



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Gültigkeit und Zuverlässigkeit von Fitness Trackern und Apps im Sport. Ein systematischer Review.“

verfasst von / submitted by

Raphael Josef Konrad Berger

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 482 884

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UF Bewegung und Sport
 UF Informatik und Informatikmanagement

Betreut von / Supervisor:

Univ.-Prof. Dipl.-Ing. Dr. Arnold Baca

Mitbetreut von / Co-Supervisor:

Mag. Dr. Roland Leser, Privatdoz.

An dieser Stelle möchte ich all jenen danken, die durch ihre fachliche und persönliche Unterstützung zum Gelingen dieser Diplomarbeit beigetragen haben. Mein besonderer Dank gilt meiner Familie, insbesondere meinen Eltern, die mir mein Studium ermöglicht und mich in all meinen Entscheidungen unterstützt haben und meinem Bruder. Ebenso gilt mein Dank meinen Freunden für die Unterstützung bei dieser Arbeit und das Korrekturlesen.

Ich erkläre, dass ich die vorliegende Arbeit selbständig verfasst und nur die ausgewiesenen Hilfsmittel verwendet habe. Diese Arbeit wurde weder an einer anderen Stelle eingereicht, noch von anderen Personen vorgelegt.

Wien, 3. Juni 2019

Raphael Berger

Zusammenfassung

Hintergrund

Wearable Fitness Tracker erfreuen sich weltweit einer immer größeren Popularität. *Fitness Tracker* (auch Activity Tracker oder Wearable) sind elektronische Geräte, die am Körper getragen werden und unterschiedliche bewegungsbezogene Daten einer Person, wie die Anzahl der Schritte, den aktuellen oder über den Tag gemittelten Puls, die zurückgelegte Distanz und den täglichen Kalorienverbrauch erfassen können.

Ziel

Wie exakt sind die Messmethoden der verfügbaren Fitness Tracker im Konsumentenbereich bezogen auf die (a) Schrittzählung, die (b) Herzfrequenzmessung und die (c) Kalorienverbrauchsberechnung?

Methoden

Gesucht wurden Studien in den Datenbanken PubMed, Elsevier und Web of Science, welche die Gültigkeit von Wearables zu direkten Messkriterien oder zu Goldstandard-Geräten maßen oder die Zuverlässigkeit dieser Wearables eruierten.

Ergebnisse

33 Studien wurden in die Metaanalyse aufgenommen. (a) wurde dabei von 23, (b) von 7 und (c) von 17 behandelt. Bei (a) liegt der Zusammenhang im Laborbereich bei $r = 0,909$ (95 % LCI: 0,855; UCI: 0,944), im Feldbereich im Vergleich zur direkten Zählung der Schritte bei $r = 0,799$ (95 % LCI: 0,727; UCI: 0,854) und im Feld beim Vergleich zum Goldstandard bei $r = 0,75$ (95 % LCI: 0,605; UCI: 0,846). Bei (b) liegt er im Labor bei $r = 0,916$ (95 % LCI: 0,88; UCI: 0,942) und bei Feldmessungen bei $r = 0,833$ (95 % LCI: 0,679; UCI: 0,917). Die Erkennung und Klassifizierung von Sportaktivitäten und -intensitäten liegt bei $r = 0,698$ (95 % LCI: 0,569; UCI: 0,793). Bei (c) liegt er im Labor bei $r = 0,39$ (95 % LCI: 0,225; UCI: 0,533), im Feld verglichen zur indirekten Kalorimetrie bei $r = 0,543$ (95 % LCI: 0,36; UCI: 0,685) und im Vergleich zum Goldstandard bei $r = 0,759$ (95 % LCI: 0,686; UCI: 0,817). Die ICC der Geräte beträgt bei (a) $r = 0,589$ (95 % LCI: 0,361; UCI: 0,751), bei (b) $r = 0,894$ (95 % LCI: 0,751; UCI: 0,957) und bei (c) $r = 0,516$ (95 % LCI: 0,0633; UCI: 0,792).

Konklusion

Wearables eignen sich im Sport zur Herzfrequenzmessung und bei Personen ohne eingeschränktem Gangbild auch für die Schrittzählung beim Gehen oder bei Ausdauersportarten. Für die Kalorienverbrauchsberechnung oder der Aktivitätserkennung sind Wearables weniger geeignet.

Abstract

Background

Wearable Fitness Tracker are gaining popularity worldwide. Fitness Tracker (so called Activity Tracker or Wearable) are electronic gadgets that are attached to the body and are able to record various different movement related data of a person like number of steps taken, the current or the over the day arithmetically averaged pulse, the distance covered and the daily calorie consumption.

Goal

The goal of this paper is to determine how exact the measuring methods of fitness trackers available on the consumer market are in terms of (a) step counter, (b) heart rate measurement and (c) energy consumption computation.

Methods

Databases like PubMed, Elsevier and Web of Science have been searched for studies which compared the validity of fitness trackers to either direct measurement or gold standard devices or studies that have tested the reliability of wearables.

Results

33 studies have been included in the meta-analysis. (a) has been covered by 23, (b) has been covered by 7 and (c) has been covered by 17 of these studies. For (a) the correlation in laboratory testing is $r = 0,909$ (95 % LCI: 0,855; UCI: 0,944), in field testing in comparison to direct counting of the steps the correlation is $r = 0,799$ (95 % LCI: 0,727; UCI: 0,854) and in field testing with comparison to the gold standard at $r = 0,75$ (95 % LCI: 0,605; UCI: 0,846). For (b) the correlation in laboratory testing is $r = 0,916$ (95 % LCI: 0,88; UCI: 0,942) and for field testing at $r = 0,833$ (95 % LCI: 0,679; UCI: 0,917). The identification and classification of activity type and intensity is $r = 0,698$ (95 % LCI: 0,569; UCI: 0,793). For (c) the correlation in laboratory testing is $r = 0,39$ (95 % LCI: 0,225; UCI: 0,533), for field testing with comparison to the indirect calorimetric method the value is $r = 0,543$ (95 % LCI: 0,36; UCI: 0,685) and in comparison, with the gold standard $r = 0,759$ (95 % LCI: 0,686; UCI: 0,817). The ICC of the devices for (a) is $r = 0,589$ (95 % LCI: 0,361; UCI: 0,751), for (b) it is $r = 0,894$ (95 % LCI: 0,751; UCI: 0,957) and for (c) it is $r = 0,516$ (95 % LCI: 0,0633; UCI: 0,792).

Conclusion

Wearables are well suited for measuring heart rate and in the case of people without restriction of their gait even for counting steps in either normal walking or endurance sport activities. However, wearables are not suited for identifying the type of activity and measuring the caloric expenditure.

Inhaltsverzeichnis

Zusammenfassung	3
Abstract	5
Inhaltsverzeichnis	7
Einleitung	9
1 Fragestellungen und Zielsetzung	11
1.1 Fitness Tracker im Konsumentenbereich	11
1.2 Goldstandards	12
2 Methode	13
2.1 Datenerhebung	15
2.2 Datenaufbereitung	16
2.3 Qualitätsbewertung	16
3 Rahmenkonzept und Theorien für die Bearbeitung der Fragestellung	19
3.1 Hardware von Fitness Tracker	19
3.1.1 Schrittzähler	21
3.1.2 Pulsmesser	24
3.1.3 Kalorienverbrauchsberechnung	27
3.2 Testmethoden in Studien	30
3.2.1 Gültigkeit	30
3.2.2 Zuverlässigkeit	31
3.2.3 Messumgebung	32
4 Aktueller Stand der Forschung	33
5 Ergebnisse	35
5.1 Schrittzählung	49
5.1.1 Gültigkeit der Schrittzählung	49
5.1.2 Zuverlässigkeit bei der Schrittzählung	72
5.2 Pulsmessung	75
5.2.1 Gültigkeit der Pulsmessung	75
5.2.2 Zuverlässigkeit bei der Pulsmessung	82

Inhaltsverzeichnis

5.3	Kalorienverbrauchsrechnung	84
5.3.1	Gültigkeit der Kalorienverbrauchsrechnung	84
5.3.2	Zuverlässigkeit bei der Kalorienverbrauchsrechnung	99
6	Diskussion	101
	Glossar	107
	Abkürzungsverzeichnis	109
	Literatur	111
	Anhang	119

Einleitung

Wearable Fitness Tracker erfreuen sich weltweit einer immer größeren Popularität. *Fitness Tracker*, auch Activity Tracker genannt, sind elektronische tragbare Geräte (Wearables), die am Handgelenk, der Hüfte oder an einem BH getragen werden und unterschiedliche bewegungsbezogene Daten einer Person, wie die Anzahl der Schritte, den aktuellen und den über einen Tag gemittelten Puls, die zurückgelegte Distanz messen und den täglichen Kalorienverbrauch schätzen können. Hierzu zählen auch „standalone personal communication devices“ wie Smartphones (Hayward, 2018). Ein Grund für diese gesteigerte Popularität könnte laut Chowdhury, Western, Nightingale, Peacock und Thompson (2017) die Möglichkeit sein, dass Personen ihre Sportaktivitäten selbst mitverfolgen können. Tagliente et al. (2016, S. 665) kamen zu dem Schluss, dass diese Art von Geräten, „wichtige Werkzeuge für die Prävention von Krankheiten und deren Intervention sind“. „Daher besteht ein enormes Interesse an der Messung der körperlichen Aktivität in der medizinischen Forschung“ (Wright, Hall Brown, Collier & Sandberg, 2017). Nach Statistiken von Statista Deutschland (2017) wurden für 2017 weltweit 121,7 Millionen Wearables prognostiziert, 2016 wurden noch 104,4 Millionen Stück verkauft und für das Jahr 2021 werden 229,5 Millionen Stück prognostiziert. In Deutschland, Österreich und Schweiz kennen nach einer Statistik vom Bundesverband Digitale Wirtschaft 65 %, 58 % bzw. 54 % den Gerätetyp Fitness Tracker, wohingegen nur 10 %, 9 % bzw. 10 % der Einwohner auch einen solchen besitzen (Jann, 2016, S. 15). Viele dieser Geräte bieten darüber hinaus die Möglichkeit die gemessenen Daten mit Apps am Smartphone oder Internetdiensten auswerten zu lassen. Smartphones können ebenfalls mithilfe der in ihnen selbst verbauten Inertial Measurement Units selbsttätig die bewegungsbezogenen Daten sportlicher Aktivität von Personen aufnehmen und interpretieren. Nach der Ausführung von Kirwan, Duncan und Vandelanotte (2013, S. 47), benutzen 59 der 400 meistdowngeloadeten Fitness-Apps in Australien solche Sensoren in Smartphones.

Das Ziel dieser Diplomarbeit ist es nun, einen systematischen Review über vorhandene Studien, welche die Gültigkeit und Zuverlässigkeit von Fitness Trackern und Apps für den Konsumentenbereich im Sport testen, zu erstellen.

Um die Gültigkeit und die Zuverlässigkeit der Wearables und der Apps zu bestimmen, sollen die Studien entweder im Labor- oder im Feldversuch durchgeführte Messungen beinhalten, welche zumindest eine aus dem Bereich gezählte Schritte, Puls und verbrannte Kalorien (siehe Kapitel 2) mit dem tatsächlichen Ausmaß oder einem Goldstandard vergleichen.

1 Fragestellungen und Zielsetzung

Wie exakt sind die Messmethoden von den im Konsumentenbereich verfügbaren Fitness Tracker bezogen auf

1. die vom Fitness Tracker gezählten Schritte?
Ergebnisse in Abschnitt 5.1.1
2. den vom Fitness Tracker gemessenen Puls?
Ergebnisse in Abschnitt 5.2.1
3. die verbrannten Kalorien, die der Fitness Tracker ermittelt?
Ergebnisse in Abschnitt 5.3.1

Die Bedingungen an Fitness Tracker aus dem *Konsumentenbereich* und die Anforderungen an *Goldstandard* Messverfahren und Geräte werden an dieser Stelle näher erklärt.

1.1 Fitness Tracker im Konsumentenbereich

Fitness Tracker werden als zum *Konsumentenbereich* gehörig angesehen, wenn sie den sechs Punkten der folgenden Aufzählung, modifiziert nach Beekman et al. (2017), entsprechen:

1. Der Tracker wird direkt am Handgelenk getragen oder durch ein extra erhältliches Armband ebenda.
2. Es besteht kein monatliches Abo für die Benutzung des Fitness Trackers oder der Webseite für den Datenzugriff.
3. Es gibt ein Echtzeitfeedback für den Benutzer, entweder per integriertem Display oder über das Smartphone durch Bluetooth oder ANT.
4. Die Herzfrequenzmessung muss direkt mit der Uhr und somit ohne zusätzlichen Brustgurt durchgeführt werden.
5. Der Fitness Tracker muss aus einem messenden Gerät bestehen, welches jedoch selbst mehrere unterschiedliche Sensoren enthalten kann.
6. Er darf kein deklarerter Prototyp sein.

1.2 Goldstandards

Goldstandards von Wearables in der Forschung (Abb. 1.1) sind beim Zählen der Schritte und auch Kalorien- bzw. Energieverbrauchsbestimmung unter anderem Akzelerometer der Hersteller ActiGraph, Actiheart, BodyMedia oder IDEEA (Bassett, 2012; Dannecker, Sazonova, Melanson, Sazonov und Browning, 2013; Dominick, Winfree, Pohlig und Papas, 2016; Hickey et al., 2016). Beim Zählen der Schritte kommt im Messumfeld von Labortestungen auch das manuelle Zählen der Studienleiter und -leiterinnen hinzu. Beim Energieverbrauch werden als Vergleichsmessungen auch spiroergometrische Messverfahren eingesetzt. Bei Herzfrequenzmessungen werden als Goldstandard medizinische Geräte wie Elektrokardiogramm (EKG) sowie auch nach Delgado-Gonzalo et al. (2015, S. 430) Wearables, welche die Herzfrequenz über einen Brustgurt bestimmen und seit 20 Jahren im Sport verwendet werden und hier eine Korrelation von $r > 0,9$ aufweisen, verwendet.

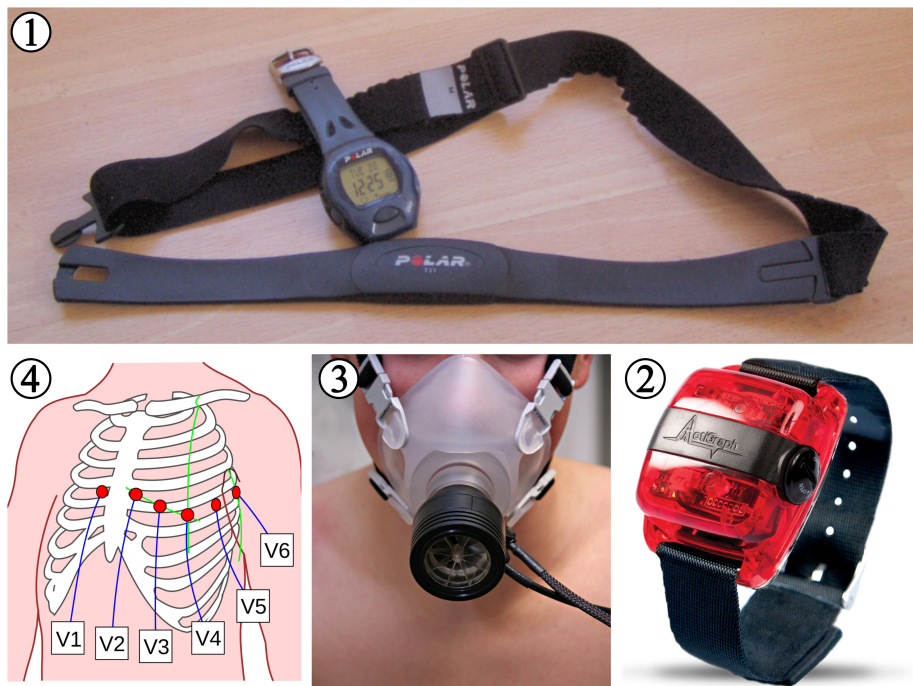


Abbildung 1.1: Goldstandardmethoden von Wearables in der Forschung.

- 1) POLAR Brustgurt für Herzfrequenzmessung (Wikimedia Commons, Public Domain), 2) ACTIGRAPH WGT3X-BT Akzelerometer(ActiGraph, LLC), 3) Spiroergometrie (Michael Heigl www.flickr.com/people/32838894@N00 CC BY 2.0), 4) Ableitungspunkte bei 12-Kanal EKG (Jmarchn commons.wikimedia.org/wiki/User:Jmarchn CC BY-SA 3.0)

2 Methode

Es wird eine systematische Übersichtsarbeit erstellt. Zu diesem Zweck wurde im Internet in unterschiedlichen Literaturdatenbanken nach Fachartikeln gesucht. Die englischen Suchbegriffe sowie Synonyme und deren deutschen Übersetzungen werden aus einzelnen Begriffen aus drei übergeordneten Kategorien, 1. ART, 2. MERKMAL und 3. SENSOR, kombiniert. Die Recherche wurde dabei aufgeteilt in eine Suche mit nur deutschsprachigen und eine mit rein englischsprachigen Wörtern.

Die drei Kategorien mit ihren Begriffen sind:

1. Art des Gerätes (ART): In der ersten Kategorie kommt einerseits a) der Überbegriff Wearable selber, aber auch für Wearable alternativ verwendete Begriffe die im Fitnessbereich verwendet werden, wie *Activity Tracker*, *Fitness Tracker* (Evenson, Goto & Furberg, 2015, S. 2) oder auch *Smartwatch* zum Einsatz. Andererseits sind auch b) die Herstellernamen der Geräte aus dem Konsumentenbereich, welche aus gelisteten Produkten von www.geizhals.eu im europäischen Raum (Stand: Oktober 2017) aus der online vorhandenen Kategorie Fitness-Tracker extrahiert wurden, angeführt.
 - a) Alle Überbegriffe für Wearables, die verwendet wurden: *Fitness Tracker*, *Activity Tracker*, *Gadget*, *Smartband*, *Smartwatch* und *Wearable(s)*. Weitere Bezeichnungen, die für Fitness Tracker nach Wright et al. (2017, S. 359) möglich wären, aber nicht verwendet wurden, sind Activity Monitor, Exercise Monitor, Exercise Tracker und Step Tracker.

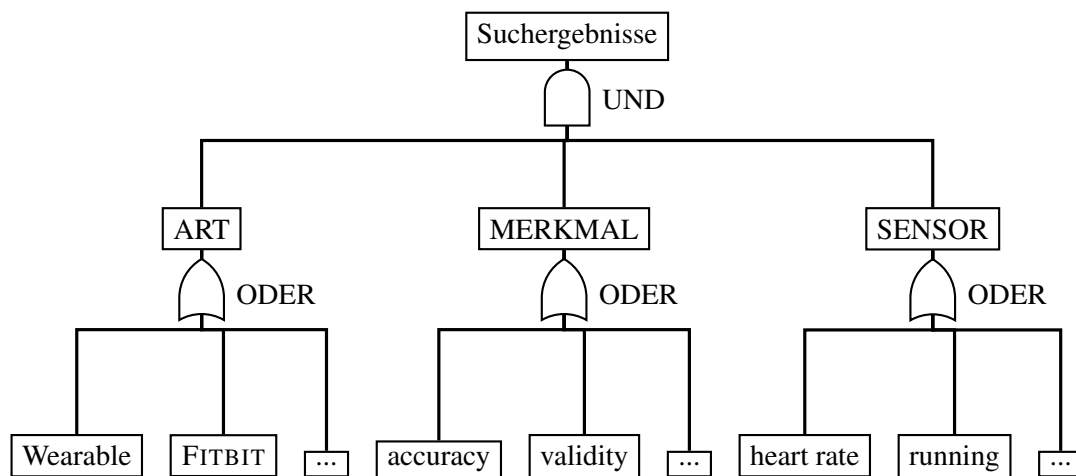


Abbildung 2.1: Kombination der einzelnen Suchbegriffe der drei Kategorien Art, Merkmal und Sensor

2 Methode

- b) Alle Firmennamen, die für die Suche verwendet wurden: ACER, ACME, ADIDAS, ALCATEL, APPLE, ASUS, BEURER, BODYMEDIA, CELLULAR LINE, COGITO, COOKOO, DENVER, FOSSIL, EPSON, FIT FLEX, FITBUG, FITBIT, GARMIN, HONOR, HUAWEI, IHEALTH, JAWBONE, JUBOURY, LENCISE, LG, LIFETrAK, MEDISANA, MICHAEL KORS, MIO, MISFIT, MOOV, MOREFIT, MOTOROLA, MYKRONOZ, NIKE, NILOX, NIXON, NOKIA, NUBAND, OMRON, OREGON SCIENTIFIC, PEBBLE, POLAR, RAZER, RUNTASTIC, SAMSUNG, SIGMA, SILVA, SIMVALLEY, SMARTLAB, SOLEUS, SONY, SPORTPLUS, STRIIV, SWISSTONE, TECHNAXX, TEEKA TAN, TIMEX, TOMTOM, TRISA, VECTOR, VITALMAXX, WAHOO FITNESS, WIKO, WITHINGS ACTIVITÉ, XIAOMI und XLYNE.
2. Merkmal der Testung der Studie (MERKMAL): In der zweiten Kategorie sind jene Begriffe gesammelt, die die Merkmale der Testung der jeweiligen Studie beschreiben. Dabei handelt es sich um die englischen Begriffe von Gültigkeit (*validity*) und Zuverlässigkeit (*reliability*) und deren Synonyme. Es wurden ebenfalls die Begriffe *accuracy* und *precision*, sowie deren Synonyme verwendet, da diese, nach Streiner und Norman (2006, S. 327), in manchen Publikationen fälschlicherweise als gleichbedeutend von *validity* bzw. *reliability* verwendet werden.
 - a) Alle Merkmale, die verwendet wurden: *accuracy, acuteness, benchmark, correctness, dependability, exactness, performance, precision, reliability, validation* und *validity*.
3. Sensorik und Art der körperlichen Bewegung (SENSOR): Der dritte Bereich umfasst a) die Sensoren und Messergebnisse der Wearables wie Kalorienverbrauch, Schritte und Puls und b) die Bewegungsformen aus dem Ausdauersport wie Laufen, Wandern und Gehen bei denen Fitness Tracker verwendet werden, sowie auch c) die Produktspezifizierung für den Konsumentenbereich.
 - a) Alle Sensorsysteme, die verwendet wurden: *calorie consumption, calory consumption, energy consumption, energy expenditure, expenditure of energy, gaits, GNSS, GPS, heart rate, PPG, pulse, steps* und *treads*.
 - b) Alle körperlichen Bewegungsformen, die verwendet wurden: *continuous exercise, continuous training, hiking, physical activity, physical exercise, run, running, sport, sports, trekking, walk* und *walking*.
 - c) Alle Spezifizierungen, die verwendet wurden: *commercial, commercially* und *consumer*. Hier könnte ebenfalls nach Wright et al. (2017, S. 359) noch der Begriff *popular* verwendet werden.

Die logische Kombination der Suchbegriffe aus den drei Kriteriengruppen ART, MERKMAL

und SENSOR wird in Abb. 2.1 dargestellt und ist wie folgt aufgebaut:

$$\begin{aligned} & (ART_1 \vee ART_2 \vee \dots \vee ART_n) \wedge \\ & (MERKMAL_1 \vee MERKMAL_2 \vee \dots \vee MERKMAL_n) \wedge \\ & (SENSOR_1 \vee SENSOR_2 \vee \dots \vee SENSOR_n) \end{aligned}$$

Als Gütekriterien für die Aufnahme der Studie gilt, dass die Studie aus dem Jahr 2012 oder jünger ist und zumindest ein getestetes Produkt aus dem Konsumentenbereich beinhaltet, welches entweder im Titel oder in der Zusammenfassung erwähnt wird. Welche Produkte als Konsumentengeräte und welche als Goldstandard angesehen werden, wird in Abschnitt 1.1 und Abschnitt 1.2 genauer ausgeführt.

2.1 Datenerhebung

Die Suche erfolgte von Februar bis Mai 2018. Es wurden dabei in den Meta-Datenbanken PubMed¹ und Web of Scienc² sowie dem Online-Portal des Elsevier-Verlags³ recherchiert.

Die Suchstrings sind beispielhaft für die Begriffe *accuracy* und *validity* (je aus der Kategorie ART), *Fitness Tracker* und *Wearable* (je aus MERKMAL) und *Photoplethysmography (PPG)* und *Global Positioning System (GPS)* (je aus SENSOR) für die drei Datenbanken:

1. PubMed

```
((accuracy[Title/Abstract]) OR (validity[Title/Abstract])) AND
((Fitness Tracker[Title/Abstract]) OR (Wearable[Title/Abstract]))
AND ((GPS[Title/Abstract]) OR (PPG[Title/Abstract])) AND ("2012"[Date
- Create] : "3000"[Date - Create])
```

2. Elsevier

```
pub-date > 2011 AND ((tak(accuracy) OR tak(validity)) AND (tak(Fitness
Tracker) OR tak(Wearable)) AND (tak(GPS) OR tak(PPG)))
```

3. Web of Science

```
PY=(2012-2018) AND TI=(((accuracy) OR (validity)) AND ((Fitness
AND Tracker) OR (Wearable)) AND ((GPS) OR (PPG)))
```

Die Ergebnisse wurden dabei nach den online bereitgestellten Möglichkeiten gefiltert, so dass nur Artikel ab 2012 oder jünger (PubMed: "2012"[Date - Create] : "3000"[Date - Create], Elsevier: pub-date > 2011 und Web of Scienc: PY=(2012-2018)) und Arti-

¹PubMed: www.ncbi.nlm.nih.gov/pubmed

²Web of Scienc: www.webofknowledge.com

³Elsevier: www.sciencedirect.com

kel, bei denen der Suchbegriff im Titel oder der Zusammenfassung aufscheint (PubMed: [Title/Abstract], Elsevier: tak und Web of Scienc: TI=), aufgelistet wurden.

2.2 Datenaufbereitung

Die Suchergebnisse, wie Titel, Autor und doi-Nummer, wurden nach dem Suchvorgang aus den unterschiedlichen Datenbanken exportiert und in MICROSOFT EXCEL (Office 365 Version 1803, Microsoft, Redmond, USA) zusammengefügt. Die Titel wurden mit NOTEPAD++ (Version 7.5.9, Don Ho, GPL-2.0) normalisiert – Sonderzeichen und Leerzeichen in den Titeln entfernt – und mit der EXCEL-Funktion Duplikate entfernen, die mehrfach vorhandenen Artikel maschinell entfernt.

Aus den Studien wurden die Korrelationen bzw. die Intra-Klassen-Korrelationen entnommen und mittels der Software MEDCALC (Version 18.11.3, MedCalc Software, Belgien) die untere und obere Grenze des 95%-Konfidenzintervalls (LCI & UCI), sowie zusätzlich eine Gewichtung pro Fall erstellt. Diese einzelnen Fälle werden in Kapitel 5 in Forest-Plots dargestellt. Mittels MEDCALC wurde für die Forest-Plots die Gesamtkorrelation und das dazugehörige 95%-Konfidenzintervall der Fälle pro Messsystem (Schritt-, Pulsmessung und Kalorienverbrauchsberechnung) und Messumgebung (Labor, Feld) berechnet. „MEDCALC uses the Hedges und Olkin (1985) method for calculating the weighted summary Correlation coefficient under the fixed effects model, using a Fisher Z transformation of the correlation coefficients“ (MedCalc Software, 2019). Verwendet wurden aber neben den gewichteten Ergebnissen, die von MEDCALC ebenfalls ausgewerteten randomisierten Ergebnisse. Die Entscheidung für diese Art der Ergebnisse wurde getroffen, da die Effekte auf die zugrundeliegende Population der Fitness Tracker interessant sind und die einzelnen Proben, welche Kombinationen von bestimmten Fitness Tracker und diversen Aktivitäten darstellen, klein sind (Gelman, 2005, S. 20, zit. n. Searle, Casella & McCulloch, 1992, Green & Tukey, 1960). Die Homogenität der einzelnen Datensätze wurde ebenfalls mit MEDCALC berechnet: „Next the heterogeneity statistic is incorporated to calculate the summary Correlation coefficient under the random effects model“ (MedCalc Software, 2019, zit. n. DerSimonian and Laird, 1986).

Zur Veranschaulichung der Ergebnisse (Kapitel 5) wurden die Forest-Plots online mit DistillerSR Forest-Plot Generator (Jänner 2019, Evidence Partners, www.evidencepartners.com/resources/forest-plot-generator) erstellt.

2.3 Qualitätsbewertung

Die Studien, die in die Meta-Analyse einfließen, werden einer Validitätskontrolle unterzogen. Die einzelnen Qualitätskriterien hierfür werden in der nachfolgenden Aufzählung,

eingegliedert in den Kategorien *Studiendesign*, *Population*, *Hardware*, *Datenauswertung* und *Diskussion*, näher beschrieben (modifiziert nach Blümle, von Elm, Antes und Meerpohl, 2014, S. 499, Tabelle 1, „Messung und Bewertung der Studienqualität“; Evenson et al., 2015; Higgins und Straub, 2006; Matthews, Hagströmer, Pober und Bowles, 2012, S. 6–9). Matthews et al. (2012, S. 17) bieten darüber hinaus eine praktikable Checkliste für Populationsstudien mit Wearables an. Die Qualitätsergebnisse, der in dieser Arbeit inkludierten Studien, werden in Kapitel 5 in Tabelle 5.2 aufgelistet.

1. Studiendesign

Die Studien sollen angeben, welche *Anzahl* an Geräten verwendet werden soll und wie viele schlussendlich auch *verwendet wurden*. Die *Testdauer* pro Gerät soll angegeben werden. Nach Evenson et al. (2015, S. 3) können die Ergebnisdaten der Tracker meist nur tageweise summiert oder mit einer passenden Programmierschnittstelle (engl.: Application Programming Interface) (API) auch auf 60 min bis 1 min reduziert werden. Matthews et al. (2012, S. 6) empfehlen für Feldstudien eine Testdauer von mindestens 10 h pro Tag. Ebenfalls soll auch der *Zeitraahmen*, in dem getestet wurde, und der *Ort* der Studiendurchführung ersichtlich sein. Da es sich, wie Evenson et al. (2015, S. 19) und Wright et al. (2017, S. 362) anmerken, bei den Fitness Trackern für Konsumenten um geheimgehaltene Algorithmen in deren proprietäre Firmware handelt, die jederzeit durch ein Update verändert werden können, kann zusätzlich der Zeitrahmen als Anhaltspunkt für die Replizierung von Studienergebnissen dienen. Auch das Kaufdatum (in der vorliegenden Arbeit nicht aufgenommen) kann hier ein wichtiges Merkmal dafür liefern. Es soll auch klar angegeben werden, dass es sich um eine *randomisierte kontrollierte Studie* (engl.: *Randomized Controlled Trial*) (*RCT*) handelt. Speziell bei Feldtests kommt noch die geplante *Tragedauer*, wie lange das Geräte getragen werden soll, zum Gegensatz der Testdauer pro spezifizierten Test, hinzu. Weiters auch ob eine *Einschulung* für Geräte durchgeführt wurde und wie *Trageunterbrechungen*, wie bei Schlaf, Wasserkontakt, etc., dokumentiert wurden.

2. Population

Es soll eine möglichst *randomisierte Population* gewählt werden. Der Vorgang der *Anwerbung* der Teilnehmer und Teilnehmerinnen sowie deren *demographischen Angaben* wie Geschlecht, Alter und Größe und Gewicht oder BMI sollen angegeben werden. Zusätzlich hierzu auch die *Eignung*, also die Ein- und Ausschlusskriterien, der Personen.

3. Hardware

Es sollen in den Studien die *Generation* der getesteten Modelle und deren *Firmwareversion* angegeben werden, um die verwendete Hardware in Kombination mit ihrer Software eindeutig identifizieren zu können. Die Wearables sollen bei den Testungen laut den offiziellen *Herstellerangaben* getragen werden. Dies soll ersichtlich sein, genauso wie die Beachtung der *Armdominanz* (Händigkeit) der Studienpartizipanten. Wenn

2 Methode

mehrere Geräte der Modellreihe und Firmwareversion getestet werden, soll angemerkt werden, ob diese *Geräte randomisiert* wurden.

4. Datenauswertung

Die *Software* und deren Version, mit denen die Daten extrahiert werden (spezielle Smartphone-Apps oder PC-Pogramme), soll neben der *Methode der Auswertung* der Ergebnisse dokumentiert sein. Auch ob diese *Auswertung verblindet* durchgeführt wurde. Wenn *unvollständige Daten* aufkommen (z.B.: Austritt einer Person, Ausfall eines Gerätes), soll dies und der Umgang mit selbigen Umstand erwähnt werden. Die Autoren sollen *Störgrößen* beachtet und bewertet haben. Beispielhaft durch Aussagen wie die folgende von von (J.-M. Lee, Kim & Welk, 2014): „Most of the consumer-based activity monitors do not provide direct access to the raw data, so estimates of EC were obtained directly from the associated Web sites for each monitor.“, welche dahinter eine mögliche Reduzierung der Datengenauigkeit vermuten.

5. Diskussion

Die *Schlussfolgerung* soll auf den erhobenen Studienergebnissen aufgebaut sein. Es soll auch der Einfluss auf die Ergebnisse durch *Störgrößen diskutiert* worden sein. Auch eine mögliche *Förderung* der Studie soll angemerkt sein.

3 Rahmenkonzept und Theorien für die Bearbeitung der Fragestellung

3.1 Hardware von Fitness Tracker

Ein Fitness Tracker besteht aus diversen Bestandteilen (Abb. 3.1), die je nach Manufaktur und Serie unterschiedlich ausfallen und unter proprietären Gesichtspunkten oft auch nicht direkt einsehbar sind. Grundsätzlich besteht ein Fitness Tracker aus folgenden Hardwarebauteilen:

Sensorsystem

Das Sensorsystem besteht je nach Fitness Tracker aus unterschiedlichen Sensoren, die diverse Bewegungs- oder physiologische Merkmale (*Motion Sensing* und *Biometric*, *Pressure*, *Temperature sensors*) messen können.

Mikrocontroller

Im Mikrocontroller (*Embedded Controller*) laufen die Messungen der Sensoren zusammen, werden ausgewertet, weiterverarbeitet und für die Benutzer des Trackers aufbereitet.

Speicher

Der Speicher (*Serial Flash*) wird benötigt, um die Sensorwerte und Ergebnisse für eine spätere Anzeige oder Weiterleitung an andere Geräte in digitalen Speicherbausteinen zwischenzuspeichern.

Stromversorgung

Die Stromversorgung (*Power Supply*) eines Fitness Trackers ist meistens ein integrierter wiederaufladbarer Akku mit einer Laufzeit von Stunden oder wenigen Tagen bis selten sogar Wochen (abhängig von der Art und Anzahl der verwendeten Sensoren, des Displays und der Kommunikationskanäle). Seltener wird auch eine Lithiumbatterie dafür verwendet. Wright et al. (2017, S. 363) fanden heraus, dass von 308 am Handgelenk getragenen Fitness Trackern für Konsumenten, 33 nicht aufladbare Batterien verwendeten, die eine Laufzeit von einem Jahr und bei einem Tracker bis zu zwei Jahren aufwiesen.

Anzeige

Die Anzeige (*Display*) ist ein häufig vorkommender, aber nicht notwendigerweise benötigter Bestandteil eines Fitness Tracker, um die Ergebnisse dem Benutzer zur Verfügung zu stellen. Die Anzeige wird durch eine oder mehrere LEDs, die die Erreichung eines vorgegebenes Zieles anzeigen, bis zu einem RGB-Bildschirm, der ausführliche Informa-

3 Rahmenkonzept und Theorien für die Bearbeitung der Fragestellung

tionen darstellen kann, realisiert. Durch einen Vibrationsmotor (*Vibrating Motor*) oder akustische Signale durch Summer (*Buzzer*), ergänzen oder ersetzen manche Tracker die visuelle Rückmeldung.

Kommunikation

Bei den meisten Fitness Trackern können die Werte auf das Smartphone oder auf den PC überspielt werden, um sie besser und ausführlicher darzustellen oder auch um sie auf Onlinesysteme hinaufladen zu können. Die Konnektivität erfolgt entweder kabelgebunden, unter anderem mittels *USB*, oder kabellos (*Wireless Connectivity*) über Bluetooth Low Energy, seltener auch ANT oder WLAN.

Eingabe

Die Eingabe und Steuerung eines Fitness Trackers kann beginnend von nur einer Taste bis hin zu einem Touchscreen (*Capacitive Keys + Touch Screen*), selten auch einer Spracheingabe (*Microphone*), erfolgen.

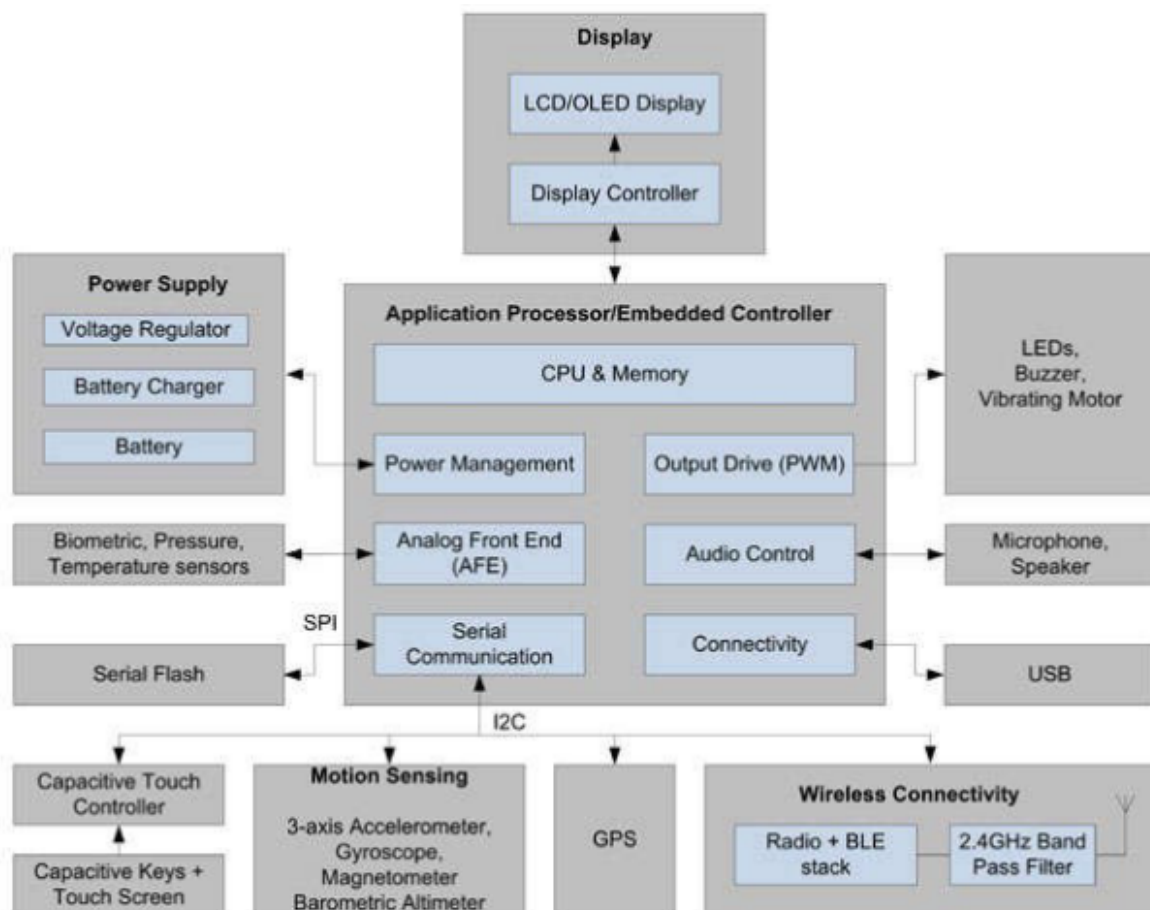


Abbildung 3.1: Hardware eines Fitness Trackers (Vairamuthu, Chethan & Sivaguru, 2018)

Die Messsysteme (Abb. 3.2) von Wearable Trackers ermöglichen eine quantitative Messung von im Sport und in Alltagsbewegungen auftretenden Werten wie Schritten, Distanz, Dauer von Bewegungen und Puls bis hin zur Zusammensetzung des Schweißes. Je nach verwendeten Messsystemen und Algorithmen sind die Ergebnisse qualitativ hochwertiger und so können

über die Dauer, Häufigkeiten und Intensitäten von Bewegungen hinausgehend auch Schlüsse über Wach-, Aktivitäts-, Ruhe- und Schlafzeiten sowie den Kalorienverbrauch gezogen werden (Alharbi, Straiton und Gallagher, 2017, S. 24–25; Seshadri, Drummond, Craker, Rowbottom und Voos, 2017, S. 40). Diese Informationen können gesammelt für den ganzen Tag oder auch für eine spezielle Dauer oder Bewegungsart angezeigt werden.

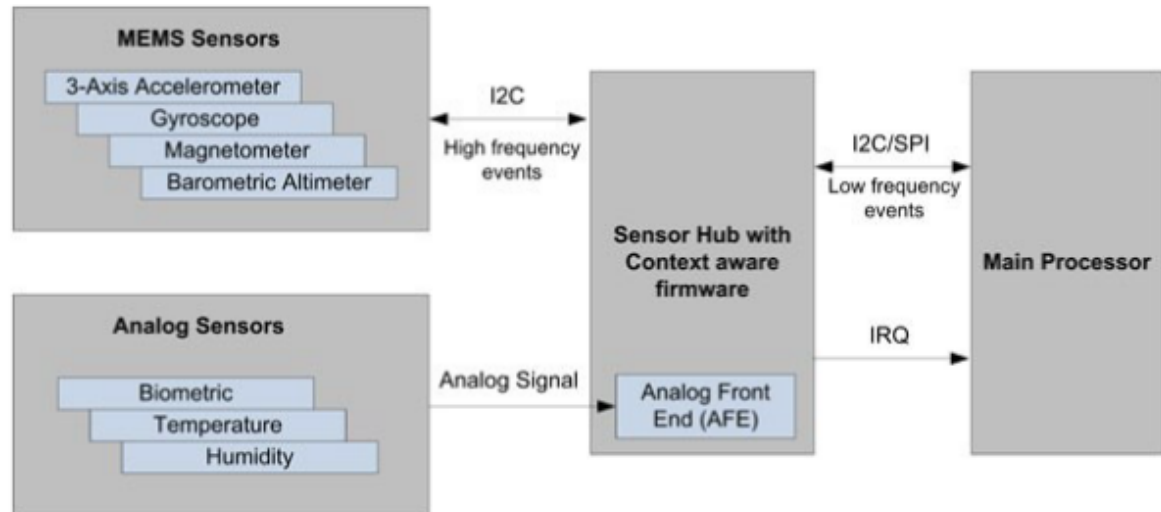


Abbildung 3.2: Sensorsysteme eines Fitness Trackers (Vairamuthu, Chethan & Sivaguru, 2018)

Sensoren der Messsysteme in dieser Arbeit sind:

Pedometer

Bewegungssensor, welcher Schritte dadurch zählt, dass die beim Auftreten eines Schrittes entstehenden vertikalen Beschleunigungen einen federgelagerten Hebel bewegen (Schrittzähler Abschnitt 3.1.1).

Akzelerometer

Bewegungssensor, welcher auftretende Beschleunigungen durch Piezoeffekte in Spannung umwandelt und so die Beschleunigung direkt erfasst (Schrittzähler Abschnitt 3.1.1 und Kalorienverbrauchsberechnung Abschnitt 5.3).

Pulsmesser

Physiologischer Sensor, welcher peripheren Puls mittels eines optischen Systems misst (Pulsmesser Abschnitt 3.1.2 und Kalorienverbrauchsberechnung Abschnitt 5.3).

3.1.1 Schrittzähler

Schrittzähler, zu denen Pedometer zählen, sind nach Cadmus-Bertram (2017, S. 34) und Li et al. (2016, S. 75) die einfachste und wohl auch meistverwendete Form eines Bewegungssensors. Fitness Tracker bieten die Möglichkeit, ein tägliches oder wöchentliches Ziel an Schritten selbst zu definieren, z.B.: Fitness Tracker der FITBIT-Reihe, oder der Tracker

gibt das Ziel anhand des vom Benutzer eingestellten Aktivitätslevels vor, z.B.: GARMIN VIVOFIT (Cadmus-Bertram, 2017, S. 35). Ein Schritt wird jedes Mal gezählt, wenn durch eine vertikale Beschleunigung, wie die Kraft, die beim Fersenauftritt entsteht, eine bestimmte Beschleunigungsgrenze übertroffen wird. Gehen ($4,5 \text{ km h}^{-1}$ bis $5,4 \text{ km h}^{-1}$ (Wikipedia Autoren, 2019b) und Jogging (Laufen langsamer als 10 km h^{-1} (Wikipedia Autoren, 2019a)) sind zwei Bewegungsformen, bei denen Pedometer häufig zum Zählen der Schritte eingesetzt werden. Die ersten Formen solcher Sensoren waren mechanisch und an der Hüfte angebracht. Heutzutage werden auch Mikrosysteme (engl.: MicroelectroMechanical Systems) (MEMS) verwendet (Ainsworth, Cahalin, Buman & Ross, 2015), die mehrere Systeme kombinieren. Die Messungen dieser MEMS können auch für die Kalorienverbrauchsberechnung genutzt werden (siehe Abschnitt 5.3).

Pedometer

Die Funktionsweise eines Pedometers wird hier anhand des Patentes eines mechanischen Schrittzählers erklärt (Bartel & Kuhn, 1904). Abb. 3.3 zeigt den Aufbau eines mechanischen Pedometers mit folgender Beschreibung der Bauteile (Woodford, 2010):

- a* Pendel, welches bei Bewegungen schwingt
- b* Gewicht am Ende des Pendels *a*, damit jenes mit einem gewissen Kraftbetrag auf und ab schwingen kann
- d* Ankerpunkt des Pendels
- e* Zahnräder (*e* und strichliert dargestellte Kreise), welche vom Pendel bewegt werden, um Schritte zu zählen (Mechanik der Anzeige nicht abgebildet)
- g* Feder, welche das Pendel zum Ursprungszustand zurück bewegt. Die Feder ist ein Teil des Mechanismus um unterschiedliche Schrittgrößen einzustellen
- m* Schraube, welche ebenfalls ein Teil des Mechanismus ist, um unterschiedliche Schrittgrößen einzustellen
- p* Skala mit Nadel *a*² am Gewicht *b* um die Zählschwelle einzustellen

Federgelagerte elektronische Pedometer (Abb. 3.4) funktionieren ähnlich wie rein mechanische (Abb. 3.3). Neben den bekannten mechanischen Anteilen Feder, Hebel und einer Masse am Ende des Hebels, besitzen diese Varianten eine zusätzliche elektronische Schaltung. Durch die Krafteinwirkung eines Schrittes wird bei dieser Schaltung ein Stromkreis geschlossen und so ein Schritt gezählt. Die Bauteile sind folgend analog zur Bauteilaufzählung des mechanischen Pedometers Abschnitt 3.1.1:

- a* Pendel, welches bei Bewegungen schwingt

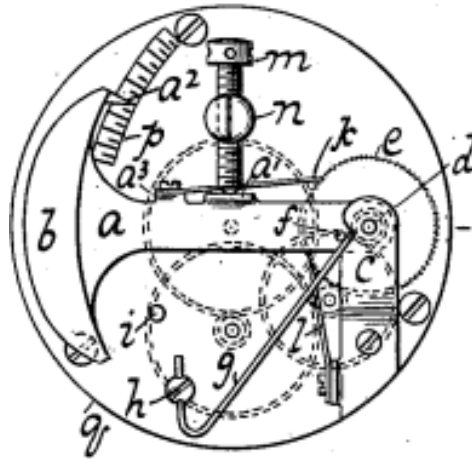


Abbildung 3.3: Mechanischer Pedometer (Bartel & Kuhn, 1904)

- b* Gewicht am Ende des Pendels *a*, damit jenes mit einem gewissen Kraftbetrag auf und ab schwingen kann
- c* Batterie
- d* Ankerpunkt des Pendels
- e* Pendel schließt bei genügend starker Bewegung einen Stromkreis wodurch ein Schritt gezählt wird
- g* Feder, welche das Pendel zum Ursprungszustand zurück bewegt

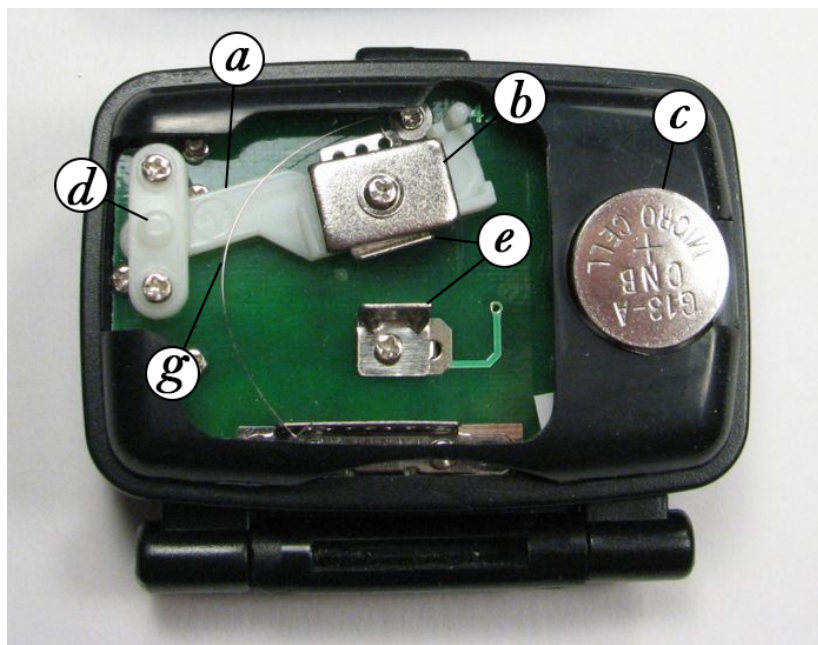


Abbildung 3.4: Pedometer mit federgelagertem Hebelarm

Modif. nach Lenore Edman www.flickr.com/photos/lenore-m CC BY 2.0

Genannte Vor- und Nachteile von Pedometern nach Ainsworth et al. (2015) sind:

Vorteile:

3 Rahmenkonzept und Theorien für die Bearbeitung der Fragestellung

- eine einfach zu verstehende Metrik, die Schritte,
- günstige Herstellung und
- Schritte als ein wichtiger Faktor im Feedback und der Motivation von Sporttreibenden.

Nachteile:

- eine Einsatzfähigkeit bei ausschließlich gehfähigen Personen,
- der fehlende direkte Rückschluss auf verbrannte Kalorien und
- deren bestehende Abhängigkeit zu proprietären Algorithmen.

Akzelerometer

Eine weitere technische Umsetzung und neuere Entwicklung für Schrittzähler sind Geräte, die sich den piezoelektrischen oder den piezoresistiven Effekt zunutze machen. Der *piezoelektrische* Effekt erzeugt bei einer Verformung eines piezoelektrischen Materials eine elektronische Spannung. Der *piezoresistive* Effekt folgt aus der Veränderung des elektrischen Widerstandes eines Materials, welches durch Zug oder Druck verformt wird. Schrittzähler, welche sich diese Sensorik zunutze machen, messen dabei direkt die auftretende Beschleunigung. Die Abtastrate beträgt dabei 40 Hz min^{-1} bis 100 Hz min^{-1} . Ein Computerchip zählt, durch spezielle Algorithmen zum Erkennen von Schritten, die je nach Bauweise mechanisch auftretenden Beschleunigungsspitzen oder den Wechsel von positiver zu negativer Beschleunigung als Schritte. Neben der Anzahl der Schritte, kann die Messung der Beschleunigung auch zur Berechnung des Kalorienverbrauches (siehe Abschnitt 5.3) genutzt werden. Die Kombination dieser mechanischen Bewegungssensorik mit dem interpretierenden Computerchip wird als MEMS bezeichnet. Die MEMS fassen dabei mehrere einzelne Sensoren zusammen und können demnach nach der Anzahl ihrer Messachsen unterteilt werden (Ainsworth et al., 2015; Bassett, 2012, S. 1770–1771; Li et al., 2016). So zählen van Remoortel et al. (2012) Uniaxial- (z. B.: POLAR ACTIVITY WATCH 200), Biaxial-, Triaxialbeschleunigungsmesser (z. B.: ACTIGRAPH GT3X) oder auch Multisensoren (z. B.: SENSEWEAR MINI ARMBAND) auf.

3.1.2 Pulsmesser

Geräte für eine Messung der Herzfrequenz werden häufig für direkte physiologische Messungen des Pulses im Alltag oder beim Sport gewählt. Die technische Umsetzung der Pulsmessung kann über einen Transducer, der über zwei integrierte Elektroden die R-Impulse im QRS-Komplexes des Elektrokardiogramms (EKG) misst (Abb. 3.5), an einem Brustgurt getragen wird und die Messdaten per Funk an eine am Handgelenk getragene Pulsuhr überträgt, erfolgen. Eine weitere Möglichkeit der Pulsmessung, ist optisch durch Photoplethysmography (PPG), mit einem direkt am Handgelenk getragenen Fitness Tracker. PPG ist

„eine einfache, kostengünstige und nichtinvasive Technik“ (Elgendi, 2012, S. 23), die auch bei der Pulsoxymetrie in der Medizin zum Einsatz kommt.

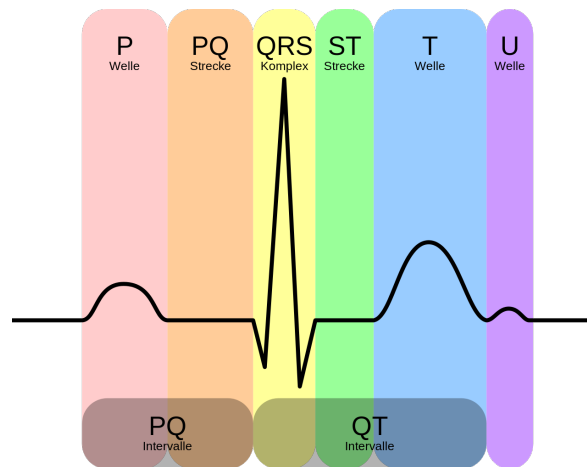


Abbildung 3.5: Schematische EKG-Darstellung mit dem vom Brustgurt gemessenen R-Impuls im QRS-Komplex
 Hazmat2 commons.wikimedia.org/wiki/File:EKG_Complex_de.svg CC BY-SA 3.0

Während bei der üblichen Pulsoxymetrie die Lichtquelle (meistens eine LED) durch einen durchbluteten peripheren Körperteil (meistens eine Fingerspitze) auf einen Lichtsensor strahlt, wird bei am Handgelenk getragenen Fitness Tracker über die Reflexion der Kapillargefäße, das ausgestrahlte Licht durch volumetrische Änderungen, die durch den Pumpvorgang des Herzens variieren, unterschiedlich stark reflektiert und durch einen Fotodetektor neben der Lichtquelle gemessen („plethysmo“ griechisch für „Erweiterung“) und daraus der Puls interpretiert (Abb. 3.6).

Wenn das Herz während der Systole Blut in den Körperkreislauf pumpt, erhöht sich vorübergehend das Volumen in den Kapillargefäßen, während bei der Diastole, wenn das Blut durch die Venen zurückfließt, sich das Volumen verringert. Die Volumendifferenzen bewirken einen kleinen Unterschied in der Ausdehnung des Gefäßsystems, welche wiederum das Licht unterschiedlich stark reflektieren. Abb. 3.7 zeigt, wie dieser durch den Puls verursachte unterschiedliche Durchscheinungsanteil (als „AC“ bezeichnet) zusätzlich zu der quasi-statischen Grundreflexion („DC“) von Gewebe, venösem Blut und pulsunabhängigem arteriellem Blut hinzukommt und somit den Herzschlag und die Herzfrequenz wiedergibt. Das Licht wird dabei von Gewebe, Knochen und Blutgefäßen unterschiedlich reflektiert. Um die Messgenauigkeit zu erhöhen, werden entweder langwelliges Infrarotlicht oder kurzwelliges grünes Licht verwendet. Langwelliges Infrarotlicht kann tiefer in den Körper eindringen, erhöht dadurch aber auch die Anfälligkeit von Bewegungsartefakten, wohingegen das kurzwellige grüne Licht weniger weit eindringt und so stabiler gegenüber von Bewegungsartefakten ist, dies aber auf Kosten der Genauigkeit bei schlecht durchbluteter Haut (z.B.: kühlere Haut) geht. (Delgado-Gonzalo et al., 2015; Sun und Thakor, 2016).

Durch die Messung am Handgelenk entfällt der Bedarf des Tragens eines dezidierten Brustgurtes, da PPG theoretisch eine Messung an jedem Punkt der Körperoberfläche ermöglicht.

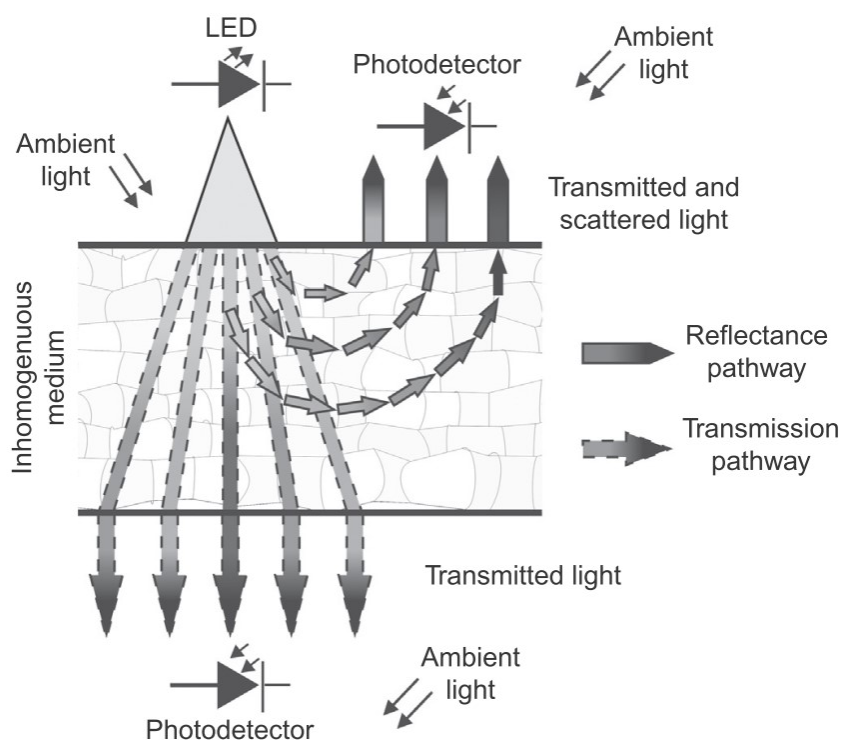


Abbildung 3.6: Transmission vs Reflektion (Reflectance) bei PPG-Messung (Lemay et al., 2014)

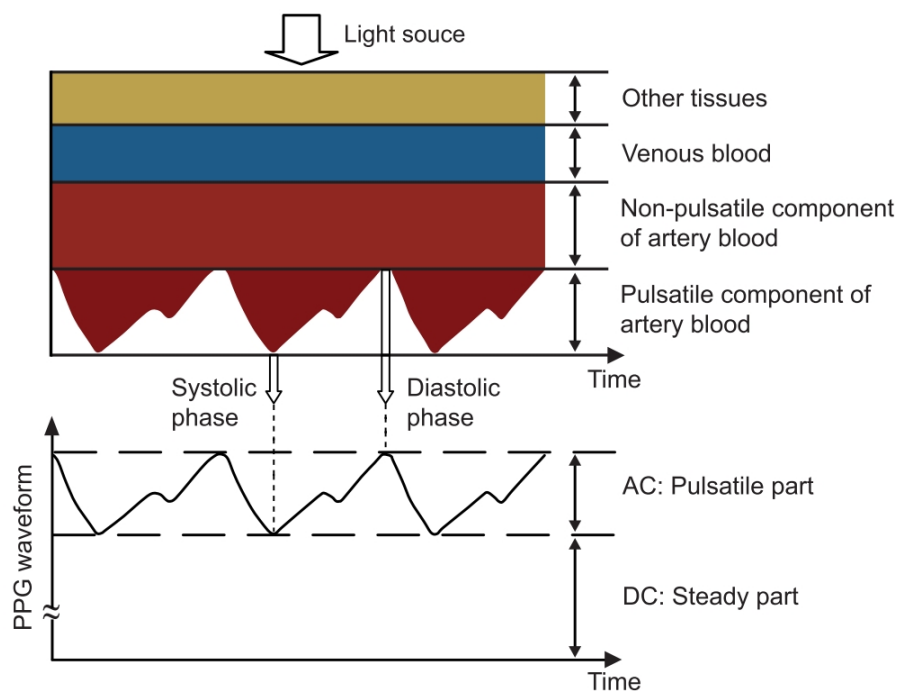


Abbildung 3.7: Durscheinungsanteile nach Gewebeart bei PPG (Sun & Thakor, 2016).

Damit jedoch PPG-Sensoren korrekt messen können, müssen jene Sensoren direkt auf der Haut aufliegen. Dies kann wiederum als unbequem erachtet werden und ist auch anfällig gegenüber von Messartefakten durch Bewegungen im Alltag und noch stärker ausgeprägt bei sportlichen Tätigkeiten. (Sun & Thakor, 2016). Weitere Messfehler können durch den unterschiedlichen Schichtaufbau der Haut (Abb. 3.8), wie „changes in skin colour, tattoos, temperature, erratic movement, and variations in skin pressure or even hairy wrists as well as fluctuations in oxygen levels“ (Carpenter & Frontera, 2016, S. 792), auftreten. Kombiniert mit einem Beschleunigungssensor können diese Messartefakte reduziert werden. Umgekehrt werden Ergebnisse durch Energieverbrauchsmessungen mittels Akzelerometer unter der zusätzlichen Berücksichtigung der Herzfrequenz verfeinert (Bai, Hibbing, Mantis und Welk, 2017; Sun und Thakor, 2016).

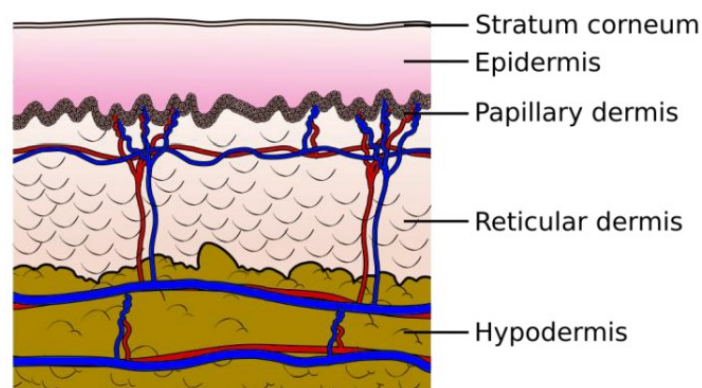


Abbildung 3.8: „Aufbau der menschlichen Haut. Die Dicke der *papillary dermis* und der *reticular dermis* variiert zwischen 0,6 mm und 3 mm. Wenn die Haut kalt ist, ist der Blutdurchfluss in der *papillary* und *reticular dermis* reduziert.“ (Delgado-Gonzalo et al., 2015)

Um quantitativ den Gesamtenergieverbrauch (siehe Abschnitt 5.3) eines Menschen oder die Einteilung der subjektiven Intensität und Dauer einer Bewegung zu messen, kann ebenfalls auf diese Pulsmessung zurückgegriffen werden. Die Herzfrequenzmessung bietet sich auch bei Bewegungen, die von einem Beschleunigungssensor nur schwer detektiert werden können, wie Radfahren und Schwimmen, an (Ainsworth et al., 2015; Delgado-Gonzalo et al., 2015).

3.1.3 Kalorienverbrauchsberechnung

Die Bestimmung des Energieverbrauchs (engl.: Energy Consumption) (EC) oder auch Kalorienverbrauchs ist komplexer als bei den oben genannten Messmethoden. Sie erfolgt durch die Messungen der verbauten Sensoren, wie Akzelerometer, PPG und Pedometer und der Berechnung durch einen je nach Hersteller variierenden proprietären Algorithmus im Fitness Tracker. Als Datengrundlage kommen meistens Beschleunigungsdaten vom Akzelerometer oder die Herzfrequenz zum Einsatz. Bassett (2012, S. 1771) und Cadmus-Bertram (2017) beschreiben, dass frühere Energieverbrauchsmessungen noch auf Pedometer und deren gezählte Schritte basierten, welche aber durch erschwinglichere Akzelerometer und deren besseren Genauigkeit

3 Rahmenkonzept und Theorien für die Bearbeitung der Fragestellung

und größeren Messumfang ersetzt wurden. So bestimmen neuere Fitness Tracker, wie z.B.: FITBIT FLEX, JAWBONE UP 24, MISFIT SHINE, NIKE+ FUELBAND und POLAR LOOP mittels triaxialer Beschleunigungssensoren den Energieverbrauch. Durch die Kombination von drei Bewegungsrichtungen bei triaxialen Sensoren, sind die Ergebnisse genauer als bei uniaxialen Sensoren (Bai et al., 2016; Stackpool, 2013; Wright et al., 2017).

Bevor Fitness Tracker die Ergebnisse darstellen, müssen zwei Prozesse durchlaufen werden. Der erste ist die Datenerfassung durch die Messsysteme und der zweite die Klassifizierung der Bewegung. Um den Energieverbrauch zu messen, greifen Fitness Tracker auf die Beschleunigungsrohdaten der Sensoren, den Zeitverlauf und die persönlichen Einstellungen des Benutzers zurück. Abb. 3.9 zeigt den schematischen Ablauf bei der POLAR A300. Die Rohdaten der Sensoren werden gefiltert, unter dem Zeitverlauf integriert und in vorab definierten Zeitintervallen, sogenannten „epochs“ (Bassett, 2012), gespeichert. Die Daten werden nach dieser Verarbeitung unterschiedlichen körperlichen Aktivitäten zugeordnet. Daraus wird ein metabolisches Äquivalent (engl.: Metabolic Equivalent Task) (MET) interpretiert und daraus der Kalorienverbrauch berechnet (Bassett, 2012; Boeselt et al., 2016; Chiauzzi, Rodarte und DasMahapatra, 2015; Kumari, Mathew und Syal, 2017).

1 MET entspricht einem Energieverbrauch von $1 \text{ kcal kg}^{-1} \text{ min}^{-1}$ und wird „verwendet, um den Energieverbrauch verschiedener Aktivitäten zu vergleichen. Es ist die Beschreibung des Stoffwechselumsatzes eines Menschen bezogen auf den Ruheumsatz im Verhältnis zu seinem Körpergewicht“ (Wikipedia Autoren, 2018). Ainsworth et al. (2015, S. 389) definieren 1 MET als den Energieverbrauch, der beim ruhigen Sitzen eines Menschen entsteht. Alharbi, Bauman, Neubeck und Gallagher (2016, S. 1479) schreiben, dass der Tracker FITBIT FLEX MET verwendet, um die tägliche Aktivität zu bestimmen. Sie teilen unterschiedliche Intensitäten in MET-Bereiche ein: leichte Bewegungsformen (engl.: Light PA) (LPA) sind jene, die bis zu 3 MET verbrauchen (z.B.: Liegen, Sitzen, langsames Gehen), moderate Bewegungsformen (engl.: Moderate PA) (MPA) liegen zwischen 3 und 6 MET (z.B.: Fahrradfahren, Gehen) und intensive Bewegungsformen (engl.: Vigorous PA) (VPA) beginnen ab 6 MET (z.B.: Jogging, Laufen, Schnurspringen).

Um die verbrauchten Kalorien eines Benutzers zufriedenstellend ermitteln zu können, benötigen die Fitness Tracker Algorithmen, welche die Intensität von moderat bis intensiv bestimmen können. Dies wird durch eine Bewegungserkennung erreicht, die sitzende, gehende oder laufende Bewegungen unterscheiden kann (Gomersall et al., 2016; Pavey, Gilson, Gomersall, Clark und Trost, 2017; Reid et al., 2017). Aber es können, wenn ausschließlich Akzelerometer verwendet werden, nicht für alle Sportarten der EC richtig berechnet werden. So erwähnen Ainsworth et al. (2015, S. 391) und Brooke et al. (2017, S. 1106), dass Radfahren, Gewichtheben oder isometrisches Training wegen der Messmethode von Beschleunigungssensoren nicht richtig erkannt werden. Cadmus-Bertram (2017, S. 35) erwähnt, dass einige Wearables so konzipiert sind, dass sie auch manuelle Eingaben von Aktivitäten zulassen.

Die Interpretierung des Kalorienverbrauches erfolgt nach Bassett (2012, zit. n. Montoye, Washburn, Servais, Ertly, Webster & Nagle, 1983) anhand einer linearen Regression, welche aber bei der Bestimmung vom EC eingeschränkt ist, da diese über die große Bandbreite von unterschiedlichen Sportarten nur beim Laufen und Gehen am Laufband genau ist.

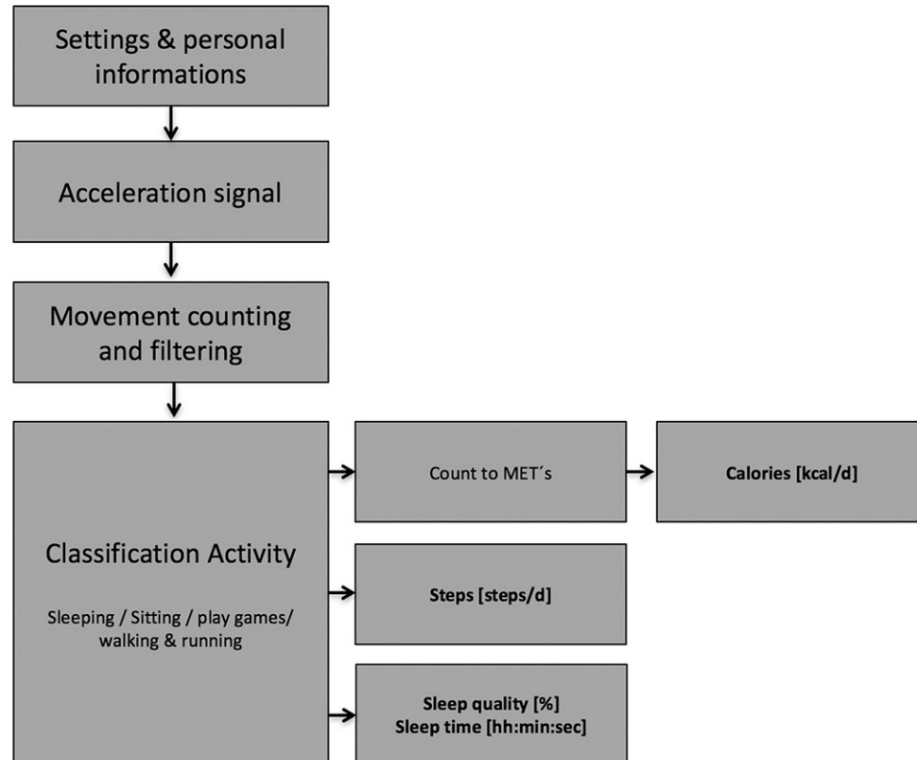


Abbildung 3.9: Schematischer Ablauf bei der Bestimmung des Kalorienverbrauches bei dem Fitness Tracker POLAR A300 (Boeselt et al., 2016)

Um die Ergebnisse der EC-Berechnung zu verbessern, können die Daten unterschiedlicher Sensoren kombiniert werden. Zum Beispiel wird nach Dominick et al. (2016, S. 7) bei dem Tracker FITBIT FLEX die Anzahl der Schritte zur Bestimmung des EC zusätzlich hinzugezogen, wie jedoch das MET berechnet wird, ist nach Alharbi et al. (2016, S. 1479) bei diesem Tracker nicht bekannt. Bei dem Goldstandard ACTIHEART werden die Beschleunigungsdaten mit der Herzfrequenz kombiniert. Bei anderen kommerziellen Fitness Trackern sind diese Berechnungsmethoden durch die Implementierung der proprietären Software nicht bekannt. Die APPLE WATCH 1 bestimmt nach Bai, Hibbing et al. (2017, S. 6) anhand der erkannten Bewegungsart selbstständig, ob die Daten vom Akzelerometer, PPG oder GPS oder einer Kombination der Sensoren zur Berechnung vom EC herangezogen werden. Nach Bai et al. (2016, S. 1) ist neben der APPLE WATCH 1 auch bei der FITBIT CHARGE HR nicht geklärt ob Daten des PPG-Sensors zur EC kombiniert werden.

3.2 Testmethoden in Studien

Die in dieser Arbeit zugrundeliegenden Studienergebnisse der Gültigkeit und Zuverlässigkeit (Abschnitt 3.2.2) von Fitness Trackern sowie deren Messumgebung (Abschnitt 3.2.3), die Labor- und Feldmessung, werden nachfolgend genauer definiert. Die Population der Testsubjekte erstreckt sich von Kindern über Erwachsene bis hin zu Personen mit unterschiedlichen körperlichen Gebrechen.

3.2.1 Gültigkeit

Der Begriff Gültigkeit (englisch: validity) beschreibt, wie gut ein spezifisches Gerät misst, was es messen soll. Die Gültigkeit wird als idealer Zustand betrachtet, der verfolgt, aber niemals erreicht wird (Higgins & Straub, 2006, S. 24). Spezifisch in dieser Arbeit ist die Construct Validity, welche versucht, die von Higgins und Straub (2006, S. 24) formulierte Frage „Is the concept accurately defined and does the instrument or tool actually measure the concept that it is supposed to measure?“ zu beantworten. Die in den Studien behandelten Gültigkeiten haben je nach Fragestellung eine unterschiedliche Ausprägung und können in zwei Gruppen gegliedert werden, die *calibration validity* und *criterion-reference validity*. Eine detaillierte Beschreibung der unterschiedlichen Gültigkeit-Ausprägungen ist in Tabelle 3.1 gelistet.

Gültigkeit bei der Schrittzählung

Bei der Schrittzählung (Abschnitt 5.1) wurden in Labor - und Feldumgebungen beim Vergleich mit einem Goldstandard-Tracker (Abschnitt 1.2) die *concurrent validity* und beim Vergleich mit den direkt gezählten Schritten die *unit calibration* eruiert.

Gültigkeit bei der Pulsmessung

Bei der Pulsmessung (Abschnitt 5.2) wurde gegen Goldstandard-Tracker und gegen EKGs die *concurrent validity* eruiert.

Gültigkeit bei der Kalorienverbrauchsberechnung

Bei der Kalorienverbrauchsberechnung wurde durch die zugrundeliegenden Algorithmen, die unterschiedliche gemessene Werte von den Sensoren in den Kalorienverbrauch umsetzen, die *value calibration* gegen spiroergometrische Messverfahren und die *concurrent validity* gegenüber Goldstandard Akzelerometer eruiert. Als Vorstufe zum Ergebnis der Kalorienverbrauchsberechnung liegt die Erkennung der Aktivität, welche in den Bereich der *predictive validity* fällt.

Tabelle 3.1: Definitionen der einzelnen Begriffe unter dem Sammelbegriff Gültigkeit

Begriff	Definition	Beispiel
Calibration validity		
Unit calibration	accuracy of the measurement being conducted	counting number of steps correctly

Tabelle 3.1: Fortsetzung

Begriff	Definition	Beispiel
Value calibration	accuracy of the conversion of the monitor's measurement into other measurement units	converting steps into energy expenditure
Criterion-referenced validity		
Concurrent validity	validity of an assessment compared with a known assessment	Is the number of steps counted on an activity monitor as accurate as steps counted on an Actigraph?
Predictive validity	extent to which a device can predict an outcome or parameter compared to the predictive gold standard device	Does the average number of weekly steps measured on an activity monitor predict cardiovascular disease risk to the same extent as measuring average steps on an Actigraph?

Tabelle 3.1: Wright et al. (2017, S. 362, zit. n. Bassett, Rowlands & Trost, 2012)

3.2.2 Zuverlässigkeit

Der Begriff Zuverlässigkeit (englisch: reliability) beschreibt, inwieweit eine Messung bei wiederholten Versuchen zu den gleichen Ergebnissen führt. Higgins und Straub (2006, S. 27) schreiben weiter, dass „ähnlich wie die Gültigkeit, die perfekte Zuverlässigkeit ein idealer Zustand ist, der niemals erreicht wird“. Die Bestimmung der Zuverlässigkeit eines Messgerätes verlangt darüber hinaus auch eine weitere Identifizierung der zugrundeliegenden Messkonzepte und statistischen Faktoren. Die Zuverlässigkeit kann weiter in eine Inter- und Intra-Instrumentale-Zuverlässigkeit aufgliedert werden:

Intra-Instrumental-Zuverlässigkeit

„Test-retest results indicating consistency within the same tracker“ (Evenson et al., 2015, S. 3).

Inter-Instrumental-Zuverlässigkeit

„Results indicating consistency across the same brand/type of tracker measured at the same time and worn in the same location“ (Evenson et al., 2015, S. 3)

Zuverlässigkeit ist notwendig, aber nicht hinreichend für die Genauigkeit. So kann ein Fitness Tracker eine hohe Zuverlässigkeit besitzen, da er beispielsweise beim Zählen von 10.000 Schritten bei wiederholten Tests konstant auf das selbige Ergebnis kommt. Wenn nun aber, der für die tägliche Schrittzählung ausgelegte Fitness Tracker, Schritte erst ab einer Beschleunigungseinwirkung bei Geschwindigkeiten über 5 km h^{-1} misst, so hat dieser weiterhin eine hohe Zuverlässigkeit aber eine niedrige Genauigkeit, da dieser Schritte bei

langsamerer Gehgeschwindigkeit (Geschwindigkeit beim Gehen zwischen $4,5 \text{ km h}^{-1}$ und $5,4 \text{ km h}^{-1}$ (Wikipedia Autoren, 2019b)) nicht beachtet hat.

3.2.3 Messumgebung

In dieser Arbeit werden die Studienergebnisse anhand unterschiedlicher unten angeführter Kriterien in die Messumgebungen *Labor* und *Feld* getrennt. Studien mit Labormessungen sind durch die genaue Beschreibung der Bewegung einfach, im Gegensatz von Feldmessungen, zu reproduzieren, wohingegen Feldmessungen nach Evenson et al. (2015, S. 20) besser geeignet sind, die Gültigkeit von Fitness Trackern zu bestimmen, da durch Alltagsbewegungen bedeutende Bewegungen im oberen Körperbereich auftreten, die vom Fitness Tracker nicht genau genug aufgenommen werden oder fehlerhafte Messungen erzeugen können.

Labormessungen

Zu den Labormessungen zählen die Aktivitäten, die anhand der Dokumentation der Studie eindeutig reproduziert werden können. Diese sind Bewegungen auf einem Laufband (aufgrund der vorgegebenen Geschwindigkeiten durch das Laufband) oder isolierte vorgegebene Bewegungsmuster, wie z.B. eine bestimmte Anzahl von Stufen zu steigen oder ein 2-Minute-Walk-Test. Laufen abseits eines Laufbandes wird auch zu den Labormessungen gezählt, sofern diese Testungen mit einer konstant vorgegebenen und dokumentierten Geschwindigkeit durchgeführt wurde. Wohingegen Laufen abseits eines Laufbandes und innerhalb eines Geschwindigkeitsbereiches bereits zu den Feldmessungen gezählt wird.

Feldmessungen

Zu den Feldmessungen werden Messungen im Alltag und auch Studien, die versuchen, den Alltag durch künstlich erschaffenen Umgebungen, zu simulieren, gezählt. Ebenfalls werden hier die Messungen hinzugenommen, die wegen mangelnder Dokumentation oder aufgrund der von den Testpersonen selbst entschiedenen Geschwindigkeiten, nicht klar reproduzierbar sind.

4 Aktueller Stand der Forschung

In diesem Kapitel werden die am Beginn der Arbeit (Oktober 2017) festgestellten Daten zur Genauigkeit, Gültigkeit und Zuverlässigkeit in den Mess- und Berechnungsbereichen der Fitness Tracker (Schrittzählung, Pulsmessung und Kalorienverbrauchsberechnung), aus einem systematischen Review (Evenson et al., 2015) und weiteren Studien (Ainsworth et al., 2015; Banerjee, Peterson, Oliver, Froehle und Lawhorne, 2017; Bassett, 2012; Case, Burwick, Volpp und Patel, 2015; Delgado-Gonzalo et al., 2015; Diaz et al., 2015; Diaz et al., 2016; Dominick et al., 2016; Lauritzen, Muñoz, Luis Sevillano und Civit, 2013; Stackpool, 2013; Takacs et al., 2014), dargestellt. Am Ende dieses Kapitels wird anhand von Evenson et al. (2015) aufgezeigt, warum zum damaligen Zeitpunkt nur eine begrenzte Auswahl von Fitness Trackern unterschiedlicher Hersteller zur Testung auf ihre Mess- und Berechnungsgültigkeit verfügbar war.

Nach Bassett (2012, S. 1770) messen Schrittzähler, die sich die Piezoeffekte zunutze machen, genauer als federgelagerte Pedometer. Nach Ainsworth et al. (2015, S. 390) funktionieren MEMS-basierte Schrittzähler über $0,9 \text{ ms}^{-1}$ sehr genau. Banerjee et al. (2017, S. 2) schreiben außerdem, dass mit steigender Geschwindigkeit die Messgenauigkeit steigt, was für sie nicht verwunderlich ist, da Wearables vornehmlich als Fitness-Geräte konzipiert sind. Ainsworth et al. (2015, S. 390) schreiben, dass, obwohl Schrittzähler für einen Einsatz an der Hüfte entwickelt wurden, diese am Sprunggelenk am genauesten messen. Evenson et al. (2015) kommen in ihrem Review, mit inkludierten 22 Studien, bezüglich Gültigkeit und Zuverlässig zu dem Schluss, dass FITBIT und JAWBONE Fitness Tracker eine generell hohe Gültigkeit bei der Schrittzählung am Laufband im Laborumfeld haben. Im Vergleich zum direkten Zählen der Schritte oder im Vergleich mit Akzelerometern weisen die Fitness Tracker FITBIT ULTRA, ONE und ZIP und JAWBONE UP generell eine hohe Korrelation ($r \geq 0,8$) beim Gehen auf dem Laufband auf, jedoch nicht beim Laufen oder beim koordinativem Training der Reaktionsschnelligkeit. Andere Studien beschreiben, dass die Tracker FITBIT ONE und FLEX sowie die JAWBONE UP beim Gehen und Laufen am Laufband zu wenig Schritte zählen, wohingegen der Tracker FITBIT ONE, in einer zweitägigen Feldstudie beim Vergleich mit einem Akzelerometer, zu viele Schritte zählte (zit. n. Diaz et al., 2015; Lauritzen et al., 2013; Case et al., 2015; Stackpool, 2013; Takacs et al., 2014). Auch Dominick et al. (2016, S. 7) kommen zu dem Ergebnis, dass die FITBIT FLEX beim Gehen zu wenig Schritte zählt gegenüber den von den Forschern direkt gezählten Schritten. Weiters schreiben sie, dass FITBIT-Tracker ebenfalls eine hohe Zuverlässigkeit aufweisen.

4 Aktueller Stand der Forschung

Evenson et al. (2015, zit. n. Fitbit Inc., 2015) betonen, dass FITBIT-Tracker eine Genauigkeit von 95-97 % beim Schrittzählen haben sollten, wenn sie laut der Empfehlung des Herstellers getragen werden. Evenson et al. (2015, zit. n. Jawbone, 2015) erwähnen, dass die Fitness Tracker von JAWBONE, in den für die am Markt befindlichen kommerziellen Trackern, eine führende Gültigkeit besitzen.

Delgado-Gonzalo et al. (2015, S. 430) schreiben, dass nicht näher bestimmte damalige Fitness Tracker mit der Möglichkeit der Herzfrequenzmessung eine komfortable Alternative bei der Bestimmung des Pulses während des Trainings darstellen, jedoch aber an „mangelnder Genauigkeit und Zuverlässigkeit leiden“.

Bei der Kalorienverbrauchsrechnung kommen Diaz et al. (2016, S. 4) mit Tests von der FITBIT FLEX gegenüber indirekter Kalorimetrie zu dem Ergebnis, dass dieser Fitness Tracker eine Fehlerrate 24,5-83,4 % aufweist.

Die Fitness Tracker von FITBIT und JAWBONE wurden von Evenson et al. (2015, S. 20) gewählt, da sie die Mehrheit der am Markt erhältlichen Geräte repräsentierten. Ein Jahr später, im ersten Quartal 2015, zeigten Verkaufsstatistiken, dass nun gleich nach den Fitness Tracker der Firma FITBIT (34 %) die Tracker von XIAOMI (25 %) kamen und erst weiter dahinter, die Tracker von GARMIN (6 %), SAMSUNG (5 %) und JAWBONE (4 %). Evenson et al. (2015, S. 20) zeigen somit die Möglichkeit eines sich schnell verändernden Markt der Fitness Tracker auf. Hinzu kommt, dass die Populationen bei Studien zu Fitness Trackern hauptsächlich aus Erwachsenen mit normalem Body Mass Index (BMI) bestehen. Bei Personen mit höherem BMI soll es bereits zu ungenaueren Messergebnissen kommen.

5 Ergebnisse

Es wurden 2759 Artikel in den Meta-Datenbanken PubMed und Web of Science sowie dem Online-Portal des Elsevier-Verlags und 23 Artikel aus anderen Quellen gefunden und zusammengefügt. Nach dem maschinell unterstützten Entfernen der Duplikate (siehe Abschnitt 2.2) verblieben 2609 Artikel.

Das Flussdiagramm von Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Abb. 5.1) zeigt die verschiedenen Phasen der Artikelauswahl für diese Arbeit. In der Vorauswahl wurden 2118 Titel als nicht geeignet empfunden. Von den übriggebliebenen 491 Artikel wurden die Zusammenfassungen von 275 Artikel als ungeeignet eingestuft und sechs Artikel waren in den Onlinequellen nicht mehr vorhanden. Bei den übrigen 210 Artikeln wurden die Volltexte auf Eignung überprüft. Bei den verworfenen Artikeln waren 38 Volltexte weder über das Netzwerk der Universität Wien noch der Technischen Universität Wien im Internet abrufbar, 31 entsprachen nicht den geforderten Kriterien (enthielten vornehmlich keine Studienergebnisse über Wearables im Konsumentenbereich) und 41 Artikel enthielten keine Fitness Tracker, die am Handgelenk getragen werden können. Einige Fitness Tracker (z.B.: FITBIT ULTRA) können wie von den Herstellern vorgesehen an einem Kleidungsstück in der Nähe der Hüfte befestigt oder mit einer extra erhältlichen Armbandhalterung am Handgelenk getragen werden. Diese galt es hier genauer anhand der Tragebeschreibung in den Studien zu trennen, um ausschließlich die am Handgelenk getragenen Tracker zu untersuchen. Weiters waren 33 Artikel nur Abstracts von Tagungen, die im Journal Medicine & Science in Sports & Exercise veröffentlicht wurden und für welche keine ausführlicheren Studiendaten vorhanden waren. Schlussendlich blieben 67 Artikel übrig, die in diese Arbeit eingeschlossen wurden und von denen weiters 33 in die Metaanalyse eingeflossen sind (Tabelle 5.1).

Tabelle 5.1 zeigt die 33 Studien, welche in die Metaanalyse eingeflossen sind. Diese Studien behandeln teilweise mehrere Messsysteme, sodass die Summe der verfügbaren Studien anhand der einzelnen Messsysteme mehr ergeben als die Anzahl der Studien in der Meta-Analyse. Die Schrittzählung wird von 23 Studien untersucht, die Pulsmessung von sieben und die Kalorienverbrauchsberechnung von 17. Die Tabelle zeigt neben den Autoren und dem Ort der Studiendurchführung bzw. der Veröffentlichung der Studienergebnisse noch folgende Informationen:

- die Messgruppen Schritte, Puls, EC und Physical Activity (PA) zusammengefasst für den Kalorienverbrauch (siehe Abschnitt 3.2),

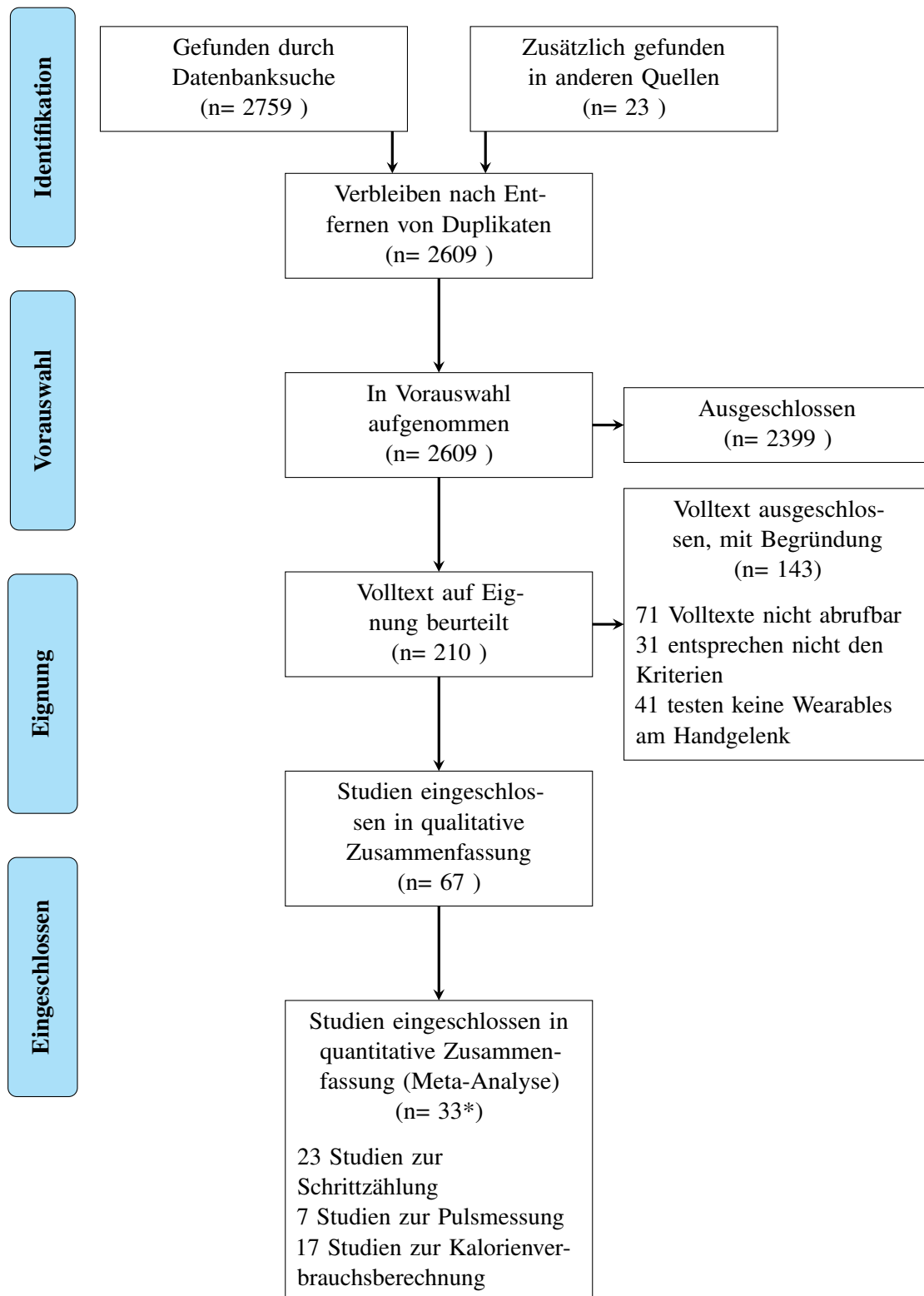


Abbildung 5.1: PRISMA Flussdiagramm

*Nachstehende Auflistung beinhaltet auch Studien, die mehrere Messsysteme behandeln.

- die Art der in der Studie verwendeten Testung, ob es sich um Zuverlässigkeit- oder Gültigkeitsstudien (R bzw. V) handelt,
- die Testumgebung Feld oder Labor (Abschnitt 3.2.3),
- die Größe der Population, geteilt auf männliche und weibliche Personen,
- das durchschnittliche Alter mit Standardabweichung (SD),
- der durchschnittliche BMI mit SD und
- die Kriterien der Teilnehmer und Teilnehmerinnen für die Aufnahme in die jeweilige Studie.

Die Güte, der in der Metaanalyse vorkommenden Studien, wird in Tabelle 5.2 durch ein Häkchen ✓ in der betreffenden Spalte des Qualitätskriteriums (siehe Abschnitt 2.3) dargelegt. In der Tabelle nicht dargestellt sind die Kriterien *randomisierte kontrollierte Studie* (engl.: *Randomized Controlled Trial*) (RCT) und *Demographie*, da diese von allen abgebildeten Studien erfüllt wurden. Weiters fehlt auch das Kriterium, ob die Störgrößen in den Ergebnissen interpretiert worden sind, da dieses von allen Artikeln, bis auf den von Diaz et al. (2016), beachtet wurde.

In der Metaanalyse kommen 31 unterschiedliche Fitness Tracker der Unternehmen APPLE, INC., BASIS SCIENCE, INC., BEURER GMBH, FITBIT, INC., GARMIN LTD., JAWBONE, XIAOMI CORPORATION, MICROSOFT CORPORATION, MEDISANA AG, MISFIT, INC., NIKE, INC., POLAR ELECTRO OY, SAMSUNG GROUP, TOMTOM N.N. und WITHINGS vor (Tabelle 5.3). Die ersten Wearables davon kamen 2011 auf den Markt, die letzten 2016. Der Verkauf von BASIS B1 BAND und PEAK, FITBIT FORCE und SURGE, NIKE+ FUELBAND und FUELBAND SE sowie POLAR LOOP wurden bereits aus unterschiedlichen Gründen eingestellt, manche davon auf Grund von Rückrufaktionen durch die Hersteller.

Bis auf zwei (MOVBAND MODEL 2 und TOMTOM RUNNER CARDIO) besitzen alle Fitness Tracker Beschleunigungssensoren, manche davon wurden von den Herstellern extra als Triaxialakzelerometer ausgewiesen. Ein Display oder eine andere visuelle Schnittstelle (LEDs) besitzen alle bis auf FITBIT FLEX, JAWBONE UP MOVE und UP24, MISFIT SHINE und NIKE+ FUELBAND SE, welche die persönlichen Messdaten dadurch nur in Kombination mit einem Smartphone und der Herstellerapp wiedergeben können. Elf Wearables können die Herzfrequenz mittels PPG messen und 21 von den 31 können die verbrauchten Kalorien bestimmen.

In Tabelle 5.3 nicht gelistet, sind eine für zumindest Schwimmen geeignete Wasserdichtheit (BASIS B1 BAND, GARMIN FORERUNNER 920XT, VIVOACTIVE, VIVOFIT und VIVOSMART, MIO ALPHA, MISFIT SHINE, POLAR A300, ACTIVE und V800 und TOMTOM RUNNER CARDIO) und die Verwendung eines globalen Navigationssatellitensystems (engl.: Global Navigation Satellite System) (GNSS) wie z.B.: GPS (BEURER AS80, FITBIT SUR-

5 Ergebnisse

GE, GARMIN FORERUNNER 920XT und VIVOACTIVE, MICROSOFT BAND und BAND 2, POLAR V800, SAMSUNG GEAR S und TOMTOM RUNNER CARDIO). Die Akkulaufzeiten der aufgenommenen Tracker reichen laut Herstellerangaben von 8 h (TOMTOM RUNNER CARDIO) bis 5 Jahre (MIO ALPHA), wobei der Großteil eine Laufzeit im Bereich von wenigen Tagen (4 d bis 8 d) besitzt.

Tabelle 5.1: In dieser Metaanalyse inkludierte Fitness Tracker Studien

Autor	Ort	Messung	Studienart	Testumgebung	Teilnehmer (männ- lich/weiblich)	Alter (SD) in Jahren	BMI (SD)	Kriterien
Abt et al. (2017)	Kingston upon Hull, UK	Puls	V/R	Labor	8/7	32 (10)	175 (8) cm, 73,5 (14,8) kg*	gesund
Alharbi et al. (2016)	Sydney, Australien	PA, Schritte	V	Labor	25/23	65,6 (6,9)	27,9 (2,4)	CHD bei 58,9% und Knochen bzw. Gelenkprobleme: 39,6%
An et al. (2017)	Nebraska, USA	Schritte	V	Labor, Feld	17/18	31 (11,8)	23,8 (3,1)	gesund
Bai et al. (2016)	Iowa, USA	EC	V	Labor	28/24	16-65	24 (17,6-39,9)	gesund
Bai, Hibbing et al. (2017)	Iowa, USA	EC, Puls, Schritte	V	Labor	23/18	32 (11)	24,7 (4)	gesund
Battenberg et al. (2017)	Los Angeles, Kalifornien, USA	Schritte	V	Labor	18/12	25,6 (2,5)	23,5 (2,9)	gesund
Boeselt et al. (2016)	Marburg, Deutschland	EC, Schritte	V	Feld	17/3	66,4 (7,4)	28,9 (5,4)	COPD
Brooke et al. (2017)	Nebraska, USA	EC	V	Feld	34/61	28,5 (9,9)	25,7 (3,4)	gesund
Diaz et al. (2016)	New York, USA	EC, Schritte	V	Labor	13/13	32 (9,2) / 32 (9,2)	24,2 (3,4) / 24,2 (3,4)	gesund
Dominick et al. (2016)	Delaware, USA	MET, PA, Schritte	V	Feld	4/15	20 (0,8) / 21,3 (4,4)	24,6 (1,0) / 21,9 (2,7)	gesund
Floegel et al. (2017)	Oxford, UK	Schritte	V	Labor	28/71	78,9 (8,6)	28,0 (6,5)	ab Ältere (Alter>62)

Tabelle 5.1: Fortsetzung

Autor	Ort	Messung	Studienart	Testumgebung	Teilnehmer (m/w)	Alter (SD)	BMI (SD)	Kriterien
Fulk et al. (2014)	USA	Schritte	V	Labor	34/16	52,9 (15,1)	n.A.	Schlaganfallspatienten: eingeschränkt gehfähig und besser gesund
Gorny et al. (2017)	Singapur, Singapur	Puls	V	Feld	7/3	25,4 (3,7)	22,9 (3,8)	gesund
Gruwez et al. (2017)	Brüssel, Belgien	EC, Schritte	V	Feld	7/13	30 (5)	n.A.	gesund
Hernández- Vicente et al. (2016)	Spanien	EC, PA (MVPA), Schritte	V	Feld	9/9	21 (1,24)	22,25 (1,75)	gesund
Jo et al. (2016)	Kalifornien, USA	Puls	V	Labor	12/12	24,8 (2,1)	166 (7) cm, 71,3 (9,9) kg*	gesund
Kooiman et al. (2015)	Groningen, Niederlande	Schritte	V/R	Labor, Feld	16 [18]/17 [38]**	39 (12,1) [37,1 (10,6)] / 25 (11,2) [30 (9,5)]	23,6 (2,2) [24,1 (2)] / 22,5 (2,1) [23,1 (2,5)]	gesund
Lauritzen et al. (2013)	Sevilla, Spanien	Schritte	V	Labor	1 - 1 - 6/4 - 6 - 0****	87,6 (3,91) - 84,14 (3,67) - 35,33 (6,53)	n.A.	Ältere mit, ohne Einschränkung und gesunde
J.-M. Lee et al. (2014)	Nebraska, USA	EC	V	Labor	30/30	28,6 (6,4) / 24,2 (4,7)	24,3 (2,6) / 21,8 (2,7)	gesund

Tabelle 5.1: Fortsetzung

Autor	Ort	Messung	Studienart	Testumgebung	Teilnehmer (m/w)	Alter (SD)	BMI (SD)	Kriterien
J. A. Lee et al. (2015)	Normal, USA	Schritte	V	Labor, Feld	22 [19]/21 [18]**	20,9 (1,9) [12,6 (2,8)] / 20,9 (2,1) [21,1 (2,5)]	25,5 (2,7) [25,5 (2,6)] / 23,1 (2,9) [23,2 (2,3)]	n.A.
Montoye et al. (2017)	Alma, USA	EC, Puls, Schritte	V	Labor	18/14	21,6 (1,1) / 26,1 (2,6)	25,6 (0,8) / 24,2 (0,6)	gesund
Reid et al. (2017)	Montreal, Kanada	PA, Schritte	V	Feld	-/22	- / 21,23 (1,63)	- / 22,35 (2,34)	gesund
Simunek et al. (2016)	Olmütz, Tschechien	Schritte	V	Feld	14/6	34 (6,3)	24,3 (4,0)	keine Mobilitätseinschränkung
Sirard et al. (2017)	Massachusetts, USA	EC, PA (MVPA)	V	Labor, Feld	7 [8]/7 [8]**	9 (2) / 8,6 (1,6)	135,3 (13,57)/133,9 (12,2) cm, 36,1 (10,6)/32,2 (8,4) kg*	gesund
Stahl et al. (2016)	Nabraska, USA	Puls	V	Labor	32/18	27,47 (6,1) / 24,17 (4,93)	25,44 (2,6) / 23,36 (3,5)	gesund
Sushames et al. (2016)	Cairns, Australien	Schritte	V/R	Labor, Feld	13/12	23,7 (5,8)	27,7 (6,1) / 22,7 (2,2)	gesund
Tam und Cheung (2018)	Hong Kong, China	Schritte	V	Labor	15/15	32,1 (8,66)	n.A.	gesund
Tophøj et al. (2018)	Aarhus, Dänemark	Schritte	R	Labor	10/10	25,6 (2)	n.A.	gesund

Tabelle 5.1: Fortsetzung

Autor	Ort	Messung	Studienart	Testumgebung	Teilnehmer (m/w)	Alter (SD)	BMI (SD)	Kriterien
Treacy et al. (2017)	Sydney, Australien	Schritte	V	Labor	91/75	80 (11)	n.A.	Gehgeschwindigkeit < 1,2m/s
Tucker et al. (2015)	Arizona, USA	EC	V	Labor	11/13	28,4 (7,8)	23,8 (3,9)	gesund
Voss et al. (2017)	Vancouver, Kanada	Schritte	V	Feld	14/16	12,8 (2,1) / 13,2 (2,3)	57,7 (39,1) / 63,6 (30,3)****	CHD, Herztransplantation, elektrophysiologische Störung
Wahl et al. (2017)	Köln, Deutschland	EC, Schritte	V	Labor	10/10	26,1 (2,8) / 24,2 (1,9)	182,3 (7,4) cm, 81,1 (11,2) kg*	gesund
Wallen et al. (2016)	Brisbane, Australien	EC, Puls	V/R	Labor	11/11	24,9 (5,6)	173,1 (9,9) cm, 72,7 (11,8) kg*	gesund

Tabelle 5.1: Auflistung der Studien die in die vorliegende Arbeit einbezogen wurden. Alter in Jahren im Mittel, BMI im Mittel.

Labor: Test unter Laborbedingungen; Feld: Test unter Feldbedingungen.

Abkürzungen: Body Mass Index (BMI), koronare Herzkrankheit (engl.: Coronary Heart Disease) (CHD), chronisch obstruktive Lungenerkrankung (engl.: Chronic Obstructive Pulmonary Disease) (COPD), Energieverbrauch (engl.: Energy Consumption) (EC), moderate bis intensive Bewegungsformen (engl.: Moderate to Vigorous PA) (MVPA), Physical Activity (PA), R Zuverlässigkeit (Reliability), Standardabweichung (SD), V Gültigkeit (Validity)

*Kein BMI, Angabe in Größe [cm] und Gewicht[kg] **Teilnehmeranzahl bei Labor und [Feld]-Messungen ***Teilnehmeranzahl von Älteren mit reduzierter Mobilität - Älteren mit normaler Mobilität - gesunden Erwachsenen ****Perzentile anstelle der Standardabweichung

Tabelle 5.2: Qualität der in diesem Review inkludierten Studien

Studie	Studiendesign					Population		Hardware				Datenauswertung			Diskussion								
	Generation	Firmware	Anzahl	Test-Dauer	Anzahl verwendet	Zeitraumen	Ort	Randomisierte Population	Anwerbung	Eignung	Herstellerangabe	Randomisierte Geräte	Arndominanz	Tragedauer*	Einschulung*	Trageunterbrechungen*	Software	Unbrauchbare Daten	Methode der Auswertung	Verblindete Auswerter	Störgrößen	Schlussfolgerung	Förderer***
Abt et al. (2017)	✓	✓		✓	✓		✓		✓	✓						✓	✓	✓			✓	✓	✓
Alharbi et al. (2016)	✓		✓	✓	✓	✓	✓		✓	✓							✓	✓			✓	✓	✓
An et al. (2017)**	✓				✓		✓		✓	✓	✓		✓	✓	✓			✓		✓	✓	✓	✓
Bai et al. (2016)	✓				✓			✓	✓	✓			✓				✓	✓		✓	✓	✓	✓
Bai, Hibbing et al. (2017)	✓				✓			✓	✓	✓							✓	✓		✓	✓	✓	✓
Battenberg et al. (2017)	✓				✓															✓	✓	✓	
Boeselt et al. (2016)**	✓				✓				✓	✓			✓			✓		✓			✓	✓	✓
Brooke et al. (2017)**	✓						✓	✓	✓	✓			✓	✓	✓			✓			✓	✓	✓
Diaz et al. (2016)	✓							✓	✓	✓													
Dominick et al. (2016)**	✓					✓	✓	✓	✓			✓	✓	✓		✓	✓	✓		✓	✓	✓	✓
Floegel et al. (2017)	✓					✓		✓	✓	✓		✓				✓		✓		✓	✓	✓	✓
Fulk et al. (2014)	✓						✓			✓							✓				✓	✓	
Gorny et al. (2017)**	✓					✓			✓	✓	✓	✓	✓	✓		✓		✓				✓	✓
Gruwez et al. (2017)**	✓									✓		✓	✓			✓	✓	✓		✓		✓	✓
Hernández-Vicente et al. (2016)**	✓					✓				✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Jo et al. (2016)	✓							✓	✓	✓	✓	✓						✓		✓	✓	✓	✓
Kooiman et al. (2015)**	✓						✓	✓	✓	✓		✓	✓				✓	✓		✓	✓	✓	✓

Tabelle 5.2: Fortsetzung

Studie	Studiendesign				Population		Hardware				Datenauswertung			Diskussion										
	Generation	Firmware	Anzahl	Test-Dauer	Anzahl verwendet	Zeitraumen	Ort	Randomisierte Population	Anwerbung	Eignung	Herstellerangabe	Randomisierte Geräte	Armdominanz	Tragedauer*	Einschulung*	Trageunterbrechungen*	Software	Unbrauchbare Daten	Methode der Auswertung	Verblindete Auswerter	Störgrößen	Schlussfolgerung	Förderer***	
Lauritzen et al. (2013)	✓						✓	✓	✓	✓		✓						✓				✓	✓	✓
J.-M. Lee et al. (2014)	✓						✓	✓	✓	✓		✓						✓		✓		✓	✓	✓
J. A. Lee et al. (2015)**	✓													✓		✓	✓	✓				✓	✓	✓
Montoye et al. (2017)	✓							✓	✓			✓				✓	✓	✓				✓	✓	✓
Reid et al. (2017)	✓							✓	✓	✓			✓		✓	✓	✓	✓					✓	✓
Simunek et al. (2016)**	✓							✓	✓	✓		✓	✓		✓	✓	✓	✓					✓	✓
Sirard et al. (2017)	✓							✓	✓	✓		✓	✓	✓	✓			✓				✓	✓	✓
Stahl et al. (2016)	✓						✓			✓	✓	✓						✓				✓	✓	✓
Sushames et al. (2016)	✓						✓	✓	✓	✓	✓		✓				✓	✓					✓	✓
Tam und Cheung (2018)	✓		✓		✓	✓				✓		✓				✓							✓	✓
Tophøj et al. (2018)	✓						✓			✓		✓				✓		✓				✓	✓	✓
Treacy et al. (2017)	✓						✓	✓	✓	✓						✓	✓	✓					✓	✓
Tucker et al. (2015)	✓							✓	✓	✓						✓		✓				✓	✓	✓
Voss et al. (2017)	✓						✓	✓	✓	✓			✓			✓		✓		✓			✓	✓
Wahl et al. (2017)	✓									✓						✓		✓				✓	✓	✓
Wallen et al. (2016)										✓	✓					✓	✓	✓				✓	✓	✓

Tabelle 5.2: Kennzeichnung der Studien in den unterschiedlichen Qualitätsmerkmalen.

*Diese Kriterien sind nur bei Studien mit Feldtest relevant. **Studien mit Feldtest ***Studien sind bei Förder nur dann mit ✓ gekennzeichnet, wenn laut den Autoren keine Förderung stattgefunden hat.

Tabelle 5.3: In dieser Metaanalyse vorkommende kommerzielle Fitness Tracker

Hersteller	Tracker	Verkaufsjahre	Beschleunigungssensor	Pulssensor	Kalorienverbrauch	Display	Studien
Apple, Inc.	APPLE WATCH 1*	ab 2014	✓Triaxial	✓(2x grün)	✗	✓	Abt et al. (2017), Bai, Hibbing et al. (2017), Wallen et al. (2016)
Basis Science, Inc.	BASIS B1 BAND	2013 - 2016**	✓Triaxial	✓(2x grün)	✓	✓	An et al. (2017), J.-M. Lee et al. (2014)
Basis Science, Inc.	BASIS PEAK	2015 - 2016**	✓	✓(2x grün)	✓	✓	Jo et al. (2016), Stahl et al. (2016)
Beurer GmbH	BEURER AS80	n.A.	✓	✗	✓	✓	Wahl et al. (2017)
Fitbit, Inc.	FITBIT CHARGE	ab 2014	✓Triaxial	✗	✗	✓	Treacy et al. (2017), Wahl et al. (2017)
Fitbit, Inc.	FITBIT CHARGE HR	ab 2015	✓Triaxial	✓	✓	✓	Bai, Hibbing et al. (2017), Brooke et al. (2017), Gorny et al. (2017), Jo et al. (2016), Montoye et al. (2017), Stahl et al. (2016), Tam und Cheung (2018), Tophøj et al. (2018), Voss et al. (2017), Wahl et al. (2017), Wallen et al. (2016)
Fitbit, Inc.	FITBIT FLEX	ab 2013	✓Triaxial	✗	✓	✗	Alharbi et al. (2016), An et al. (2017), Bai et al. (2016), Brooke et al. (2017), Diaz et al. (2016), Dominick et al. (2016), Floegel et al. (2017), Kooiman et al. (2015), Reid et al. (2017), Sushames et al. (2016)
Fitbit, Inc.	FITBIT FORCE	2013 - 2014***	✓Triaxial	✗	✓	✓	Battenberg et al. (2017)

Tabelle 5.3: Fortsetzung

Hersteller	Tracker	Verkaufsjahre	Beschleunigungssensor	Pulssensor	Kalorienverbrauch	Display	Studien
Fitbit, Inc.	FITBIT SURGE	2015 - 2017	✓	✓	✗	✓	Tophøj et al. (2018)
Fitbit, Inc.	FITBIT ULTRA	ab 2011	✓	✗	✗	✓	Lauritzen et al. (2013)
Garmin Ltd.	GARMIN FORERUNNER 920XT	ab 2014	✓Triaxial	✗	✓	✓	Wahl et al. (2017)
Garmin Ltd.	GARMIN VIVOACTIVE	ab 2015	✓Triaxial	✗	✓	✓	Wahl et al. (2017)
Garmin Ltd.	GARMIN VIVOFIT	ab 2014	✓Triaxial	✗	✓	✓	An et al. (2017), Brooke et al. (2017), Simunek et al. (2016), Treacy et al. (2017), Wahl et al. (2017)
Garmin Ltd.	GARMIN VIVOSMART	ab 2014	✓Triaxial	✗	✓	✓	Wahl et al. (2017)
Jawbone	JAWBONE UP MOVE	ab 2014	✓Triaxial	✗	✓	✗	Gruwez et al. (2017)
Jawbone	JAWBONE UP24	ab 2013	✓Triaxial	✗	✓	✗	An et al. (2017), Bai et al. (2016), Brooke et al. (2017)
Xiaomi Corporation	MI BAND 2	ab 2016	✓	✓	✗	✓	Tam und Cheung (2018)
Microsoft Corporation	MICROSOFT BAND	ab 2014	✓Triaxial	✓	✗	✓	Stahl et al. (2016)
Microsoft Corporation	MICROSOFT BAND 2	ab 2015	✓Triaxial	✓	✗	✓	Tophøj et al. (2018)
Medisana AG	MIO ALPHA	n.A.	✓	✓(2x)	✗	✓	Stahl et al. (2016), Wallen et al. (2016)

Tabelle 5.3: Fortsetzung

Hersteller	Tracker	Verkaufsjahre	Beschleunigungssensor	Pulssensor	Kalorienverbrauch	Display	Studien
Misfit, Inc.	MISFIT SHINE	ab 2013	✓Triaxial	✗	✓	✗	An et al. (2017), Bai et al. (2016), Brooke et al. (2017)
n.A.	MOVBAND MODEL 2****	n.A.	✗	✗	✓	✓	Sirard et al. (2017)
Nike, Inc.	NIKE+ FUELBAND	2012 - 2015*****	✓Triaxial	✗	✓	✓	Fulk et al. (2014), J.-M. Lee et al. (2014), Tucker et al. (2015)
Nike, Inc.	NIKE+ FUELBAND SE	2013 - 2015*****	✓Triaxial	✗	✓	✗	An et al. (2017), Bai et al. (2016), Battenberg et al. (2017), Brooke et al. (2017), Kooiman et al. (2015)
Polar Electro Oy	POLAR A300	ab 2015	✓	✗	✓	✓	Boeselt et al. (2016)
Polar Electro Oy	POLAR ACTIVE	n.A.	✓Uniaxial	✗	✓	✓	J. A. Lee et al. (2015)
Polar Electro Oy	POLAR LOOP	2015 - 2017	✓	✗	✓	✓	An et al. (2017), Brooke et al. (2017), Simunek et al. (2016), Wahl et al. (2017)
Polar Electro Oy	POLAR V800	ab 2014	✓	✗	✓	✓	Hernández-Vicente et al. (2016)
Samsung Group	SAMSUNG GEAR S	ab 2014	✓	✗	✗	✓	Wallen et al. (2016)
TomTom N.N.	TOMTOM RUNNER CARDIO	n.A.	✗	✓	✗	✓	Stahl et al. (2016)

Tabelle 5.3: Fortsetzung

Hersteller	Tracker	Verkaufsjahre	Beschleunigungssensor	Pulssensor	Kalorienverbrauch	Display	Studien
Withings	WITHINGS PULSE O2	ab 2013	✓Triaxial	✓	✓	✓	Gruwez et al. (2017)

Tabelle 5.3: Spezifikationen der in dieser Arbeit vorkommenden Fitness Tracker, mit Angabe der sie enthaltenen Studien. Die Spalte *Beschleunigungssensor* und *Pulssensor* geben, wenn bekannt, die Art des Akzelerometers bzw. die Anzahl und Lichtfarbe der PPG LEDs an.

Unbekannt ob APPLE WATCH 1ST GENERATION oder Apple WATCH SERIES 1 (2nd Generation), da eine Umbenennung dieser Tracker nach Veröffentlichung der Studie stattfand. **Rückruf: www.intel.com/content/www/us/en/support/articles/000025310/emerging-technologies/wearable-devices.html ***Rückruf: www.fitbit.com/at/forcesupport ****Keine genaueren Spezifikationen, nachdem Hersteller MOVABLE von DHS GROUP übernommen wurde. *****Rückruf: www.nike.com/help/a/why-cant-i-sync

Die Darstellung der Ergebnisse der zwei Messsysteme Schrittzählung (Abschnitt 5.1) und Pulsmessung (Abschnitt 5.2) sowie die Ergebnisse der Kalorienverbrauchsberechnung (Abschnitt 5.3), erfolgt in der jeweiligen Kategorie durch einen Forest-Plot. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

5.1 Schrittzählung

Das Zählen der Schritte ist, wie Tabelle 5.1 auflistet, in 23 Studien behandelt worden. Die Gültigkeit wurde davon in 20 Studien behandelt, deren Ergebnisse in Abschnitt 5.1.1 zusammengefasst werden. Eine Studie, Reid et al. (2017) wurde aus der Metaanalyse entnommen, da sie bei ihrem Ergebnis einen p-Wert von 0,376 besitzt. Aus der Studie von Dominick et al. (2016) wurden zwei Fälle übernommen und sechs gestrichen, da diese sechs einen p-Wert zwischen 0,108 und 0,619 besitzen. Die Zuverlässigkeit wurde aus 3 Studien zusammengefasst, deren Ergebnisse in Abschnitt 5.1.2 ersichtlich sind.

5.1.1 Gültigkeit der Schrittzählung

Für die Metaanalyse sind aus 20 Studien die Ergebnisse zu 123 unterschiedlichen Fällen (unterschiedliche Fitness Tracker und verschiedene Aktivitäten) in Tabelle 5.4 zusammengefasst. Die Metaanalyse trennt dabei zwischen Labortestungen (Abschnitt 5.1.1) und bei den Feldtestungen (Abschnitt 5.1.1) zwischen *unit calibration* und *concurrent validity*.

Neben manuellem Mitzählen von einer Videoaufnahme oder während des Tests, kamen bei der *concurrent validity* die Geräte ACTIGRAPH ACTITRAINER, GT3X und GT3X+, AKTIVPAL, BODYMEDIA SENSEWEAR und SENSEWEAR PRO ARMBAND, NEW LIFESTYLE NL-1000, OMRON HJ 323U PEDOMETER, OPTOGAIT, ORTHOCARE STEPWATCH und YAMAX SW-200 und SW-701 PEDOMETER als Goldstandard zum Einsatz.

Labormessung

Bei den Labormessungen für die Gültigkeit der Schrittzählung wurden anhand der 67 Fälle ein hoher Zusammenhang ($r = 0,909$ (95 % unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI): 0,855; oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI): 0,944)) bei einer Heterogenität der Fälle von $I^2 = 96,34$ ($P < 0,0001$) festgestellt (Forest-Plot: Abb. 5.2). Es wurden bis auf zwei Studien (4 Fälle) die von den Fitness Trackern gezählten Schritte mit direkt gezählten verglichen. Bei einem reinen Vergleich der *unit calibration* (ohne Bai, Hibbing et al. (2017) und Kooiman et al. (2015)), ohne Fälle mit Aktivitäten im Bereich

5 Ergebnisse

des Gehens oder langsamer ($<5,4 \text{ km h}^{-1}$), verändert sich der Koeffizient auf $r = 0,941$ (95 % LCI: 0,9; UCI: 0,965) und $I^2 = 93,73$ ($P < 0,0001$) für die übriggebliebenen 42 Fälle.

Die schlechteste Gültigkeit bei den Fitness Trackern im Zusammenhang mit der Schrittzählung besitzt die POLAR LOOP, welche bei einem Test am Laufband ($7,2 \text{ km h}^{-1}$) für 5 min von Wahl et al. (2017) [69] eine Korrelation von $r = -0,27$ (95 % LCI: -0,64; UCI: 0,20) gegenüber den manuell gezählten Schritten aufweist. Sie hat aber beim Laufen mit konstanter Geschwindigkeit ($10,1 \text{ km h}^{-1}$) auch eine Gültigkeit von $r = 0,82$ (95 % LCI: 0,59; UCI: 0,93) (Wahl et al., 2017 [30]). Am genauesten misst die MI BAND 2 beim Gehen am Laufband ($4,8 \text{ km h}^{-1}$) für 30 min. Tam und Cheung (2018) [8] eruierten hier eine Gültigkeit von $r = 0,997$ (95 % LCI: 0,994; UCI: 0,999). Die schlechteste Gültigkeit der MI BAND 2 beträgt immer noch $r = 0,991$ (95 % LCI: 0,981; UCI: 0,996), diesmal mit $6,4 \text{ km h}^{-1}$ am Laufband.

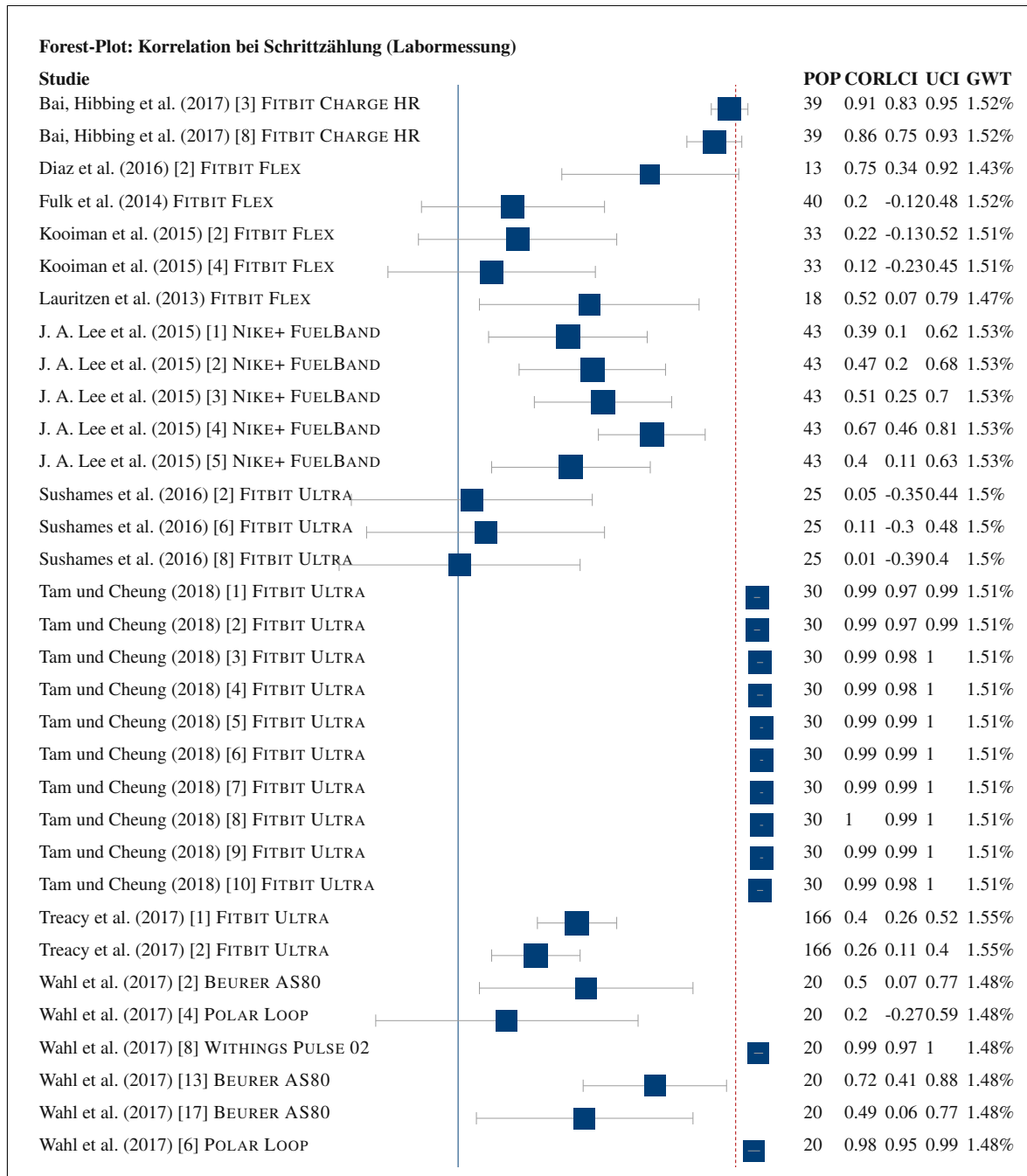


Abbildung 5.2: Fortsetzung mit Abb. 5.3

5 Ergebnisse

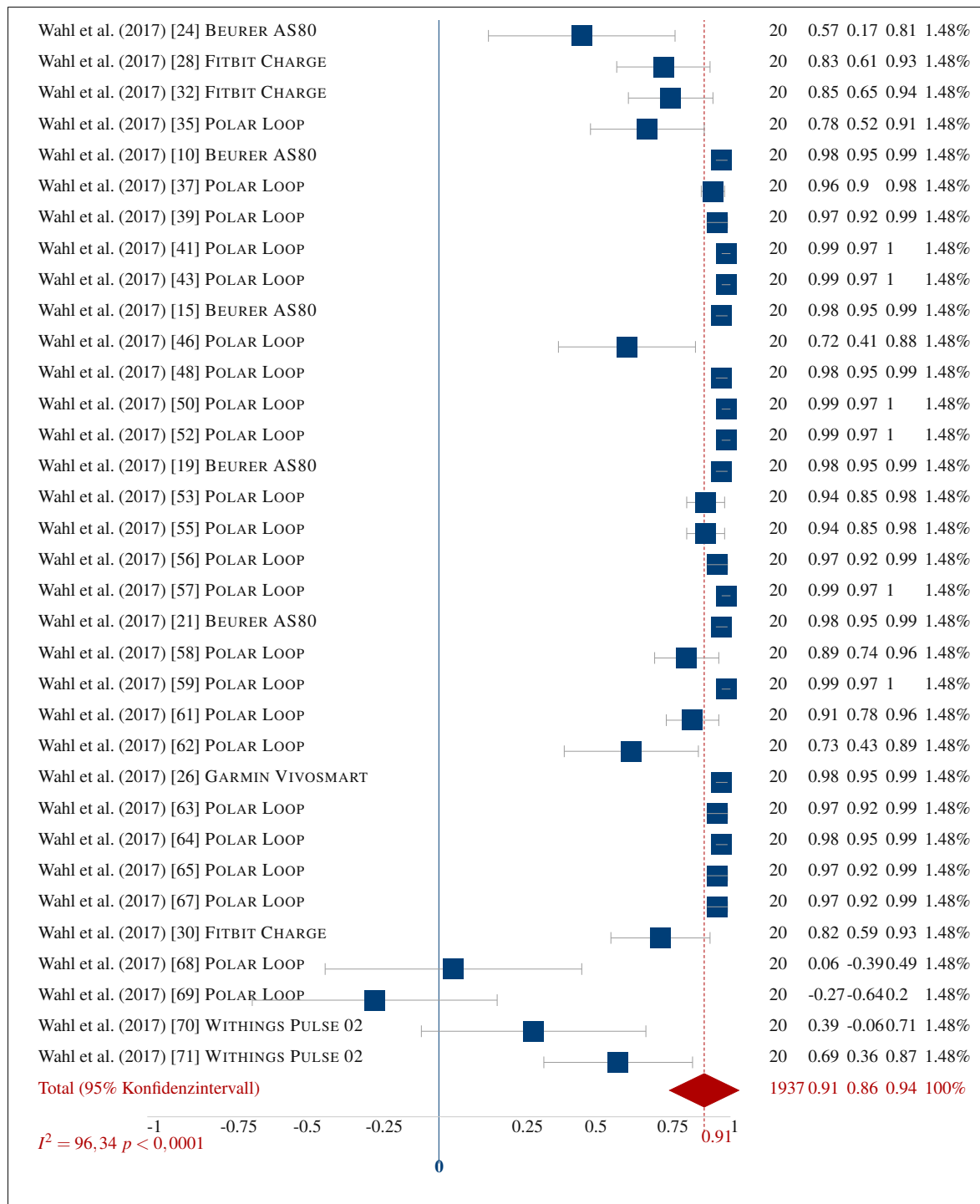


Abbildung 5.3: Fortsetzung von Abb. 5.2

Forest-Plot der Gültigkeit bei der Schrittzählung von Fitness Trackern im Laborumfeld verglichen mit manuell erfassten Schritten und Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Feldmessung

Der *unit calibration* Teil der Schrittzählung bei der Feldmessung hat einen ebenfalls hohen Zusammenhang ($r = 0,799$ (95 % LCI: 0,727; UCI: 0,854)) bei einer Heterogenität von $I^2 = 82,18$ ($P < 0,0001$) (Forest-Plot: Abb. 5.4) bei 27 Fällen. Wenn hier um die Fälle reduziert wird, welche Gehen oder langsamere Geschwindigkeiten ($< 5,4 \text{ km h}^{-1}$) behandeln, bleiben 20 Fälle übrig mit $r = 0,802$ (95 % LCI: 0,698; UCI: 0,87) und $I^2 = 82,23$ ($P < 0,0001$).

Bei den Feldmessungen war die geringste Gültigkeit ($r = 0,30$ (95 % LCI: -0,04; UCI: 0,58)) beim Tracker NIKE+ FUEL BAND SE. An et al. (2017) [19] führten hier einen Test auf einer 200 m Laufstrecke mit Geschwindigkeiten von 4 km h^{-1} bis $6,5 \text{ km h}^{-1}$ durch. NIKE+ FUEL BAND SE erreicht aber bei Battenberg et al. (2017) auch die beste Gültigkeit von $r = 0,956$ (95 % LCI: 0,91; UCI: 0,98) bei einer 400 m Laufstrecke mit variablem Tempo.

Im Vergleich mit den Goldstandards im Feld hat die Schrittzählung bei 29 Fällen einen hohen Zusammenhang ($r = 0,75$ (95 % LCI: 0,605; UCI: 0,846)) mit $I^2 = 91,41$ ($P < 0,0001$) zwischen den Geräten ergeben (Forest-Plot: Abb. 5.5). Bei diesen Ergebnissen gab es wiederum einzelne Fälle mit Geschwindigkeiten unter $< 5,4 \text{ km h}^{-1}$. Wenn diese weggelassen werden, haben die übriggebliebenen 18 Fälle ebenfalls einen hohen Zusammenhang mit $r = 0,833$ (95 % LCI: 0,715; UCI: 0,905) bei $I^2 = 84,9$ ($P < 0,0001$).

Mit $r = -0,03$ (95 % LCI: -0,43; UCI: 0,38) schnitt die FITBIT FLEX bei Floegel et al. (2017) [4], bei einer 100 m Strecke, die mit einem Rollator zurückgelegt wurde, am schlechtesten im Vergleich mit einer ORTHOCARE STEPWATCH ab. Die langsamen Bewegungsformen wieder ausgenommen, ist es die FITBIT FLEX mit $r = 0,10$ (95 % LCI: -0,31; UCI: 0,48), diesmal bei Sushames et al. (2016) [1] bei einem Feldtest kombiniert mit Gehen am Laufband (5 km h^{-1} bis $6,5 \text{ km h}^{-1}$) im Vergleich zu einer ACTIGRAPH GT3X, die am ungenauesten misst. Die FITBIT FLEX hat jedoch im Vergleich mit der AKTIVPAL eine Gültigkeit von $r = 0,96$ (95 % LCI: 0,93; UCI: 0,98) bei Büroaktivitäten während eines eintägigen Arbeitstages (Kooiman et al., 2015 [1]). Am besten schnitt die POLAR A300 ab. Die Gültigkeit beträgt $r = 0,986$ (95 % LCI: 0,964; UCI: 0,995) bei einem von Boeselt et al. (2016) [2] durchgeführten dreitägigen Test inklusive den Nächten.

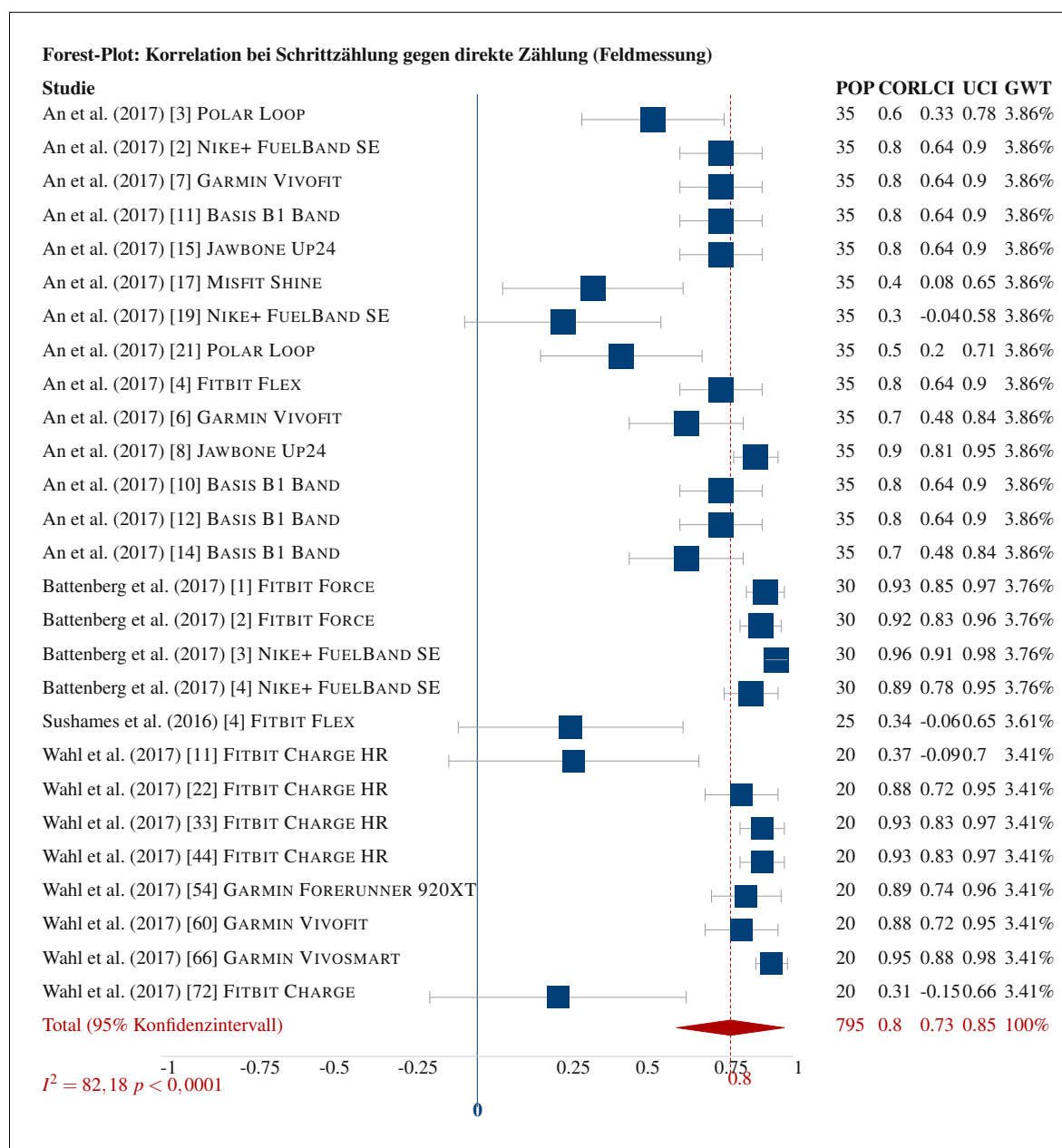


Abbildung 5.4: Forest-Plot der Gültigkeit bei der Schrittzählung von Fitness Trackern in der Feldmessung verglichen mit manuell erfassten Schritten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

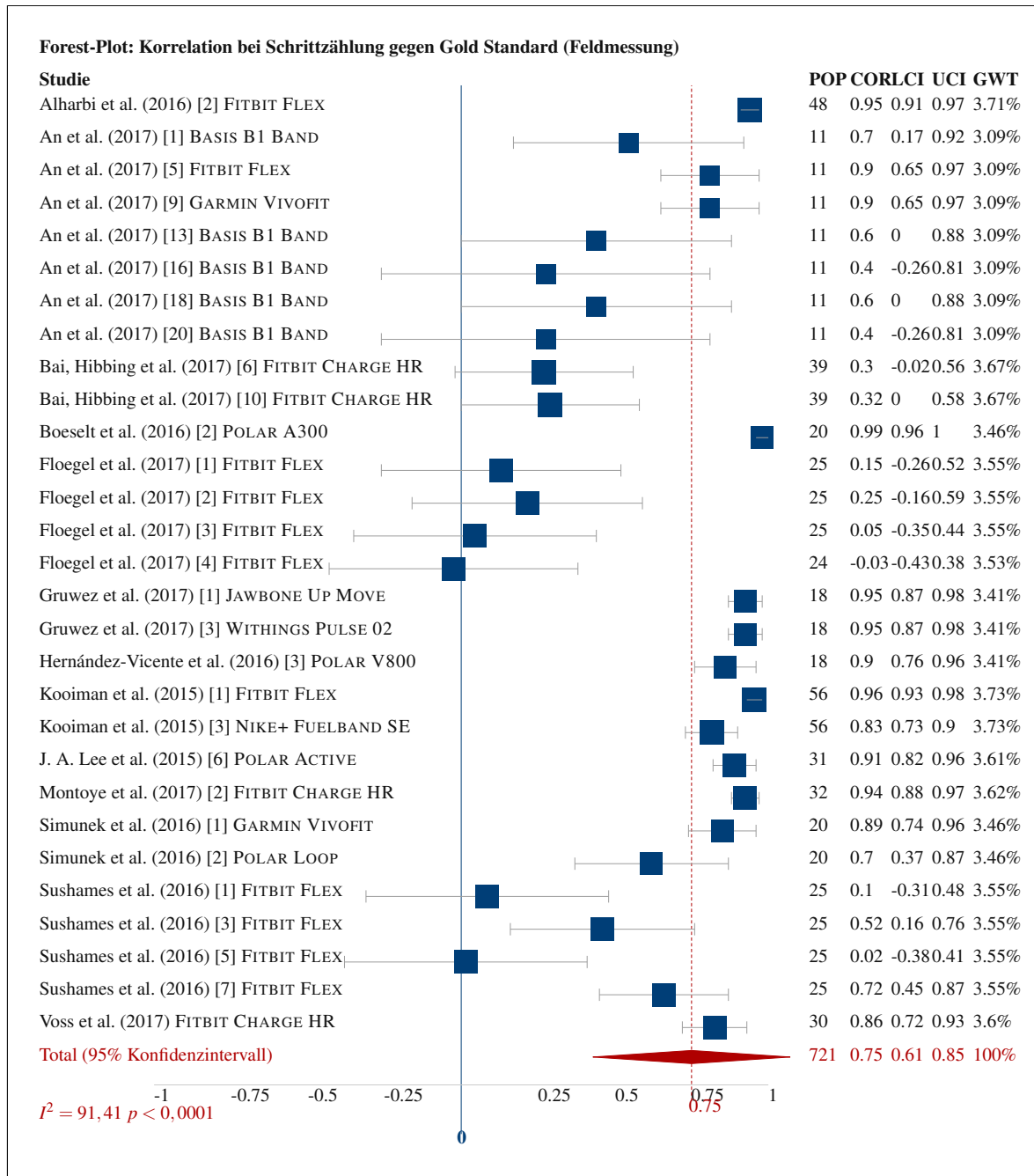


Abbildung 5.5: Forest-Plot der Gültigkeit bei der Schrittzählung von Fitness Trackern in der Feldmessung verglichen mit Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Tabelle 5.4: Studienergebnisse zur Gültigkeit der Schrittmessungen

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Alharbi et al. (2016) [2]	FITBIT FLEX	ActiGraph GT3X	Feld	normaler Alltag	4 d (exkl. Nacht)	48	0,948	0,909	0,971
An et al. (2017) [1]	BASIS B1 BAND	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,7	0,173	0,915
An et al. (2017) [10]	MISFIT SHINE	direkt gezählt (Video)	Feld	Laufband (3,2 km h ⁻¹ bis 8 km h ⁻¹)	3 min	35	0,8	0,636	0,895
An et al. (2017) [11]	GARMIN VIVOFIT	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h ⁻¹ bis 6,5 km h ⁻¹)	200 m	35	0,8	0,636	0,895
An et al. (2017) [12]	NIKE+ FUEL-BAND SE	direkt gezählt (Video)	Feld	Laufband (3,2 km h ⁻¹ bis 8 km h ⁻¹)	3 min	35	0,8	0,636	0,895
An et al. (2017) [13]	JAWBONE UP24	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,6	0	0,882
An et al. (2017) [14]	POLAR LOOP	direkt gezählt (Video)	Feld	Laufband (3,2 km h ⁻¹ bis 8 km h ⁻¹)	3 min	35	0,7	0,478	0,838

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
An et al. (2017) [15]	JAWBONE UP24	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,8	0,636	0,895
An et al. (2017) [16]	MISFIT SHINE	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,4	-0,263	0,806
An et al. (2017) [17]	MISFIT SHINE	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,4	0,077	0,647
An et al. (2017) [18]	NIKE+ FUEL-BAND SE	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,6	0	0,882
An et al. (2017) [19]	NIKE+ FUEL-BAND SE	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,3	-0,036	0,576
An et al. (2017) [2]	BASIS B1 BAND	direkt gezählt (Video)	Feld	Laufband ($3,2 \text{ km h}^{-1}$ bis 8 km h^{-1})	3 min	35	0,8	0,636	0,895
An et al. (2017) [20]	POLAR LOOP	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,4	-0,263	0,806

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
An et al. (2017) [21]	POLAR LOOP	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,5	0,2	0,714
An et al. (2017) [3]	BASIS B1 BAND	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,6	0,333	0,778
An et al. (2017) [4]	FITBIT FLEX	direkt gezählt (Video)	Feld	Laufband ($3,2 \text{ km h}^{-1}$ bis 8 km h^{-1})	3 min	35	0,8	0,636	0,895
An et al. (2017) [5]	FITBIT FLEX	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,9	0,652	0,974
An et al. (2017) [6]	GARMIN VIVOFIT	direkt gezählt (Video)	Feld	Laufband ($3,2 \text{ km h}^{-1}$ bis 8 km h^{-1})	3 min	35	0,7	0,478	0,838
An et al. (2017) [7]	FITBIT FLEX	direkt gezählt (Klickzähler)	Feld	Laufstrecke (4 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	200 m	35	0,8	0,636	0,895
An et al. (2017) [8]	JAWBONE UP24	direkt gezählt (Video)	Feld	Laufband ($3,2 \text{ km h}^{-1}$ bis 8 km h^{-1})	3 min	35	0,9	0,81	0,949

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
An et al. (2017) [9]	GARMIN VIVOFIT	New Lifestyle NL-1000	Feld	normaler Alltag	24	11*	0,9	0,652	0,974
Bai, Hibbing et al. (2017) [10]	FITBIT CHARGE HR	Yamax SW-200	Feld	Alltagstätigkeiten im Bereich LPA	25 min	39	0,32	0,004	0,577
Bai, Hibbing et al. (2017) [3]	APPLE WATCH 1	Yamax SW-200	Labor	Laufband (selbst gewähltes Tempo)	25 min	39	0,91	0,834	0,952
Bai, Hibbing et al. (2017) [6]	APPLE WATCH 1	Yamax SW-200	Feld	Alltagstätigkeiten im Bereich LPA	25 min	39	0,3	-0,017	0,562
Bai, Hibbing et al. (2017) [8]	FITBIT CHARGE HR	Yamax SW-200	Labor	Laufband (selbst gewähltes Tempo)	25 min	39	0,86	0,747	0,925
Battenberg et al. (2017) [1]	FITBIT FORCE	direkt gezählt (Klickzähler)	Feld	400 m Strecke laufen	400 m	30	0,927	0,851	0,965

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Battenberg et al. (2017) [2]	FITBIT FORCE	direkt gezählt (Klickzähler)	Feld	400 m Strecke laufen	400 m	30	0,916	0,829	0,96
Battenberg et al. (2017) [3]	NIKE+ FUEL-BAND SE	direkt gezählt (Klickzähler)	Feld	400 m Strecke laufen	400 m	30	0,956	0,909	0,979
Battenberg et al. (2017) [4]	NIKE+ FUEL-BAND SE	direkt gezählt (Klickzähler)	Feld	400 m Strecke laufen	400 m	30	0,891	0,782	0,947
Boeselt et al. (2016) [2]	POLAR A300	Bodymedia SenseWear Pro Armband	Feld	normaler Alltag	3 d á 24	20	0,986	0,964	0,995
Diaz et al. (2016) [2]	FITBIT FLEX	direkt gezählt (Video)	Labor	Laufband (á 6 min bei 3 km h^{-1} , $4,8 \text{ km h}^{-1}$, $6,4 \text{ km h}^{-1}$ und $8,5 \text{ km h}^{-1}$ + á 3 min Pause)	31 min	13	0,75	0,339	0,921
Floegel et al. (2017) [1]	FITBIT FLEX	OrthoCare StepWatch	Feld	Gehen uneingeschränkt	100 m	25	0,15	-0,261	0,515
Floegel et al. (2017) [2]	FITBIT FLEX	OrthoCare StepWatch	Feld	Gehen eingeschränkt	100 m	25	0,25	-0,161	0,587

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Floegel et al. (2017) [3]	FITBIT FLEX	OrthoCare StepWatch	Feld	Gehen mit Stock	100 m	25	0,05	-0,352	0,437
Floegel et al. (2017) [4]	FITBIT FLEX	OrthoCare StepWatch	Feld	Gehen mit Rollator	100 m	24	-0,03	-0,428	0,378
Fulk et al. (2014)	NIKE+ FUEL-BAND	direkt gezählt (Video)	Labor	2 min Walk Test	2 min	40	0,2	-0,119	0,482
Gruwez et al. (2017) [1]	JAWBONE UP MOVE	Bodymedia SenseWear Pro Armband	Feld	normaler Alltag	1 d (exkl. Nacht)	18	0,95	0,868	0,982
Gruwez et al. (2017) [3]	WITHINGS PULSE 02	Bodymedia SenseWear Pro Armband	Feld	normaler Alltag	1 d (exkl. Nacht)	18	0,95	0,868	0,982
Hernández-Vicente et al. (2016) [3]	POLAR V800	ActiGraph ActiTrainer	Feld	normaler Alltag	7 d (exkl. Nacht)	18	0,904	0,756	0,964
Kooiman et al. (2015) [1]	FITBIT FLEX	AtivPAL	Feld	Aktivitäten während Büroarbeit	1 d (09:00 - 16:30)	56	0,96	0,932	0,976

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Kooiman et al. (2015) [2]	FITBIT FLEX	Optogait	Labor	Laufband ($4,8 \text{ km h}^{-1}$)	30 min	33	0,22	-0,133	0,524
Kooiman et al. (2015) [3]	NIKE+ FUEL-BAND SE	AtivPAL	Feld	Aktivitäten während Büroarbeit	1 d (09:00 - 16:30)	56	0,83	0,725	0,897
Kooiman et al. (2015) [4]	NIKE+ FUEL-BAND SE	Optogait	Labor	Laufband ($4,8 \text{ km h}^{-1}$)	30 min	33	0,12	-0,233	0,445
Lauritzen et al. (2013)	FITBIT ULTRA	direkt gezählt (Video)	Labor	gehen	20 m	18	0,518	0,067	0,793
J. A. Lee et al. (2015) [1]	POLAR ACTIVE	direkt gezählt	Labor	Laufband ($3,2 \text{ km h}^{-1}$)	3 min	43	0,39	0,102	0,618
J. A. Lee et al. (2015) [2]	POLAR ACTIVE	direkt gezählt	Labor	Laufband (4 km h^{-1})	3 min	43	0,47	0,198	0,675
J. A. Lee et al. (2015) [3]	POLAR ACTIVE	direkt gezählt	Labor	Laufband ($4,8 \text{ km h}^{-1}$)	3 min	43	0,51	0,248	0,703

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
J. A. Lee et al. (2015) [4]	POLAR ACTIVE	direkt gezählt	Labor	Laufband ($5,6 \text{ km h}^{-1}$)	3 min	43	0,67	0,463	0,808
J. A. Lee et al. (2015) [5]	POLAR ACTIVE	direkt gezählt	Labor	Laufband ($6,4 \text{ km h}^{-1}$)	3 min	43	0,4	0,113	0,625
J. A. Lee et al. (2015) [6]	POLAR ACTIVE	Yamax SW-701 Pedometer	Feld	normaler Alltag	3 d	31	0,91	0,82	0,956
Montoye et al. (2017) [2]	FITBIT CHARGE HR	Omron HJ 323u Pedometer	Feld	diverse Aktivitäten	89,6 min (SD 5,5)	32	0,94	0,88	0,971
Simunek et al. (2016) [1]	GARMIN VIVOFIT	ActiGraph GT3X	Feld	normaler Alltag	7 d (exkl. Nacht)	20	0,89	0,738	0,956
Simunek et al. (2016) [2]	POLAR LOOP	ActiGraph GT3X	Feld	normaler Alltag	7 d (exkl. Nacht)	20	0,7	0,373	0,872
Sushames et al. (2016) [1]	FITBIT FLEX	ActiGraph GT3X	Feld	Feld inkl. Laufband gehen (5 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	1 d	25	0,1	-0,307	0,476
Sushames et al. (2016) [2]	FITBIT FLEX	direkt gezählt	Labor	Laufband gehen (5 km h^{-1} bis $6,5 \text{ km h}^{-1}$)	8 min bis 89 min	25	0,05	-0,352	0,437

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Sushames et al. (2016) [3]	FITBIT FLEX	ActiGraph GT3X	Feld	Feld inkl. Laufband jogging (8 km h^{-1} bis 10 km h^{-1})	1 d	25	0,52	0,157	0,759
Sushames et al. (2016) [4]	FITBIT FLEX	direkt gezählt	Feld	Laufband jogging (8 km h^{-1} bis 10 km h^{-1})	8 min bis 89 min	25	0,34	-0,063	0,648
Sushames et al. (2016) [5]	FITBIT FLEX	ActiGraph GT3X	Feld	Feld inkl. gehen Steigung (5 %)	1 d	25	0,02	-0,378	0,412
Sushames et al. (2016) [6]	FITBIT FLEX	direkt gezählt	Labor	gehen Steigung (5 %)	8 min bis 89 min	25	0,11	-0,298	0,484
Sushames et al. (2016) [7]	FITBIT FLEX	ActiGraph GT3X	Feld	Feld inkl. Stufensteigen (15 cm)	1 d	25	0,72	0,454	0,868
Sushames et al. (2016) [8]	FITBIT FLEX	direkt gezählt	Labor	Stufensteigen (15 cm)	8 min bis 89 min	25	0,01	-0,387	0,404
Tam und Cheung (2018) [1]	FITBIT CHARGE HR	direkt gezählt (Video)	Labor	Laufband gehen ($3,2 \text{ km h}^{-1}$)	30 min	30	0,986	0,97	0,993
Tam und Cheung (2018) [10]	MI BAND 2	direkt gezählt (Video)	Labor	Laufband gehen ($6,4 \text{ km h}^{-1}$)	30 min	30	0,991	0,981	0,996

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Tam und Cheung (2018) [2]	FITBIT CHARGE HR	direkt gezählt (Video)	Labor	Laufband gehen (4 km h^{-1})	30 min	30	0,985	0,968	0,993
Tam und Cheung (2018) [3]	FITBIT CHARGE HR	direkt gezählt (Video)	Labor	Laufband gehen ($4,8 \text{ km h}^{-1}$)	30 min	30	0,991	0,981	0,996
Tam und Cheung (2018) [4]	FITBIT CHARGE HR	direkt gezählt (Video)	Labor	Laufband gehen ($5,5 \text{ km h}^{-1}$)	30 min	30	0,992	0,983	0,996
Tam und Cheung (2018) [5]	FITBIT CHARGE HR	direkt gezählt (Video)	Labor	Laufband gehen ($6,4 \text{ km h}^{-1}$)	30 min	30	0,993	0,985	0,997
Tam und Cheung (2018) [6]	MI BAND 2	direkt gezählt (Video)	Labor	Laufband gehen ($3,2 \text{ km h}^{-1}$)	30 min	30	0,993	0,985	0,997
Tam und Cheung (2018) [7]	MI BAND 2	direkt gezählt (Video)	Labor	Laufband gehen (4 km h^{-1})	30 min	30	0,993	0,985	0,997
Tam und Cheung (2018) [8]	MI BAND 2	direkt gezählt (Video)	Labor	Laufband gehen ($4,8 \text{ km h}^{-1}$)	30 min	30	0,997	0,994	0,999
Tam und Cheung (2018) [9]	MI BAND 2	direkt gezählt (Video)	Labor	Laufband gehen ($5,5 \text{ km h}^{-1}$)	30 min	30	0,993	0,985	0,997

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Treacy et al. (2017) [1]	FITBIT CHARGE	direkt gezählt	Labor	6 min Walk Test	6 min	166	0,399	0,263	0,52
Treacy et al. (2017) [2]	GARMIN VIVOFIT	direkt gezählt	Labor	6 min Walk Test	6 min	166	0,259	0,111	0,396
Voss et al. (2017)	FITBIT CHARGE HR	ActiGraph GT3X+	Feld	normaler Alltag	7 d (exkl. Nacht)	30	0,855	0,715	0,929
Wahl et al. (2017) [10]	FITBIT CHARGE HR	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [11]	BEURER AS80	direkt gezählt	Feld	Intermettierendes Laufen($\emptyset 9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,37	-0,086	0,698
Wahl et al. (2017) [13]	BEURER AS80	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,72	0,407	0,882
Wahl et al. (2017) [15]	GARMIN FORERUNNER 920XT	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [17]	BEURER AS80	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,49	0,06	0,766
Wahl et al. (2017) [19]	GARMIN VIVOACTIVE	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [2]	BEURER AS80	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,5	0,073	0,772

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [21]	GARMIN VIVOFIT	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [22]	FITBIT CHARGE	direkt gezählt	Feld	Intermettierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,88	0,716	0,952
Wahl et al. (2017) [24]	FITBIT CHARGE	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,57	0,17	0,809
Wahl et al. (2017) [26]	GARMIN VIVOS-MART	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [28]	FITBIT CHARGE	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,83	0,612	0,931
Wahl et al. (2017) [30]	POLAR LOOP	direkt gezählt	Labor	Laufen ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,82	0,592	0,926
Wahl et al. (2017) [32]	FITBIT CHARGE	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,85	0,653	0,939
Wahl et al. (2017) [33]	FITBIT CHARGE HR	direkt gezählt	Feld	Intermettierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,93	0,828	0,972
Wahl et al. (2017) [35]	FITBIT CHARGE	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,78	0,515	0,909
Wahl et al. (2017) [37]	FITBIT CHARGE HR	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,96	0,9	0,984

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [39]	FITBIT CHARGE HR	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,97	0,924	0,988
Wahl et al. (2017) [4]	BEURER AS80	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,2	-0,266	0,59
Wahl et al. (2017) [41]	FITBIT CHARGE HR	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [43]	FITBIT CHARGE HR	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [44]	GARMIN FORERUNNER 920XT	direkt gezählt	Feld	Intermettierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,93	0,828	0,972
Wahl et al. (2017) [46]	GARMIN FORERUNNER 920XT	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,72	0,407	0,882
Wahl et al. (2017) [48]	GARMIN FORERUNNER 920XT	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,98	0,949	0,992

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [50]	GARMIN FORERUNNER 920XT	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [52]	GARMIN FORERUNNER 920XT	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [53]	GARMIN VIVOACTIVE	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,94	0,852	0,976
Wahl et al. (2017) [54]	GARMIN VIVOACTIVE	direkt gezählt	Feld	Intermettierendes Laufen($\varnothing 9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,89	0,738	0,956
Wahl et al. (2017) [55]	GARMIN VIVOACTIVE	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,94	0,852	0,976
Wahl et al. (2017) [56]	GARMIN VIVOACTIVE	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,97	0,924	0,988
Wahl et al. (2017) [57]	GARMIN VIVOACTIVE	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [58]	GARMIN VIVOFIT	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,89	0,738	0,956

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [59]	GARMIN VIVOFIT	direkt gezählt	Labor	Laufband (7,2 km h ⁻¹)	5 min	20	0,99	0,974	0,996
Wahl et al. (2017) [6]	FITBIT CHARGE	direkt gezählt	Labor	Laufen (10,1 km h ⁻¹)	2,4 km	20	0,98	0,949	0,992
Wahl et al. (2017) [60]	GARMIN VIVOFIT	direkt gezählt	Feld	Intermettierendes Laufen(Ø 9,1 km h ⁻¹ Max.: 22,4 km h ⁻¹)	5 min	20	0,88	0,716	0,952
Wahl et al. (2017) [61]	GARMIN VIVOFIT	direkt gezählt	Labor	Laufband (10,1 km h ⁻¹)	5 min	20	0,91	0,783	0,964
Wahl et al. (2017) [62]	GARMIN VIVOFIT	direkt gezählt	Labor	Laufband (13 km h ⁻¹)	5 min	20	0,73	0,425	0,886
Wahl et al. (2017) [63]	GARMIN VIVOS-MART	direkt gezählt	Labor	Laufband (4,3 km h ⁻¹)	5 min	20	0,97	0,924	0,988
Wahl et al. (2017) [64]	GARMIN VIVOS-MART	direkt gezählt	Labor	Laufband (7,2 km h ⁻¹)	5 min	20	0,98	0,949	0,992
Wahl et al. (2017) [65]	GARMIN VIVOS-MART	direkt gezählt	Labor	Laufband (10,1 km h ⁻¹)	5 min	20	0,97	0,924	0,988
Wahl et al. (2017) [66]	GARMIN VIVOS-MART	direkt gezählt	Feld	Intermettierendes Laufen(Ø 9,1 km h ⁻¹ Max.: 22,4 km h ⁻¹)	5 min	20	0,95	0,876	0,98
Wahl et al. (2017) [67]	GARMIN VIVOS-MART	direkt gezählt	Labor	Laufband (13 km h ⁻¹)	5 min	20	0,97	0,924	0,988

Tabelle 5.4: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [68]	POLAR LOOP	direkt gezählt	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,06	-0,393	0,49
Wahl et al. (2017) [69]	POLAR LOOP	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	-0,27	-0,636	0,196
Wahl et al. (2017) [70]	POLAR LOOP	direkt gezählt	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,39	-0,063	0,71
Wahl et al. (2017) [71]	POLAR LOOP	direkt gezählt	Labor	Laufband (13 km h^{-1})	5 min	20	0,69	0,356	0,868
Wahl et al. (2017) [72]	POLAR LOOP	direkt gezählt	Feld	Intermettierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,31	-0,154	0,662
Wahl et al. (2017) [8]	BEURER AS80	direkt gezählt	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,99	0,974	0,996

Tabelle 5.4: Ergebnisse einzelner Fitness Tracker Studien als Korrelation zum Gültigkeitskriterium direktes Zählen oder einem Goldstandard, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar. Abkürzungen: Korrelation (COR), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

Wert von Autor definiert, da je 3 von verfügbaren 10 Wearables (inkl. Pedometer) zufällig ausgeteilt wurden.

5.1.2 Zuverlässigkeit bei der Schrittzählung

Die Tests der Zuverlässigkeit bei der Schrittzählung umfassen sechs Fälle aus dem Laborbereich und zwei aus einer Feldmessung (Tabelle 5.5).

Labormessung

Die sechs Fälle der Labormessung besitzen eine ICC mit einer mittleren Zuverlässigkeit ($r = 0,589$ (95 % LCI: 0,361; UCI: 0,751)), bei einer Heterogenität der Fälle mit $I^2 = 65,48$ ($P = 0,0128$) (Forest-Plot: Abb. 5.6).

Die FITBIT FLEX schnitt bei den wiederholten Tests je nach Anforderung am schlechtesten, aber auch am besten ab. Beim Gehen einer 5 % Steigung bei einer Dauer von 8 min bis 89 min wurde eine niedrige ICC von $r = 0,22$ (95 % LCI: -0,19; UCI: 0,57) eruiert (Sushames et al., 2016 [R3]). Wohingegen beim Gehen auf einem Laufband ($4,8 \text{ km h}^{-1}$ für 30 min) eine hohe ICC mit $r = 0,81$ (95 % LCI: 0,64; UCI: 0,91) von Kooiman et al. (2015) [R1] angegeben wurde.

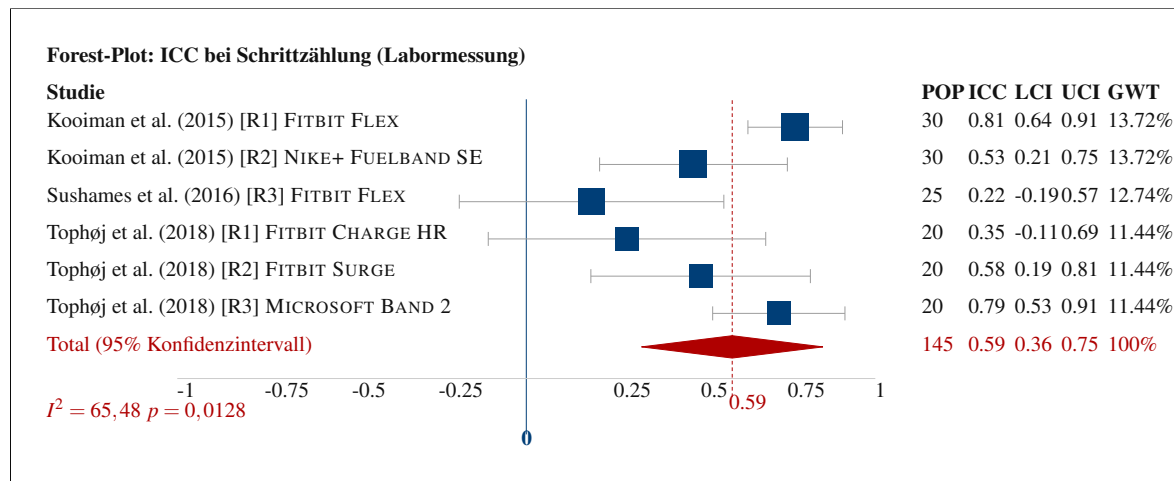


Abbildung 5.6: Forest-Plot der Zuverlässigkeit bei der Schrittzählung von Fitness Trackern im Laborumfeld. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar. Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)
Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Feldmessung

Der Feldmessung entsprachen zwei Fälle. Beide aus der Studie von Sushames et al. (2016) und behandeln den Fitness Tracker FITBIT FLEX. Dieser weist beim Gehen (5 km h^{-1} bis $6,5 \text{ km h}^{-1}$) eine mittlere Zuverlässigkeit auf ($r = 0,57$ (95 % LCI: 0,226; UCI: 0,788)).

Wohingegen die Zuverlässigkeit beim Joggen (8 km h^{-1} bis 10 km h^{-1}) nur noch im niedrigen Zuverlässigkeitsbereich gegeben ist ($r = 0,26$ (95 % LCI: $-0,151$; UCI: $0,954$)).

Tabelle 5.5: Studienergebnisse zur Zuverlässigkeit von Fitness Tracker bei der Schrittzählung

Autor	Fitness Tracker	Aktivität	Messdauer	Anzahl	ICC	LCI (95 %)	UCI (95 %)
Kooiman et al. (2015) [R1]	FITBIT FLEX	Laufband (4,8 km h ⁻¹)	30 min	30	0,81	0,635	0,906
Kooiman et al. (2015) [R2]	NIKE+ FUELBAND SE	Laufband (4,8 km h ⁻¹)	30 min	30	0,53	0,21	0,748
Sushames et al. (2016) [R1]	FITBIT FLEX	Laufband gehen (5 km h ⁻¹ bis 6,5 km h ⁻¹)	8 min bis 89 min	25	0,57	0,226	0,788
Sushames et al. (2016) [R2]	FITBIT FLEX	Laufband jogging (8 km h ⁻¹ bis 10 km h ⁻¹)	8 min bis 89 min	25	0,26	-0,151	0,594
Sushames et al. (2016) [R3]	FITBIT FLEX	gehen Steigung (5 %)	8 min bis 89 min	25	0,22	-0,192	0,566
Tophøj et al. (2018) [R1]	FITBIT CHARGE HR	Laufband gehen (2x 2 km h ⁻¹ , 4 km h ⁻¹ , 4,5 km h ⁻¹ und 5,5 km h ⁻¹ für 100 Schritte)	6 min	20	0,35	-0,109	0,686
Tophøj et al. (2018) [R2]	FITBIT SURGE	Laufband gehen (2x 2 km h ⁻¹ , 4 km h ⁻¹ , 4,5 km h ⁻¹ und 5,5 km h ⁻¹ für 100 Schritte)	6 min	20	0,58	0,185	0,814
Tophøj et al. (2018) [R3]	MICROSOFT BAND 2	Laufband gehen (2x 2 km h ⁻¹ , 4 km h ⁻¹ , 4,5 km h ⁻¹ und 5,5 km h ⁻¹ für 100 Schritte)	6 min	20	0,79	0,534	0,913

Tabelle 5.5: Ergebnisse einzelner Fitness Tracker Studien als Intra-Klassen-Korrelation, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

5.2 Pulsmessung

Das Ergebnis der Gültigkeit der Herzfrequenzmessung (Abschnitt 5.2.1) von Fitness Trackern, wurde aus den Daten von sieben Studien berechnet. Für die Zuverlässigkeit der Pulsmessung (Abschnitt 5.2.2), standen die Daten aus zwei Studien (Abt et al., 2017; Wallen et al., 2016) zur Verfügung.

5.2.1 Gültigkeit der Pulsmessung

Für die Metaanalyse sind sieben Studienergebnisse zu 17 unterschiedlichen Fällen (unterschiedliche Fitness Tracker und verschiedene Aktivitäten) in Tabelle 5.6 zusammengefasst. Die Metaanalyse ist dabei zwischen Labormessung (Abschnitt 5.2.1) und Feldmessung (Abschnitt 5.2.1) getrennt.

Beim Messen der Herzfrequenz wurde gegen ein 12-Kanal EKG oder gegen die validierten Tracker NONIN PURESAT PULSE OXIMETER, POLAR H6 HR-MONITOR, HR SENSOR, RS400 und T31 CHEST STRAP getestet.

Labormessung

Es wurden 14 Fälle ausgemacht, welche bei den Testungen der Gültigkeit der Herzfrequenzmessung in einem Laborumfeld stattfanden. In Laborumgebungen besitzen Fitness Tracker mit PPG einen hohen Zusammenhang ($r = 0,916$ (95 % LCI: 0,88; UCI: 0,942)) bei einer Heterogenität der Fälle von $I^2 = 72,09$ ($P < 0,0001$) (Forest-Plot: Abb. 5.7). Als Kriterium wurden entweder ein 12-Kanal EKG oder Sportuhren mit einer Pulsmessung mittels Hautelektroden über einen Brustgurt (POLAR HR SENSOR, RS400 und T31 CHEST STRAP) verwendet.

Die geringste Gültigkeit wurde von Wallen et al. (2016) [7] bei der SAMSUNG GEAR S am Laufband (9 min) und Ergometer (18 min bei 25 W) mit $r = 0,67$ (95 % LCI: 0,35; UCI: 0,85) gemessen, die höchste von Abt et al. (2017) [2] bei der APPLE WATCH ebenfalls am Laufband (Start mit 3 km h^{-1} , Erhöhung um $0,5 \text{ km h}^{-1}$ alle 30 s mit Messung von 30 s in für Person individueller höchster Herzfrequenz) mit $r = 0,98$ (95 % LCI: 0,94; UCI: 0,99) und bei einer Trageweise am rechten Handgelenk. Die APPLE WATCH erreicht im schlechtesten Fall noch immer eine Gültigkeit von $r = 0,87$ (95 % LCI: 0,63; UCI: 0,96), diesmal beim Tragen am linken Handgelenk (Abt et al., 2017 [1]).

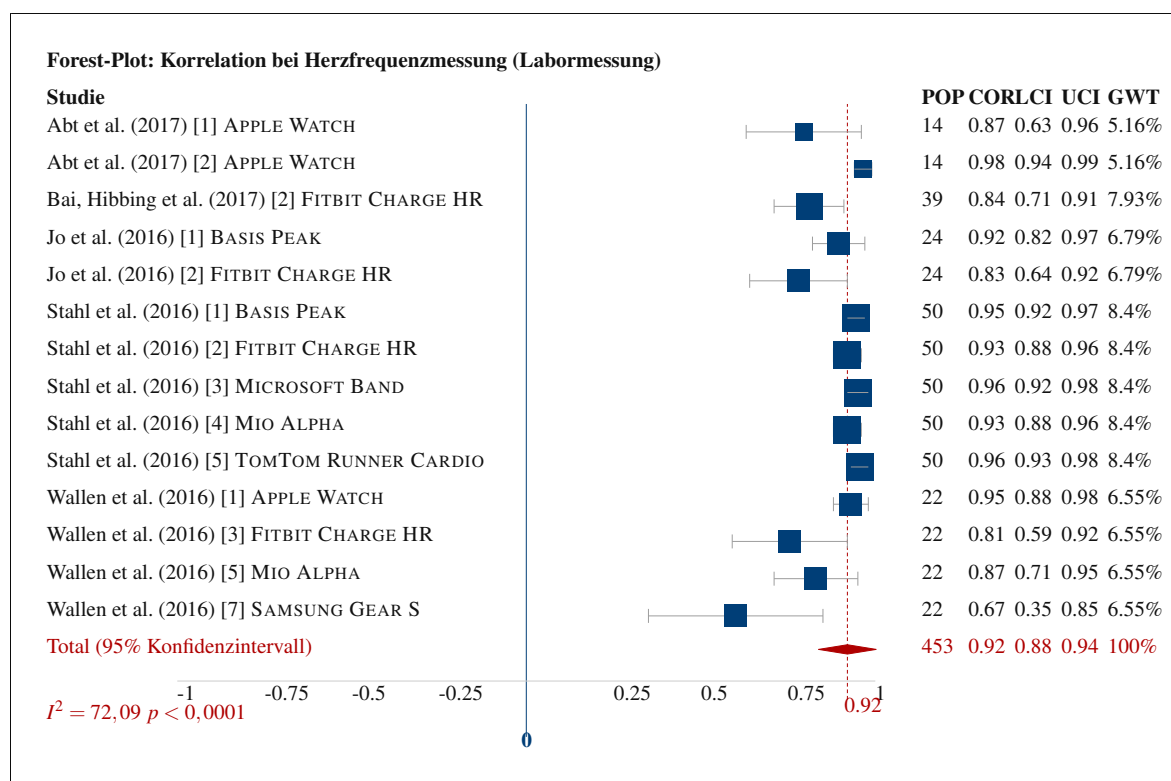


Abbildung 5.7: Forest-Plot der Gültigkeit bei der Herzfrequenzmessung von Fitness Trackern im Laborumfeld verglichen mit EKGs und Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Feldmessung

FITBIT CHARGE HR ist der einzige Tracker, welcher in drei unterschiedlichen Studien vorkommt und somit in drei Fälle zusammengefasst wurde. Als Kriterium wurden entweder ein Oximeter (NONIN PURESAT PULSE OXIMETER) oder Sportuhren, mit einer Pulsmessung mittels Hautelektroden über einen Brustgurt (POLAR H6 HR-MONITOR und HR SENSOR), verwendet. Auch bei Feldmessungen weisen Wearables, wenn auch schon geringer als im Laborumfeld, einen hohen Zusammenhang ($r = 0,833$ (95 % LCI: 0,679; UCI: 0,917)) bei $I^2 = 54,26$ ($P < 0,0001$) auf (Forest-Plot: Abb. 5.8).

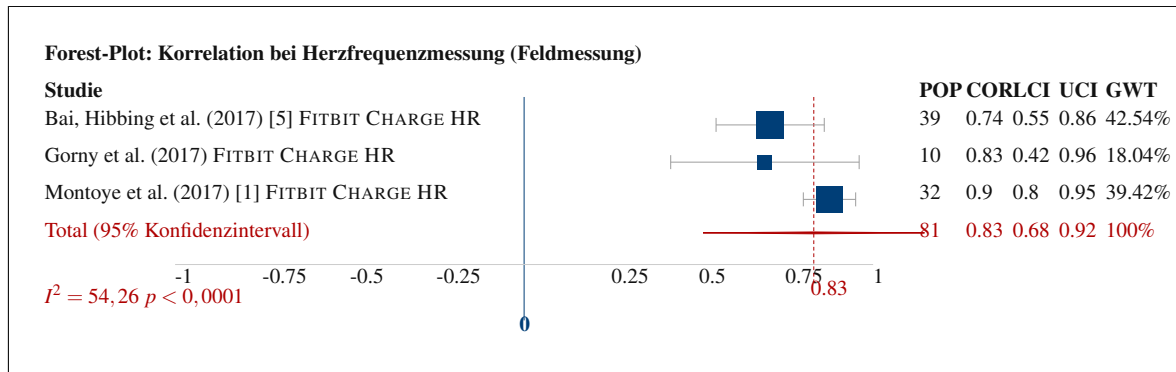


Abbildung 5.8: Forest-Plot der Gültigkeit bei der Herzfrequenzmessung von Fitness Trackern in der Feldmessung verglichen mit einem Pulsoximeter und Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Tabelle 5.6: Studienergebnisse zur Gültigkeit der Herzfrequenzmessung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Abt et al. (2017) [1]	APPLE WATCH	Polar T31 chest strap	Labor	Laufband (Start mit 3 km h^{-1} , Erhöhung um $0,5 \text{ km h}^{-1}$ alle 30 s)	30 s in HR-max	14	0,87	0,63	0,958
Abt et al. (2017) [2]	APPLE WATCH	Polar T31 chest strap	Labor	Laufband (Start mit 3 km h^{-1} , Erhöhung um $0,5 \text{ km h}^{-1}$ alle 30 s)	30 s in HR-max	14	0,98	0,936	0,994
Bai, Hibbing et al. (2017) [2]	FITBIT CHARGE HR	Polar HR Sensor	Labor	Laufband (selbst gewähltes Tempo)	25 min	39	0,84	0,714	0,913
Bai, Hibbing et al. (2017) [5]	FITBIT CHARGE HR	Polar HR Sensor	Feld	Alltagstätigkeiten im Bereich LPA	25 min	39	0,74	0,554	0,856
Gorny et al. (2017)	FITBIT CHARGE HR	Polar H6 HR-Monitor	Feld	normaler Alltag	1 m zu 3 d ⁻¹ bis 6 d ⁻¹	10	0,83	0,42	0,959

Tabelle 5.6: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Jo et al. (2016) [1]	BASIS PEAK	12-Kanal EKG	Labor	Ergometer (60 W und 120 W); gehen ($4,8 \text{ km h}^{-1}$ bis $5,6 \text{ km h}^{-1}$), joggen ($6,4 \text{ km h}^{-1}$ bis 8 km h^{-1}), laufen ($8,9 \text{ km h}^{-1}$ bis $11,2 \text{ km h}^{-1}$) á 5 min; Hanteltraining (12 Wiederholungen); Lunges (12 Wiederholungen); 60 s Planking + 5 min Pause	77 min	24	0,92	0,821	0,965
Jo et al. (2016) [2]	FITBIT CHARGE HR	12-Kanal EKG	Labor	Ergometer (60 W und 120 W); gehen ($4,8 \text{ km h}^{-1}$ bis $5,6 \text{ km h}^{-1}$), joggen ($6,4 \text{ km h}^{-1}$ bis 8 km h^{-1}), laufen ($8,9 \text{ km h}^{-1}$ bis $11,2 \text{ km h}^{-1}$) á 5 min; Hanteltraining (12 Wiederholungen); Lunges (12 Wiederholungen); 60 s Planking + 5 min Pause	77 min	24	0,83	0,641	0,924
Montoye et al. (2017) [1]	FITBIT CHARGE HR	Nonin PureSAT Pulse Oximeter	Feld	diverse Aktivitäten	89,6 min (SD 5,5)	32	0,9	0,803	0,95
Stahl et al. (2016) [1]	BASIS PEAK	Polar RS400	Labor	Laufband ($3,2 \text{ km h}^{-1}$ bis $9,6 \text{ km h}^{-1}$)	30 min	50	0,954	0,92	0,974

Tabelle 5.6: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Stahl et al. (2016) [2]	FITBIT CHARGE HR	Polar RS400	Labor	Laufband (3,2 km h ⁻¹ bis 9,6 km h ⁻¹)	30 min	50	0,933	0,884	0,962
Stahl et al. (2016) [3]	MICROSOFT BAND	Polar RS400	Labor	Laufband (3,2 km h ⁻¹ bis 9,6 km h ⁻¹)	30 min	50	0,956	0,923	0,975
Stahl et al. (2016) [4]	MIO ALPHA	Polar RS400	Labor	Laufband (3,2 km h ⁻¹ bis 9,6 km h ⁻¹)	30 min	50	0,929	0,878	0,959
Stahl et al. (2016) [5]	TOMTOM RUNNER CARDIO	Polar RS400	Labor	Laufband (3,2 km h ⁻¹ bis 9,6 km h ⁻¹)	30 min	50	0,959	0,929	0,977
Wallen et al. (2016) [1]	APPLE WATCH	ECG CASE exercise	Labor	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,95	0,881	0,979
Wallen et al. (2016) [3]	FITBIT CHARGE HR	ECG CASE exercise	Labor	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,81	0,59	0,918
Wallen et al. (2016) [5]	MIO ALPHA	ECG CASE exercise	Labor	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,87	0,708	0,945
Wallen et al. (2016) [7]	SAMSUNG GEAR S	ECG CASE exercise	Labor	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,67	0,346	0,851

Tabelle 5.6: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messungsbung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
-------	-----------------	-----------	--------------	-----------	-----------	--------	-----	------------	------------

Tabelle 5.6: Ergebnisse einzelner Fitness Tracker Studien als Korrelation zum Gültigkeitskriterium direktes Zählen oder einem Goldstandard, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar. Abkürzungen: Korrelation (COR), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

Bei Abt et al. (2017) [1] wurde die APPLE WATCH am linken Handgelenk, bei Abt et al. (2017) [2] am rechten getragen.

5.2.2 Zuverlässigkeit bei der Pulsmessung

Die fünf ausgemachten Fälle behandeln die Zuverlässigkeit bei der Herzfrequenzmessung in einer Laborumgebung. Die hohe Korrelation der ICC der Fälle beträgt $r = 0,894$ (95 % LCI: 0,751; UCI: 0,957) bei einer Heterogenität der Fälle von $I^2 = 79,84$ ($P = 0,0005$).

Die schlechteste und beste ICC (beide mit hohem Zusammenhang) wurden von Wallen et al. (2016) bei jeweils einem Test am Laufband (9 min) und am Ergometer (18 min bei 25 W) bei einer Dauer von 25 min ausgemacht. Die FITBIT CHARGE HR hat mit $r = 0,78$ (95 % LCI: 0,53; UCI: 0,90) (Wallen et al., 2016 [R3]) eine schlechtere ICC als die APPLE WATCH mit $r = 0,98$ (95 % LCI: 0,95; UCI: 0,99) (Wallen et al., 2016 [R1]).

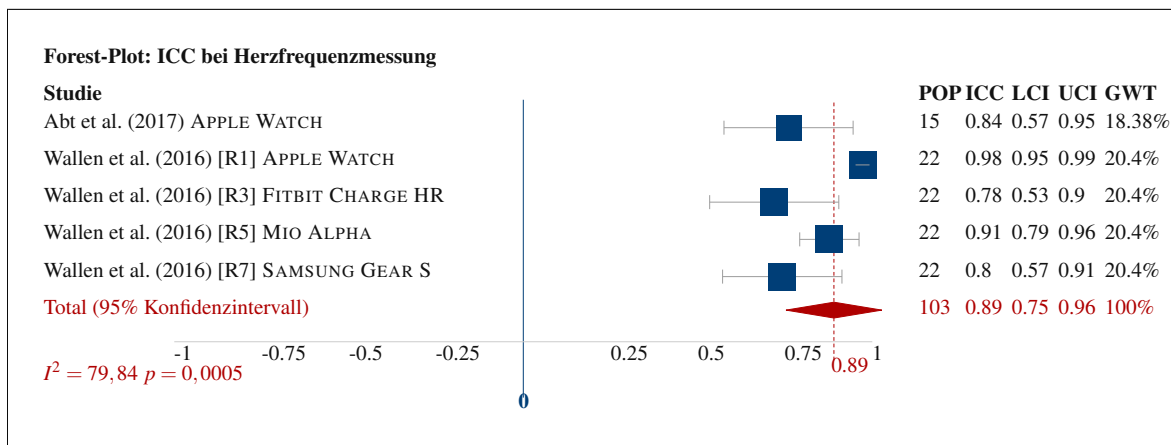


Abbildung 5.9: Forest-Plot der Zuverlässigkeit bei der Herzfrequenzmessung von Fitness Trackern im Laborumfeld. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar. Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)
Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Tabelle 5.7: Studienergebnisse zur Zuverlässigkeit von Fitness Tracker bei der Pulsmessung

Autor	Fitness Tracker	Aktivität	Messdauer	Anzahl	ICC	LCI (95 %)	UCI (95 %)
Abt et al. (2017)	APPLE WATCH	Laufband (Start mit 3 km h^{-1} , Erhöhung um $0,5 \text{ km h}^{-1}$ alle 30 s)	30 s in HRmax	15	0,84	0,575	0,945
Wallen et al. (2016) [R3]	FITBIT CHARGE HR	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,78	0,534	0,904
Wallen et al. (2016) [R7]	SAMSUNG GEAR S	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,8	0,571	0,913
Wallen et al. (2016) [R5]	MIO ALPHA	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,91	0,792	0,962
Wallen et al. (2016) [R1]	APPLE WATCH	Laufband (9 min) & Ergometer (18 min bei 25 W)	25 min	22	0,98	0,952	0,992

Tabelle 5.7: Ergebnisse einzelner Fitness Tracker Studien als Intra-Klassen-Korrelation, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

5.3 Kalorienverbrauchsberechnung

Die Korrelationen zwischen den einzelnen Fällen, wurden bei der Gültigkeit (Abschnitt 5.3.1) von 16 Studien und bei der Zuverlässigkeit (Abschnitt 5.3.2) von einer Studie in Laborumgebung (Wallen et al., 2016) behandelt. Bei den Ergebnissen der Gültigkeit kommt, neben den Ergebnissen zur Kalorienverbrauchsberechnung in Labor- und Feldumgebung, auch die Erkennungsrate von unterschiedlichen Bewegungsintensitäten bei Fitness Trackern hinzu.

5.3.1 Gültigkeit der Kalorienverbrauchsberechnung

Für die Metaanalyse sind aus 16 Studien die Ergebnisse als 62 unterschiedliche Fälle (unterschiedliche Fitness Tracker und verschiedene Aktivitäten) in Tabelle 5.8 dargestellt. Neben der Einteilung in die Erkennung der Intensitäten bei Feldmessungen (Abschnitt 5.3.1) wird die Gültigkeit bei der Messung vom Kalorienverbrauch auf eine Labormessung (Abschnitt 5.3.1), welche ausschließlich gegen die indirekte Kalorimetrie getestet wurde, und in eine Feldmessung (Abschnitt 5.3.1) unterteilt. In der Feldmessung wird nochmals getrennt zwischen Tests gegen Goldstandard-Geräte und wiederum gegen die indirekte Kalorimetrie.

Bei der indirekten Kalorimetrie wurden die spiroergometrischen Geräte CORTEX METAMAX 3B, OXYCON MOBILE und PARVO TRUEONE 2400 METABOLIC ANALYZER verwendet. Als Goldstandard kamen die Akzelerometer ACTIGRAPH ACTITRAINER und GT3X, BODY-MEDIA SENSEWEAR, SENSEWEAR ARMBAND MINI und SENSEWEAR PRO ARMBAND zum Einsatz.

Erkennung der Bewegungsintensität

In diesem Kapitel wird die Korrelation bei der Erkennung von unterschiedlichen Intensitätsstufen von leichte Bewegungsformen (engl.: Light PA) (LPA) über moderate Bewegungsformen (engl.: Moderate PA) (MPA) bis zu intensive Bewegungsformen (engl.: Vigorous PA) (VPA) von Fitness Trackern zu Goldstandard-Geräten aufgezeigt. Bei den Studien wurde dafür das jeweilige berechnete MET eines Konsumentengerätes, mit dem MET des Goldstandard-Gerätes verglichen.

Die Erkennung der Intensität beherrschen die Fitness Tracker mit einem mittleren Zusammenhang ($r = 0,698$ (95 % LCI: 0,569; UCI: 0,793)) bei einer Heterogenität der Fälle von $I^2 = 39,11$ ($P = 0,1310$) verglichen mit Goldstandard-Geräten (Forest-Plot: Abb. 5.10).

Am geringsten wurden die Bewegungen (im Feldtest über 7 d, exkl. Nacht) von der POLAR V800 (Hernández-Vicente et al., 2016 [2]) erkannt ($r = 0,337$ (95 % LCI: -0,15; UCI: 0,70)). Am genauesten schnitt bei Reid et al. (2017) [2] die FITBIT FLEX mit $r = 0,872$ (95 % LCI: 0,71; UCI: 0,95) bei siebentägigen Tests, exkl. Nacht, mit min. 10,000 Schritte/d und VPA

20 min d^{-1} bis 30 min d^{-1} ab. Mit MPA 20 min d^{-1} bis 30 min d^{-1} fällt die FITBIT FLEX auf ihre ungenaueste Messung mit $r = 0,567$ (95 % LCI: 0,191; UCI: 0,798) zurück.

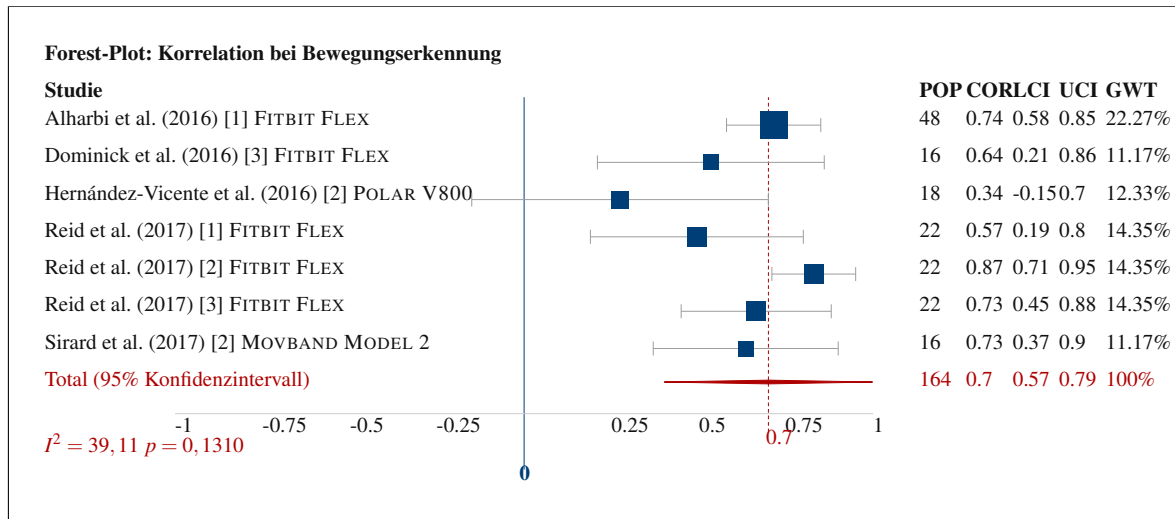


Abbildung 5.10: Forest-Plot der Gültigkeit bei der Intensitätsbewertung von Fitness Trackern in der Feldmessung verglichen mit Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Labormessung

Bei der Kalorienverbrauchsberechnung im Labor kommen Fitness Tracker auf eine niedrige Gültigkeit ($r = 0,39$ (95 % LCI: 0,225; UCI: 0,533)) bei $I^2 = 77,89$ ($P < 0,0001$) im Vergleich zur indirekten Kalorimetrie (Forest-Plot: Abb. 5.11).

Die geringste Gültigkeit, sogar mit negativer Korrelation von $r = -0,25$ (95 % LCI: -0,62; UCI: 0,22), wurde am Laufband (5 min mit 13 km h^{-1}) bei der POLAR LOOP gemessen (Wahl et al., 2017 [47]). Das beste Ergebnis der POLAR LOOP liegt noch immer bei $r = 0,02$ (95 % LCI: -0,43; UCI: 0,46) bei $7,2 \text{ km h}^{-1}$ am Laufband für 5 min (Wahl et al., 2017 [42]). Die höchste Gültigkeit wurde von der APPLE WATCH auf einem Laufband mit einem selbst gewählten Tempo für 25 min erreicht ($r = 0,88$ (95 % LCI: 0,78; UCI: 0,94)) (Bai, Hibbing et al., 2017 [1]). Die APPLE WATCH hat als geringstes Ergebnis eine Gültigkeit von $r = 0,16$ (95 % LCI: -0,28; UCI: 0,55) bei einem Test von Wallen et al. (2016) [2] am Laufband (9 min) und am Ergometer (18 min bei 25 W).

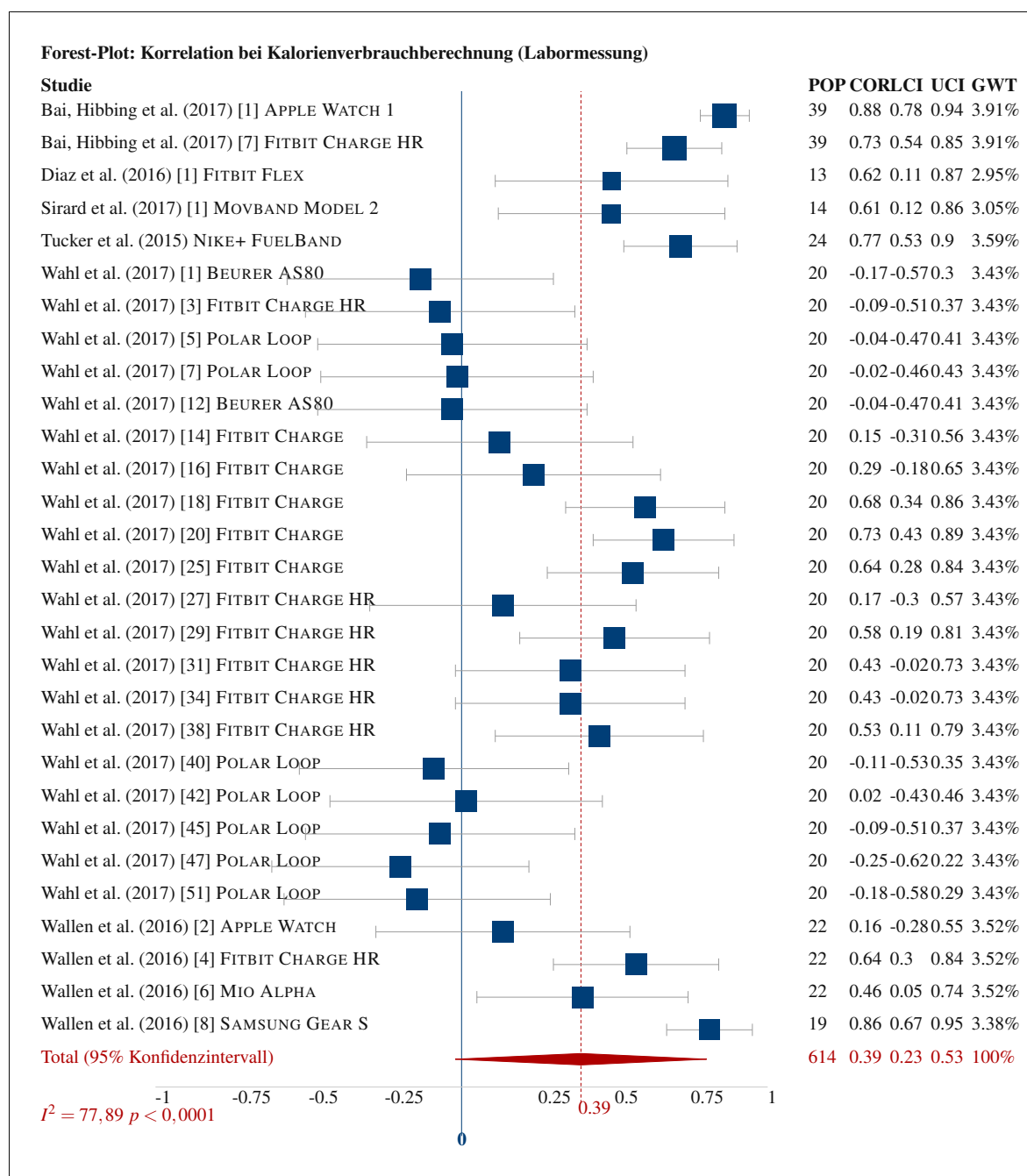


Abbildung 5.11: Forest-Plot der Gültigkeit bei der Kalorienverbrauchsberechnung von Fitness Tracker in der Labormessung verglichen mit indirekter Kalorimetrie. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Feldmessung

Im Feld besteht ein mittlerer Zusammenhang ($r = 0,543$ (95 % LCI: 0,36; UCI: 0,685)) bei einer Heterogenität $I^2 = 84,29$ ($P < 0,0001$), wenn EC mit spiroergometrischen Geräten verglichen wird (Forest-Plot: Abb. 5.12) und bei Goldstandard-Geräten dagegen ein hoher Zusammenhang ($r = 0,759$ (95 % LCI: 0,686; UCI: 0,817)) bei $I^2 = 43,3$ ($P = 0,0480$) (Forest-Plot: Abb. 5.13).

Im Vergleich der Gültigkeit der Kalorienverbrauchsschätzung gegen die indirekte Kalorimetrie (in diesem Fall CORTEX METAMAX 3B) schnitt die POLAR LOOP bei einem fünfminütigen intermittierenden Lauftest ($\varnothing 9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$) am schlechtesten mit $r = -0,30$ (95 % LCI: -0,66; UCI: 0,16) ab (Wahl et al., 2017 [49]). Am besten schnitt die FITBIT CHARGE HR mit $r = 0,84$ (95 % LCI: 0,70; UCI: 0,92) ab. Der Test dauerte durchschnittlich 89,6 min (SD 5,5 min) mit den unterschiedlichen Aktivitäten von Liegen, Sitzen, Gehen und Laufen bis 16 km h^{-1} (Montoye et al., 2017 [3]).

Bei der Messung gegen ein Goldstandard Kriterium (ACTIGRAPH ACTITRAINER) schnitt am schlechtesten die POLAR V800 ab ($r = 0,483$ (95 % LCI: 0,02; UCI: 0,78)). Hernández-Vicente et al. (2016) [1] führten einen 7 d (exkl. Nacht) dauernden Feldtest durch. Am genauesten war die POLAR LOOP bei einem zweitägigen Test inkl. einer Nacht mit $r = 0,9$ (95 % LCI: 0,80; UCI: 0,95) (Brooke et al., 2017 [7]).

5 Ergebnisse

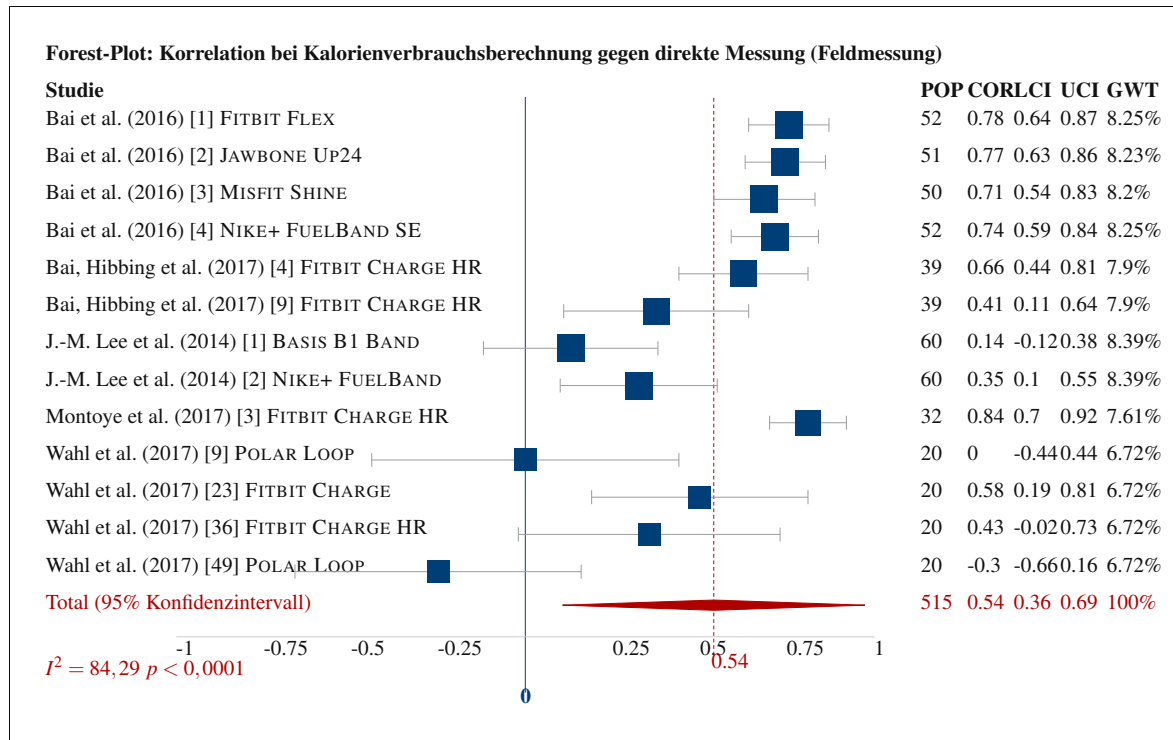


Abbildung 5.12: Forest-Plot der Gültigkeit bei der Kalorienverbrauchsberechnung von Fitness Trackern in der Feldmessung verglichen mit indirekter Kalorimetrie. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

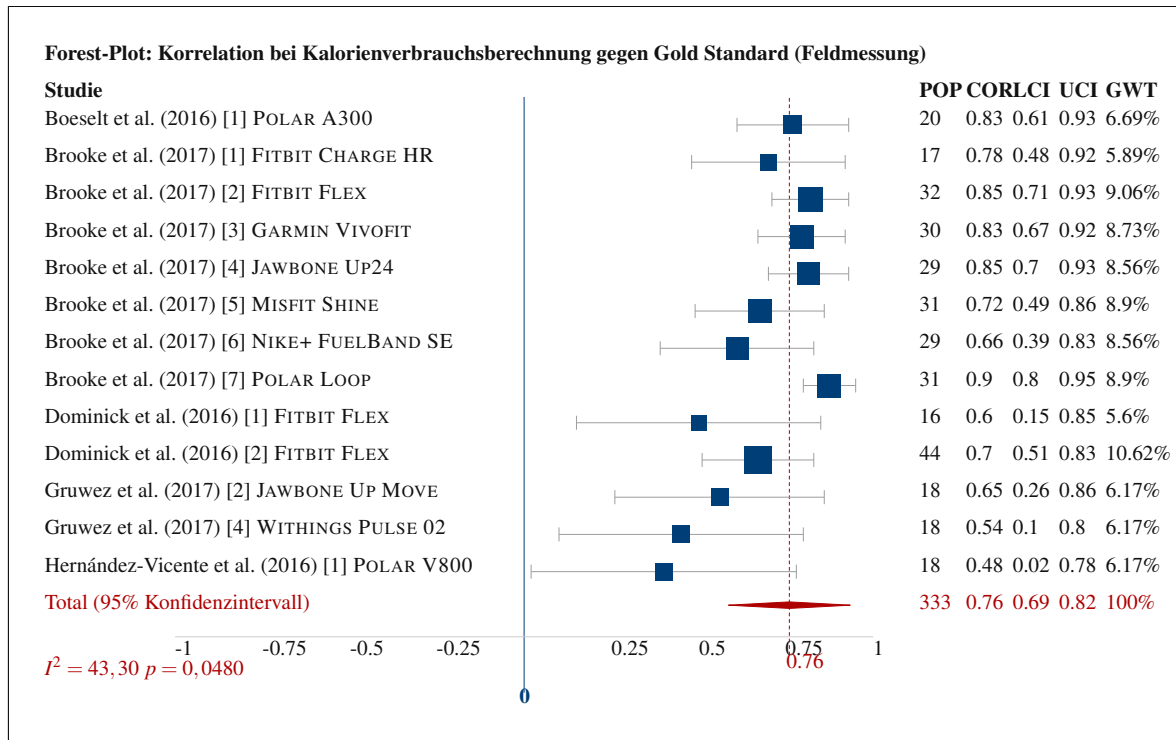


Abbildung 5.13: Forest-Plot der Gültigkeit bei der Kalorienverbrauchsberechnung von Fitness Tracker in der Labormessung verglichen mit Goldstandard-Geräten. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Tabelle 5.8: Studienergebnisse zur Gültigkeit der Kalorienverbrauchsberechnung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Alharbi et al. (2016) [1]	FITBIT FLEX	Actigraph GT3X	Feld	normaler Alltag	4 d (exkl. Nacht)	48	0,743	0,582	0,848
Bai et al. (2016) [1]	FITBIT FLEX	Oxycon Mobile v5.0	Feld	ruhige Tätigkeit (20 min), Laufband (25 min), Krafttraining (25 min), Pause (10 min)	80 min	52	0,78	0,644	0,868
Bai et al. (2016) [2]	JAWBONE UP24	Oxycon Mobile v5.0	Feld	ruhige Tätigkeit (20 min), Laufband (25 min), Krafttraining (25 min), Pause (10 min)	80 min	51	0,77	0,628	0,863
Bai et al. (2016) [3]	MISFIT SHINE	Oxycon Mobile v5.0	Feld	ruhige Tätigkeit (20 min), Laufband (25 min), Krafttraining (25 min), Pause (10 min)	80 min	50	0,71	0,538	0,825
Bai et al. (2016) [4]	NIKE+ FUEL-BAND SE	Oxycon Mobile v5.0	Feld	ruhige Tätigkeit (20 min), Laufband (25 min), Krafttraining (25 min), Pause (10 min)	80 min	52	0,74	0,585	0,843
Bai, Hibbing et al. (2017) [1]	APPLE WATCH 1	Oxycon Mobile v5.0	Labor	Laufband (selbst gewähltes Tempo)	25 min	39	0,88	0,781	0,936
Bai, Hibbing et al. (2017) [4]	APPLE WATCH 1	Oxycon Mobile v5.0	Feld	Alltagstätigkeiten im Bereich LPA	25 min	39	0,66	0,435	0,807

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Bai, Hibbing et al. (2017) [7]	FITBIT CHARGE HR	Oxycon Mobile v5.0	Labor	Laufband (selbst gewähltes Tempo)	25 min	39	0,73	0,539	0,85
Bai, Hibbing et al. (2017) [9]	FITBIT CHARGE HR	Oxycon Mobile v5.0	Feld	Alltagstätigkeiten im Bereich LPA	25 min	39	0,41	0,109	0,642
Boeselt et al. (2016) [1]	POLAR A300	Bodymedia SenseWear	Feld	normaler Alltag	3 d á 24	20	0,829	0,61	0,93
Brooke et al. (2017) [1]	FITBIT CHARGE HR	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	17	0,78	0,479	0,917
Brooke et al. (2017) [2]	FITBIT FLEX	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	32	0,85	0,712	0,925
Brooke et al. (2017) [3]	GARMIN VIVOFIT	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	30	0,83	0,67	0,916

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Brooke et al. (2017) [4]	JAWBONE UP24	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	29	0,85	0,702	0,928
Brooke et al. (2017) [5]	MISFIT SHINE	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	31	0,72	0,491	0,856
Brooke et al. (2017) [6]	NIKE+ FUEL-BAND SE	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	29	0,66	0,387	0,827
Brooke et al. (2017) [7]	POLAR LOOP	Bodymedia SenseWear Armband Mini	Feld	normaler Alltag	1 d (exkl. Nacht) + 1 d (inkl. Nacht)	31	0,9	0,801	0,951
Diaz et al. (2016) [1]	FITBIT FLEX	Indirekte Kalorimetrie	Labor	Laufband (á 6 min bei 3 km h^{-1} , $4,8 \text{ km h}^{-1}$, $6,4 \text{ km h}^{-1}$ und $8,5 \text{ km h}^{-1}$ + á 3 min Pause)	31 min	13	0,62	0,105	0,873
Dominick et al. (2016) [1]	FITBIT FLEX	Actigraph GT3X	Feld	Gehen Outdoor (als zielgerichtetes Training)	4 d von 7 d	16	0,6	0,148	0,845
Dominick et al. (2016) [2]	FITBIT FLEX	Actigraph GT3X	Feld	Laufen oder (aerober) Sport (als zielgerichtetes Training)	4d von 1w	44	0,7	0,509	0,825

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Dominick et al. (2016) [3]	FITBIT FLEX	Actigraph GT3X	Feld	Gehen Outdoor (als zielgerichtetes Training)	4d von 1w	16	0,64	0,211	0,862
Gruwez et al. (2017) [2]	JAWBONE UP MOVE	Bodymedia SenseWear Pro Armband	Feld	normaler Alltag	1 d (exkl. Nacht)	18	0,65	0,263	0,857
Gruwez et al. (2017) [4]	WITHINGS PULSE 02	Bodymedia SenseWear Pro Armband	Feld	normaler Alltag	1 d (exkl. Nacht)	18	0,54	0,097	0,804
Hernández-Vicente et al. (2016) [1]	POLAR V800	ActiGraph ActiTrainer	Feld	normaler Alltag	7 d (exkl. Nacht)	18	0,483	0,02	0,775
Hernández-Vicente et al. (2016) [2]	POLAR V800	ActiGraph ActiTrainer	Feld	normaler Alltag	7 d (exkl. Nacht)	18	0,337	-0,154	0,695
J.-M. Lee et al. (2014) [1]	BASIS B1 BAND	Oxycon Mobile v5.0	Feld	diverse Aktivitäten	69 min	60	0,136	-0,122	0,377
J.-M. Lee et al. (2014) [2]	NIKE+ FUEL-BAND	Oxycon Mobile v5.0	Feld	diverse Aktivitäten	69 min	60	0,346	0,101	0,551

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Montoye et al. (2017) [3]	FITBIT CHARGE HR	Parvo TrueOne 2400 metabolic analyzer	Feld	diverse Aktivitäten	89,6 min (SD 5,5)	32	0,84	0,695	0,919
Reid et al. (2017) [1]	FITBIT FLEX	ActiGraph GT3X	Feld	Min. 10,000 Schritte/d, MPA 20 mind ⁻¹ bis 30 mind ⁻¹	7 d (exkl. Nacht)	22	0,567	0,191	0,798
Reid et al. (2017) [2]	FITBIT FLEX	ActiGraph GT3X	Feld	Min. 10,000 Schritte/d, VPA 20 mind ⁻¹ bis 30 mind ⁻¹	7 d (exkl. Nacht)	22	0,872	0,712	0,946
Reid et al. (2017) [3]	FITBIT FLEX	ActiGraph GT3X	Feld	Min. 10,000 Schritte/d, MVPA 20 mind ⁻¹ bis 30 mind ⁻¹	7 d (exkl. Nacht)	22	0,731	0,447	0,881
Sirard et al. (2017) [1]	MOVBAND MODEL 2	Oxycon Mobile	Labor	10 Aktivitäten von Sitzen bis VPA je 7 min	70 min	14	0,61	0,117	0,862
Sirard et al. (2017) [2]	MOVBAND MODEL 2	ActiGraph GT3X	Feld	normaler Alltag	4 d (exkl. Nacht)	16	0,73	0,367	0,9
Tucker et al. (2015)	NIKE+ FUEL-BAND	Oxycon Mobile	Labor	12 Aktivitäten (4 LPA, 4 MPA, 4 VPA)	60 min	24	0,77	0,532	0,895
Wahl et al. (2017) [1]	BEURER AS80	Cortex MetaMax 3B	Labor	Laufband (4,3 km h ⁻¹)	5 min	20	-0,17	-0,57	0,295
Wahl et al. (2017) [12]	BEURER AS80	Cortex MetaMax 3B	Labor	Laufband (10,1 km h ⁻¹)	2,4 km	20	-0,04	-0,474	0,41

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [14]	FITBIT CHARGE	Cortex MetaMax 3B	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,15	-0,313	0,556
Wahl et al. (2017) [16]	FITBIT CHARGE	Cortex MetaMax 3B	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,29	-0,175	0,649
Wahl et al. (2017) [18]	FITBIT CHARGE	Cortex MetaMax 3B	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	0,68	0,34	0,863
Wahl et al. (2017) [20]	FITBIT CHARGE	Cortex MetaMax 3B	Labor	Laufband (13 km h^{-1})	5 min	20	0,73	0,425	0,886
Wahl et al. (2017) [23]	FITBIT CHARGE	Cortex MetaMax 3B	Feld	Intermittierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0,58	0,185	0,814
Wahl et al. (2017) [25]	FITBIT CHARGE	Cortex MetaMax 3B	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	2,4 km	20	0,64	0,275	0,844
Wahl et al. (2017) [27]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband ($4,3 \text{ km h}^{-1}$)	5 min	20	0,17	-0,295	0,57
Wahl et al. (2017) [29]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband ($7,2 \text{ km h}^{-1}$)	5 min	20	0,58	0,185	0,814

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [3]	BEURER AS80	Cortex MetaMax 3B	Labor	Laufband (7,2 km h ⁻¹)	5 min	20	-0,09	-0,512	0,367
Wahl et al. (2017) [31]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband (10,1 km h ⁻¹)	5 min	20	0,43	-0,015	0,733
Wahl et al. (2017) [34]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband (13 km h ⁻¹)	5 min	20	0,43	-0,015	0,733
Wahl et al. (2017) [36]	FITBIT CHARGE HR	Cortex MetaMax 3B	Feld	Intermittierendes Laufen(Ø 9,1 km h ⁻¹ Max.: 22,4 km h ⁻¹)	5 min	20	0,43	-0,015	0,733
Wahl et al. (2017) [38]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband (10,1 km h ⁻¹)	2,4 km	20	0,53	0,114	0,788
Wahl et al. (2017) [40]	POLAR LOOP	Cortex MetaMax 3B	Labor	Laufband (4,3 km h ⁻¹)	5 min	20	-0,11	-0,527	0,35
Wahl et al. (2017) [42]	POLAR LOOP	Cortex MetaMax 3B	Labor	Laufband (7,2 km h ⁻¹)	5 min	20	0,02	-0,426	0,458
Wahl et al. (2017) [45]	POLAR LOOP	Cortex MetaMax 3B	Labor	Laufband (10,1 km h ⁻¹)	5 min	20	-0,09	-0,512	0,367

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wahl et al. (2017) [47]	POLAR LOOP	Cortex MetaMax 3B	Labor	Laufband (13 km h^{-1})	5 min	20	-0,25	-0,624	0,216
Wahl et al. (2017) [49]	POLAR LOOP	Cortex MetaMax 3B	Feld	Intermittierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	-0,3	-0,656	0,164
Wahl et al. (2017) [5]	BEURER AS80	Cortex MetaMax 3B	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	5 min	20	-0,04	-0,474	0,41
Wahl et al. (2017) [51]	POLAR LOOP	Cortex MetaMax 3B	Labor	Laufband ($10,1 \text{ km h}^{-1}$)	2,4 km	20	-0,18	-0,577	0,285
Wahl et al. (2017) [7]	BEURER AS80	Cortex MetaMax 3B	Labor	Laufband (13 km h^{-1})	5 min	20	-0,02	-0,458	0,426
Wahl et al. (2017) [9]	BEURER AS80	Cortex MetaMax 3B	Feld	Intermittierendes Laufen($\emptyset$ $9,1 \text{ km h}^{-1}$ Max.: $22,4 \text{ km h}^{-1}$)	5 min	20	0	-0,443	0,443
Wallen et al. (2016) [2]	APPLE WATCH	Cortex MetaMax 3B	Labor	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,16	-0,281	0,545
Wallen et al. (2016) [4]	FITBIT CHARGE HR	Cortex MetaMax 3B	Labor	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,64	0,299	0,836

Tabelle 5.8: Fortsetzung

Autor	Fitness Tracker	Kriterium	Messumgebung	Aktivität	Messdauer	Anzahl	COR	LCI (95 %)	UCI (95 %)
Wallen et al. (2016) [6]	MIO ALPHA	Cortex MetaMax 3B	Labor	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,46	0,047	0,738
Wallen et al. (2016) [8]	SAMSUNG GEAR S	Cortex MetaMax 3B	Labor	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	19	0,86	0,666	0,945

Tabelle 5.8: Ergebnisse einzelner Fitness Tracker Studien als Korrelation zum Gültigkeitskriterium direktes Zählen oder einem Goldstandard, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar. Abkürzungen: Korrelation (COR), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

5.3.2 Zuverlässigkeit bei der Kalorienverbrauchsberechnung

Die vier Fälle, welche die Zuverlässigkeit bei der Eruierung des Kalorienverbrauches behandeln, stammen aus der Studie von Wallen et al. (2016). Die Zuverlässigkeit ist dabei mittel ($r = 0,516$ (95 % LCI: 0,0633; UCI: 0,792)) bei einer Heterogenität der Fälle von $I^2 = 79,49$ ($P = 0,0022$). Der Test erfolgte am Laufband (9 min und am Ergometer 18 min bei 25 W) für 25 min.

Die schlechteste ICC hat hier die APPLE WATCH mit $r = 0,05$ (95 % LCI: -0,38; UCI: 0,46) (Wallen et al., 2016 [R2]), die beste die SAMSUNG GEAR S mit $r = 0,86$ (95 % LCI: 0,67; UCI: 0,95) (Wallen et al., 2016 [R8]).

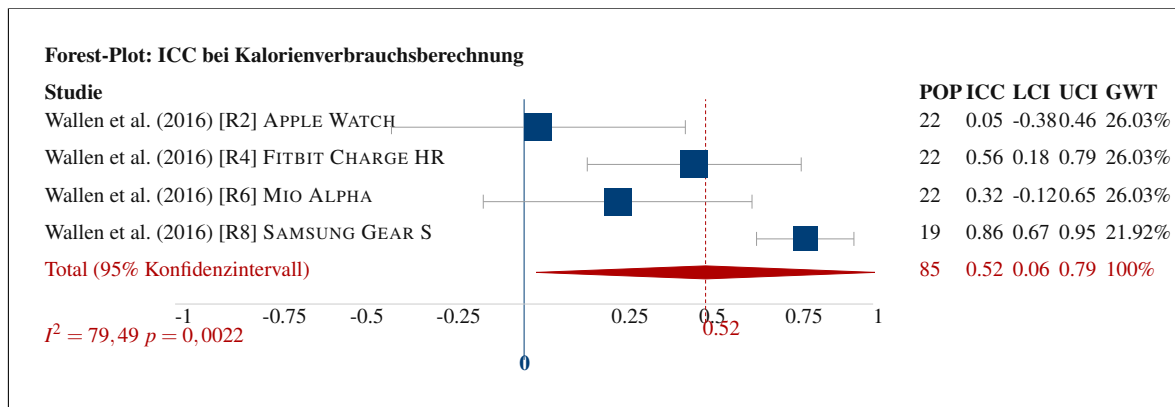


Abbildung 5.14: Forest-Plot der Zuverlässigkeit bei der Kalorienverbrauchsberechnung von Fitness Trackern im Laborumfeld. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein R vorangestellt) eindeutig identifizierbar.

Abkürzungen: Population (POP), Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI), Gewichtung (GWT)

Erzeugt mit DistillerSR Forest-Plot Generator von Evidence Partners

Tabelle 5.9: Studienergebnisse zur Zuverlässigkeit von Fitness Tracker bei der Kalorienverbrauchsberechnung

Autor	Fitness Tracker	Aktivität	Messdauer	Anzahl	ICC	LCI (95 %)	UCI (95 %)
Wallen et al. (2016) [R2]	APPLE WATCH	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,05	-0,38	0,462
Wallen et al. (2016) [R4]	FITBIT CHARGE HR	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,56	0,181	0,794
Wallen et al. (2016) [R6]	MIO ALPHA	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	22	0,32	-0,117	0,653
Wallen et al. (2016) [R8]	SAMSUNG GEAR S	Laufband (9 min & Ergometer 18 min bei 25 W)	25 min	19	0,86	0,666	0,945

Tabelle 5.9: Ergebnisse einzelner Fitness Tracker Studien als Intra-Klassen-Korrelation, innerhalb der Studien nach einzelnen Aktivitäten getrennt und nummeriert. Die einzelnen Fälle (Aktivitäten) in den Forest-Plots, in den Tabellen und im Text, sind mittels der Kombination von Studienautoren und einer nachfolgend in eckigen Klammern stehender Zahl (bei Zuverlässigkeitsstudien wird der Zahl ein *R* vorangestellt) eindeutig identifizierbar.

Abkürzungen: Intra-Klassen-Korrelation (engl.: Intra Class Correlation) (ICC), unteres Konfidenzintervall (engl.: Lower Confidence Interval) (LCI), oberes Konfidenzintervall (engl.: Upper Confidence Interval) (UCI)

6 Diskussion

Die Fitness Tracker in diesem Review bieten bei der Gültigkeit der Schrittzählung einen hohen Zusammenhang verglichen mit den Kriterien in den Labormessungen ($r = 0,909$ (95 % LCI: 0,855; UCI: 0,944)), sowie bei Feldtestungen niedrigere, aber immer noch hohe Zusammenhänge im Vergleich zu von den Forschern direkt gezählten Schritten ($r = 0,799$ (95 % LCI: 0,727; UCI: 0,854)), als auch beim Vergleich mit Goldstandard Pedometer oder Akzelerometer ($r = 0,75$ (95 % LCI: 0,605; UCI: 0,846)). Die bei $r = 0,75$ liegende Korrelation von Fitness Trackern zu Goldstandard-Geräten, könnte nach Evenson et al. (2015, S. 17) davon herführen, dass die in der Forschung verwendeten Akzelerometer oftmals aus mehreren Beschleunigungssensoren bestehen, die über den Körper aufgeteilt oder hüftnah angebracht sind im Gegensatz zu den am Handgelenk getragenen Fitness Trackern. Auch nach Evenson et al. (2015, S. 17) und Wright et al. (2017, S. 363) messen Fitness Tracker, die nahe am Körperstamm getragen werden, die Schritte besser. Hingegen kamen Ainsworth et al. (2015, S. 390) zu dem Schluss, dass Vergleichsstudien zwischen an der Hüfte und am Handgelenk getragenen Wearables nur eine geringe Differenz in deren Gültigkeit aufweisen.

Für Cadmus-Bertram (2017) sind die am Handgelenk getragenen Fitness Tracker, trotz ihrer höheren Anzahl an Messfehlern geeignet, in der Gesundheitsvorsorge eingesetzt zu werden. Meines Erachtens gilt das jedoch nur so lange, wie das Gehen uneingeschränkt möglich ist. Denn ansonsten weisen die Fitness Tracker bei Testfällen mit eingeschränktem Gehen, teilweise mit Hilfsmittel und dem dadurch veränderten Gangbild nur eine niedrige ($r = 0,399$) bis fast gar keine Gültigkeit auf ($r = -0,030$) auf (Floegel et al., 2017; Fulk et al., 2014; Treacy et al., 2017). Gerade hier wäre eine hohe Gültigkeit interessant, da Fitness Tracker auch zur Unterstützung in Rehabilitation und Therapie eingesetzt werden (Cadmus-Bertram, 2017; Kumari et al., 2017; Matthews et al., 2012; Seshadri et al., 2017; Wright et al., 2017).

Werden langsamere Bewegungsarten wie Gehen ($<5,4 \text{ kmh}^{-1}$) oder durch etwa Rollator eingeschränkte Bewegungsformen weggenommen, besitzen die Wearables eine geringfügig höhere Gültigkeit (Labormessungen: $r = 0,941$ (95 % LCI: 0,9; UCI: 0,965; Feldmessungen mit direktem Zählvergleich: $r = 0,802$ (95 % LCI: 0,698; UCI: 0,87) und Feldmessungen im Vergleich mit Goldstandard-Geräten: $r = 0,833$ (95 % LCI: 0,715; UCI: 0,905)).

Diese aufgezeigte hohe Gültigkeit besitzen die Wearables trotz nur einer mittleren ICC mit $r = 0,589$ (95 % LCI: 0,361; UCI: 0,751), die in Labormessungen eruiert wurde. In der Feldumgebung bei nur zwei Fällen mit der FITBIT FLEX besitzt diese eine mittlere ICC ($r = 0,57$ (95 % LCI: 0,226; UCI: 0,788)) bei Gehgeschwindigkeiten von 5 kmh^{-1} bis

$6,5 \text{ km h}^{-1}$ und bei höheren Geschwindigkeiten wie die des Joggings (8 km h^{-1} bis 10 km h^{-1}) sogar nur eine niedrige ICC ($r = 0,26$ (95 % LCI: $-0,151$; UCI: $0,594$)). Alinia et al. (2017, S. 7) vermuten, dass die im Laborumfeld besseren Ergebnisse bei der Gültigkeit davon herführen können, dass sich Menschen beim Gehen auf einem Laufband anders bewegen. So kann es sein, dass sie zusätzliche Ausgleichs- und Balancebewegungen beim Gehen machen, welche die Zählmechanik für Schritte im Tracker unterstützen.

Insgesamt haben Fitness Tracker bei der Herzfrequenzmessung eine hohe Gültigkeit von $r = 0,92$ (95 % LCI: $0,88$; UCI: $0,942$). Im Vergleich mit Herzfrequenzmessungen über den Brustgurt von handelsüblichen Geräten wie bei Abt et al. (2017), Bai, Ellingson und Welk (2017) und Stahl et al. (2016), liegt die Gültigkeit in einem hohen Bereich von $r = 0,84$ (95 % LCI: $0,71$; UCI: $0,91$) bei der FITBIT CHARGE HR (Bai, Ellingson & Welk, 2017) bis $r = 0,98$ (95 % LCI: $0,94$; UCI: $0,99$) bei der APPLE WATCH (Abt et al., 2017). Interessant ist hier zu bemerken, dass dabei die Apple-Smartwatch am rechten Handgelenk getragen wurde. Bei der Trageweise am linken Handgelenk, liegt die Gültigkeit nur noch bei $r = 0,87$ (95 % LCI: $0,63$; UCI: $0,96$) (Abt et al., 2017). Die Fitness Tracker BASIS PEAK (Jo et al., 2016) und APPLE WATCH (Wallen et al., 2016) besitzen im Vergleich zu einem 12-Kanal-EKG eine Korrelation über $r > 0,9$ ($r = 0,92$ (95 % LCI: $0,82$; UCI: $0,97$) bzw. $r = 0,95$ (95 % LCI: $0,88$; UCI: $0,98$)). Ebenfalls wurde von (Wallen et al., 2016) die SAMSUNG GEAR S getestet, die nur noch eine Gültigkeit mit $r = 0,67$ (95 % LCI: $0,82$; UCI: $0,35$) aufweist. Die Diskrepanz zwischen der hohen Korrelation der APPLE WATCH und die nur mittlere der SAMSUNG GEAR S, kann vermutlich auf die Geräte selbst zurückgeführt werden, da beide Geräte aus der gleichen Studie, bei gleicher Test-Aktivität und identem 12-Kanal-EKG als Vergleichskriterium, stammen.

Bei den Pulsmessung in Feldumgebungen, wurde in drei Studien ausschließlich die FITBIT CHARGE HR getestet. Hier ist zu bemerken, dass dieser Tracker bei Vergleichsmessungen mit einem in der Medizin verwendeten Pulsoximeter (NONIN PURESAT PULSE OXIMETER) eine höhere Gültigkeit besitzt ($r = 0,9$ (95 % LCI: $0,8$; UCI: $0,95$) (Montoye et al., 2017)) als bei Studien, die den Tracker mit handelsüblichen Herzfrequenzmessungen über einen Brustgurt (POLAR HR SENSOR) verglichen (Bai, Ellingson und Welk, 2017; Gorny et al., 2017).

Carpenter und Frontera (2016) vermuten, dass sich die Messgenauigkeiten der Herzfrequenz verbessern ließe, wenn die PPG mittels Elektrokardiogramm über eine Elektrode kombiniert wird. Zusätzlich könnte hier auch die Herzfrequenzvariabilität, also die Änderung der Herzfrequenz durch unterschiedliche Belastungsstufen, in Echtzeit aufgenommen werden. Eine weitere Möglichkeit, die von Delgado-Gonzalo et al. (2015) aufgezeigt wurde, wäre die Gewichtsverringerung der am Handgelenk getragenen Geräte. Obwohl die Daten von der Brustgurtmessung bei höherem Puls genauer und weniger sensibel für Bewegungsartefakte sind (Li et al., 2016, S. 76), verwenden nach Parak und Korhonen (2014) und Carpenter und

Frontera (2016) immer mehr aktuelle Fitness Tracker die optische Technik der PPG zum Messen der Herzfrequenz an der Hautoberfläche. LEDs und Photodetektoren würden, durch die reduzierte Masse der Tracker, weniger den Messungen verfälschenden Bewegungen, durch hervorgerufene Kräfte auf der Hautoberfläche, ausgesetzt sein. Diese Verbesserung könnte sich auch positiv auf die integrierten Beschleunigungssensoren auswirken.

Während Fitness Tracker bei den unterschiedlichen Kategorien der Schritt- und Herzfrequenzmessung immer einen hohen Zusammenhang der Gültigkeiten (Beginnend bei $r = 0,75$ bei der Schrittmessung im Feld bis zu $r = 0,92$ bei der Pulsmessgenauigkeit in Laborumgebungen) aufweisen, sieht es bei den Bestimmungen des Kalorienverbrauches anders aus. Hier weist gerade in der Feldmessung, bei der die beiden anderen Testmethoden ihre schlechtesten Gültigkeiten aufweisen, die Kalorienverbrauchsberechnung ihren besten Zusammenhang der Gültigkeit im Vergleich zu Goldstandard-Geräten $r = 0,759$ (95 % LCI: 0,686; UCI: 0,817) auf. Neben der mittleren Gültigkeit im Vergleich zu spiroergometrischen Messverfahren im Feld ($r = 0,543$ (95 % LCI: 0,36; UCI: 0,685)) weist die Gültigkeit bei der Kalorienverbrauchsberechnung im Laborbereich, mit Spiroergometrie als Vergleichskriterium, nur einen niedrigen Zusammenhang ($r = 0,39$ (95 % LCI: 0,225; UCI: 0,533)) auf.

Ein Grund für diese niedrigere Korrelationen bei Kalorienverbrauchsberechnungen zu spiroergometrischen Verfahren (Feld: $r = 0,543$ und Labor: $r = 0,39$), kann die bei einer Korrelation von $r = 0,698$ (95 % LCI: 0,569; UCI: 0,793) liegenden Erkennungsrate der Bewegungsintensität und Bewegungsart, als Berechnungsgrundlage für den Kalorienverbrauch bei Fitness Trackern, sein. Kumari et al. (2017) merken an, dass sich einerseits nicht einmal die selbe Person bei gleichen Aktivitäten gleich bewegt und andererseits, dass manche unterschiedliche Aktivitäten ähnliche Bewegungsmuster aufweisen. Dies kann zu einer falschen Interpretierung bei Fitness Trackern führen, vornehmlich bei denjenigen, die ausschließlich Daten von Akzelerometern verwenden. Vermutlich liegt darin auch der Grund, der nur in der Kalorienverbrauchsberechnung bestehenden besseren Korrelation in einer Feldumgebung zur Laborumgebung (Feld: $r = 0,759$ versus Labor: $r = 0,39$), da die als Vergleichskriterium verwendeten Goldstandard-Geräte selber ausschließlich Akzelerometer sind.

Abhilfe können andere Regressionsgleichungen schaffen, die unterschiedliche Arten von körperlicher Bewegungen erkennen sollen. So erkennen nach Kate, Swartz, Welch und Strath (2016, S. 373) die statistischen Techniken Support Vector Machine, neuronale Netzwerke und logistische Regression typische Alltagsbewegungen wie Staubsaugen, Aufwischen, Umzugskartons tragen und Stiegen steigen besser als lineare Regressionsmodelle.

Auch die Zuverlässigkeit der Fitness Tracker hat bei EC nur eine ICC von $r = 0,516$ (95 % LCI: 0,0633; UCI: 0,792). Die Gründe für die geringe Gültigkeit der Tracker bei Kalorienverbrauchsberechnungen könnten zum Teil durch diese mittlere ICC erklärbar sein. Hingegen wurden bei der Schrittmessung, bei ebenfalls mittlerer ICC ($r = 0,589$), hohe Korrelationen bei der Gültigkeit errechnet.

Mit 23 Studien zur Schrittzählung, ist dies die häufigste Klasse der Messmethoden, zum Gegensatz der Kalorienverbrauchsberechnung mit 17 und Pulsmessung mit sieben Studien, in dieser Arbeit. Ein Grund hierfür könnte die bereits seit längerem bestehende Möglichkeit sein, die Schrittzählung mit Pedometer in kleine Gehäuse zu verpacken.

Insgesamt konnten 33 Studien in die Metaanalyse einbezogen werden. Dies ist nur ein Teil der Studien, die verfügbar gewesen wären. Immerhin mussten weitere 102 Studien für diese Arbeit, bei denen der Titel anfangs vielversprechend war, verworfen werden. Die Gründe hierfür waren unterschiedlich. Einerseits konnten 71 Volltexte nicht abgerufen werden. Davon sind 33, mit für die Fragestellungen passend wirkenden Angaben, Zusammenfassungen von Tagungen, gesammelt in dem Journal *Medicine & Science in Sports & Exercise*. Aus den restlichen 31 Studien wurden einige nicht verwendet, da sie die Ergebnisse als MAPE (Prozentsatz des absoluten Fehlers des Erwartungswertes, engl.: Mean Absolute Percentage Error) und nicht als Korrelation darstellten. Die systematische Betrachtung des MAPEs wäre eine Möglichkeit die Genauigkeit von Fitness Trackern zu bestimmen.

Bei der Wahl eines Fitness Trackers ist die Überlegung wichtig, welche Informationen gesammelt werden sollen (Stackpool, 2013). So kann, wenn jemandem die gezählten Schritte wichtig sind, die MI BAND 2 (Tam & Cheung, 2018) empfehlenswert sein. Wenn es jedoch die Messung der Herzfrequenz ist, die APPLE WATCH (Abt et al., 2017) oder wenn es der tägliche Kalorienverbrauch ist, kann wiederum die POLAR LOOP (Brooke et al., 2017) empfehlenswerter sein.

Eine Möglichkeit für die unterschiedlichen Messergebnisse zwischen den Geräten können die Algorithmen zur Berechnung der Daten, die proprietär und somit oftmals nicht einsehbar sind (Cadmus-Bertram, 2017; Hernández-Vicente et al., 2016; Wright et al., 2017), sein. So vermuten Dominick et al. (2016) aufgrund einer Überschätzung des MET bei FITBIT-Tracker beim Gehen und Laufen zum Vergleich von ACTIGRAPH-Daten unterschiedliche Algorithmen. Auch Bassett, Troiano, McClain und Wolff (2015) sehen in den unterschiedlichen Algorithmen Probleme, Studien untereinander vergleichen zu können. Evenson et al. (2015, S. 19) bringen für diesen Umstand den Vorschlag an die Hersteller von Wearables, mehr Daten, auch die Roh-Daten aus den Fitness Trackern bereitzustellen. So wird nach Evenson et al. (2015) z.B.: der Datenexport bei FITBIT-Trackern eingeschränkt. Weiters meinen sie auch, wäre es schon ein guter Anfang, die Information bereitzustellen, welche Sensoren und Daten für die Ergebnisse miteinbezogen werden. Zum Beispiel ist nach Bai, Hibbing et al. (2017) bei den Fitness Trackern APPLE WATCH 1 und FITBIT CHARGE HR nicht bekannt, welche Sensoren für die Messgrundlage beim Kalorienverbrauch herangezogen werden. Bei JAWBONE-Trackern schreiben Evenson et al. (2015), dass in dessen Benutzerhandbuch zumindest eine Angabe gemacht wird, welche demographischen Daten und Sensoren, wenn vorhanden, in die Kalorienmessung miteinbezogen werden. Beg, Gupta, Stewart und Rethorst (2017, S. 86–87) meinen, dass kommerzielle Wearables mit proprietärer Software, wenn sie

miteinander verglichen oder in der Forschung eingesetzt werden sollen, firmenübergreifende Standards bräuchten.

Wright et al. (2017, S. 362) und Evenson et al. (2015, S. 19) bemerken, dass durch Firmwareupdates, welche die Fitness Tracker jederzeit erhalten könnten, z. B.: wenn sie die Fitnessdaten mit einer Datenbank im Internet synchronisieren, die Algorithmen zur Berechnung der Daten geändert werden können. In der in dieser Arbeit verwendeten Qualitätsbewertung wurde aus diesem Grund der Punkt der Angabe einer Firmwareversion hinzugezogen, jedoch fehlen bei allen in dieser Metaanalyse einbezogenen Studien, bis auf einer von Abt et al. (2017), Angaben über die Version der Firmware oder der Software des Fitness Trackers.

Aber nicht nur kurzfristige Softwareupdates können die Messsysteme der Fitness Tracker verändern, sondern auch neue Produkte mit verbesserter Hardware. Diese zukünftigen Geräte wären dann z.B. in der Lage den Schweiß durch chemische Sensoren auf der Haut zu messen und physiologische Rückschlüsse zu ziehen (Seshadri et al., 2017) oder durch bessere Mustererkennungsalgorithmen den Kalorienverbrauch besser zu bestimmen (Wright et al., 2017). Seshadri et al. (2017) schreiben, dass Fitness Tracker im Sport prophylaktisch wirken können, da sie missgünstige Bewegungen frühzeitig erkennen und den Sportler warnen können.

Die in dieser Arbeit anzutreffende Aufteilung in Labor- und Feldtestungen bei den drei unterschiedlichen Messtypen birgt die Möglichkeit, dass einzelne spezielle Sportarten bzw. Eingrenzungen von Sportarten (Intensität, Geschwindigkeiten, etc.) in der Menge der Daten untergehen. So kommen die Sportarten, außer in wenigen Fällen und da wiederum nur als Teil der Gesamtaktivität, nur im Bereich der Ausdauersportarten vor. Auch kann die Aufteilung der Studien auf einzelne Fälle, die in den jeweiligen Studien getestet wurden, die Ergebnisse verfälschen, da einzelne Studien mit mehreren Fällen so verstärkt in die Metaanalyse eingeflossen sind. Spezielle Aktivitätsbereiche können, wenn diese Arbeit als PDF-Datei vorliegt, aus den Rechercherothdaten im Anhang (Kapitel 6) als CSV-Tabellen eingebettet, nochmals genauer betrachtet werden.

Zusammengefasst sei erwähnt, dass sich nach den Ergebnissen in dieser Arbeit, Wearables zur Herzfrequenzmessung im Sport eignen. Bei Personen ohne eingeschränktem Gangbild eignen sie sich auch für die Schrittzählung beim Gehen oder bei moderaten Ausdauersportarten. Für die Kalorienverbrauchsberechnung oder die Aktivitätserkennung sind Fitness Tracker dagegen noch weniger geeignet.

Glossar

Body Mass Index

Der Body Mass Index (BMI) ist eine Relation zwischen Körpergröße l in m und Körpermasse m in kg ($m \div l^2$). Normwerte liegen bei Erwachsenen bei $18,5 \text{ kg m}^{-2}$ bis $<25 \text{ kg m}^{-2}$.

Feldstudie

Feldmessungen sind Messungen im Alltag und auch Studien, die versuchen den Alltag durch künstlich erschaffenen Umgebungen zu simulieren, oder Studien, die Bewegungsbereiche angeben, die nicht klar zu reproduzieren sind.

Fitness Tracker

Das sind jene Geräte, welche unterschiedliche physikalische und physiologische Daten bei der Bewegung einer Person messen.

Goldstandard

Darunter versteht man die zum Zeitpunkt dieser Arbeit in der Forschung mehrfach validierten Geräte, die als Vergleichskriterium zu den kommerziellen Fitness Trackern zum Einsatz kamen.

Korrelationskoeffizient

Der Korrelationskoeffizient r zeigt bei $r = 0$ keinen, bei $0 < r \leq 0,4$ einen niedrigen, bei $0,4 < r \leq 0,7$ einen mittleren, bei $0,7 < r \leq 1,0$ einen hohen und bei $r = 1$ einen vollständigen Zusammenhang.

Laborstudie

Laborstudien sind jene Studien, die eine klar vorgegebene Aktivität (z.B.: Geschwindigkeit durch ein Laufband) aufweisen.

leichte Bewegungsformen (engl.: Light PA)

Unter 3 MET liegen leichte Bewegungsformen (engl.: Light PA) (LPA). Hierzu gehören z.B. Schlafen, sitzende Tätigkeiten (Fernsehen, Arbeit am Computer) und langsames Gehen (bis etwa 4 km h^{-1}).

metabolisches Äquivalent (engl.: Metabolic Equivalent Task)

Das metabolische Äquivalent (engl.: Metabolic Equivalent Task) (MET) wird verwendet um die Intensität von Bewegungen zu bestimmen. 1 MET entspricht einem Energieverbrauch von $1 \text{ kcal kg}^{-1} \text{ min}^{-1}$ und entspricht den Ruheumsatz eines Menschen.

moderate Bewegungsformen (engl.: Moderate PA)

In einem Bereich von 3 bis 6 MET liegen moderate Bewegungsformen (engl.: Moderate PA) (MPA). Hierzu zählen z.B.: Fahrradfahren (100 W oder $<16 \text{ km h}^{-1}$).

Physical Activity

Physical Activity (PA) ist der englische Begriff für körperliche Aktivität und kann nochmals in die unterschiedlichen Intensitäten LPA, MPA und VPA unterteilt werden.

Photoplethysmography

Photoplethysmography (PPG) ist eine Technik, welche mittels einer Lichtbestrahlung die Durchflussrate von Blutgefäßen bestimmt.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) wird verwendet um eine Reihe konzeptueller und praktischer Fortschritte bei systematischen Übersichten zu adressieren (Moher, Liberati, Tetzlaff & Altman, 2009).

intensive Bewegungsformen (engl.: Vigorous PA)

Über 6 MET liegen intensive Bewegungsformen (engl.: Vigorous PA) (VPA). Hierzu zählen Callisthenics, Jogging, Laufen und Schnurspringen.

Wearables

siehe Fitness Tracker oben

Abkürzungsverzeichnis

API	Programmierschnittstelle (engl.: Application Programming Interface)
BMI	Body Mass Index
EC	Energieverbrauch (engl.: Energy Consumption)
EKG	Elektrokardiogramm
GNSS	globales Navigationssatellitensystem (engl.: Global Navigation Satellite System)
GPS	Global Positioning System
ICC	Intra-Klassen-Korrelation (engl.: Intra Class Correlation)
LCI	unteres Konfidenzintervall (engl.: Lower Confidence Interval)
LPA	leichte Bewegungsformen (engl.: Light PA)
MEMS	Mikrosysteme (engl.: MicroelEctroMechanical Systems)
MET	metabolisches Äquivalent (engl.: Metabolic Equivalent Task)
MPA	moderate Bewegungsformen (engl.: Moderate PA)
MVPA	moderate bis intensive Bewegungsformen (engl.: Moderate to Vigorous PA)
PA	Physical Activity
PPG	Photoplethysmography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	randomisierte kontrollierte Studie (engl.: Randomiced Controlled Trial)
SD	Standardabweichung
UCI	oberes Konfidenzintervall (engl.: Upper Confidence Interval)
VPA	intensive Bewegungsformen (engl.: Vigorous PA)

Literatur

- Abt, G., Bray, J. & Benson, A. C. (2017). The validity and inter-device variability of the Apple WatchTM for measuring maximal heart rate. *Journal of sports sciences*, 1–6. doi:10.1080/02640414.2017.1397282
- Ainsworth, B., Cahalin, L., Buman, M. & Ross, R. (2015). The current state of physical activity assessment tools. *Progress in cardiovascular diseases*, 57(4), 387–395. doi:10.1016/j.pcad.2014.10.005
- Alharbi, M., Bauman, A., Neubeck, L. & Gallagher, R. (2016). Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. *European journal of preventive cardiology*, 23(14), 1476–1485. doi:10.1177/2047487316634883
- Alharbi, M., Straiton, N. & Gallagher, R. (2017). Harnessing the Potential of Wearable Activity Trackers for Heart Failure Self-Care. *Current heart failure reports*, 14(1), 23–29. doi:10.1007/s11897-017-0318-z
- Alinia, P., Cain, C., Fallahzadeh, R., Shahrokni, A., Cook, D. & Ghasemzadeh, H. (2017). How Accurate Is Your Activity Tracker? A Comparative Study of Step Counts in Low-Intensity Physical Activities. *JMIR mHealth and uHealth*, 5(8), e106. doi:10.2196/mhealth.6321
- An, H.-S., Jones, G. C., Kang, S.-K., Welk, G. J. & Lee, J.-M. (2017). How valid are wearable physical activity trackers for measuring steps? *European journal of sport science*, 17(3), 360–368. doi:10.1080/17461391.2016.1255261
- Bai, Y., Ellingson, L. & Welk, G. (2017). The Validity of Fitbit Charge in Free Living Conditions. *Medicine & Science in Sports & Exercise*, 49, 529–530. doi:10.1249/01.mss.0000518364.62548.c9
- Bai, Y., Hibbing, P., Mantis, C. & Welk, G. J. (2017). Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *Journal of sports sciences*, 1–8. doi:10.1080/02640414.2017.1412235
- Bai, Y., Welk, G. J., Nam, Y. H., Lee, J. A., Lee, J.-M., Kim, Y., ... Dixon, P. M. (2016). Comparison of Consumer and Research Monitors under Semistructured Settings. *Medicine and science in sports and exercise*, 48(1), 151–158. doi:10.1249/MSS.0000000000000727
- Banerjee, T., Peterson, M., Oliver, Q., Froehle, A. & Lawhorne, L. (2017). Validating a commercial device for continuous activity measurement in the older adult population for dementia management. *Smart Health*. doi:10.1016/j.smhl.2017.11.001

- Bartel, O. & Kuhn, E. (1904). Pedometer. *USA Patent*, (US758405A).
- Bassett, D. R. (2012). Device-based monitoring in physical activity and public health research. *Physiological measurement*, 33(11), 1769–1783. doi:10.1088/0967-3334/33/11/1769
- Bassett, D. R., Troiano, R. P., McClain, J. J. & Wolff, D. L. (2015). Accelerometer-based physical activity: Total volume per day and standardized measures. *Medicine and science in sports and exercise*, 47(4), 833–838. doi:10.1249/MSS.0000000000000468
- Battenberg, A. K., Donohoe, S., Robertson, N. & Schmalzried, T. P. (2017). The accuracy of personal activity monitoring devices. *Seminars in Arthroplasty*, 28(2), 71–75. doi:10.1053/j.sart.2017.07.006
- Beekman, E., Braun, S. M., Ummels, D., van Vijven, K., Moser, A. & Beurskens, A. J. (2017). Validity, reliability and feasibility of commercially available activity trackers in physical therapy for people with a chronic disease: A study protocol of a mixed methods research. *Pilot and feasibility studies*, 3(64), 2–10. doi:10.1186/s40814-017-0200-5
- Beg, M. S., Gupta, A., Stewart, T. & Rethorst, C. D. (2017). Promise of Wearable Physical Activity Monitors in Oncology Practice. *Journal of oncology practice*, 13(2), 82–89. doi:10.1200/JOP.2016.016857
- Blümle, A., von Elm, E., Antes, G. & Meerpohl, J. J. (2014). Messung und Bewertung der Studienqualität und Berichtsqualität. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 108(8-9), 495–503. doi:10.1016/j.zefq.2014.09.022
- Boeselt, T., Spielmanns, M., Nell, C., Storre, J. H., Windisch, W., Magerhans, L., . . . Koczulla, A. R. (2016). Validity and Usability of Physical Activity Monitoring in Patients with Chronic Obstructive Pulmonary Disease (COPD). *PloS one*, 11(6), e0157229. doi:10.1371/journal.pone.0157229
- Brooke, S. M., An, H.-S., Kang, S.-K., Noble, J. M., Berg, K. E. & Lee, J.-M. (2017). Concurrent Validity of Wearable Activity Trackers Under Free-Living Conditions. *Journal of strength and conditioning research*, 31(4), 1097–1106. doi:10.1519/JSC.0000000000001571
- Cadmus-Bertram, L. (2017). Using Fitness Trackers in Clinical Research: What Nurse Practitioners Need to Know. *The journal for nurse practitioners : JNP*, 13(1), 34–40. doi:10.1016/j.nurpra.2016.10.012
- Carpenter, A. & Frontera, A. (2016). Smart-watches: A potential challenger to the implantable loop recorder? *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*, 18(6), 791–793. doi:10.1093/europace/euv427
- Case, M. A., Burwick, H. A., Volpp, K. G. & Patel, M. S. (2015). Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA*, 313(6), 625–626. doi:10.1001/jama.2014.17841

- Chiauzzi, E., Rodarte, C. & DasMahapatra, P. (2015). Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC medicine*, 13(77), 1–6. doi:10.1186/s12916-015-0319-2
- Chowdhury, E. A., Western, M. J., Nightingale, T. E., Peacock, O. J. & Thompson, D. (2017). Assessment of laboratory and daily energy expenditure estimates from consumer multi-sensor physical activity monitors. *PloS one*, 12(2), e0171720. doi:10.1371/journal.pone.0171720
- Dannecker, K. L., Sazonova, N. A., Melanson, E. L., Sazonov, E. S. & Browning, R. C. (2013). A comparison of energy expenditure estimation of several physical activity monitors. *Medicine and science in sports and exercise*, 45(11), 2105–2112. doi:10.1249/MSS.0b013e318299d2eb
- Delgado-Gonzalo, R., Parak, J., Tarniceriu, A., Renevey, P., Bertschi, M. & Korhonen, I. (2015). Evaluation of accuracy and reliability of PulseOn optical heart rate monitoring device. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2015*, 430–433. doi:10.1109/EMBC.2015.7318391
- Diaz, K. M., Krupka, D. J., Chang, M. J., Peacock, J., Ma, Y., Goldsmith, J., ... Davidson, K. W. (2015). Fitbit®: An accurate and reliable device for wireless physical activity tracking. *International journal of cardiology*, 185, 138–140. doi:10.1016/j.ijcard.2015.03.038
- Diaz, K. M., Krupka, D. J., Chang, M. J., Shaffer, J. A., Ma, Y., Goldsmith, J., ... Davidson, K. W. (2016). Validation of the Fitbit One® for physical activity measurement at an upper torso attachment site. *BMC research notes*, 9, 213. doi:10.1186/s13104-016-2020-8
- Dominick, G. M., Winfree, K. N., Pohlig, R. T. & Papas, M. A. (2016). Physical Activity Assessment Between Consumer- and Research-Grade Accelerometers: A Comparative Study in Free-Living Conditions. *JMIR mHealth and uHealth*, 4(3), 1–11. doi:10.2196/mhealth.6281
- Elgendi, M. (2012). On the Analysis of Fingertip Photoplethysmogram Signals. *Current Cardiology Reviews*, 8(1), 14–25. doi:10.2174/157340312801215782
- Evenson, K. R., Goto, M. M. & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The international journal of behavioral nutrition and physical activity*, 12(159), 1–22. doi:10.1186/s12966-015-0314-1
- Floegel, T. A., Florez-Pregonero, A., Hekler, E. B. & Buman, M. P. (2017). Validation of Consumer-Based Hip and Wrist Activity Monitors in Older Adults With Varied Ambulatory Abilities. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 72(2), 229–236. doi:10.1093/gerona/glw098

- Fulk, G. D., Combs, S. A., Danks, K. A., Nirider, C. D., Raja, B. & Reisman, D. S. (2014). Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. *Physical therapy*, 94(2), 222–229. doi:10.2522/ptj.20120525
- Gelman, A. (2005). Analysis of variance? Why it is more important than ever. *The Annals of Statistics*, 33(1), 1–53. doi:10.1214/009053604000001048
- Gomersall, S. R., Ng, N., Burton, N. W., Pavey, T. G., Gilson, N. D. & Brown, W. J. (2016). Estimating Physical Activity and Sedentary Behavior in a Free-Living Context: A Pragmatic Comparison of Consumer-Based Activity Trackers and ActiGraph Accelerometry. *Journal of medical Internet research*, 18(9), e239. doi:10.2196/jmir.5531
- Gorny, A. W., Liew, S. J., Tan, C. S. & Müller-Riemenschneider, F. (2017). Fitbit Charge HR Wireless Heart Rate Monitor: Validation Study Conducted Under Free-Living Conditions. *JMIR mHealth and uHealth*, 5(10), e157. doi:10.2196/mhealth.8233
- Gruwez, A., Libert, W., Ameye, L. & Bruyneel, M. (2017). Reliability of commercially available sleep and activity trackers with manual switch-to-sleep mode activation in free-living healthy individuals. *International journal of medical informatics*, 102, 87–92. doi:10.1016/j.ijmedinf.2017.03.008
- Hayward, J. (2018). Wearable Technology 2018-2028: Markets, Players, Forecasts: IDTechEx. Zugriff unter <https://www.idtechex.com/research/reports/wearable-technology-2018-2028-markets-players-forecasts-000606.asp>
- Hedges, L. V. & Olkin, I. (1985). *Statistical Method for Meta-Analysis*. Orlando: Academic Press.
- Hernández-Vicente, A., Santos-Lozano, A., de Cocker, K. & Garatachea, N. (2016). Validation study of Polar V800 accelerometer. *Annals of translational medicine*, 4(15), 278. doi:10.21037/atm.2016.07.16
- Hickey, A., Newham, J., Slawinska, M. M., Kwasnicka, D., McDonald, S., Del Din, S., ... Godfrey, A. (2016). Estimating cut points: A simple method for new wearables. *Maturitas*, 83, 78–82. doi:10.1016/j.maturitas.2015.10.003
- Higgins, P. A. & Straub, A. J. (2006). Understanding the error of our ways: Mapping the concepts of validity and reliability. *Nursing outlook*, 54(1), 23–29. doi:10.1016/j.outlook.2004.12.004
- Jann, J. (2016). DACH-Studie 2016: Digitale Mediennutzung in der DACH-Region 2016. Zugriff 29. November 2017 unter www.bvdw.org
- Jo, E., Lewis, K., Directo, D., Kim, M. J. & Dolezal, B. A. (2016). Validation of Biofeedback Wearables for Photoplethysmographic Heart Rate Tracking. *Journal of sports science & medicine*, 15(3), 540–547.
- Kate, R. J., Swartz, A. M., Welch, W. A. & Strath, S. J. (2016). Comparative evaluation of features and techniques for identifying activity type and estimating energy cost from accelerometer data. *Physiological measurement*, 37(3), 360–379. doi:10.1088/0967-3334/37/3/360

- Kirwan, M., Duncan, M. & Vandelanotte, C. (2013). Smartphone apps for physical activity: A systematic review. *Journal of science and medicine in sport*, 16, e47. doi:10.1016/j.jsams.2013.10.113
- Kooiman, T. J. M., Dontje, M. L., Sprenger, S. R., Krijnen, W. P., van der Schans, C. P. & de Groot, M. (2015). Reliability and validity of ten consumer activity trackers. *BMC sports science, medicine & rehabilitation*, 7, 24. doi:10.1186/s13102-015-0018-5
- Kumari, P., Mathew, L. & Syal, P. (2017). Increasing trend of wearables and multimodal interface for human activity monitoring: A review. *Biosensors & bioelectronics*, 90, 298–307. doi:10.1016/j.bios.2016.12.001
- Lauritzen, J., Muñoz, A., Luis Sevillano, J. & Civit, A. (2013). The usefulness of activity trackers in elderly with reduced mobility: A case study. *Studies in health technology and informatics*, 192, 759–762.
- Lee, J. A., Williams, S. M., Brown, D. D. & Laurson, K. R. (2015). Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *Journal of sports sciences*, 33(10), 991–1000. doi:10.1080/02640414.2014.981848
- Lee, J.-M., Kim, Y. & Welk, G. J. (2014). Validity of consumer-based physical activity monitors. *Medicine and science in sports and exercise*, 46(9), 1840–1848. doi:10.1249/MSS.0000000000000287
- Lemay, M., Bertschi, M., Sola, J., Renevey, P., Parak, J. & Korhonen, I. (2014). Application of Optical Heart Rate Monitoring. In E. Sazonov & M. R. Neuman (Hrsg.), *Wearable sensors* (S. 105–129). doi:10.1016/B978-0-12-418662-0.00023-4
- Li, R. T., Kling, S. R., Salata, M. J., Cupp, S. A., Sheehan, J. & Voos, J. E. (2016). Wearable Performance Devices in Sports Medicine. *Sports health*, 8(1), 74–78. doi:10.1177/1941738115616917
- Matthews, C. E., Hagströmer, M., Poher, D. M. & Bowles, H. R. (2012). Best practices for using physical activity monitors in population-based research. *Medicine and science in sports and exercise*, 44(1 Suppl 1), S68–76. doi:10.1249/MSS.0b013e3182399e5b
- Meta-analysis: correlation: Manual » Meta-analysis. (2019). Zugriff 30. Januar 2019 unter <https://www.medcalc.org/manual/meta-analysis-correlation.php>
- Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS medicine*, 6(7), e1000097. doi:10.1371/journal.pmed.1000097
- Montoye, A. H. K., Mitzyk, J. R. & Molesky, M. J. (2017). Comparative Accuracy of a Wrist-Worn Activity Tracker and a Smart Shirt for Physical Activity Assessment. *Measurement in Physical Education and Exercise Science*, 21(4), 201–211. doi:10.1080/1091367X.2017.1331166
- Parak, J. & Korhonen, I. (2014). Evaluation of wearable consumer heart rate monitors based on photoplethysmography. *Conference proceedings : ... Annual International Conference of*

- the IEEE Engineering in Medicine and Biology Society. *IEEE Engineering in Medicine and Biology Society. Annual Conference, 2014*, 3670–3673. doi:10.1109/EMBC.2014.6944419
- Pavey, T. G., Gilson, N. D., Gomersall, S. R., Clark, B. & Trost, S. G. (2017). Field evaluation of a random forest activity classifier for wrist-worn accelerometer data. *Journal of science and medicine in sport*, 20(1), 75–80. doi:10.1016/j.jsams.2016.06.003
- Reid, R. E. R., Insogna, J. A., Carver, T. E., Comptour, A. M., Bewski, N. A., Sciortino, C. & Andersen, R. E. (2017). Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. *Journal of science and medicine in sport*, 20(6), 578–582. doi:10.1016/j.jsams.2016.10.015
- Seshadri, D. R., Drummond, C., Craker, J., Rowbottom, J. R. & Voos, J. E. (2017). Wearable Devices for Sports: New Integrated Technologies Allow Coaches, Physicians, and Trainers to Better Understand the Physical Demands of Athletes in Real time. *IEEE pulse*, 8(1), 38–43. doi:10.1109/MPUL.2016.2627240
- Simunek, A., Dygryn, J., Gaba, A., Jakubec, L., Stelzer, J. & Chmelik, F. (2016). Validity of Garmin Vivofit and Polar Loop for measuring daily step counts in free-living conditions in adults. *Acta Gymnica*. doi:10.5507/ag.2016.014
- Sirard, J. R., Masteller, B., Freedson, P. S., Mendoza, A. & Hickey, A. (2017). Youth Oriented Activity Trackers: Comprehensive Laboratory- and Field-Based Validation. *Journal of medical Internet research*, 19(7), e250. doi:10.2196/jmir.6360
- Stackpool, C. M. (2013). *The accuracy of various activity trackers in estimating steps taken and energy expenditure* (Masterthesis, University of Wisconsin–La Crosse, La Crosse, Wisconsin).
- Stahl, S. E., An, H.-S., Dinkel, D. M., Noble, J. M. & Lee, J.-M. (2016). How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ open sport & exercise medicine*, 2(1), e000106. doi:10.1136/bmjsem-2015-000106
- Statistiken zu Wearables. (2017). Zugriff 29. November 2017 unter <https://de.statista.com/>
- Streiner, D. L. & Norman, G. R. (2006). "Precision" and "accuracy": Two terms that are neither. *Journal of clinical epidemiology*, 59(4), 327–330. doi:10.1016/j.jclinepi.2005.09.005
- Sun, Y. & Thakor, N. (2016). Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. *IEEE transactions on bio-medical engineering*, 63(3), 463–477. doi:10.1109/TBME.2015.2476337
- Sushames, A., Edwards, A., Thompson, F., McDermott, R. & Gebel, K. (2016). Validity and Reliability of Fitbit Flex for Step Count, Moderate to Vigorous Physical Activity and Activity Energy Expenditure. *PloS one*, 11(9), e0161224. doi:10.1371/journal.pone.0161224
- Tagliente, I., Solvoll, T., Trieste, L., de Cecco, C. N., Murgia, F. & Bella, S. (2016). Which indicators for measuring the daily physical activity? An overview on the challenges

- and technology limits for Telehealth applications. *Technology and health care : official journal of the European Society for Engineering and Medicine*, 24(5), 665–672. doi:10.3233/THC-161216
- Takacs, J., Pollock, C. L., Guenther, J. R., Bahar, M., Napier, C. & Hunt, M. A. (2014). Validation of the Fitbit One activity monitor device during treadmill walking. *Journal of science and medicine in sport*, 17(5), 496–500. doi:10.1016/j.jsams.2013.10.241
- Tam, K. M. & Cheung, S. Y. (2018). Validation of Electronic Activity Monitor Devices During Treadmill Walking. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*. doi:10.1089/tmj.2017.0263
- Tophøj, K. H., Petersen, M. G., Sæbye, C., Baad-Hansen, T. & Wagner, S. (2018). Validity and Reliability Evaluation of Four Commercial Activity Trackers' Step Counting Performance. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*. doi:10.1089/tmj.2017.0264
- Treacy, D., Hassett, L., Schurr, K., Chagpar, S., Paul, S. S. & Sherrington, C. (2017). Validity of Different Activity Monitors to Count Steps in an Inpatient Rehabilitation Setting. *Physical therapy*, 97(5), 581–588. doi:10.1093/ptj/pzx010
- Tucker, W. J., Bhammar, D. M., Sawyer, B. J., Buman, M. P. & Gaesser, G. A. (2015). Validity and reliability of Nike + Fuelband for estimating physical activity energy expenditure. *BMC sports science, medicine & rehabilitation*, 7, 14. doi:10.1186/s13102-015-0008-7
- Vairamuthu, R., Chethan, G. & Sivaguru, N. (2018). The basics of designing wearable electronics with microcontrollers. Zugriff 23. März 2018 unter <https://www.embedded.com/print/4431259>
- van Remoortel, H., Giavedoni, S., Raste, Y., Burtin, C., Louvaris, Z., Gimeno-Santos, E., ... Troosters, T. (2012). Validity of activity monitors in health and chronic disease: A systematic review. *The international journal of behavioral nutrition and physical activity*, 9, 84. doi:10.1186/1479-5868-9-84
- Voss, C., Gardner, R. F., Dean, P. H. & Harris, K. C. (2017). Validity of Commercial Activity Trackers in Children With Congenital Heart Disease. *The Canadian journal of cardiology*, 33(6), 799–805. doi:10.1016/j.cjca.2016.11.024
- Wahl, Y., Düking, P., Droszez, A., Wahl, P. & Mester, J. (2017). Criterion-Validity of Commercially Available Physical Activity Tracker to Estimate Step Count, Covered Distance and Energy Expenditure during Sports Conditions. *Frontiers in physiology*, 8, 725. doi:10.3389/fphys.2017.00725
- Wallen, M. P., Gomersall, S. R., Keating, S. E., Wisløff, U. & Coombes, J. S. (2016). Accuracy of Heart Rate Watches: Implications for Weight Management. *PloS one*, 11(5), e0154420. doi:10.1371/journal.pone.0154420
- Wikipedia Autoren. (2018). Metabolisches Äquivalent. Zugriff 6. Februar 2019 unter https://de.wikipedia.org/w/index.php?title=Metabolisches_%C3%84quivalent&oldid=182746546

Literatur

- Wikipedia Autoren. (2019a). Jogging. Zugriff 31. Januar 2019 unter <https://en.wikipedia.org/w/index.php?title=Jogging&oldid=880036621>
- Wikipedia Autoren. (2019b). Walking. Zugriff 31. Januar 2019 unter <https://en.wikipedia.org/w/index.php?title=Walking&oldid=878539048>
- Woodford, C. (2010). Pedometers. Zugriff 22. Januar 2019 unter <https://www.explainthatstuff.com/how-pedometers-work.html>
- Wright, S. P., Hall Brown, T. S., Collier, S. R. & Sandberg, K. (2017). How consumer physical activity monitors could transform human physiology research. *American journal of physiology. Regulatory, integrative and comparative physiology*, 312(3), 358–367. doi:10.1152/ajpregu.00349.2016

Anhang

Tabellen sind hier, wenn die Arbeit als PDF-Datei vorliegt, im CSV-Format eingebettet (Zeichenkodierung: ANSI; Trennzeichen: Semikolon):

Tabelle 5.1

 Studien in Metaanalyse

Tabelle 5.2

 Qualität der Studien in Metaanalyse

Tabelle 5.3

 Inkludierte Fitness Tracker

Tabelle 5.4

 Ergebnisse der Gültigkeit der Schrittzählung

Tabelle 5.5

 Ergebnisse der Zuverlässigkeit der Schrittzählung

Tabelle 5.6

 Ergebnisse der Gültigkeit der Pulsmessung

Tabelle 5.7

 Ergebnisse der Zuverlässigkeit der Pulsmessung

Tabelle 5.8

 Ergebnisse der Gültigkeit der Kalorienverbrauchsberechnung

Tabelle 5.9

 Ergebnisse der Zuverlässigkeit der Kalorienverbrauchsberechnung