with walking, rising from a chair, climbing stairs, and fall).

[e]BMD: bone mineral density.

Table 3. Comparison of gait parameters between patients with and without physical frailty.

Variable	No physical frailty, mean (SD)	Physical frailty, mean (SD)	*P* value	SMD[a]
Mean gait speed (m/s)	1.09 (0.28)	0.69 (0.19)	<.001	–1.637
TUG[b] time (s)	8.52 (1.93)	15.79 (5.50)	<.001	1.765
Mean stride length (m)	1.12 (0.19)	0.85 (0.17)	<.001	–1.450
Mean gait cadence (strides/min)	59.72 (8.83)	49.37 (8.21)	<.001	–1.214
Mean gait cycle time (s)	1.05 (0.16)	1.27 (0.20)	<.001	1.199
Mean double support time (s)	0.40 (0.13)	0.51 (0.14)	.003	0.843
Number of steps (n)	15.32 (6.05)	20.04 (5.67)	.005	0.804
Mean acceleration over gait cycle right (*g*)	0.03 (0.89)	0.59 (0.74)	.02	0.695
COP[c] trace length right (m)	5.25 (1.96)	7.06 (3.22)	.02	0.680
Mean acceleration over gait cycle right (*g*)	–2.36 (1.32)	–1.39 (1.54)	.02	0.672
Mean length width of gait line right (mm)	131.10 (21.20)	142.66 (19.05)	.04	0.574
Variance of acceleration over gait cycle (m/s^2)	1.66 (0.86)	1.21 (0.78)	.05	–0.552

[a]SMD: standardized mean difference.
[b]TUG: Timed-Up-and-Go.
[c]COP: center of pressure.

Figure 1. Comparison of the receiver operating characteristic (ROC) curves of the classification properties of the sarcopenia index SARC-F (A), Timed-Up-and-Go (TUG) test (B), and the random forest (C) and k-nearest neighbor (D) algorithms. AUC: area under the ROC curve.

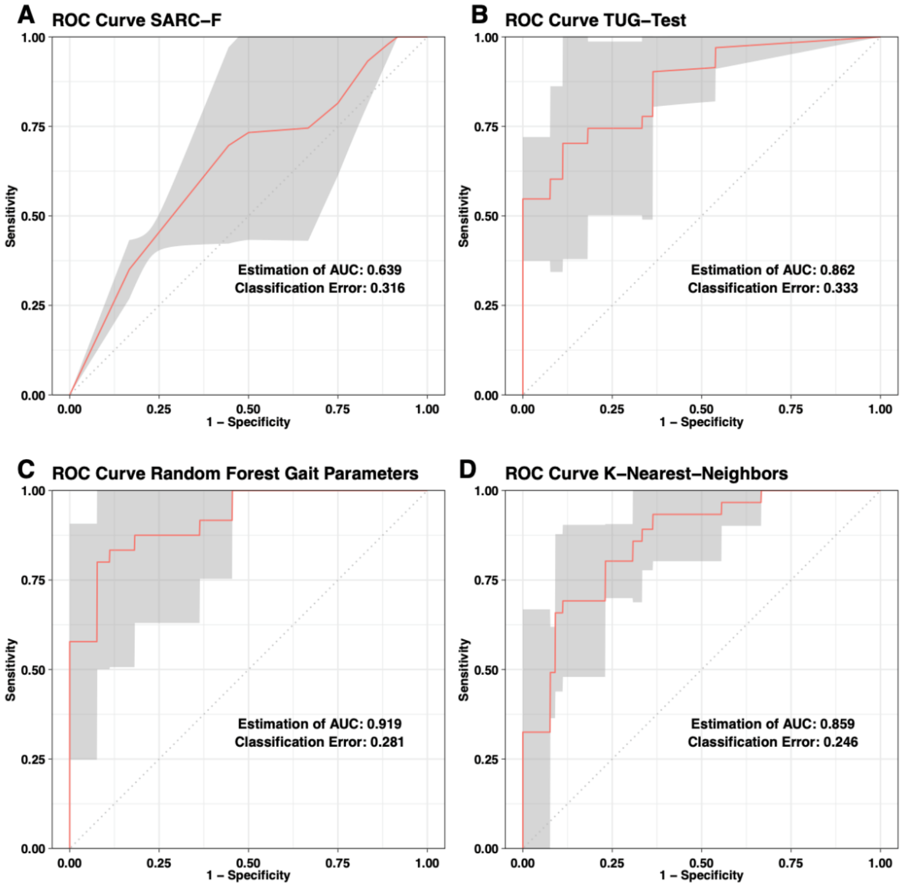


Figure 2. Selected parameters based on the recursive feature elimination algorithm, ordered by their importance for reduction of classification error ranked by Gini-Impurity [29].

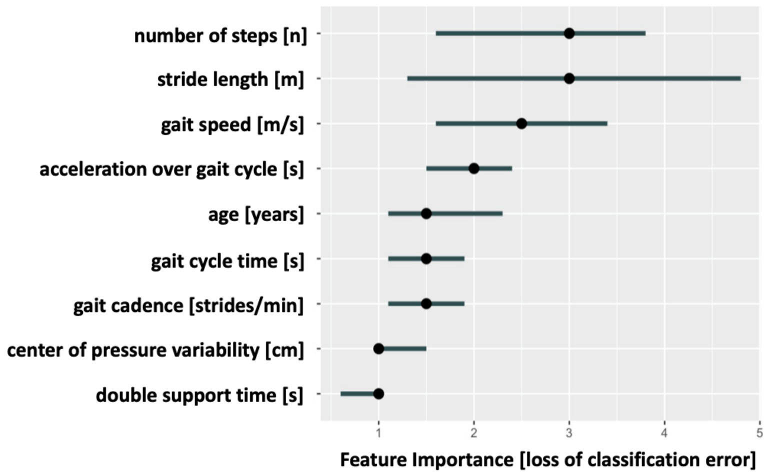


Table 4. Comparison of physical frailty prediction methods.

Performance metric	SARC-F[a] LR[b]	TUG[c] test LR	KNN[d] classifier	RF[e] classifier
Accuracy	0.684	0.667	0.719	0.724
AUROC[f]	0.639	0.862	0.919	0.859

[a]SARC-F: sarcopenia test (strength, assistance with walking, rising from a chair, climbing stairs, and fall).

[b]LR: logistic regression.

[c]TUG: Timed-Up-and-Go.

[d]KNN: K-nearest neighbor.

[e]RF: random forest.

[f]AUROC: area under the receiver operating characteristic curve.

Discussion

Principal Findings

Based on a sample of 57 patients and advanced statistics, this study shows that gait parameters assessed by digital insoles during the TUG test outperformed both the benchmark tests (the TUG physical assessment and SARC-F questionnaire) to identify patients with physical frailty.

Patients identified as physically frail classified by their SPPB scores (≤8) were on average 5 years older than patients that were not classified as physically frail, with no significant difference in BMI or body composition. By contrast, previous studies have reported a decreased muscle mass and increased fat percentage in patients with physical frailty [36]. Despite the considerable amount of physical frailty–related data collected (Tables 1 and 2), the vast majority (8 out of 9) of the parameters selected by the recursive feature elimination algorithm were insole gait parameters collected during the TUG test. Although the temporal gait variables such as gait speed, double support time, and gait cadence can be considered dependent variables, they all reflect different aspects of gait. For this reason, it makes sense to integrate several of these aspects into the machine learning algorithms to better map the gait pattern of an individual patient and derive the best possible classification.

Previous studies have proposed that gait speed is the most relevant parameter to identify patients with physical frailty [4]. It has been shown that a slow gait speed is associated with an increased fall risk [37], as well as a higher mortality rate [38]. Interestingly, the advanced modeling used in this study weighted stride length equally important as gait speed to differentiate between physical frailty and no physical frailty in patients, in terms of their classification importance measured by the Gini impurity (Figure 2). Although gait speed is easily assessed, it might be biased by patients' motivation. One can hypothesize a "white coat effect," in this case a higher level of motivation during medical gait speed examinations. Stride length might be a more robust (ie, harder to influence consciously) parameter in such settings, which might explain its superiority in the herein applied modeling. Espy et al [39] provided a possible explanation for the higher robustness of stride length compared to gait speed. They were able to show that a slow gait leads to instability, which again is compensated for by a small-stepped gait pattern [39]. It appears reasonable that patients with physical frailty would therefore compensate for their unstable gait pattern by a reduction of their stride length [39]. Overall, stride length and gait speed were found to be the two most relevant parameters for the model (Figure 2), and could only be slightly increased by adding additional gait parameters such as cadence, double support time, and acceleration over gait cycle. Consequently, stride length in addition to gait speed might be a valuable clinical parameter to identify patients with physical frailty. Their early identification is essential to reduce the number of falls [37] and possibly mortality rates [38], as well as to increase further health outcomes [40]. These considerable implications are not only important in an orthogeriatric setting but also for almost all medical specialties.

In line with previous studies, the SARC-F as well as the TUG test were found to be suitable for estimating the physical frailty status [41]. The slightly better results for the TUG test compared with the SARC-F might be explained by their different natures. The SARC-F is a patient-reported outcome measure, whereas the TUG test is a more objective score. Older patients have been shown to overestimate their physical abilities [42,43], which might result in false negative SARF-F scores. Complementing the SARC-F by an objective measurement such as the TUG test, handgrip strength, or a gait analysis might increase its accuracy and therefore screening value.

Nevertheless, the combination of machine learning algorithms and digital gait analysis outperformed the TUG test and SARC-F in the detection of physical frailty. The digital insoles used in this study can easily be applied and have proven to be reliable [25]. Furthermore, they could be integrated into health assessment apps, such as on a smartphone. This can facilitate both the collection of longitudinal data and remote monitoring of at-risk patients, and potentially even guide rehabilitation. Consequently, gait analysis by digital insoles might become another valuable part of the growing body of digital health devices.

Limitations and Strengths

An obvious limitation of this study is the limited number of patients. The smaller the number of patients the algorithm is trained on, the more limited is its generalizability. Therefore, the herein proposed algorithm must be validated in a larger cohort. In the setting of a longitudinal, multicenter trial, the applied statistics could be extended to deep learning methods such as neural networks, which could further increase the accuracy of the predictions. Another limitation is the definition of physical frailty. Due to the current setup, it was only possible to define physical frailty by the SPPB. Although the SPBB is considered one of the benchmark tests for physical frailty [44], it would be even more meaningful to directly assess the occurrence of various health impairments such as falls, fractures, progression to impaired ambulation, or death. Nonetheless, these parameters can only be assessed in a longitudinal study setup.

Despite these limitations, several strengths of this study are noteworthy. First, the combined use of modern wearables and data analysis strategies from the field of data science to complement the classic statistical analysis is an advantage of this study. Due to the increasing amount of data points collected by digital devices, advanced statistics will become the primary working horse to analyze the data. Second, the meta-modeling approach applied represents a pessimistic estimation of the models' performance in a larger cohort. Nevertheless, the resulting AUROC values of 0.801 and 0.841 can be judged as excellent [45]. These excellent results argue for the value of digital insole gait parameters. For application in clinical practice, it is conceivable that a doctor will receive an analysis on their terminal device in real time during the test, which can provide time-efficient support in clinical decision-making for or against prescribing fall prevention training, certain medications, or other therapeutic interventions. Finally, this study also indicates that gait parameters might be a promising target for physical frailty therapies. It can by hypothesized that focused physiotherapy or fall risk minimization counseling could counteract physical frailty and thereby increase the patient's health-related quality of life.

Conclusion

Machine learning algorithms–based gait analysis using mobile insoles appears to be a promising approach to screen for physical frailty in an outpatient setting. Due to the increasing amount of data collected, high-performance data processing will become increasingly important. Future large-scale, longitudinal, and multicenter screening trials should collect as many data points as possible, including from digital devices such as wearables, and apply advanced statistics to increase the diagnostic sensitivity and accuracy of physical frailty diagnosis.

Conflicts of Interest

None declared.

References

1. Cruz-Jimenez M. Normal changes in gait and mobility problems in the elderly. Phys Med Rehabil Clin N Am 2017 Nov;28(4):713-725. [doi: 10.1016/j.pmr.2017.06.005] [Medline: 29031338]
2. Tieland M, Trouwborst I, Clark BC. Skeletal muscle performance and ageing. J Cachexia Sarcopenia Muscle 2018 Feb;9(1):3-19. [doi: 10.1002/jcsm.12238] [Medline: 29151281]
3. Marquez A, Queirós C. Frailty, sarcopenia and falls. In: Hertz K, Santy-Tomlinson J, editors. Fragility fracture nursing: holistic care and management of the orthogeriatric patient. Cham: Springer; 2018:15-26.

4. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. JAMA 2011 Jan 05;305(1):50-58 [FREE Full text] [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
5. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
6. de Morton NA, Davidson M, Keating JL. The de Morton Mobility Index (DEMMI): an essential health index for an ageing world. Health Qual Life Outcomes 2008 Aug 19;6:63 [FREE Full text] [doi: 10.1186/1477-7525-6-63] [Medline: 18713451]
7. Thompson MQ, Theou O, Tucker GR, Adams RJ, Visvanathan R. FRAIL scale: Predictive validity and diagnostic test accuracy. Australas J Ageing 2020 Dec;39(4):e529-e536. [doi: 10.1111/ajag.12829] [Medline: 32748992]
8. Freiberger E, de Vreede P, Schoene D, Rydwik E, Mueller V, Frändin K, et al. Performance-based physical function in older community-dwelling persons: a systematic review of instruments. Age Ageing 2012 Nov;41(6):712-721. [doi: 10.1093/ageing/afs099] [Medline: 22885845]
9. Treacy D, Hassett L. The Short Physical Performance Battery. J Physiother 2018 Jan;64(1):61 [FREE Full text] [doi: 10.1016/j.jphys.2017.04.002] [Medline: 28645532]
10. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
11. Bleijenberg N, Drubbel I, Neslo RE, Schuurmans MJ, Ten Dam VH, Numans ME, et al. Cost-effectiveness of a proactive primary care program for frail older people: a cluster-randomized controlled trial. J Am Med Dir Assoc 2017 Dec 01;18(12):1029-1036.e3 [FREE Full text] [doi: 10.1016/j.jamda.2017.06.023] [Medline: 28801235]
12. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
13. De Luca V, Muaremi A, Giggins O, Walsh L, Clay I. Towards fully instrumented and automated assessment of motor function tests. 2018 Mar 01 Presented at: IEEE EMBS Int Conf Biomed Heal Informatics; 2018; Las Vegas p. 83. [doi: 10.1109/bhi.2018.8333375]
14. Keppler AM, Nuritidinow T, Mueller A, Hoefling H, Schieker M, Clay I, et al. Validity of accelerometry in step detection and gait speed measurement in orthogeriatric patients. PLoS One 2019;14(8):e0221732 [FREE Full text] [doi: 10.1371/journal.pone.0221732] [Medline: 31469864]
15. Pfeufer D, Becker CA, Faust L, Keppler AM, Stagg M, Kammerlander C, et al. Load-bearing detection with insole-force sensors provides new treatment insights in fragility fractures of the pelvis. J Clin Med 2020 Aug 06;9(8):2551 [FREE Full text] [doi: 10.3390/jcm9082551] [Medline: 32781675]
16. Keppler AM, Holzschuh J, Pfeufer D, Neuerburg C, Kammerlander C, Böcker W, et al. Postoperative physical activity in orthogeriatric patients - new insights with continuous monitoring. Injury 2020 Mar;51(3):628-632. [doi: 10.1016/j.injury.2020.01.041] [Medline: 32033808]
17. Mueller A, Hoefling HA, Muaremi A, Praestgaard J, Walsh LC, Bunte O, et al. Continuous digital monitoring of walking speed in frail elderly patients: noninterventional validation study and longitudinal clinical trial. JMIR Mhealth Uhealth 2019 Nov 27;7(11):e15191 [FREE Full text] [doi: 10.2196/15191] [Medline: 31774406]
18. Garg A, Mago V. Role of machine learning in medical research: a survey. Comput Sci Rev 2021 May;40:100370. [doi: 10.1016/j.cosrev.2021.100370]
19. Hernandez G, Garin O, Dima AL, Pont A, Martí Pastor M, Alonso J, ASTRO-LAB Group. EuroQol (EQ-5D-5L) validity in assessing the quality of life in adults with asthma: cross-sectional study. J Med Internet Res 2019 Jan 23;21(1):e10178 [FREE Full text] [doi: 10.2196/10178] [Medline: 30672744]
20. Hax V, do Espírito Santo RC, Dos Santos LP, Farinon M, de Oliveira MS, Três GL, et al. Practical screening tools for sarcopenia in patients with systemic sclerosis. PLoS One 2021;16(1):e0245683 [FREE Full text] [doi: 10.1371/journal.pone.0245683] [Medline: 33481872]
21. Yuguchi S, Saitoh M, Oura K, Tahara M, Kamisaka K, Kawamura T, et al. Impact of preoperative frailty on regaining walking ability in patients after cardiac surgery: Multicenter cohort study in Japan. Arch Gerontol Geriatr 2019;83:204-210. [doi: 10.1016/j.archger.2019.04.003] [Medline: 31082565]
22. Lim YJ, Ng YS, Sultana R, Tay EL, Mah SM, Chan CHN, et al. Frailty assessment in community-dwelling older adults: a comparison of 3 diagnostic instruments. J Nutr Health Aging 2020;24(6):582-590. [doi: 10.1007/s12603-020-1396-2] [Medline: 32510110]
23. Miyoshi S, Morita T, Kadoi Y, Goto F. Analysis of the factors related to a decrease in jugular venous oxygen saturation in patients with diabetes mellitus during normothermic cardiopulmonary bypass. Surg Today 2005;35(7):530-534. [doi: 10.1007/s00595-004-2977-0] [Medline: 15976948]
24. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/steadi/pdf/TUG_Test-print.pdf [accessed 2021-05-10]
25. Braun BJ, Veith NT, Hell R, Döbele S, Roland M, Rollmann M, et al. Validation and reliability testing of a new, fully integrated gait analysis insole. J Foot Ankle Res 2015;8(1):54 [FREE Full text] [doi: 10.1186/s13047-015-0111-8] [Medline: 26396594]

26. Braun BJ, Bushuven E, Hell R, Veith NT, Buschbaum J, Holstein JH, et al. A novel tool for continuous fracture aftercare - Clinical feasibility and first results of a new telemetric gait analysis insole. Injury 2016 Feb;47(2):490-494. [doi: 10.1016/j.injury.2015.11.004] [Medline: 26626806]
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 2018 Dec 05;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
28. Wright MN, Ziegler A. ranger: a fast implementation of random forests for high dimensional data in C++ and R and. J Stat Soft 2017 Mar 01;77(1):1-17. [doi: 10.18637/jss.v077.i01]
29. Archer KJ, Kimes RV. Empirical characterization of random forest variable importance measures. Comput Stat Data Anal 2008 Jan;52(4):2249-2260. [doi: 10.1016/j.csda.2007.08.015]
30. Keller JM, Gray MR, Givens JA. A fuzzy K-nearest neighbor algorithm. IEEE Trans Syst Man Cybern 1985 Jul;SMC-15(4):580-585. [doi: 10.1109/tsmc.1985.6313426]
31. van Rijn J, Hutter F. Hyperparameter importance across datasets. : ACM; 2018 Aug 19 Presented at: 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; August 19-23, 2018; London URL: https://dl.acm.org/doi/10.1145/3219819.3220058 [doi: 10.1145/3219819.3220058]
32. Bischl B, Mersmann O, Trautmann H, Weihs C. Resampling methods for meta-model validation with recommendations for evolutionary computation. Evol Comput 2012 Jun;20(2):249-275. [doi: 10.1162/evco_a_00069]
33. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
34. Lang M, Binder M, Richter J, Schratz P, Pfisterer F, Coors S, et al. mlr3: A modern object-oriented machine learning framework in R. JOSS 2019 Dec;4(44):1903. [doi: 10.21105/joss.01903]
35. Kraus M. Models for gait analysis evaluation Internet. GitHub. URL: https://github.com/KrausMoritz/Gait_Analysis [accessed 2021-10-01]
36. Cesari M, Leeuwenburgh C, Lauretani F, Onder G, Bandinelli S, Maraldi C, et al. Frailty syndrome and skeletal muscle: results from the Invecchiare in Chianti study. Am J Clin Nutr 2006 May 01;83(5):1142-1148 [FREE Full text] [doi: 10.1093/ajcn/83.5.1142] [Medline: 16685058]
37. Verghese J, Holtzer R, Lipton RB, Wang C. Quantitative gait markers and incident fall risk in older adults. J Gerontol A Biol Sci Med Sci 2009 Aug 06;64(8):896-901 [FREE Full text] [doi: 10.1093/gerona/glp033] [Medline: 19349593]
38. Stanaway FF, Gnjidic D, Blyth FM, Le Couteur DG, Naganathan V, Waite L, et al. How fast does the Grim Reaper walk? Receiver operating characteristics curve analysis in healthy men aged 70 and over. BMJ 2011 Dec 15;343(dec15 1):d7679-d7679 [FREE Full text] [doi: 10.1136/bmj.d7679] [Medline: 22174324]
39. Espy D, Yang F, Bhatt T, Pai Y. Independent influence of gait speed and step length on stability and fall risk. Gait Posture 2010 Jul;32(3):378-382 [FREE Full text] [doi: 10.1016/j.gaitpost.2010.06.013] [Medline: 20655750]
40. Pandey A, Kitzman D, Whellan DJ, Duncan PW, Mentz RJ, Pastva AM, et al. Frailty among older decompensated heart failure patients: prevalence, association with patient-centered outcomes, and efficient detection methods. JACC Heart Fail 2019 Dec;7(12):1079-1088 [FREE Full text] [doi: 10.1016/j.jchf.2019.10.003] [Medline: 31779931]
41. Si H, Jin Y, Qiao X, Tian X, Liu X, Wang C. Comparison of 6 frailty screening tools in diagnostic properties among Chinese community-dwelling older people. Geriatr Nurs 2021 Jan;42(1):276-282. [doi: 10.1016/j.gerinurse.2020.08.017] [Medline: 32948340]
42. Kawasaki T, Tozawa R. Motor function relating to the accuracy of self-overestimation error in community-dwelling older adults. Front Neurol 2020;11:599787. [doi: 10.3389/fneur.2020.599787] [Medline: 33329358]
43. Martin P, Keppler AM, Alberton P, Neuerburg C, Drey M, Böcker W, et al. Self-Assessment of mobility of people over 65 years of age. Medicina (Kaunas) 2021 Sep 17;57(9):980 [FREE Full text] [doi: 10.3390/medicina57090980] [Medline: 34577903]
44. Chen J, Liang C, Chang Q. Comparison of fallers and nonfallers on four physical performance tests: a prospective cohort study of community-dwelling older indigenous Taiwanese women. Int J Geront 2018 Mar;12(1):22-26. [doi: 10.1016/j.ijge.2017.04.006]
45. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010 Sep;5(9):1315-1316 [FREE Full text] [doi: 10.1097/JTO.0b013e3181ec173d] [Medline: 20736804]

Abbreviations

AUROC: area under the receiver operating characteristic curve
EQ-5D-5L: European Quality of Life 5-dimension
KNN: K-nearest neighbor
RF: random forest
ROC: receiver operating characteristic
SARC-F: sarcopenia questionnaire (strength, assistance with walking, rising from a chair, climbing stairs, and falls)

SPPB: Short Physical Performance Battery
TUG: Timed Up and Go

Edited by C Lovis, J Hefner; submitted 07.08.21; peer-reviewed by I Clay, SD Boie; comments to author 19.09.21; revised version received 29.10.21; accepted 10.11.21; published 05.01.22

Please cite as:
Kraus M, Saller MM, Baumbach SF, Neuerburg C, Stumpf UC, Böcker W, Keppler AM
Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor Insole–Based Gait Analysis and Machine Learning Algorithms: Cross-sectional Study
JMIR Med Inform 2022;10(1):e32724
URL: https://medinform.jmir.org/2022/1/e32724
doi: 10.2196/32724
PMID:

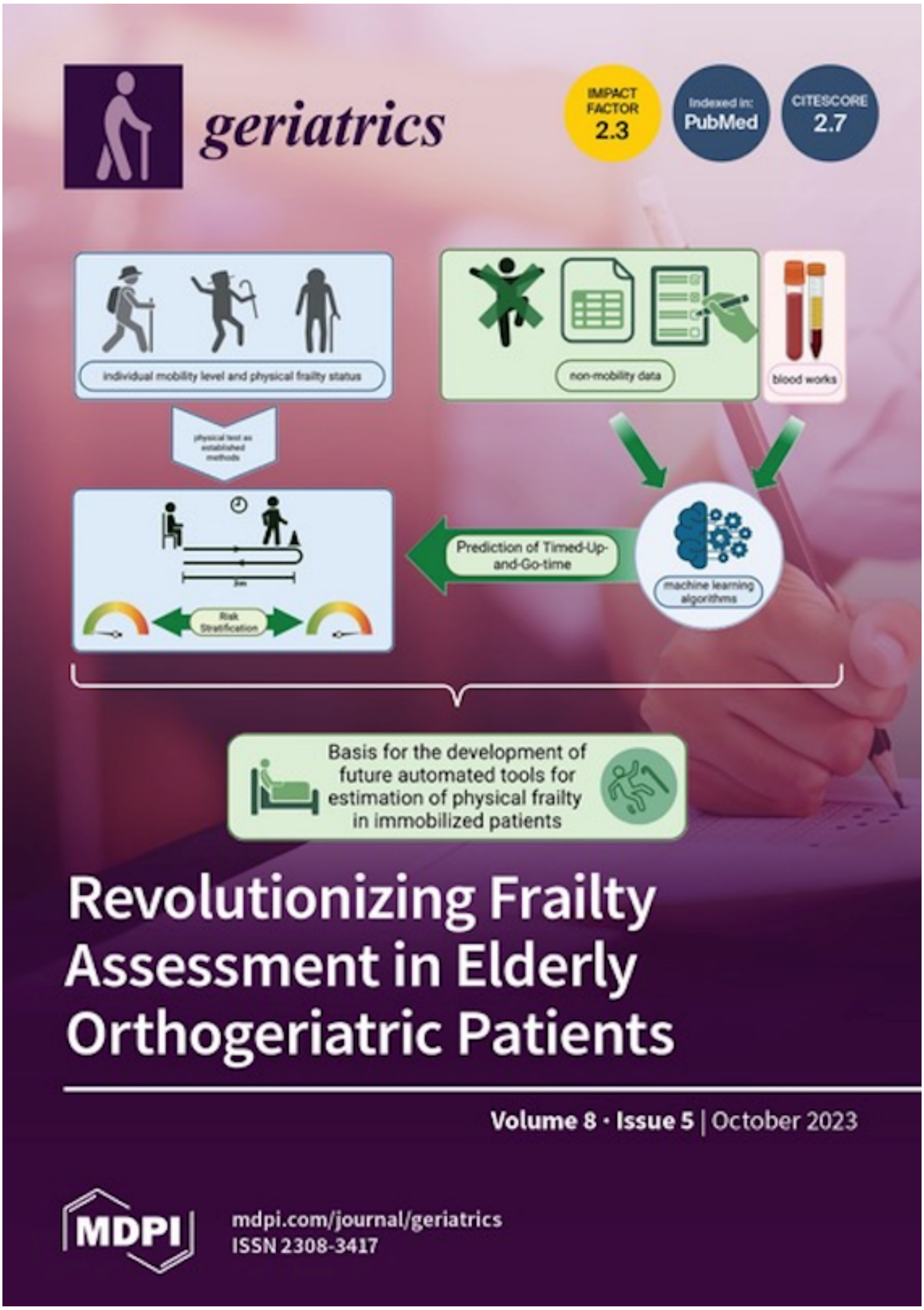
Publication II

Veröffentlicht am: 07.10.2023

DOI: 10.3390/geriatrics8050099

Aim of publication II

The study aimed to redefine physical frailty assessment in bedridden patient groups by utilizing multifactorial non-mobility data. The main objective was to create and validate strong machine learning models that could precisely predict Timed-Up-and-Go test times. This innovative approach seeks to bypass the use of traditional mobility-related measures and provides an objective and inclusive system for evaluating physical ability and locomotion results in immobilized individuals. The study aimed to develop a nuanced understanding of Timed-Up-and-Go test outcomes using various biological, inflammatory, and physiological markers, prioritizing systemic inflammation and physiological markers over chronological age. Ultimately, the study aimed to develop a precise, objective, and detailed assessment of physical frailty among bedridden patients. This will help to improve clinical decision-making and customize interventions in orthogeriatric care settings.


geriatrics
IMPACT FACTOR 2.3
Indexed in: PubMed
CITESCORE 2.7
individual mobility level and physical frailty status
physical test as established methods
non-mobility data
blood works
Prediction of Timed-Up-and-Go-time
machine learning algorithms
Risk Stratification
Basis for the development of future automated tools for estimation of physical frailty in immobilized patients
Revolutionizing Frailty Assessment in Elderly Orthogeriatric Patients
Volume 8 · Issue 5 | October 2023
MDPI
mdpi.com/journal/geriatrics
ISSN 2308-3417

geriatrics

MDPI

Article

Development of a Machine Learning-Based Model to Predict Timed-Up-and-Go Test in Older Adults

Moritz Kraus [1,*], Ulla Cordula Stumpf [1], Alexander Martin Keppler [1], Carl Neuerburg [1], Wolfgang Böcker [1], Henning Wackerhage [2], Sebastian Felix Baumbach [1] and Maximilian Michael Saller [1]

[1] Department of Orthopedics and Trauma Surgery, Musculoskeletal University Center Munich, University Hospital of Ludwig-Maximilians-University (LMU), 81377 Munich, Germany; ulla.stumpf@med.uni-muenchen.de (U.C.S.); alexander.keppler@med.uni-muenchen.de (A.M.K.); carl.neuerburg@med.uni-muenchen.de (C.N.); wolfgang.boecker@med.uni-muenchen.de (W.B.); sebastian.baumbach@med.uni-muenchen.de (S.F.B.); maximilian.saller@med.uni-muenchen.de (M.M.S.)

[2] Faculty of Sport and Health Sciences, Technical University of Munich, 80809 Munich, Germany; henning.wackerhage@tum.de

* Correspondence: moritz.kraus@med.uni-muenchen.de; Tel.: +49-89-4400-0

Abstract: Introduction: The measurement of physical frailty in elderly patients with orthopedic impairments remains a challenge due to its subjectivity, unreliability, time-consuming nature, and limited applicability to uninjured individuals. Our study aims to address this gap by developing objective, multifactorial machine models that do not rely on mobility data and subsequently validating their predictive capacity concerning the Timed-up-and-Go test (TUG test) in orthogeriatric patients. Methods: We utilized 67 multifactorial non-mobility parameters in a pre-processing phase, employing six feature selection algorithms. Subsequently, these parameters were used to train four distinct machine learning algorithms, including a generalized linear model, a support vector machine, a random forest algorithm, and an extreme gradient boost algorithm. The primary goal was to predict the time required for the TUG test without relying on mobility data. Results: The random forest algorithm yielded the most accurate estimations of the TUG test time. The best-performing algorithm demonstrated a mean absolute error of 2.7 s, while the worst-performing algorithm exhibited an error of 7.8 s. The methodology used for variable selection appeared to exert minimal influence on the overall performance. It is essential to highlight that all the employed algorithms tended to overestimate the time for quick patients and underestimate it for slower patients. Conclusion: Our findings demonstrate the feasibility of predicting the TUG test time using a machine learning model that does not depend on mobility data. This establishes a basis for identifying patients at risk automatically and objectively assessing the physical capacity of currently immobilized patients. Such advancements could significantly contribute to enhancing patient care and treatment planning in orthogeriatric settings.

Keywords: frailty; clinical assessment; machine learning; TUG test; age; osteoporosis

check for updates

Citation: Kraus, M.; Stumpf, U.C.; Keppler, A.M.; Neuerburg, C.; Böcker, W.; Wackerhage, H.; Baumbach, S.F.; Saller, M.M. Development of a Machine Learning-Based Model to Predict Timed-Up-and-Go Test in Older Adults. *Geriatrics* **2023**, *8*, 99. https://doi.org/10.3390/geriatrics8050099

Academic Editors: Märta Sund Levander and Ewa Grodzinsky

Received: 3 August 2023
Revised: 29 September 2023
Accepted: 5 October 2023
Published: 7 October 2023

1. Introduction

A key challenge in geriatric medicine is to develop objective measures that report a patient's physical capability. Such biomarkers would help to base treatment decisions more on evidence. The Timed-Up-and-Go test (TUG test) [1] is a commonly used tool to assess the physical performance of orthogeriatric patients over 60 years of age. It is important for the long-term care of patients to have reproducible examinations available for an evaluation of the therapy success. Conducting physical tests such as the TUG test to objectively measure the physical segment of frailty is crucial, as recent research has revealed that simple clinical evaluation correlates poorly with objective geriatric assessment. [2,3]. The majority of individuals with numerous geriatric deficits subjectively underestimated their actual frailty in comparison to an objective assessment [2,3]. Consequently, there is a pressing need to more objectively assess the physical capacities of geriatric patients [4]. The current tools used to assess physical capacity are mainly based on standardized patient-reported questionnaires, such as the Barthel Index [5], DeMorton Mobility Index [6], or screening questionnaires, such as the sarcopenia and physical frailty screening questionnaire (SARC-F) [7]. The main disadvantages of these questionnaires are that they are very time-consuming to administer and are influenced by the subjective self-assessment and/or assessment by the caretakers. In addition, older patients tend to overestimate their physical activity [2] and patients treated in trauma surgery are often immobilized, which limits their capacity to undertake physical testing. ML-based fall detection and prevention systems are evaluated in a review by Usmani et al., with a focus on the impact of old age on increased fall risk. The frequent use of support vector machines is an often-used algorithm, and wearables for these applications are common. However, limitations arise from primarily conducting studies in controlled environments with adults, and future research directions such as energy efficiency, sensor fusion, context awareness, and wearable design are highlighted [8].

A review on the latest research trends on fall risk prediction including over 1000 studies showed that below 5% of the studies evaluated the quality of fall risk prediction models. These models used patient assessment data related to physical and cognitive function, but often did not consider post-admission factors or interventions, as well as cross-sectional blood-work data. The reporting quality was generally poor, but it has improved over the past decade. The review recommends exploring artificial intelligence and machine learning with high-dimensional data from digital hospital systems to enhance fall risk prediction in hospitals [9].

In the future, telemedicine systems will play an important role in this automation to close the gap between inpatient monitoring and outpatient care. This may help to address the unique needs of patients and their environmental contexts [10]. A user-friendly portable digital system for sarcopenia assessment, following the EWGSOP2 algorithm, has already been established by Teixeira et al. in 2022. This system not only facilitates the diagnosis and monitoring of sarcopenia but also holds potential for increasing public awareness about sarcopenia's characteristics and risk factors [11].

For future applications, the principles and interventions set out by Petretto et al. address the potential digital paradox, where individuals who could benefit the most from telemedicine may be inadvertently excluded, particularly individuals with disabilities and the elderly. These principles encompass structural considerations, knowledge and skill requirements, and necessary adaptations, with a focus on accommodating diverse user needs. The needs and specificities of all stakeholders, including healthcare professionals and caregivers, are regarded as integral to the discussion [12].

Because of these limitations, we aimed to develop a more objective test to obtain a measure of physical performance in elderly patients by generating multifactorial data without mobility data. Second, we validated that test by comparing it to TUG test data by utilizing supervised machine learning methods [13]. It is challenging to evaluate the pertinent influencing factors holistically with reference to the individual risk, especially when evaluating multi-factorial disorders, such as physical frailty. Here, machine learning algorithms have a lot of potential to assist human judgment and enhance patient care. The long-term objective is to utilize the pilot study's findings to create clinical decision support tools that can be linked into hospital information systems to automatically identify patients at risk.

2. Materials and Methods

2.1. Patient Recruitment

We recruited patients attending our orthogeriatric outpatient clinic, with a primary focus on osteoporosis treatment, during the period spanning December 2020 to March 2021. Our inclusion criteria encompassed individuals aged over 60 years who demonstrated independent ambulation without reliance on walking aids and exhibited no signs of mental or neurological impairments. Conversely, we excluded patients with dementia, those currently undergoing acute tumor treatment, or individuals who had sustained significant lower extremity injuries, such as fractures or joint replacements, within the preceding 6 months, to ensure the validity of the investigations, ensure the reliability of patient-reported questionnaires, and limit the impact of concurrent illnesses, as well as acute regenerative processes of the musculoskeletal system, on the laboratory values. Prior to their participation in the study, all participants provided informed consent, which encompassed the anonymized evaluation and publication of collected data. Ethical approval for the study was granted by the local ethics committee of Ludwig Maximilians University Munich (Protocol #19 177).

2.2. General Data Assessment

A single, properly trained investigator collected all data, including age, weight, height, BMI, body composition, blood draw, general health-related quality of life as measured by the European Quality of Life 5-dimension (EQ-5D-5L) questionnaire [14], and SARC-F [15]. To ensure data quality, we completed all surveys with the patients. A clinically approved body composition monitor was used to determine body composition regarding body fat and muscle percentages (BF511, Omron-Healthcare, Kyoto, Japan).

2.3. Data Collection

Data collection for each individual patient was conducted following their regular appointment at the geriatric traumatology osteoporosis outpatient clinic, typically between 9 am and 1 pm. This timing was chosen to minimize the potential impact of circadian fluctuations in the measured parameters. A single examiner conducted the data collection to ensure consistency and reduce inter-observer variability. When patients met the inclusion criteria for the study, they were provided with information about the potential study participation and given the autonomy to decide whether they wished to take part in the examinations. During data acquisition, our foremost objective was to gather a comprehensive set of parameters pertinent to physical frailty. These parameters were obtained within the confines of routine clinical practices. To uphold methodological precision, we referred to established guidelines and the pertinent literature recommendations. In particular, laboratory values from a standardized osteological screening laboratory, according to the current DVO guideline, were included as an essential component of data collection. [16] It was expanded to include the muscle markers myoglobin, LDH, and muscle-specific creatine kinase. In addition, demographic data, such as age, weight, height, BMI, were collected, and a BIA (bioelectrical impedance analysis) was used to measure body fat and muscle percentage. EQ-5D-5L was surveyed as an index for health-related quality of life. SARC-F [15] was completed with assistance given to the patients to ensure the greatest possible data quality. Patients were asked if they can lift 5 kg, walk across the room, struggle to get out of a chair, climb 10 flights of stairs, and how many times they have fallen in the previous year. Together with handgrip strength, measured using a digital dynamometer (EH101, Kuptone, London, UK) and Timed-Up-and-Go time measurements, 65 variables were collected for each patient. In shaping the parameter selection, we conducted a thorough evaluation by comparing guidelines and the current literature within the context of an expert panel, while also taking into careful consideration the available resources for data collection.

2.4. Timed-Up-and-Go Test

Subjects were instructed to walk from a seated position on a regular chair to a marker 3 m away, turn around, and return to the starting position in the TUG test. For all subjects, the same iPhone application (Apple Inc., Cupertino, CA) was used to record timings.

2.5. Clinical Laboratory Data

To minimize biochemical alterations of the blood, the samples were evaluated immediately after blood collection in the hospital's central laboratory. An extended osteological basic laboratory [16], broadened to include muscle markers, was obtained, including sodium, potassium, glucose, creatinine clearance, creatinine, serum calcium, protein-corrected serum calcium phosphate in serum, total protein, c-reactive protein electrophoresis, albumins, beta globulins, gamma globulins, alpha-1 globulins, alpha-2 globulins alkaline phosphatase, gamma-glutamyl transferase, count of red blood cells, erythrocytes, leukocytes hematocrit, hemoglobin, average corpuscular volume mean corpuscular hemoglobin concentration, mean corpuscular hemoglobin, platelets hormone parathyroid, thyroid stimulating hormone, 25-hydroxyvitamin D3, lactate dehydrogenase, creatine-kinase, glomerular filtration rate (GFR), and myoglobin. A detailed list can be found in the supplementary data (Table S1) and on the projects GitHub repository [17].

2.6. Machine Learning Model Construction

The data analysis and modeling was carried out after data collection was completed using the open source programming language R (version 4.2.0), utilizing library mlr3 [18] and its dependent packages. To perform a dimensionality reduction for the machine learning algorithms, we used six different feature selection methods of the praznik package [19], each applying a threshold of 0.8 on the mutual information score (mi-score) [20] to select the most relevant variables. When the ground truth is unknown, the mi-score may be used to assess the agreement of two independent label assignment strategies on the same dataset. Comparing feature selection methods helps to make informed decisions about which method to use for specific data and objectives, considering mathematical underpinnings and trade-offs between information gain and redundancy reduction [21].

Therefore, the following six methods were selected based on their suitability for the present dataset: impurity (imp), which evaluates variables based on Gini impurity, which is used to split data in decision trees [22]; A minimum redundancy maximal relevancy filter (mrmr), which aims to minimize redundancy among selected features while maximizing their relevance to the target variable [23]; A minimal conditional mutual information maximization filter (cmim), which seeks to maximize conditional mutual information, focusing on the dependence of a feature on the target variable given the other selected features [24]; a minimal joint mutual information maximization filter (jmim), which focuses on maximizing joint mutual information, considering the mutual information of a feature with all other selected features [25]; a minimal normalized joint mutual information maximization filter (njmim), which is similar to the jmim and njmim and also maximizes joint mutual information but with the additional step of normalizing the mutual information values [26]; and a joint mutual information filter (jmi), which maximizes joint mutual information but without normalization [22]. As described, these methods differ from a mathematical point of view in how they evaluate the variables in terms of entropy, either minimizing redundancy or maximizing information gain, and whether they normalize the input data or directly use the data structure of the raw data.

The process of variable selection was followed in our analyses by training four different algorithms: the random forest algorithm [27], one generalized linear model [28], a support vector machine (SVM) [29], and an XG-Boost-algorithm [30]. During the training process, we performed resampling by five-fold internal cross-validation to increase the reliability of our models. The data were split into training and validation data in a ratio of 80/20.

Subsequently, we evaluated and compared the models with respect to their training and testing error. For this purpose, the error measures mean squared error (MSE), root mean squared error (RM0SE), and mean absolute error (MAE) were used, as a combination of metrics is often required to best assess the performance of a model [31].

Based on these results, boxplots, correlation, residual plots, and Taylor diagrams [32] were created to visualize the results.

2.7. Statistical Analysis

To enhance comprehensibility of the dataset, the analysis was initiated with a comprehensive descriptive statistical examination. This initial phase involved the computation of mean values and standard deviations for all numerical variables. Concurrently, categorical and binary variables were presented in terms of their respective percentage frequencies. In addition to our ML approach, a multivariate ANOVA analysis was performed to discover the optimal combination of variable extraction and algorithm selection by determining statistical differences in the training and testing errors between the utilized learners and feature selection approaches. To maximize traceability, the complete code used can be viewed in the project's GitHub repository [17].

3. Results

Of the 115 eligible patients in our outpatient clinic, 103 agreed to participate in this study. In five instances, participants declined to take part in the assessments, citing that their subjective physical capacity was insufficient to complete all the tests. Additionally, seven patients declined participation due to scheduling commitments. Table 1 shows the general demographic data of these patients. See the Supplemental Materials Table S1 for a comprehensive exploratory data analysis.

Table 1. Demographic patient data (n = 103, IQR = inter-quartal range).

Variable	N	Median	IQR
Age	103	76	(71, 80)
Handgrip strength		22.4	(18.8, 25.2)
TUG test time		9.5	(8.0, 13.8)
Weight		64	(58, 70)
Height		162	(158, 166)
BMI		24.4	(21.7, 25.9)

3.1. Feature Selection Process

The cut-off of 0.8 for the mutual information score resulted in 10 selected features for each of the six methods.

The evaluation of the feature frequency in our six different feature selection methods showed that age and leukocytes were the two most frequently selected variables for the regression analysis. They were selected by all six methods. By five methods, the variables EQ-5D index, GFR, grip strength of the dominant hand, and patient-reported health state were selected. The frequencies of all the selected features are shown in Figure 1 by proportion.

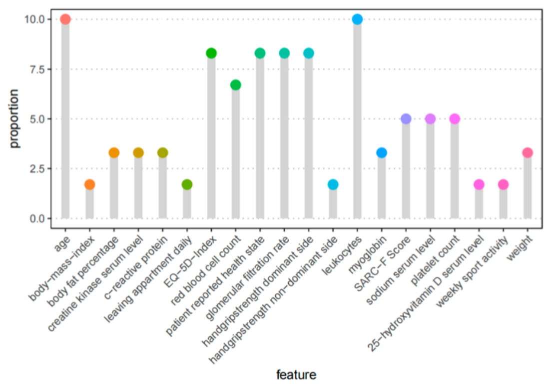


Figure 1. Feature frequency of the top 20 automated selected features. The process included the 6 feature selection methods described before. Each approach picked 10 features, for a total of 60, yielding a proportion of 10% if a variable is selected by all six methods and 1.6 percent if it is chosen by one method.

3.2. Validation of the Model

To obtain an initial overview of the performance of the different models, we created a Taylor diagram [32] (Figure 2) in which the used algorithms are color-coded in A and the feature selection methods in B. This graphic, published first by Taylor in 2001 [32], aids in the comparison of several models. It measures the degree of agreement between modeled and observed behavior using three statistics: the Pearson correlation coefficient, the root mean square error (RMSE), and the standard deviation.

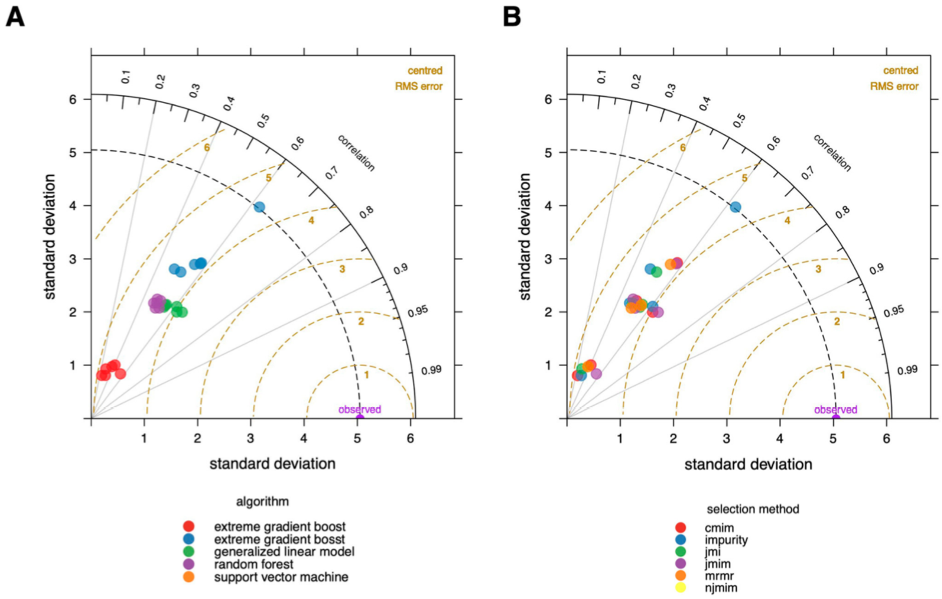


Figure 2. Taylor diagram of modelling results. The degree of agreement between modeled and observed behavior are visualized using three statistics: the Pearson correlation coefficient, the root-mean-square error (RMSE), and the standard deviation. (**A**) Colors correspond to the used algorithms; (**B**) colors correspond to the feature selecting methods. It is evident that the random forest algorithm is the best fit, and algorithm selection has a higher impact on the ultimate performance of the model than feature selection approaches.

The Taylor diagram provides a clear summary of how the models differ in terms of performance, as assessed by the root mean square error (RMSE). The choice of algorithm clearly has a considerably bigger influence on the overall performance when compared to the feature selection method. The random forest method outperforms the other algorithms, and xgboost seems to perform the worst on our data.

To dissect the differences between the models used in more detail, we have created a summary table with three different error measures that differ, particularly in terms of their penalization of the outliers. The three testing error measures MSE, RMSE and MAE are listed in Table S2, broken down by the feature selection methods and algorithms used.

When comparing the models created using the root mean squared error (RMSE), the combination of the cmim feature selection and the random forest algorithm performed best with 3.7 s, whereas the RMSE of the xgboost is more than twice as large with 8.9 s. The MAE, representing the average of all the absolute errors, was lowest for the combination of random forest algorithm and a mrmr at 2.7 s and highest for the combination of xgboost and an njmim at 7.9 s.

The MSE is visualized in Figure 3, where we show the MSE split into the training data and the test data. The MSE is significantly higher for the test data than for the training data across all the algorithms and feature selection methods, except for xgboost, where the training and testing errors are almost identical.

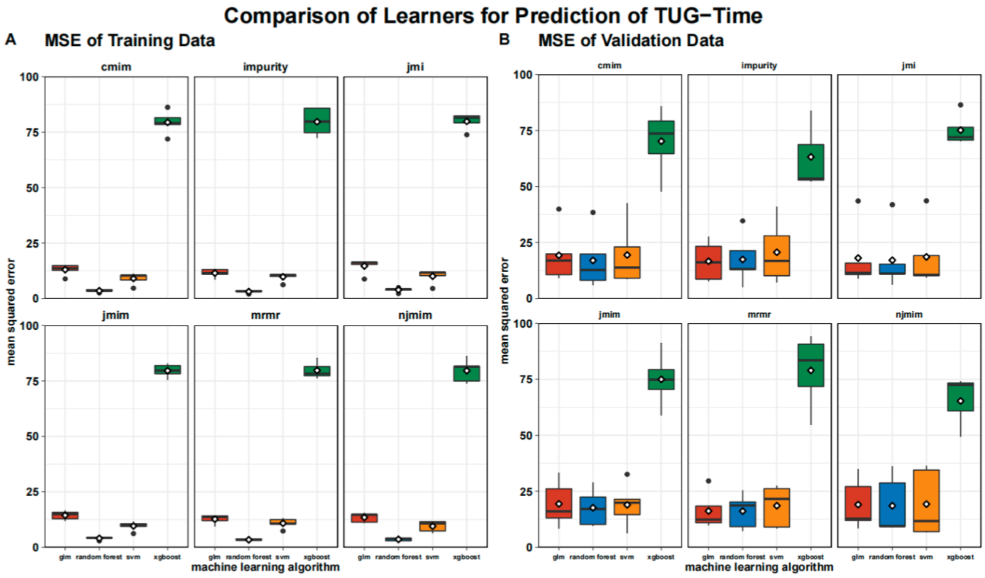


Figure 3. Boxplot of mean squared error of testing (**A**) and training data (**B**) grouped by algorithms and feature selection techniques. Means are shown as white rhombus, outliers are presented as black dots.

Figure 4 shows the individual results of the training process as a correlation plot to the actual values of the subjects after cross-validation, with only the test data in each case. Next to it is the corresponding residual plot, which revealed a significant increase in the actual time required. This pattern can be seen in all the algorithms used, whereby it is evident that these erroneous deviations are significantly greater with longer TUG test times with the generalized linear model and the support vector machine than with the random forest algorithm. Accordingly, the random forest algorithm overestimates the slow patients less when compared to all the other algorithms. The xgboost algorithm performed worst in the training and testing processes. With respect to the feature selection methods, only a few differences can be identified.

To comprehensively assess the outcomes achieved through the implementation of machine learning techniques, a test statistic was applied to the presented results. For comparison, the models were categorized based on their respective algorithm types, followed by an ANOVA-based pairwise comparison. The ANOVA analysis, with Tukey's multiple pairwise comparisons of the mean squared error in TUG time estimation, revealed a significant inferiority of xgboost compared to the other three algorithms ($p < 0.001$). No statistically significant disparities were observed among the remaining three algorithms, namely the random forest algorithm, generalized linear model, and support vector machine.

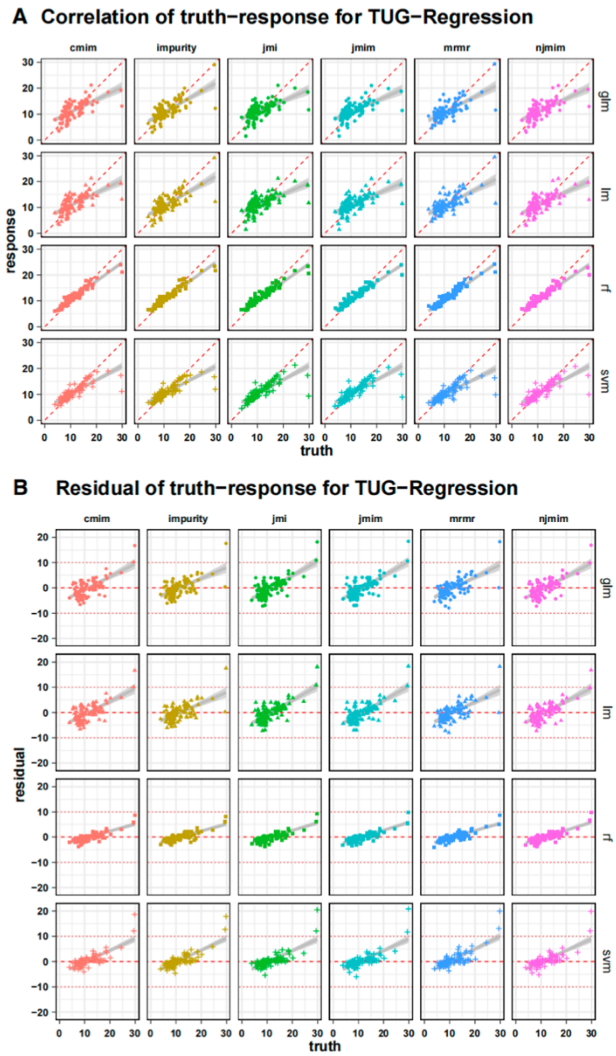


Figure 4. Individual training process outcomes as a correlation plot (**A**) to the subjects' actual values after cross-validation, only showing the testing data in each instance. The dashed red line represents a perfect correlation of 1. The matching residual plot (**B**) shows a considerable increase in predicting error when the true time increases across all models. The horizontal line shows a residual value of 0 and ± 10 as reference.

4. Discussion

Using multifactorial non-mobility data from over 100 patients, we were able to successfully develop machine learning models that predict TUG test times relatively reliably. We only used data that can be collected from bedridden patients. Our findings should help to better stratify acutely immobile patients in terms of their risk of physical frailty, allowing clinicians to make more appropriate therapeutic decisions [33]. It is crucial to bear in mind that machine learning models are founded on correlations and not causations [34]. This aspect must be considered when interpreting our results. The aim of developing these models is to provide clinical practitioners with valuable support in their assessment of frail patients, ultimately optimizing patient care.

The outcomes of our study not only advance the accuracy of TUG test time predictions but also shed light on algorithmic behavior in different patient mobility contexts. These insights are invaluable for optimizing predictive models in orthogeriatric care and have broader implications for enhancing clinical decision support systems across various healthcare domains. The achievement of a mean absolute error as low as 2.7 s underscores the potential of machine learning in refining the accuracy of TUG test time estimations. Increasing the level of accuracy is pivotal, as even small discrepancies in TUG test time predictions can have substantial clinical repercussions, affecting patient care plans and interventions [35].

Our findings reveal an important nuance in the behavior of the algorithms—the tendency to overestimate the TUG test time for quicker patients and underestimate it for slower patients. Addressing this issue in further studies is paramount to ensuring the predictive models' clinical utility across a wide spectrum of patients with varying mobility levels.

The broader applicability of our findings extends beyond the specific context of orthogeriatric patients. The machine learning methodologies employed in this study can serve as a foundation for predictive modeling in various clinical scenarios where mobility or frailty assessment plays a pivotal role. These scenarios encompass not only fall risk estimation but also patient rehabilitation planning, resource allocation, and personalized care strategies.

The field of feature selection plays a critical role in data analysis and machine learning, aiding in the identification of relevant variables for predictive modeling [36]. Common approaches include variable filtering, which ranks variables based on their relevance to the target using predefined criteria. Other methods, such as wrapper and embedded techniques, optimize feature sets based on the performance of subsequent learning algorithms [37]. Filtering is often favored for its computational efficiency, reduced risk of overfitting, and generic applicability across various inference models. Information-theoretic measures, such as mutual information (MI) and conditional mutual information (CMI), are popular criteria for variable selection due to their model-independent nature and capacity to capture variable dependencies of arbitrary order [38]. The existing definitions of feature relevancy and redundancy fail to rigorously address interactions among variables, impeding practical feature selection methods. Discrimination power analysis is a different method for feature selection, firmly rooted in the principles of inter-class and intra-class variation, and excels in discerning the discriminatory capacity of individual features within a dataset [39]. It is particularly adept at identifying features characterized by low correlation and high discrimination, making it invaluable when dealing with complex databases comprising multiple classes and abundant training samples. DPA's ability to balance inter-class differences and intra-class consistency ensures the selection of features that contribute significantly to predictive accuracy while reducing redundancy, which is especially important in feature extraction from multi-faceted data, such as images or shapes [40]. Since our work focuses on how multiple variables contribute information to the target of TUG test time and comparing different machine learning algorithms, our study utilized information-based methods to identify feature relevancy and redundancy in information-theoretic terms [41].

The final performance of the models was only slightly affected by the usage of various feature selection techniques. We attempted to identify the most helpful variables for machine learning using a variety of feature selection approaches, all of which were information-based, due to the high dimensionality of the dataset created by our investigations. Since there were no appreciable performance differences between the strategies and the evaluated feature selection procedures, all approaches may be approximately compared for our purposes.

Age and inflammatory parameters seem to be crucial factors for the estimation of the TUG test. To generate valuable information from the results of the feature selection methods, we tried to evaluate the frequency with which the individual variables were selected. The two variables chosen from all the selection methods, age and leukocyte count, appear to be key influencing factors for physical frailty syndrome [42]. Reviews over the past few years have shown that a high leukocyte count is a sign of systemic inflammation, illness progression, and a poor prognosis [43], and all-cause mortality can be predicted by systemic inflammation [44].

Aging is a process that happens at wildly varying speeds in various people. It appears to be a highly significant and trustworthy indicator when it comes to physical performance. This may be due to the fact that peak muscle and bone mass deterioration begins in the 20s and 30s [45]. As a result, the age attained provides critical information on how much of the musculoskeletal structures remain. A limiting note here is that muscle mass alone is not a determinant of preserved function, and degradation is subject to interindividual variation. If chronological age was extended to include biological age, the accuracy of the results achieved would most likely increase, since it is well known that biological methods of determining age are even more consistent with functional resilience than chronological age [46].

In addition, existing analyses on composite biomarker predictors for biological age also found that, for example, CRP and hemoglobin serum levels are meaningful predictors of biological age, which were also deemed relevant in our analyses [47].

The two described variables were followed in terms of importance by self-assessed health status, GFR, EQ-5D index, and handgrip strength of the dominant hand. These variables are already used in existing scores such as the Fried Frailty Scale or functional age estimators, among others. The fact that we were able to reproduce these results underlines the reliability of the factors found.

The most commonly used tool, the frailty index [48] offers the advantage that only external, physical appearance has to be assessed, and no aperitive diagnostics are necessary. Its only drawback is that the decision is made solely based on a personal assessment of external factors. Because we intended to generate the highest level of objectivity and reliability, we opted against using the Frailty Index in our investigations. Recent research in constrained patient groups has demonstrated that the TUG test and handgrip strength are also excellent tools for estimating mortality risk. This implies that the TUG test could function as a reliable gauge of biological age. Additional functional and molecular level research is required to test this theory [49].

The random forest algorithm yields the best results in the estimation of the TUG test in the utilized dataset. The algorithm we used for variable selection appears to play only a minor role in the final performance. While all algorithms, except the xgboost, start to overestimate the TUG test time of relatively fast patients and underestimate the TUG test time of slow patients, which should be improved in the further development of the algorithms, only the xgboost dramatically underestimates the time of all subjects. Our pilot study thus showed that it is possible to create relatively reliable models for estimating the TUG test time without directly using mobility data. Statements about the most important influencing factors in the utilized models could also be made, thus fulfilling the demand for explainable AI in clinical decision support systems [50].

In the present study, only classic supervised machine learning algorithms were used due to the fact that classic AI algorithms perform similarly well to deep learning approaches with the present small number of subjects, and the explainability of the algorithms used is significantly better than with deep learning approaches [51]. This is because deep learning approaches, such as deep neural networks, obscure the decision cut-offs, which makes it much more difficult to understand the decision-making process. Considering that the random forest algorithm just had a mean absolute error of 2.7 s and the utilized variables are solely based on non-mobility data, this can be considered a good result, especially when considering that many of the patients to be evaluated need a TUG test time of more than 20 s, which means that the mean absolute error is over 10%. If an imprecision of more than 10% must be expected when estimating functional outcomes, its use as a valid diagnostic tool is limited. Since the models we have developed are mainly intended to be used for risk stratification, the deviation does not have very serious direct consequences.

It should be highlighted that the subject we address, estimating mobility using non-mobility data, is dependent on complicated linkages that are challenging to answer more precisely.

Larger differences could be found between the used algorithms when compared to the feature selection methods. The combination of the impurity filter and the tree-based random forest algorithm was the best-performing algorithm in our evaluations. The reason for this could be that the random forest algorithm achieves good results, especially with diverse data structures. It should be noted that the training error of the random forest algorithm is significantly lower, when compared to the other algorithms, which leads to the risk of overfitting [52] and thus limits the generalizability. The SVM, for example, has a higher validation error in our evaluations, but at the same time, the training error deviates less from the validation error, which suggests a better transferability of the results to a larger patient population.

For very slower TUG pace, the predictions of our model become significantly less accurate. This is since we have a few subjects with very extreme TUG test times in the training data, and, at the same time, the parameters used take on very extreme values, which makes it difficult for the algorithm to make precise predictions with the relatively limited number of individual datasets. Furthermore, it is possible that additional factors, such as current motivation or other factors that we did not collect, play a relevant role in the longer TUG test times.

5. Summary

- Multifactorial non-mobility data from over 100 patients enabled the development of reliable machine learning models for predicting TUG (Time-Up-and-Go) test times in bedridden patients.
- The choice of feature selection techniques minimally impacted the final model performance.
- Age and inflammatory parameters, particularly leukocyte count, emerged as crucial factors in TUG estimation, indicative of systemic inflammation and mortality risk.
- Biological age, incorporating factors such as CRP and hemoglobin levels, correlated with the TUG outcomes.
- Variables such as self-assessed health, GFR, EQ-5D index, and handgrip strength were identified as influential, aligning with existing frailty assessment tools.
- The random forest algorithm outperformed the other ML algorithms in TUG estimation
- The study achieved a mean absolute error of 2.7 s in TUG estimation, though limitations existed for TUG test times over 20 s, potentially due to limited extreme data and uncollected factors such as motivation.
- Estimating mobility from non-mobility data involves complex relationships, posing challenges.
- The impurity filter combined with the random forest algorithm showed the best performance, although overfitting risk and lower validation errors were noted.

6. Limitations

The number of subjects included in the analysis is relatively low for a machine learning approach. However, it is only an exploratory pilot study investigating the special patient population of orthogeriatric patients. Another limitation within the confines of our preliminary investigation pertains to its monocentric study design. This particular design imposes constraints on the extrapolation of findings, particularly in the context of applying machine learning algorithms, due to the inability to aggregate structural attributes specific to the study center across multiple centers. Therefore, the results should only serve as a basis for further studies. The measures proposed here, which appear to be relevant for assessing physical frailty, should be evaluated in larger-scale, ideally multicenter research.

Since our study was designed in a single-stage, single-center setting, during the model creation, an internal five-fold cross-validation was conducted to create more generalizability. We recognize the importance of prospective validation to corroborate the robustness of our findings. Future research initiatives should focus on validating our predictive models in independent cohorts of orthogeriatric patients to assess their generalizability and clinical applicability. We made the models open-source to enable validation across patient populations.

The very-slow-walking patients were especially difficult to estimate correctly. According to our findings, the slower the patients get, the more difficult the correct prediction becomes. In the future, investigations of only these slower patients will be necessary to better understand the underlying relationships and thus be able to make better assessments.

Machine learning studies are always based on correlation analyses, which take a closer look at the data structure. Therefore, the results must not be considered causal, but only represent a possibility to understand the correlations and patterns in the data and to be able to draw clinically relevant correlations from them, which are not necessarily subject to direct causalities.

No sample size calculation was performed for our study as it was conducted as a pilot investigation. The predetermined target sample size of 100 individuals was selected primarily to facilitate a fundamental correlation analysis.

7. Conclusions

Our results demonstrate that non-mobility data can be used effectively to forecast the time required for the TUG test in orthogeriatric patients using machine learning models, although the more time patients needed, the less accurate the predictions became.

This is a building block to automate the detection of patients at risk and to create the possibility of also objectively assessing immobilized patients regarding their physical capacity. Statements regarding the most influential aspects of the employed models could also be made, thus meeting the requirement for explainable AI in medicine and at the same time gaining new insights into physical frailty and related factors. Future research is required to confirm our findings and adopt clinical decision support systems based on the developed algorithms.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/geriatrics8050099/s1, Table S1: Explanatory Data Analysis of complete variable set presented in mean and inter-quartile range (IQR) in parentheses, Table S2: Results of model evaluation, ordered in descending order.

Author Contributions: Conceptualization, M.K., M.M.S., A.M.K., U.C.S., S.F.B. and H.W.; methodology, M.K., W.B., M.M.S., S.F.B. and A.M.K.; software, M.K. and M.M.S.; validation, H.W., C.N., M.M.S., S.F.B. and A.M.K.; formal analysis, M.K. and M.M.S.; investigation, M.K. and U.C.S.; resources, M.K. and M.M.S.; data curation: M.K. and M.M.S.; writing—original draft preparation, M.K.; writing—review and editing, M.M.S., S.F.B. and H.W.; visualization, M.K. and M.M.S.; supervision, W.B. and C.N.; project administration, M.M.S.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The Institutional Review Board of the LMU Munich Medical Faculty gave its approval to the study after it was carried out in accordance with the Declaration of Helsinki's principles. The university's ethics committee gave the study their seal of approval, and it was filed as AZ 19-177.

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The corresponding author can provide the data described in this study upon request. Due to laws governing data protection and privacy, the data are not accessible to the general public.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. CDC. STEADI Assessment Timed Up & Go (TUG). Available online: https://www.cdc.gov/steadi/pdf/TUG_Test-print.pdf (accessed on 28 July 2023).
2. Martin, P.; Keppler, A.M.; Alberton, P.; Neuerburg, C.; Drey, M.; Böcker, W.; Kammerlander, C.; Saller, M.M. Self-Assessment of Mobility of People over 65 Years of Age. *Medicina* **2021**, *57*, 980. [CrossRef] [PubMed]
3. van Walree, I.C.; Scheepers, E.R.M.; van den Bos, F.; van Huis-Tanja, L.H.; Emmelot-Vonk, M.H.; Hamaker, M.E. Clinical Judgment versus Geriatric Assessment for Frailty in Older Patients with Cancer. *J. Geriatr. Oncol.* **2020**, *11*, 1138–1144. [CrossRef] [PubMed]
4. Beaudart, C.; Rolland, Y.; Cruz-Jentoft, A.J.; Bauer, J.M.; Sieber, C.; Cooper, C.; Al-Daghri, N.; Araujo de Carvalho, I.; Bautmans, I.; Bernabei, R.; et al. Assessment of Muscle Function and Physical Performance in Daily Clinical Practice: A Position Paper Endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). *Calcif. Tissue Int.* **2019**, *105*, 1–14. [CrossRef] [PubMed]
5. Mahoney, F.I.; Barthel, D.W. Functional Evaluation: The Barthel Index. *MD State Med. J.* **1965**, *14*, 61–65. [PubMed]
6. de Morton, N.A.; Davidson, M.; Keating, J.L. The de Morton Mobility Index (DEMMI): An Essential Health Index for an Ageing World. *Health Qual. Life Outcomes* **2008**, *6*, 63. [CrossRef] [PubMed]
7. Hax, V.; Do Espirito Santo, R.C.; Dos Santos, L.P.; Farinon, M.; De Oliveira, M.S.; Três, G.L.; Gasparin, A.A.; De Andrade, N.P.B.; Bredemeier, M.; Xavier, R.M.H.; et al. Practical Screening Tools for Sarcopenia in Patients with Systemic Sclerosis. *PLoS ONE* **2021**, *16*, e0245683. [CrossRef] [PubMed]
8. Usmani, S.; Saboor, A.; Haris, M.; Khan, M.A.; Park, H. Latest Research Trends in Fall Detection and Prevention Using Machine Learning: A Systematic Review. *Sensors* **2021**, *21*, 5134. [CrossRef]
9. Parsons, R.; Blythe, R.D.; Cramb, S.M.; McPhail, S.M. Inpatient Fall Prediction Models: A Scoping Review. *Gerontology* **2023**, *69*, 14–29. [CrossRef]
10. Lippi, L.; D'Abrosca, F.; Folli, A.; Dal Molin, A.; Moalli, S.; Maconi, A.; Ammendolia, A.; de Sire, A.; Invernizzi, M. Closing the Gap between Inpatient and Outpatient Settings: Integrating Pulmonary Rehabilitation and Technological Advances in the Comprehensive Management of Frail Patients. *Int. J. Environ. Res. Public Health* **2022**, *19*, 9150. [CrossRef]
11. Teixeira, E.; Bohn, L.; Guimarães, J.P.; Marques-Aleixo, I. Portable Digital Monitoring System for Sarcopenia Screening and Diagnosis. *Geriatrics* **2022**, *7*, 121. [CrossRef]
12. Petretto, D.R.; Gaviano, L.; Carrogu, G.P.; Berti, R.; Pinna, M.; Pili, R. Telemedicine: Issues in the Analysis of Its Use in Elderly People and in People with Disabilities, According to the Perspective of the Clinical Psychology of Disability. *Geriatrics* **2022**, *8*, 5. [CrossRef]
13. Chen, L.P.; Mohri, M.; Rostamizadeh, A.; Talwalkar, A. Foundations of Machine Learning, Second Edition. *Stat. Pap.* **2019**, *60*, 1793–1795. [CrossRef]
14. Ludwig, K.; von der Schulenburg, J.-M.G.; Greiner, W. German Value Set for the EQ-5D-5L. *Pharmacoeconomics* **2018**, *36*, 663–674. [CrossRef]
15. Bahat, G.; Ozkok, S.; Kilic, C.; Karan, M.A. SARC-F Questionnaire Detects Frailty in Older Adults. *J. Nutr. Health Aging* **2021**, *25*, 448–453. [CrossRef]
16. Dachverband Osteologie e.V. Prophylaxe, Diagnostik Und Therapie Der OSTEOPOROSE. Available online: https://www.dv-osteologie.org/uploads/Leitlinie%202017/Finale%20Version%20Leitlinie%20Osteoporose%202017_end.pdf (accessed on 28 July 2023).
17. Kraus, M. TUG Regression Pipeline. Available online: https://github.com/KrausMoritz/KI_Workstation/blob/main/TUG_Regression.Rmd (accessed on 25 May 2022).
18. Lang, M.; Binder, M.; Richter, J.; Schratz, P.; Pfisterer, F.; Coors, S.; Au, Q.; Casalicchio, G.; Kotthoff, L.; Bischl, B. Mlr3: A Modern Object-Oriented Machine Learning Framework in R. *J. Open Source Softw.* **2019**, *4*, 1903. [CrossRef]
19. Kursa, M.B. Praznik: High Performance Information-Based Feature Selection. *SoftwareX* **2021**, *16*, 100819. [CrossRef]
20. Li, K.; Yu, M.; Liu, L.; Li, T.; Zhai, J. Feature Selection Method Based on Weighted Mutual Information for Imbalanced Data. *Int. J. Softw. Eng. Knowl. Eng.* **2018**, *28*, 1177–1194. [CrossRef]
21. Chormunge, S.; Jena, S. Correlation Based Feature Selection with Clustering for High Dimensional Data. *J. Electr. Syst. Inf. Technol.* **2018**, *5*, 542–549. [CrossRef]
22. Yang, H.; Moody, J. Data Visualization and Feature Selection: New Algorithms for Nongaussian Data. *Adv. Neural Inf. Process. Syst.* **1999**, *12*, 687–693.
23. Peng, H.; Long, F.; Ding, C. Feature Selection Based on Mutual Information Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef]
24. Fleuret, F. Fast Binary Feature Selection with Conditional Mutual Information. *J. Mach. Learn. Res.* **2004**, *5*, 1531–1555.
25. Sechidis, K.; Azzimonti, L.; Pocock, A.; Corani, G.; Weatherall, J.; Brown, G. Efficient Feature Selection Using Shrinkage Estimators. *Mach. Learn.* **2019**, *108*, 1261–1286. [CrossRef]

26. Bennasar, M.; Hicks, Y.; Setchi, R. Feature Selection Using Joint Mutual Information Maximisation. *Expert Syst. Appl.* **2015**, *42*, 8520–8532. [CrossRef]
27. Wright, M.N.; Ziegler, A. Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [CrossRef]
28. Wang, M.; Tian, G.L. Adaptive Group Lasso for High-Dimensional Generalized Linear Models. *Stat. Pap.* **2019**, *60*, 1469–1486. [CrossRef]
29. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
30. Chen, T.; Guestrin, C. XGBoost. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: New York, NY, USA, 2016; pp. 785–794.
31. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
32. Taylor, K.E. Summarizing Multiple Aspects of Model Performance in a Single Diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]
33. Leng, S.; Chen, X.; Mao, G. Frailty Syndrome: An Overview. *Clin. Interv. Aging* **2014**, *2014*, 433–441. [CrossRef]
34. Obermeyer, Z.; Emanuel, E.J. Predicting the Future—Big Data, Machine Learning, and Clinical Medicine. *N. Engl. J. Med.* **2016**, *375*, 1216–1219. [CrossRef]
35. Larsson, B.A.M.; Johansson, L.; Johansson, H.; Axelsson, K.F.; Harvey, N.; Vandenput, L.; Magnusson, P.; McCloskey, E.; Liu, E.; Kanis, J.A.; et al. The Timed up and Go Test Predicts Fracture Risk in Older Women Independently of Clinical Risk Factors and Bone Mineral Density. *Osteoporos. Int.* **2021**, *32*, 75–84. [CrossRef] [PubMed]
36. Battiti, R. Using Mutual Information for Selecting Features in Supervised Neural Net Learning. *IEEE Trans. Neural Networks* **1994**, *5*, 537–550. [CrossRef] [PubMed]
37. Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A.W.; O'Sullivan, J.M. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Front. Bioinform.* **2022**, *2*, 927312. [CrossRef] [PubMed]
38. Brown, G.; Pocock, A.; Zhao, M.J.; Luján, M. Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection. *JMLR* **2012**, *13*, 27–66.
39. Leng, L.; Zhang, J. PalmHash Code vs. PalmPhasor Code. *Neurocomputing* **2013**, *108*, 1–12. [CrossRef]
40. Leng, L.; Li, M.; Kim, C.; Bi, X. Dual-Source Discrimination Power Analysis for Multi-Instance Contactless Palmprint Recognition. *Multimed. Tools Appl.* **2017**, *76*, 333–354. [CrossRef]
41. Wollstadt, P.; Schmitt, S.; Wibral, M. A Rigorous Information-Theoretic Definition of Redundancy and Relevancy in Feature Selection Based on (Partial) Information Decomposition. *J. Mach. Learn. Res.* **2023**, *24*, 1–44.
42. Welstead, M.; Jenkins, N.D.; Russ, T.C.; Luciano, M.; Muniz-Terrera, G. A Systematic Review of Frailty Trajectories: Their Shape and Influencing Factors. *Gerontologist* **2021**, *61*, e463–e475. [CrossRef] [PubMed]
43. Chmielewski, P.P.; Strzelec, B. Elevated Leukocyte Count as a Harbinger of Systemic Inflammation, Disease Progression, and Poor Prognosis: A Review. *Folia Morphol.* **2018**, *77*, 171–178. [CrossRef]
44. Proctor, M.J.; McMillan, D.C.; Horgan, P.G.; Fletcher, C.D.; Talwar, D.; Morrison, D.S. Systemic Inflammation Predicts All-Cause Mortality: A Glasgow Inflammation Outcome Study. *PLoS ONE* **2015**, *10*, e0116206. [CrossRef]
45. Klibanski, A.; Adams-Campbell, L.; Bassford, T.; Blair, S.N.; Boden, S.D.; Dickersin, K.; Gifford, D.R.; Glasse, L.; Goldring, S.R.; Hruska, K.; et al. Osteoporosis Prevention, Diagnosis, and Therapy. *JAMA J. Am. Med. Assoc.* **2001**, *285*, 785–795. [CrossRef]
46. Baker, G.T.; Sprott, R.L. Biomarkers of Aging. *Exp. Gerontol.* **1988**, *23*, 223–239. [CrossRef] [PubMed]
47. Lohman, T.; Bains, G.; Berk, L.; Lohman, E. Predictors of Biological Age: The Implications for Wellness and Aging Research. *Gerontol. Geriatr. Med.* **2021**, *7*, 1–13. [CrossRef] [PubMed]
48. Tracy, B.M.; Wilson, J.M.; Smith, R.N.; Schenker, M.L.; Gelbard, R.B. The 5-Item Modified Frailty Index Predicts Adverse Outcomes in Trauma. *J. Surg. Res.* **2020**, *253*, 167–172. [CrossRef]
49. Chua, K.Y.; Lim, W.S.; Lin, X.; Yuan, J.-M.; Koh, W.-P. Handgrip Strength and Timed Up-and-Go (TUG) Test Are Predictors of Short-Term Mortality among Elderly in a Population-Based Cohort in Singapore. *J. Nutr. Health Aging* **2020**, *24*, 371–378. [CrossRef]
50. Amann, J.; Vetter, D.; Blomberg, S.N.; Christensen, H.C.; Coffee, M.; Gerke, S.; Gilbert, T.K.; Hagendorff, T.; Holm, S.; Livne, M.; et al. To Explain or Not to Explain?—Artificial Intelligence Explainability in Clinical Decision Support Systems. *PLOS Digit. Health* **2022**, *1*, e0000016. [CrossRef]
51. Markus, A.F.; Kors, J.A.; Rijnbeek, P.R. The Role of Explainability in Creating Trustworthy Artificial Intelligence for Health Care: A Comprehensive Survey of the Terminology, Design Choices, and Evaluation Strategies. *J. Biomed. Inform.* **2021**, *113*, 103655. [CrossRef]
52. Belkin, M.; Hsu, D.; Mitra, P.P. Overfitting or Perfect Fitting? Risk Bounds for Classification and Regression Rules That Interpolate. *arXiv* **2018**, arXiv:1806.05161.

5. Literaturverzeichnis

Es empfiehlt sich für das Erstellen des Literaturverzeichnis Programme, wie beispielsweise Endnote, Citavi oder **Mendeley** zu verwenden.

Formatierung des Literaturverzeichnisses:

- Zitieren gem. Harvard-Konvention bzw. APA (Autor-Jahr-System): Angabe von Autor und Erscheinungsjahr im Fließtext (z.B. Mustermann, 2000).
- Beim Autor-Jahr-System werden in einem Fließtext neben dem Erscheinungsjahr der Erst- und Zweitautor (Mustermann, 2000; Mustermann und Musterfrau, 2000) genannt, bei mehreren Autoren der Erstautor gefolgt von „et al." (Mustermann et al., 2000), dies wird als Kurzbeleg bezeichnet. Werden an einer Stelle mehrere Werke gleichzeitig zitiert, sind diese mit Kommata getrennt anzugeben (Mustermann, 2000, Musterfrau 2001).
  Im Literaturverzeichnis werden die Artikel mit vollständiger Autorenliste, Erscheinungsjahr, Titel, Journal, Band, Seitenzahl angegeben. Die Artikel werden alphabetisch nach dem Erstautor gelistet.
- Zitieren gem. Vancouver-Konvention bzw. IEEE (Nummernsystem): fortlaufende eingeklammerte Ziffern im Fließtext und analog dazu auch im Literaturverzeichnis die numerische Auflistung. Wird ein Literaturbeleg im Fließtext mehrfach zitiert, so ist immer dieselbe Nummer dafür zu verwenden (z.B. [17] oder alternativ hochgestellt [17]).
  Beim Nummernsystem kann auf mehrere Quellen gleichzeitig verwiesen werden, indem diese dann innerhalb der Klammer per Kommata getrennt werden ([17, 19, 24-26] bzw. [17, 19, 24-26]).
  Im Literaturverzeichnis werden die Artikel mit Nummer, vollständiger Autorenliste, Erscheinungsjahr, Titel, Journal, Band, Seitenzahl angegeben. Die Artikel werden numerisch gelistet.

Bitte verwenden Sie nur eine der möglichen Zitierweisen und seien Sie konsistent.

[1] World Health Organization, "WHO Global Report on Falls Prevention in Older Age.," *Community Health (Bristol).*, p. 53, 2007.

[2] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, "Medical Costs of Fatal and Nonfatal Falls in Older Adults," *J. Am. Geriatr. Soc.*, vol. 66, no. 4, pp. 693–698, Apr. 2018.

[3] S. S. Paul *et al.*, "Patterns of health service use before and after a statewide fall prevention initiative for older adults at risk of falls," *Australas. J. Ageing*, vol. 41, no. 4, pp. 542–553, Dec. 2022.

[4] Weisenfluh, A. Morrison, T. Fan, and Sen, "Epidemiology of falls and osteoporotic fractures: a systematic review," *Clin. Outcomes Res.*, p. 9, Dec. 2012.

[5] S. Subramaniam, A. I. Faisal, and M. J. Deen, "Wearable Sensor Systems for Fall Risk Assessment: A Review," *Front. Digit. Heal.*, vol. 4, Jul. 2022.

[6] H. P. Hirschfeld, R. Kinsella, and G. Duque, "Osteosarcopenia: where bone, muscle, and fat collide," *Osteoporosis International*, vol. 28, no. 10. Springer London, pp. 2781–2790, 01-Oct-2017.

[7] B. Kirk, J. Zanker, and G. Duque, "Osteosarcopenia: epidemiology, diagnosis, and treatment—facts and numbers," *J. Cachexia. Sarcopenia Muscle*, vol. 11, no. 3, pp. 609–618, Jun. 2020.

[8] J. Reiss *et al.*, "Sarcopenia and osteoporosis are interrelated in geriatric inpatients," *Z. Gerontol. Geriatr.*, vol. 52, no. 7, pp. 688–693, Nov. 2019.

[9] J. Zanker and G. Duque, "Osteosarcopenia: the Path Beyond Controversy," *Current*

Osteoporosis Reports, vol. 18, no. 2. Springer, pp. 81–84, 01-Apr-2020.

[10] J. Il Yoo, H. Kim, Y. C. Ha, H. Bin Kwon, and K. H. Koo, “Osteosarcopenia in patients with hip fracture is related with high mortality,” *J. Korean Med. Sci.*, vol. 33, no. 4, 2018.

[11] C. Neuerburg *et al.*, “Investigation and management of osteoporosis in aged trauma patients: A treatment algorithm adapted to the German guidelines for osteoporosis,” *Journal of Orthopaedic Surgery and Research*, vol. 12, no. 1. BioMed Central Ltd., 08-Jun-2017.

[12] D. Schray *et al.*, “Value of a coordinated management of osteoporosis via Fracture Liaison Service for the treatment of orthogeriatric patients,” *Eur. J. Trauma Emerg. Surg.*, vol. 42, no. 5, pp. 559–564, Oct. 2016.

[13] L. S. Toh, P. S. M. Lai, S. Othman, K. T. Wong, B. Y. Low, and C. Anderson, “An analysis of inter-professional collaboration in osteoporosis screening at a primary care level using the D’Amour model,” *Res. Soc. Adm. Pharm.*, vol. 13, no. 6, pp. 1142–1150, Nov. 2017.

[14] J. A. Kanis *et al.*, “Algorithm for the management of patients at low, high and very high risk of osteoporotic fractures,” *Osteoporos. Int.*, vol. 31, no. 1, pp. 1–12, Jan. 2020.

[15] B. Kirk, S. Phu, S. L. Brennan-Olsen, E. Bani Hassan, and G. Duque, “Associations between osteoporosis, the severity of sarcopenia and fragility fractures in community-dwelling older adults,” *Eur. Geriatr. Med.*, vol. 11, no. 3, pp. 443–450, Jun. 2020.

[16] J. A. Kanis *et al.*, “A decade of FRAX: how has it changed the management of osteoporosis?,” *Aging Clinical and Experimental Research*, vol. 32, no. 2. Springer, pp. 187–196, 01-Feb-2020.

[17] A. J. Cruz-Jentoft *et al.*, “Sarcopenia: Revised European consensus on definition and diagnosis,” *Age and Ageing*, vol. 48, no. 1. Oxford University Press, pp. 16–31, 01-Jan-2019.

[18] J. A. Kanis, A. Odén, E. V. McCloskey, H. Johansson, D. A. Wahl, and C. Cooper, “A systematic review of hip fracture incidence and probability of fracture worldwide,” *Osteoporosis International*, vol. 23, no. 9. pp. 2239–2256, Sep-2012.

[19] Dachverband Osteologie e.V., “Prophylaxe, Diagnostik und Therapie der OSTEOPOROSE,” *Prophylaxe, Diagnostik und Therapie der OSTEOPOROSE*, 2017. [Online]. Available: https://www.dv-osteologie.org/uploads/Leitlinie 2017/Finale Version Leitlinie Osteoporose 2017_end.pdf.

[20] J. Compston *et al.*, “UK clinical guideline for the prevention and treatment of osteoporosis,” *Arch. Osteoporos.*, vol. 12, no. 1, p. 43, Dec. 2017.

[21] M. S. LeBoff *et al.*, “The clinician’s guide to prevention and treatment of osteoporosis,” *Osteoporos. Int.*, vol. 33, no. 10, pp. 2049–2102, Oct. 2022.

[22] C. L. Gregson *et al.*, “UK clinical guideline for the prevention and treatment of osteoporosis,” *Arch. Osteoporos.*, vol. 17, no. 1, p. 58, Dec. 2022.

[23] L. Vandenput *et al.*, “Update of the fracture risk prediction tool FRAX: a systematic review of potential cohorts and analysis plan,” *Osteoporos. Int.*, vol. 33, no. 10, pp. 2103–2136, Oct. 2022.

[24] M. Keppler et al., “Identifikation von alterstraumatologischen „patients at risk“ für ein postoperatives Mobilisationsdefizit mittels Handkraftmessung,” *Dtsch. Kongress für Orthopädie und Unfallchirurgie (DKOU 2021*, 2021.

[25] M. Kraus *et al.*, “Entwicklung eines KI-gestützten Entscheidungstools zur Bewertung der Physical Frailty bei immobilisierten orthogeriatrischen Patienten,” 2022.

[26] M. Kraus et al., “Prospektive Einschätzung des Sturzrisikos anhand Machine-Learning-basierter Ganganalyse während des TUG-Tests mit 2 Jahres Follow-Up,” 2023.

[27] M. Kraus et al., “Prospektiver Vergleich zur Einschätzung des Sturzrisikos anhand etablierter Scores in der Alterstraumatologie mit 2 Jahres „real-world“ Evaluation,” 2023.

[28] M. H. Oh, H. E. Shin, K. S. Kim, C. W. Won, and M. Kim, “Combinations of Sarcopenia Diagnostic Criteria by Asian Working Group of Sarcopenia (AWGS) 2019 Guideline and Incident Adverse Health Outcomes in Community-Dwelling Older Adults,” *J. Am. Med. Dir. Assoc.*, May 2023.

[29] S. Studenski *et al.*, “Gait speed and survival in older adults,” *JAMA - J. Am. Med. Assoc.*, vol.

305, no. 1, pp. 50–58, Jan. 2011.

[30] J. C. Ayena, L. Chioukh, M. J. D. Otis, and D. Deslandes, “Risk of falling in a timed up and go test using an UWB radar and an instrumented insole,” *Sensors (Switzerland)*, vol. 21, no. 3, pp. 1–23, Feb. 2021.

[31] T. Amundsen, M. Rossman, I. Ahmad, A. Clark, and M. Huber, “Fall risk assessment and visualization through gait analysis,” *Smart Heal.*, vol. 25, p. 100284, Sep. 2022.

[32] H. Simila, M. Immonen, J. Merilahti, and T. Petakoski-Hult, “Gait analysis and estimation of changes in fall risk factors,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 6939–6942.

[33] J. Zanker *et al.*, “Mortality, falls and slow walking speed are predicted by different muscle strength and physical performance measures in women and men,” *Arch. Gerontol. Geriatr.*, vol. 114, p. 105084, Nov. 2023.

[34] S. Jiang, B. Zhang, and D. Wei, “The Elderly Fall Risk Assessment and Prediction Based on Gait Analysis,” in *2011 IEEE 11th International Conference on Computer and Information Technology*, 2011, pp. 176–180.

[35] A. M. Keppler *et al.*, “Validity of accelerometry in step detection and gait speed measurement in orthogeriatric patients.,” *PLoS One*, vol. 14, no. 8, p. e0221732, 2019.

[36] A. M. Keppler *et al.*, “Mobility improvement in the first 6 postoperative weeks in orthogeriatric fracture patients,” *Eur. J. Trauma Emerg. Surg.*, vol. 48, no. 4, pp. 2867–2872, Aug. 2022.

[37] A. M. Keppler *et al.*, “Postoperative physical activity in orthogeriatric patients - new insights with continuous monitoring.,” *Injury*, vol. 51, no. 3, pp. 628–632, Mar. 2020.

[38] J. H. Hollman, E. M. McDade, and R. C. Petersen, “Normative spatiotemporal gait parameters in older adults,” *Gait Posture*, vol. 34, no. 1, pp. 111–118, May 2011.

[39] J. Verghese, R. Holtzer, R. B. Lipton, and C. Wang, “Quantitative Gait Markers and Incident Fall Risk in Older Adults,” *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.*, vol. 64A, no. 8, pp. 896–901, Aug. 2009.

[40] B. J. Braun *et al.*, “Validation and reliability testing of a new, fully integrated gait analysis insole,” *J. Foot Ankle Res.*, vol. 8, no. 1, p. 54, Dec. 2015.

[41] C. Giordano, M. Brennan, B. Mohamed, P. Rashidi, F. Modave, and P. Tighe, “Accessing Artificial Intelligence for Clinical Decision-Making,” *Front. Digit. Heal.*, vol. 3, Jun. 2021.

[42] S. P. Lalehzarian, A. K. Gowd, and J. N. Liu, “Machine learning in orthopaedic surgery,” *World J. Orthop.*, vol. 12, no. 9, pp. 685–699, Sep. 2021.

[43] F. Cabitza, A. Locoro, and G. Banfi, “Machine Learning in Orthopedics: A Literature Review,” *Front. Bioeng. Biotechnol.*, vol. 6, Jun. 2018.

[44] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction,” *BMC Med. Res. Methodol.*, vol. 19, no. 1, p. 64, Dec. 2019.

[45] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science (80-.).*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.

[46] I. Say, Y. E. Chen, M. Z. Sun, J. J. Li, and D. C. Lu, “Machine learning predicts improvement of functional outcomes in traumatic brain injury patients after inpatient rehabilitation,” *Front. Rehabil. Sci.*, vol. 3, Sep. 2022.

[47] M. Tschuggnall, V. Grote, M. Pirchl, B. Holzner, G. Rumpold, and M. J. Fischer, “Machine learning approaches to predict rehabilitation success based on clinical and patient-reported outcome measures,” *Informatics Med. Unlocked*, vol. 24, p. 100598, 2021.

[48] B. Vasey *et al.*, “Association of Clinician Diagnostic Performance With Machine Learning–Based Decision Support Systems,” *JAMA Netw. Open*, vol. 4, no. 3, p. e211276, Mar. 2021.

[49] E. J. Harris, I.-H. Khoo, and E. Demircan, “A Survey of Human Gait-Based Artificial Intelligence Applications,” *Front. Robot. AI*, vol. 8, Jan. 2022.

[50] P. Khera and N. Kumar, “Role of machine learning in gait analysis: a review,” *J. Med. Eng. Technol.*, vol. 44, no. 8, pp. 441–467, Nov. 2020.

[51] C.-H. Liu, Y.-H. Hu, and Y.-H. Lin, “A Machine Learning–Based Fall Risk Assessment Model for Inpatients,” *CIN Comput. Informatics, Nurs.*, vol. 39, no. 8, pp. 450–459, Aug. 2021.

[52] N. Eichler, S. Raz, A. Toledano-Shubi, D. Livne, I. Shimshoni, and H. Hel-Or, “Automatic and Efficient Fall Risk Assessment Based on Machine Learning,” *Sensors*, vol. 22, no. 4, p. 1557, Feb. 2022.

[53] K. Rockwood, “A global clinical measure of fitness and frailty in elderly people,” *Can. Med. Assoc. J.*, vol. 173, no. 5, pp. 489–495, Aug. 2005.

[54] P. Martin *et al.*, “Self-Assessment of Mobility of People over 65 Years of Age.,” *Medicina (Kaunas).*, vol. 57, no. 9, Sep. 2021.

[55] C. Beaudart *et al.*, “Assessment of Muscle Function and Physical Performance in Daily Clinical Practice: A position paper endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO),” *Calcified Tissue International*, vol. 105, no. 1. Springer New York LLC, 15-Jul-2019.

[56] E. R. Burns, R. Lee, S. E. Hodge, V. J. Pineau, B. Welch, and M. Zhu, “Validation and comparison of fall screening tools for predicting future falls among older adults,” *Arch. Gerontol. Geriatr.*, vol. 101, p. 104713, Jul. 2022.

[57] R. S. Kamper *et al.*, “Feasibility of Assessing Older Patients in the Acute Setting: Findings From the Copenhagen PROTECT Study,” *J. Am. Med. Dir. Assoc.*, vol. 24, no. 12, pp. 1898–1903, Dec. 2023.

[58] X. Wu *et al.*, “The association between major complications of immobility during hospitalization and quality of life among bedridden patients: A 3 month prospective multi-center study,” *PLoS One*, vol. 13, no. 10, p. e0205729, Oct. 2018.

[59] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electron. Mark.*, vol. 31, no. 3, pp. 685–695, Sep. 2021.

[60] A. Haleem, M. Javaid, and R. Vaishya, “Industry 4.0 and its applications in orthopaedics,” *J. Clin. Orthop. Trauma*, vol. 10, no. 3, pp. 615–616, May 2019.

[61] M. Javaid, A. Haleem, R. Pratap Singh, R. Suman, and S. Rab, “Significance of machine learning in healthcare: Features, pillars and applications,” *Int. J. Intell. Networks*, vol. 3, pp. 58–73, 2022.

Anhang B:

Danksagung

Ein Besonderer Dank gilt meinem wissenschaftlichen Betreuer und Mentor Dr. rer. nat. Maximilian Saller. Von ihm das grundlegende wissenschaftliche Handwerkszeug lernen zu dürfen, hat bei mir eine große Begeisterung für wissenschaftliches Arbeiten ausgelöst. Während der gesamten Zusammenarbeit war er immer unterstützend für mich bestens erreichbar und ich schätze sein Feedback auch über das eigentliche Projekt meiner Dissertation hinaus. Er hat mich durch seine supportive, fördernde und fordernde Art in meiner Einstellung zur Forschung sehr positiv geprägt und mich dazu gebracht Herausforderungen auf ganz neuen Gebieten anzunehmen. Dies habe ich auch als prägend für andere Lebensbereiche empfunden. Besonders dankbar bin ich ihm für seine kooperative Zusammenarbeit, mit anderen Lehrstühlen und Arbeitsgruppen, da ich denke, dass Forschung heutzutage nur in interdisziplinärer Zusammenarbeit das volle Potential ausschöpfen kann.

Ein weiterer Dank gilt meinen ärztlichen Betreuern. Allen voran Frau Dr. Ulla Stumpf. Sie hat es möglich gemacht, dass ich ihre Patienten in der Osteoporosesprechstunde über mehr als 8 Monate wöchentlich untersuchen durfte und mich auch beim 2 Jahres-Follow-up unterstützt.

Prof. Dr. Wolfgang Böcker, Prof. Dr. Henning Wackerhage und Prof. Dr. Dr. Eric Hesse danke ich dafür, dass sie Teil meines Thesis-Advisory-Committees sind. Sie haben mich bei der Projektplanung ebenso unterstützt wie bei den Zwischen- und Endevaluationen wertvollen Input gegeben. Allen dreien danke ich für eine sehr gute Betreuung, da sie sich Zeit für mich genommen haben und immer erreichbar waren.

Außerdem möchte ich mich bei allen Co-Autoren bedanken, die die Durchführung der gemeinsamen Projekte ermöglichten.

Ich danke der deutschen Gesellschaft für Orthopädie und Unfallchirurgie dafür, dass der ersten Publikation meiner Dissertation der Digitalisierungspreis der Fachgesellschaft im Jahr 2022 zuerkannt wurde. Hier bin ich meinem Co-Autor Dr. Alexander Keppler für seine Unterstützung beim Bewerbungsprozess sehr dankbar.

Ich danke der ARCUS-Klinik Pforzheim, dass ich während einem Teil meiner Zeit durch das Orthopädiestipendium der Klinik gefördert wurde.

Ich danke meinen langjährigen Schulfreunden Dennis und Sebastian, dass sie mich mit ihren informatischen Kenntnissen beim Aufsetzten der Studiendatenbank unterstützt haben.

Zuletzt danke ich meiner Familie für die Unterstützung während der Dissertation.