

ความตรง และความแม่นยำในการวัดอัตราการเต้นของหัวใจขณะออกกำลังกายของอุปกรณ์ วัดอัตราการเต้นของหัวใจชนิดสวมใส่ที่แขน : การทบทวนวรรณกรรมอย่างเป็นระบบ

ธีรวัฒน์ นิธิอรธวานนท์ นิธินันท์ ชัยศิริ รมภา บุญสินสุข

สาขาวิชากายภาพบำบัด คณะกายภาพบำบัด มหาวิทยาลัยศรีนครินทรวิโรฒ

Received: November 25, 2022

Revised: December 20, 2022

Accepted: February 13, 2023

บทคัดย่อ

การศึกษานี้เป็นการทบทวนวรรณกรรมอย่างเป็นระบบ โดยมีวัตถุประสงค์เพื่อรวบรวมหลักฐานเชิงประจักษ์เกี่ยวกับความตรง และความแม่นยำของอุปกรณ์ในการวัดอัตราการเต้นของหัวใจขณะออกกำลังกาย โดยสืบค้นรายงานการวิจัยที่เกี่ยวข้อง และได้รับการตีพิมพ์ตั้งแต่ ค.ศ. 2014 - 2021 จากฐานข้อมูล ScienceDirect, PubMed, Scopus และ Web of science และคัดเลือกเฉพาะเอกสารงานวิจัยที่มีเอกสารฉบับเต็ม และเป็นภาษาอังกฤษ และประเมินบทความวิจัยโดยใช้แบบประเมิน COSMIN และ QUADAS จากนั้นเข้าสู่กระบวนการวิเคราะห์เนื้อหา สรุป และสังเคราะห์ข้อมูล โดยจำแนกผลลัพธ์เป็น 2 กลุ่ม (ยอมรับ และไม่ยอมรับ) และ 3 กลุ่ม (สูง ปานกลาง และต่ำ) สำหรับความแม่นยำ และความตรง ตามลำดับ ผลการศึกษาพบงานวิจัยจำนวน 33 ฉบับ มีคุณภาพของบทความอยู่ในระดับที่ 2 ตามเกณฑ์ของ OCEBM โดยทดสอบคุณสมบัติการวัดของอุปกรณ์วัดอัตราการเต้นของหัวใจที่แขนจำนวน 15 ตราสินค้า 41 อุปกรณ์ พบหลักฐานสนับสนุนว่าอุปกรณ์หลายตราสินค้า สามารถวัดอัตราการเต้นของหัวใจได้สอดคล้องกับอุปกรณ์มาตรฐาน โดยมีค่าความแม่นยำที่ยอมรับได้ ($CCC = 0.86 - 0.96$; $MAPE < 5$) และความตรงในระดับปานกลางถึงสูง ($r = 0.59 - 0.99$) ในการวัดอัตราการเต้นของหัวใจขณะออกกำลังกายแบบใช้ออกซิเจน และแบบมีแรงต้าน ดังนั้น นักกีฬา นักวิจัย หรือบุคคลที่สนใจ สามารถนำข้อมูลเหล่านี้มาใช้ในการตัดสินใจเลือกอุปกรณ์ในการติดตามอัตราการเต้นของหัวใจขณะออกกำลังกายได้ อย่างไรก็ตาม การศึกษานี้พบว่าอุปกรณ์มีการพัฒนา และออกแบบใหม่ตลอดเวลา ส่งผลให้อุปกรณ์มีราคาที่สูง และใช้เทคโนโลยีที่ทันสมัยขึ้น ซึ่งบ่งชี้ถึงความจำเป็นของการทบทวนวรรณกรรม และการวิจัยในอนาคต

คำสำคัญ: คุณสมบัติการวัดเชิงจิตวิทยา ความเที่ยง เครื่องมือประเมินติดตามผล การออกกำลังกายแบบใช้ออกซิเจน

ผู้สนับสนุนประสานงาน:

ธีรวัฒน์ นิธิอรธวานนท์

สาขาวิชากายภาพบำบัด คณะกายภาพบำบัด มหาวิทยาลัยศรีนครินทรวิโรฒ

63 หมู่ 7 ถนนรังสิต-นครนายก ตำบลองครักษ์ อำเภองครักษ์ จังหวัดนครนายก 26120

อีเมล: Teerawtn@g.swu.ac.th

Concurrent validity and accuracy of arm-wearable devices to detect heart rate during exercise: A systematic review

Teerawat Nithiatthawanon, Nithinun Chaikereee, Rumpa Boonsinsukh

Division of Physical Therapy, Faculty of Physical Therapy, Srinakharinwirot University

Abstract

This study was systematically conducted to review research article regarding the accuracy and validity of commercial wearables in measuring heart rate during exercise. Studies published between 2014 - 2021 were searched in several databases, including ScienceDirect, PubMed, Scopus and the Web of science. Relevant studies with full text published in Thai and English were retrieved. The COSMIN and QUADAS were then used to assess the quality and risk of bias of each study. Later, data from selected studies were extracted, analyzed, and synthesized. The outcomes were divided into two categories (acceptable, unacceptable) and three (high, moderate, low) groups for accuracy and validity, respectively. There were 33 publications with Level 2, and based on the OCEBM examining 15 brands and 41 different arm-wearable devices. The data included studies illustrated that several arm-wearable devices were accurate ($CCC = 0.86 - 0.96$; $MAPE < 5$) and valid at a moderate to high level ($r = 0.59 - 0.99$) in measuring heart rate during both aerobic and resistance exercises. Thus, athletes, researchers, or those who were interested could apply this information on choosing a heart rate monitor for tracking heart rate while exercising. However, the study suggested that devices are consistently being developed and redesigned resulting in lower prices and using more accurate and valid technology, suggesting the need for more current reviews and further research.

Keywords: psychometric property, reliability, follow-up tool, aerobic exercise

Corresponding Author:

Teerawat Nithiatthawanon

Faculty of Physical Therapy, Srinakharinwirot University

63 Moo 7, Rangsit-Nakhon Nayok Road, Ongkharak District, Nakhon Nayok 26120, Thailand

E-mail: Teerawatn@g.swu.ac.th

บทนำ

กิจกรรมทางกายที่ลดลง (physical inactivity) กำลังกลายเป็นปัญหาที่เกิดขึ้นทั่วโลก โดยพบว่าในกลุ่มผู้ใหญ่ และวัยรุ่น (จำนวนร้อยละ 23 และ 81 ตามลำดับ) มีกิจกรรมทางกายไม่เป็นไปตามเกณฑ์ที่กำหนดในกลุ่มนั้นๆ¹ ดังนั้น ภาครัฐจึงมีนโยบายในการกระตุ้นให้กลุ่มคนในช่วงวัยต่างๆ เกิดกิจกรรมทางกายเพิ่มมากขึ้น โดยองค์การอนามัยโลก (World Health Organization: WHO) ได้แนะนำการออกกำลังกาย ด้วยความหนักระดับปานกลาง เพราะสามารถช่วยลดอุบัติการณ์การเกิดโรคไม่ติดต่อเรื้อรัง ลดภาวะวิตกกังวล และช่วยเพิ่มทักษะการรู้คิดอีกด้วย² อย่างไรก็ตาม นอกจากรูปแบบของการออกกำลังกายที่สำคัญ การติดตามระดับความหนักในการออกกำลังกายก็เป็นสิ่งสำคัญเช่นกัน เนื่องจากการออกกำลังกายในช่วงความหนักที่เหมาะสมจะทำให้เกิดประสิทธิภาพได้สูงสุด โดยการวัดอัตราการเต้นของหัวใจ (heart rate measurement) เป็นวิธีการวัดระดับความหนักของการออกกำลังกายที่นิยมใช้อย่างแพร่หลายในปัจจุบัน เนื่องจากสามารถสะท้อนการเปลี่ยนแปลงของระบบหัวใจและหลอดเลือดได้อย่างแม่นยำ³ อีกทั้งยังสามารถวัดได้ง่ายอีกด้วย⁴

ดังนั้น อุตสาหกรรมการผลิตและพัฒนาอุปกรณ์ที่ใช้วัดอัตราการเต้นของหัวใจจึงมีการเติบโตขึ้นอย่างต่อเนื่องเพื่อตอบสนองต่อความต้องการของผู้ใช้งาน⁵ โดยสายคาดหน้าอกวัดชีพจร (chest strap heart-rate monitor) ที่มีคุณสมบัติการวัดเชิงจิตวิทยาที่สูงเมื่อเทียบกับเครื่องมือมาตรฐาน หรือ Electrocardiogram (ECG) (CCC = 0.99)⁵ จึงกลายมาเป็นอุปกรณ์ที่มีความนิยมเพิ่มมากขึ้น แต่เนื่องจากอุปกรณ์ชนิดนี้ต้องคาดสายรอบอกและจำเป็นต้องใช้โทรศัพท์ในการแสดงผล จึงไม่สะดวกในการสวมใส่และวัดผลต่างๆ โดยเฉพาะอย่างยิ่งในเพศหญิง⁶

ในปัจจุบัน อุปกรณ์วัดอัตราการเต้นของหัวใจชนิดสวมใส่ที่แขนได้ถูกพัฒนาขึ้นมาให้สามารถประเมินอัตราการเต้นของหัวใจได้ และด้วยวิธีการใช้งานที่ง่าย

ไม่ซับซ้อน ขนาดเล็ก กะทัดรัด ส่งผลให้อุปกรณ์ได้รับความนิยมเป็นอย่างมาก โดยจากการทบทวนวรรณกรรมพบว่ามีการวิจัยก่อนหน้านี้ได้ทดสอบคุณสมบัติการวัดเชิงจิตวิทยาของอุปกรณ์ต่างๆ ในหลากหลายตราสินค้าที่มีจำหน่ายโดยทั่วไปในท้องตลาด โดยทดสอบคุณสมบัติการวัดอัตราการเต้นของหัวใจในมนุษย์ขณะออกกำลังกายทั้งแบบใช้ออกซิเจน (aerobic exercise) และแบบมีแรงต้าน (resistance exercise) โดยพบว่ามีบางตราสินค้าที่มีคุณสมบัติในการวัดเชิงจิตวิทยาที่สูงเทียบได้กับเครื่องมือมาตรฐาน⁶ อย่างไรก็ตาม จากการทบทวนวรรณกรรมพบว่าข้อมูลต่างๆ ที่เกี่ยวข้องได้กระจายอยู่ในที่ต่างๆ ยังไม่มีการรวบรวมข้อมูลอย่างเป็นระบบ หรือจัดเรียงข้อมูลของค่าคุณสมบัติการวัดทางจิตวิทยาในรูปแบบการออกกำลังกายที่แตกต่างกัน รวมถึงบทความปริทัศน์เชิงระบบ (systematic review) ยังไม่มีข้อสรุปอย่างชัดเจนเกี่ยวกับแนวทางการนำไปใช้ และยังไม่สามารถรวบรวมอุปกรณ์ต่างๆ ไว้ได้อย่างครบถ้วน โดยผู้วิจัยเชื่อว่าข้อมูลต่างๆ เหล่านี้หากมีการทบทวนอย่างเป็นระบบจะทำให้ได้ข้อสรุปเกี่ยวกับคุณสมบัติในการวัดเชิงจิตวิทยาของเครื่องมือ และแนวทางการนำความรู้ไปประยุกต์ใช้ได้

วัตถุประสงค์

การทบทวนวรรณกรรมอย่างเป็นระบบในครั้งนี้มีวัตถุประสงค์เพื่อสืบค้นหลักฐานเชิงประจักษ์และสรุปข้อมูลเกี่ยวกับความตรงและความแม่นยำของอุปกรณ์วัดอัตราการเต้นของหัวใจชนิดสวมใส่ที่แขนในการวัดอัตราการเต้นของหัวใจในกลุ่มอาสาสมัครสุขภาพดี และนักกีฬา ขณะออกกำลังกายในรูปแบบต่างๆ

วิธีการศึกษา

การรวบรวมวรรณกรรมอย่างเป็นระบบนี้ได้จัดทำขึ้นภายใต้วิธีการของ the Preferred Reporting Items for Systematic Reviews and Meta - Analysis guidelines หรือ PRISMA⁷ โดยรวบรวม

บทความต้นฉบับ (original article) ในช่วง 10 ปีย้อนหลัง เพื่อให้ได้บทความวิจัยที่ทันสมัย และครอบคลุมการศึกษาเกี่ยวกับการทดสอบความตรงและความแม่นยำของอุปกรณ์ชนิดสวมใส่ที่แขนทั้งหมด

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ มี 3 ประเภท ได้แก่ 1) แบบคัดกรองงานวิจัยตามเกณฑ์ที่กำหนด โดยเป็นเกณฑ์คัดเข้า (inclusion criteria) และเกณฑ์คัดออก (exclusion criteria) โดยผู้วิจัยสร้างขึ้นตาม

เกณฑ์การคัดเลือกที่กำหนดไว้ 2) แบบประเมินคุณภาพการวิจัย (critical appraisal form) โดยอ้างอิงจาก The Consensus - based Standards for the selection of health Measurement Instruments (COSMIN)⁸ และ Quality Assessment of Diagnostic Accuracy Studies (QUADAS)⁹ รูปที่ 1 (Figure 1) เพื่อใช้ในการประเมินคุณภาพเชิงระเบียบวิธีวิจัยของรายงานวิจัยที่ผ่านการคัดเลือก 3) แบบบันทึกผลการสกัดข้อมูล เพื่อบันทึกข้อมูลงานวิจัยที่คัดเลือกไว้ ตารางที่ 1 (Table 1)

Evaluated measurement properties in the article			
Internal consistency			
Reliability			
Measurement error			
Content validity			
Structural validity			
Hypotheses testing			
Cross-cultural validity			
Criterion validity			
Responsiveness			
Interpretability			

Box H. Criterion validity			
Design requirements			
1	Was the percentage of missing items given?	yes	no ?
2	Was there a description of how missing items were handled?	yes	no ?
3	Was the sample size included in the analysis adequate?	yes	no ?
4	Can the criterion used or employed be considered as a reasonable 'gold standard'?	yes	no ?
5	Were there any important flaws in the design or methods of the study?	yes	no ?
Statistical methods			
6	for continuous scores: Were correlations, or the area under the receiver operating curve calculated?	yes	no NA
7	for dichotomous scores: Were sensitivity and specificity determined?	yes	no NA

Box Generalisability box			
Was the sample in which the HR-PRO instrument was evaluated adequately described? In terms of:			
1	median or mean age (with standard deviation or range)?	yes	no NA
2	distribution of sex?	yes	no NA
3	important disease characteristics (e.g. severity, status, duration) and description of treatment?	yes	no NA
4	setting(s) in which the study was conducted? e.g. general population, primary care or hospital/rehabilitation care	yes	no NA
5	countries in which the study was conducted?	yes	no NA
6	language in which the HR-PRO instrument was evaluated?	yes	no NA
7	Was the method used to select patients adequately described? e.g. convenience, consecutive, or random	yes	no NA
8	Was the percentage of missing responses (response rate) acceptable?	yes	no NA

QUADAS-2 tool: Risk of bias and applicability judgments	
Domain 1: Patient selection	
A. Risk of bias	
Describe methods of patient selection:	
• Did the study avoid inappropriate exclusions?	Yes/No/Unclear
• Could the selection of patients have introduced bias?	RISK: LOW/HIGH/UNCLEAR
• Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of index test and setting):	
B. Is there concern that the included patients do not match the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 2: Index test(s) (if more than 1 index test was used, please complete for each test)	
Risk of bias	
Describe the index test and how it was conducted and interpreted:	
A. Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/No/Unclear
• Could the conduct or interpretation of the index test have introduced bias?	RISK: LOW/HIGH/UNCLEAR
• Concerns regarding applicability	
Is there concern that the index test, its conduct, or interpretation differ from the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 3: Reference standard	
Risk of bias	
Describe the reference standard and how it was conducted and interpreted:	
A. Is the reference standard likely to correctly classify the target condition?	Yes/No/Unclear
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes/No/Unclear
• Could the reference standard, its conduct, or its interpretation have introduced bias?	RISK: LOW/HIGH/UNCLEAR
• Concerns regarding applicability	
Is there concern that the target condition as defined by the reference standard does not match the review question?	CONCERN: LOW/HIGH/UNCLEAR
Domain 4: Flow and timing	
Risk of bias	
Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):	
Describe the time interval and any interventions between index test(s) and reference standard:	
A. Was there an appropriate interval between index test(s) and reference standard?	Yes/No/Unclear
Did all patients receive a reference standard?	Yes/No/Unclear
• Did patients receive the same reference standard?	Yes/No/Unclear
• Were all patients included in the analysis?	Yes/No/Unclear
• Could the patient flow have introduced bias?	RISK: LOW/HIGH/UNCLEAR

Figure 1 Critical appraisal form

ขั้นตอนและวิธีการรวบรวมข้อมูล

1. วิธีการ และกลยุทธ์ในการสืบค้นรายงานวิจัย (search strategy)

การค้นหาค้นหาความได้อ้างอิงจากฐานข้อมูลทั้งสี่ 4 ฐาน ดังนี้ ScienceDirect, PubMed, Scopus และ Web of science ตั้งแต่เริ่มต้นการสืบค้น จนถึงวันที่ 31 มกราคม 2565 โดยกำหนดเงื่อนไขทางตรรกะ ได้แก่ การใช้คำเชื่อม AND/OR ร่วมกับการใช้คำค้นที่เกี่ยวข้องตามหลักการ PICO คือ *Healthy adults, Smartwatch, Wrist worn, Heart rate, ECG, Polar, Psychometric properties, Accuracy และ Validity* ภายหลังจากการค้นหาวีธีวิจัยรวบรวม บทความย่อ รายงานประชุมวิชาการ (conference proceeding) และบทความอื่นๆ ในช่วงเริ่มต้นของการค้นหา หลังจากนั้นนำบทความต่างๆ ที่ได้เข้าสู่โปรแกรม EndNote20 software (Clarivate Analytics: 3092075960 [SWU-license]) เพื่อตรวจสอบความซ้ำซ้อนของบทความวิจัย หลังจากนั้นบทความที่เหลือถูกนำเข้าสู่กระบวนการตรวจสอบความสอดคล้องตามเกณฑ์การคัดเลือกบทความในขั้นตอนถัดไป

2. การคัดเลือกบทความวิจัย (study selection)

กลยุทธ์การคัดเลือกบทความ (study selection strategy) โดยบทความการศึกษาที่รวบรวมได้จากการค้นหาฐานข้อมูลในเบื้องต้นจะถูกคัดกรองเบื้องต้นอีก 1 ครั้ง ผ่านการพิจารณาโดยใช้ชื่อเรื่อง และบทความย่อจากผู้รวบรวมข้อมูลอย่างน้อย 2 คนจากทีม โดยเป็นการประเมินอย่างอิสระ หากพบว่าการประเมินจากผู้รวบรวมข้อมูลทั้ง 2 คนขัดแย้งกัน จะใช้การประเมินจากผู้รวบรวมคนที่ 3 เป็นผู้ตัดสิน หลังจากการผ่านการคัดกรองเบื้องต้น บทความฉบับเต็มจะถูกนำมาพิจารณาโดยทีมผู้ตรวจสอบอย่างน้อย 2 คน โดยผู้ตรวจสอบทุกคนปฏิบัติตามเกณฑ์การคัดเลือกที่กำหนดไว้อย่างเคร่งครัด หากเกิดข้อขัดแย้งของผลการประเมิน บทความฉบับนั้นจะถูกนำมาพิจารณาผ่านการ

อภิปราย (discussion) จนเกิดความเอกฉันท์ (consensus) จากทีมผู้เขียน 2 ท่าน (TN, NC) โดยเกณฑ์ที่กำหนดไว้มีดังนี้

เกณฑ์การคัดเข้า: (1) การศึกษาในประชากรผู้มีสุขภาพดี (2) วัดอัตราการเต้นของหัวใจขณะออกกำลังกายแบบใช้ออกซิเจน (3) มีเครื่องมือมาตรฐาน (gold standard) ในการเปรียบเทียบ (4) ทำการวัดความเที่ยง (reliability) และ/หรือ ความตรง (validity) โดยอุปกรณ์วัดอัตราการเต้นของหัวใจชนิดสวมใส่ที่ข้อมือโดยคำนึงถึงความตรง และความเที่ยงในการศึกษานี้คือ ความตรงตามเกณฑ์สัมพัทธ์ (criterion validity) และความแม่นยำในการวัดของเครื่องมือ (measurement agreement หรือ accuracy)

เกณฑ์การคัดออก: (1) งานวิจัยที่ไม่ใช่ภาษาไทยหรือภาษาอังกฤษ (2) ไม่มีงานวิจัยฉบับเต็ม (3) งานวิจัยที่ศึกษาในเด็กเล็ก (4) งานวิจัยที่ใช้เครื่องมือมาตรฐานอื่นนอกจาก ECG หรือ Polar (5) เป็นบทความปริทัศน์แบบปกติ (narrative review)

3. การประเมินคุณภาพงานวิจัย

ก่อนเริ่มกระบวนการสกัดข้อมูล และประเมินคุณภาพ ทีมผู้ประเมินจะเข้ากระบวนการเตรียมความพร้อม กล่าวคือศึกษาเกณฑ์การประเมินคุณภาพตามคำอธิบายที่แสดงในแต่ละเครื่องมือ จากนั้นอภิปรายเกณฑ์ร่วมกันและทำการทดสอบความเที่ยงระหว่างผู้ประเมินโดยกระบวนการดึงข้อมูล ประกอบด้วย: (1) การจัดหมวดหมู่บทความฉบับเต็มที่เลือกไว้ในการศึกษาความตรงหรือความแม่นยำ; (2) ใช้ The COSMIN⁸ และ QUADAS⁹ เพื่อประเมินคุณภาพความเสี่ยงของอคติ (quality and risk of bias) ของแต่ละการศึกษา และเป็นแนวทางในการระบุระดับความน่าเชื่อถือของหลักฐานเชิงประจักษ์ร่วมกับเกณฑ์การจัดอันดับของ Oxford Centre for Evidence-Based Medicine (OCEBM) ซึ่งจัดระดับความน่าเชื่อถือไว้ 5 ระดับ (3) แยกลักษณะ และข้อมูลสำคัญจากบทความที่ได้รับคัดเลือก และรวบรวมรายงานเป็นตาราง (4) เปรียบเทียบรายละเอียดจาก

ผู้ตรวจสอบแต่ละคน และแก้ไขความไม่สอดคล้องกัน
ผ่านการอภิปรายก่อนรวบรวมผลลัพธ์

4. กระบวนการสกัดข้อมูล (data extraction)

ผู้ช่วยวิจัย 3 คน ได้ทำการสกัดข้อมูลต่างๆ และถูกตรวจสอบอีกครั้งจากทีมผู้เขียน 2 ท่าน (TN, NC) และรวบรวมผลที่ได้จากการสกัดเนื้อหาจากบทความวิจัยที่ผ่านการคัดเลือกประกอบไปด้วยลักษณะของผู้เข้าร่วม ประเภทของอุปกรณ์มาตรฐานตราสินค้า และรุ่นของอุปกรณ์ที่ศึกษา วิธีการทดสอบ และผลลัพธ์ที่วัดได้ โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ (r) ค่าสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (ICC) ค่าสัมประสิทธิ์สหสัมพันธ์ความสอดคล้อง (CCC หรือ r_c) ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) และค่าจากกราฟ Bland-Altman plot ที่รายงานในแต่ละการศึกษา

การวิเคราะห์ข้อมูลทางสถิติ

เนื่องจากการรวบรวมและทบทวนวรรณกรรมอย่างเป็นระบบในครั้งนี้มีองค์ประกอบที่อาจจะแตกต่างกันหลายอย่าง เช่น สถานที่ทำการทดสอบ วิธีการทดสอบ ตัวแปร อุปกรณ์มาตรฐาน และอุปกรณ์ที่สนใจศึกษา จึงทำให้ยังไม่สามารถทำ meta - analyses ได้ ดังนั้น ข้อมูลทั้งหมดที่ได้จึงถูกนำมาวิเคราะห์ด้วยการพรรณนาอย่างเป็นระบบ และสังเคราะห์ข้อมูลด้านคุณสมบัติการวัดทางจิตวิทยาต่างๆ โดยใช้เกณฑ์ที่เหมาะสมเพื่อคัดแยกอุปกรณ์ที่มีความสามารถในการวัดที่แม่นยำ และตรงตามเกณฑ์ที่กำหนด^{10,11}

โดยงานวิจัยนี้ได้แปรผลค่าสัมประสิทธิ์สหสัมพันธ์โดยการใช้เกณฑ์ตามการศึกษาที่ผ่านมา ดังนี้ น้อย (0.0 ถึง < 0.4) ปานกลาง (0.4 ถึง < 0.6) และ มาก (0.6 ถึง 1.0)^{10,11} รวมถึงค่า MAPE สามารถแปลผลได้เมื่อค่า < 5 = ยอมรับได้ 5 - 25 = ยอมรับได้ปานกลาง > 25 = ยอมรับไม่ได้¹² ในขณะที่ค่าที่ยอมรับได้ของค่าสัมประสิทธิ์สหสัมพันธ์ความสอดคล้องคือ > 0.8¹¹

ผลการศึกษา

มีรายงานการวิจัยตามเกณฑ์ที่กำหนดจำนวนทั้งสิ้น 3,637 ฉบับ ส่วนใหญ่เป็นการศึกษาในรูปแบบภาคตัดขวาง (cross - sectional study) โดยภายหลังการคัดกรองด้วยบทคัดย่อ และชื่อบทความ มีงานที่ถูกคัดออกทั้งสิ้น 3,534 ฉบับ หลังจากนั้นนำบทความวิจัยที่เหลือทั้งหมดนำเข้าสู่อุปกรณ์ EndNote เพื่อคัดบทความที่มีความซ้ำซ้อนของบทความวิจัย จึงเหลืองานวิจัยที่ผ่านเกณฑ์การคัดเข้าทั้งสิ้นจำนวน 39 บทความ จากนั้นถูกคัดออกจากเกณฑ์การคัดออก 6 ฉบับ เนื่องจากไม่มีเอกสารฉบับเต็ม 1 ฉบับ การศึกษาที่ไม่ใช่เครื่องมือมาตรฐานเป็น ECG หรือ Polar 2 ฉบับ และเป็นบทความปริทัศน์จำนวน 3 ฉบับ คงเหลือรายงานการวิจัยที่นำมาใช้ในการศึกษานี้ จำนวนทั้งสิ้น 33 ฉบับ^{5,13-44} รูปที่ 2 (Figure 2)

ลักษณะของการศึกษา และอาสาสมัคร (study and participant characteristics)

ข้อมูลที่สกัดได้จากบทความดังแสดงในตารางที่ 1 (Table 1) พบว่า บทความวิจัยทั้งหมด 33 ฉบับที่ตีพิมพ์ในช่วงปี ค.ศ. 2014 - 2022 โดยมีแนวโน้มสูงขึ้นหลังจากปี 2016 และมีความสนใจสูงที่สุดในปี 2017 มีอาสาสมัครส่วนใหญ่เป็นผู้ที่มีสุขภาพดี ตามด้วยนักกีฬา โดยมีอายุส่วนใหญ่อยู่ในช่วงระหว่าง 18 - 65 ปี อุปกรณ์ที่ใช้ในการวัดอัตราการเต้นของหัวใจในการศึกษาทุกฉบับใช้เทคโนโลยี Photoplethysmogram (PPG) โดยแบ่งออกเป็น optical sensor และ light emitting diode (LED) ทั้งหมด 15 ตราสินค้า (มีราคาอยู่ในช่วง 5,590 - 31,900 บาท) โดยสินค้าจากตราสินค้า Fitbit (n = 17, 7 โมเดล) ถูกนำมาทำการทดสอบมากที่สุด (รุ่น Fitbit charge HR ได้รับการทดสอบมากที่สุด [n = 6]) ตามด้วย Garmin (n = 12, 5 โมเดล), Apple (n = 11, 4 โมเดล) และ Polar electro (n = 9, 6 โมเดล) ในขณะที่ Xiaomi, Tempo HR, Sunto oy และ fifth zone fitness laboratory company ได้รับการทดสอบน้อยที่สุดจากการ

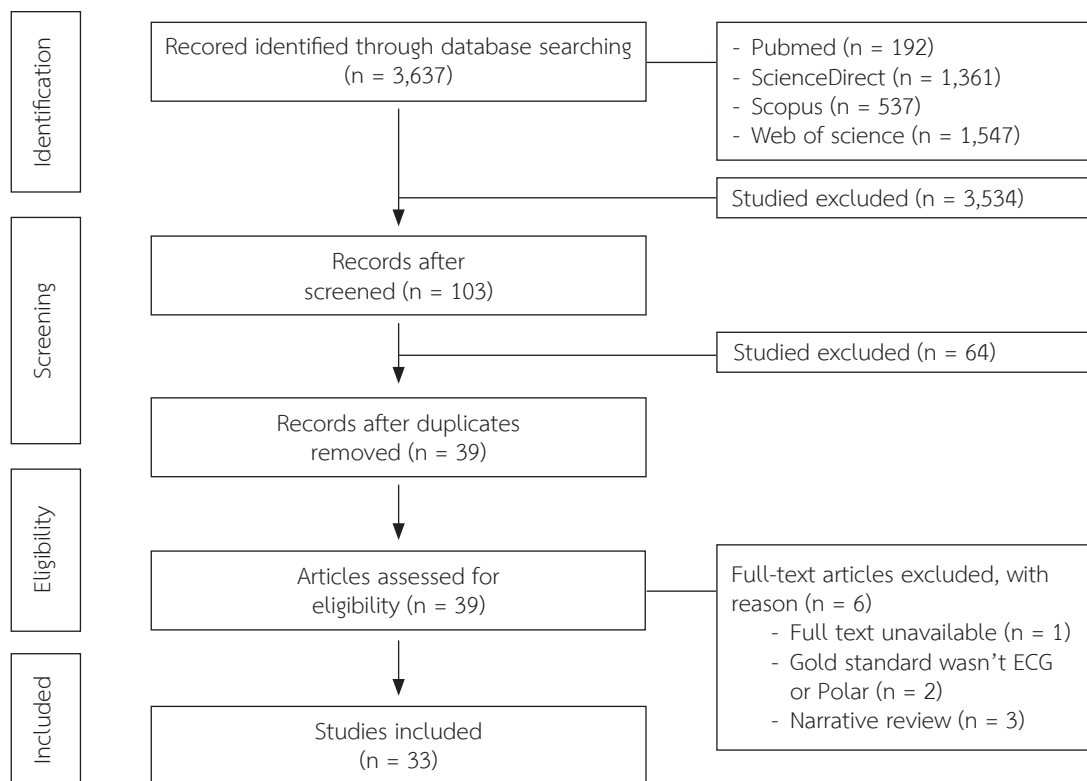


Figure 2 Flow diagram of literature search and study selection procedure according to PRISMA

รวบรวมบทความ (n = 1) โดยอุปกรณ์เหล่านี้ถูกนำมาเปรียบเทียบความแม่นยำ และความตรงกับเครื่องมือมาตรฐาน ECG (n = 16) ตามด้วย Polar H7 (n = 7), Polar H10 (n = 4) ตามลำดับ และมีกิจกรรมที่ใช้ในการทดสอบมีทั้งสิ้น 9 กิจกรรม โดยมีความถี่ในการเลือกใช้กิจกรรม คือ การวิ่ง และเดิน (n = 28) ปั่นจักรยาน (n = 13) กิจกรรมทั่วไป (n = 4) ยกน้ำหนัก (n = 3) เดินบนเครื่องเดินวงรีร่วมกับการแกว่ง และไม่แกว่งแขน (n = 2) และว่ายน้ำ (n = 1) ตามลำดับ และมีระดับความหนักส่วนใหญ่อยู่ที่ระดับปานกลางถึงระดับหนัก โดยมีการทำการทดสอบภายในห้องปฏิบัติการ (n = 28) และในสภาพแวดล้อมจริง (n = 5)

ความแม่นยำ (accuracy)

จากรายงานการวิจัยแบบการทบทวนอย่างเป็นระบบ ตารางที่ 1 (Table 1) แสดงให้เห็นว่ามีการรายงานเกี่ยวกับความแม่นยำในนาฬิกาอัจฉริยะ

จำนวน 32 ฉบับ (ร้อยละ 97.22) พบว่า Apple watch เป็นนาฬิกาที่มีค่าความแม่นยำสูงที่สุดในทุกๆ กิจกรรมการเคลื่อนไหวทางกาย (CCC = 0.86 - 0.99)

อย่างไรก็ตาม เมื่อพิจารณาในแต่ละกิจกรรมจะพบว่า Polar OH 1 (CCC = 0.86-0.96) มีค่าความแม่นยำมากที่สุด ในขณะที่ Fitbit charge 2 (CCC = 0.23 - 0.81) มีค่าความแม่นยำน้อยที่สุดเมื่อใช้ในขณะที่ออกกำลังกายด้วยการเดินหรือวิ่งสำหรับการปั่นจักรยานพบว่า Schosche Rhythm 24 (CCC = 0.99) มีค่าความแม่นยำสูงที่สุด และ Fitbit charge 2 (CCC = 0.58) มีค่าความแม่นยำต่ำที่สุด แต่กลับพบว่า Fitbit charge 2 มีค่า CCC สูงเมื่อออกกำลังกายแบบมีแรงต้าน (CCC = 0.72) ทางด้านการออกกำลังกายด้วยเครื่องเดินวงรี Apple watch (CCC = 0.94) มีค่า CCC สูงที่สุดและลดลงเมื่อมีการเคลื่อนที่ของแขนร่วมด้วย ในขณะที่ Garmin Forerunner 235

มีค่าความแม่นยำต่ำที่สุด ($CCC = 0.54$) และกิจกรรมสุดท้ายคือการดำเนินกิจกรรมในชีวิตประจำวัน พบว่า Apple watch ยังคงเป็นอุปกรณ์ที่วัดได้ค่าความแม่นยำสูงที่สุด ($CCC = 0.96$)

นอกจากนั้น เมื่อพิจารณาค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) จากบทความทั้งหมด ($n = 12$) พบว่าอุปกรณ์ที่มีค่าน้อยกว่า 5 ในกิจกรรมการเดิน และวิ่ง คือ Apple watch, Schosche Rhythm, Mio Alpha, Basis peak, TomTom runner cardio และ Fizzo และการปั่นจักรยานพบอุปกรณ์ที่มีค่า >5 คือ Apple watch, Schosche Rhythm 24, Basis peak, Fitbit surge, Microsoft band, Mio alpha 2 รูปที่ 3 (Figure 3)

ความตรงตามสภาพปัจจุบัน (Concurrent validity)

จากตารางที่ 1 (Table 1) แสดงให้เห็นค่าความตรงในการวัดอัตราการเต้นของหัวใจขณะออกกำลังกายของอุปกรณ์ต่างๆ เมื่อเทียบกับเครื่องมือมาตรฐาน โดยพบการรายงานจำนวน 32 ฉบับ พบว่าตราสินค้า Apple watch ($r = 0.59 - 0.99$) มีค่าความตรงมากที่สุดในทุกๆ กิจกรรมการออกกำลังกาย และ Fitbit charge HR ($r = 0.16 - 0.99$) มีค่าความตรงที่แปรปรวนมาก และพบว่ามีค่าน้อยที่สุดในการออกกำลังกายด้วยการเดิน หรือวิ่ง โดยพบค่าต่ำที่สุดเมื่อออกกำลังกายที่ระดับความหนักสูง (vigorous intensity) ในขณะที่การปั่นจักรยาน

พบว่า Apple watch ($r = 0.80 - 0.95$) มีค่าความตรงสูงที่สุด และ Fitbit charge 2 ($r = 0.46$) มีค่าความตรงต่ำที่สุด แต่กลับพบว่ามีความตรงสูงเมื่อออกกำลังกายแบบมีแรงต้าน ($r = 0.88$)

คุณภาพ ความเสี่ยงของอคติ (quality and risk of bias)

จากบทความทั้งหมดพบ ร้อยละ 57.58 ที่มีความเสี่ยงสูงด้านจำนวนของอาสาสมัครที่ใช้ในการศึกษาเนื่องจากมีจำนวนอาสาสมัครน้อยกว่า 30 ราย และร้อยละ 42.42 ที่มีจำนวนอาสาสมัครที่มีความเสี่ยงต่ำเนื่องจากมีจำนวนมากกว่าหรือเท่ากับ 30 ราย เมื่อประเมินความเสี่ยงที่จะมีอคติโดยแบบประเมิน QUADAS-2 รูปที่ 4 (Figure 4) พบว่ามีอคติมากที่สุด ในหัวข้อ ลำดับ และระยะเวลา (flow and timing) และมีอคติน้อยที่สุดในหัวข้อการคัดเลือกอาสาสมัคร (patient selection) รวมถึงเมื่อประเมินความกังวลในการนำไปใช้ (concern of regarding applicability) ด้วยแบบประเมิน QUADAS-2 พบว่าความกังวลในการนำไปใช้ระดับน้อย (low concern of regarding applicability) มีสัดส่วนมากที่สุด ในทั้ง 3 หัวข้อ รูปที่ 5 (Figure 5) จากผลประเมินระดับความน่าเชื่อถือทั้งหมดตามเกณฑ์ OCEBM เป็นบทความวิจัยที่มีความน่าเชื่อถือระดับ 2 จำนวน 27 บทความ (ร้อยละ 81.82) และระดับ 3 จำนวน 6 บทความ (ร้อยละ 18.18) แสดงให้เห็นว่าบทความวิจัยส่วนใหญ่อยู่ในระดับที่ 2 จาก 5 ระดับ



Figure 3 Arm-wearable heart rate monitors (A) Polar OH1[®] (B) Garmin Forerunner 235[®] (C) Schosche Rhythm 24[®] (D) Basis peak[®] (E) TomTom runner cardio[®] (F) Apple watch 4[®] (G) Mio Alpha[®] (H) Fitbit charge 2[®]

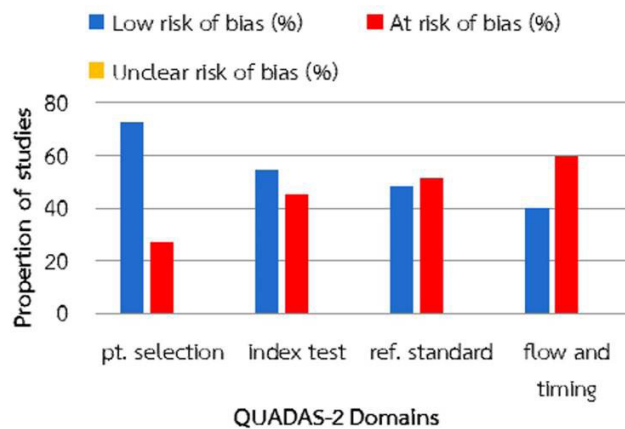


Figure 4 Assessment for risk of bias according to QUADAS-2

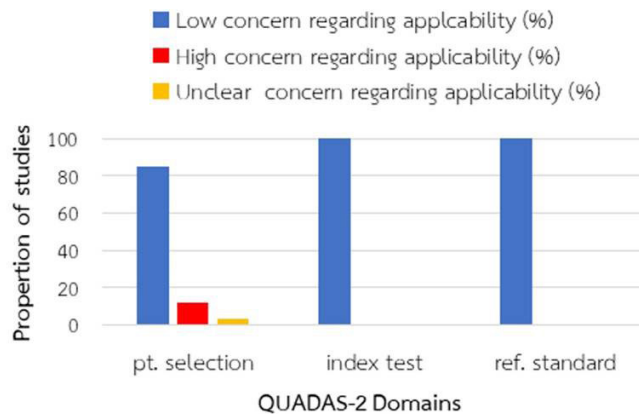


Figure 5 Assessments for applicability concern according to QUADAS-2

Table 1 Summarize the target population, protocol and validity and accuracy of arm-wearable heart rate monitors in previous studies (n=33)

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Baek et al. (2020) ¹³	Healthy subjects (n = 15) Age 23.7 ± 3.0 yrs	ECG	Fitbit (Fitbit Charge 2) PPG based	- Conventional walking (14 min) - Nordic Walking (14 min)	CCC = 0.89, MRD = -1.26 bpm 95% LoA = 9.90 to -12.42 bpm CCC = 0.76, MRD = 0.18 bpm 95% LoA = 18.8 to -18.45 bpm
Bai et al. (2017) ¹⁴	Healthy subjects (n = 39) Age 31.5 ± 11.0 yrs	Polar H7	Fitbit (Fitbit Charge HR) PPG based (2 LED)	- Aerobic exercises (25 min) (walk or jog on treadmill) - Self-selected pace	r = 0.84
Benedetto et al. (2018) ¹⁵	Healthy subjects (n = 15) Age 31.5 ± 4.0 yrs	ECG	Fitbit (Fitbit Charge HR) PPG based (2 LED)	- Cycle ergometer (10 min) - Maximum-selected level	ICC = 0.21 (0.09 to 0.34) 95% LoA = 16.8 to -28.5 bpm
Bondreaux et al. (2018) ¹⁶	Healthy subjects (n = 50) Age 22.0 ± 2.67 yrs	ECG	Apple (Apple Watch 2) PPG based Fitbit (Fitbit Blaze) PPG based Fitbit (Fitbit charge 2) PPG based Polar electro (Polar A360) PPG based Garmin Viviosmart HR PPG based TomTom (TomTom touch) PPG based	- Graded cycling (12 min) - Exhaustion level	ICC = 0.90 (MAPE = 4.14) r = 0.80 ICC = 0.50 (MAPE = 21.06) r < 0.75 ICC = 0.58 (MAPE = 21.36) r < 0.75 ICC = 0.52 (MAPE = 19.48) r < 0.75 ICC = 0.36 (MAPE = 25.38) r < 0.75 ICC = 0.59 (MAPE = 12.33) r < 0.75
Camilla et al. (2017) ¹⁷	Healthy subjects (n = 22) Age 22.0 ± 2.0 yrs	ECG	Fitbit (Fitbit charge HR) PPD based (2 LED)	- Walking on treadmill with progressed incline	ICC _{3,1} = 0.8 (0.67-0.9); MRD = -6.04 bpm 95% LoA = 14.35 to -26.42 bpm

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Chow et al. (2020) ¹⁸	Healthy subjects Age: 20-26 yrs (n = 20)	Polar H7	Xiaomi (Mi Band 2)	- Walking on treadmill, jogging, stationary bike and elliptical machine [24 min]	CCC = 0.76 (อายุ 20-26 ปี) MAPE = 7.69
			PPG based (2 LED)		CCC = 0.73 (อายุ ≥65 ปี) MAPE = 6.04
			Garmin (Vivosmart HR+)	- Moderate HR intensity	CCC = 0.92 (อายุ 20-26 ปี) MAPE = 3.77
Claes et al. (2017) ¹⁹	Healthy subjects (n = 12) Age 28.0 ± 4.79 yrs	ECG	PPG based (3 LED)		CCC = 0.80 (อายุ ≥65 ปี) MAPE = 4.73
			Garmin (Forerunner 225)	- Walking on treadmill with progressed incline (30 min)	r = 0.650 to 0.868; MRD = -1.57 bpm
Damasceno et al. (2021) ²⁰	Healthy subjects (n = 28) Age 23.41 ± 3.43 yrs	ECG	PPG based	- Walk or run on treadmill	r = 0.96; MRD = 0.7 bpm
			Garmin (Forerunner 735Xt)	(20 min) & self-selected intensity	95% LoA = 29.4 to -32.53 bpm
Dooley et al. (2017) ²¹	Healthy subjects (n = 62) Age 22.55 ± 4.34 yrs	Polar T31	Apple (Apple watch)	- Aerobic exercises (16 min) (walk or jog on treadmill)	r = 0.59 - 0.99
			PPG		
			Fitbit (Fitbit Charge HR)	- Light-vigorous intensity	r = 0.16 - 0.99
			PPG based (LED)		
Duking et al. (2020) ²²	Healthy subjects (n = 25) Age 26.0 ± 7.0 yrs	Polar H7	Garmin (Forerunner 225): PPG based (LED)		r = 0.05 - 0.75
			Apple (Apple watch 4)	- Walking and running at 1.1, 1.9, 2.7, 3.6, and 4.1 m/s (5 min for each speed) (25 min in total)	- r = 0.97 (light intensity), 0.85 - 1.00 (vigorous intensity), 0.95 (average)
			PPG based		
			Polar electro (Vantage V)		- r = 0.89 (light intensity), 0.75 - 0.91 (vigorous intensity), 0.88 (average)
			PPG based		
			Garmin (Fenix)	- Light-vigorous intensity	- r = 0.85 (light intensity), 0.58 - 0.83 (vigorous intensity), 0.77 (average)
			PPG based		
			Fitbit (Fitbit versa)		- r = 0.57 (light intensity), 0.52 - 0.82 (vigorous intensity), 0.65 (average)
			PPG		

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Gillinov et al. (2017) ²³	Healthy subjects (n = 25) Age 38.0 ± 12.0 yrs	ECG	Apple (Apple Watch) PPG based	- Circuit Training (Treadmill, Stationary bicycle, Elliptical trainer with or without-arm levers) (18 min) - Light to vigorous intensity	- $r_c = 0.92$ (circuit exercise), 0.93 (treadmill), 0.88 (bicycle) - elliptical trainer - $r_c = 0.94$ (without arm), 0.75 (with arm) - 95% LoA = 20 to -17 bpm
			Fitbit (Fitbit Blaze) PPG based		- $r_c = 0.67$ (circuit exercise), 0.76 (treadmill), 0.41 (bicycle) - elliptical trainer - $r_c = 0.58$ (without arm), 0.48 (with arm) - 95% LoA = 45 to -30 bpm
			TomTom International BV (TomTom Spark Cardio) PPG based		- $r_c = 0.83$ (circuit exercise), 0.88 (treadmill), 0.77 (bicycle) - elliptical trainer - $r_c = 0.55$ (without arm), 0.76 (with arm) - 95% LoA = 31 to -24 bpm
			Garmin (Forerunner 235) PPG based		- $r_c = 0.81$ (circuit exercise), 0.92 (treadmill), 0.91 (bicycle) - elliptical trainer - $r_c = 0.54$ (without arm), 0.31 (with arm) - 95% LoA = 33 to -27 bpm
Hettiarachchi et al. (2019) ²⁴	Healthy subjects (n = 24) Age 28.09 ± 5.50 yrs	ECG	Polar electro (Polar OH1) optical HR sensor	- Walking on a treadmill and riding a stationary spin bike (24 min) and low-vigorous intensity	- ICC = 0.95 - 0.99 - 95% LoA = 5.22 to -4.68 bpm

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Horton et al. (2017) ²⁵	Healthy subjects (n = 36) Age 40.5 ± 9.60 yrs	ECG	Polar electro (Polar M600) optical HR sensor	- Circuit Training [Cycling, Circuit weight training, Treadmill] (76 min with rest) - Low-vigorous intensity	OHR percent accuracy = 91.8 % (cycle intervals) OHR percent accuracy = 34.5 % (circuit weight training)
Khushhal et al. (2017) ²⁶	Healthy subjects (n = 21) Age 31.4 ± 7.20 yrs	Polar S810i	Apple (Apple Watch) PPG based	- Walking and running at 4, 7, and 10 km/h (5 min for each speed) (15 min in total) - Low-vigorous intensity	r = 0.97; 95%; LoA = -8 to 8 bpm (walking) r = 0.92 – 0.93 (jogging); 95% LoA = -19 to 21 bpm r = 0.81 – 0.86 (running); 95% LoA = -21 to 25 bpm
Muggeridge et al. (2021) ²⁷	Healthy subjects (n = 20) Age 40.0 ± 10.0 yrs	Polar H10	Polar electro (Polar OH1) PPG based (6 LED) Fitbit (Fitbit Charge 3) PPG (PurePulse tech.)	- Trial 1 sedentary activities, cycling on a bicycle ergometer, and incremental exercise (18-42 min) on high intensity - Trial 2; 4 x 15-second maximal sprints on a cycle ergometer and 4 x 30-second sprints on high intensity	r = 0.95 - 95% LoA = -20 to 19 bpm r = 0.8 - 95% LoA = -46 to 33 bpm
Muller et al. (2019) ²⁸	Healthy subjects (n = 55) Age 30.5 ± 9.8 yrs	Polar H10	Tempo HR (NSC) (-) Polar electro (Polar A370) PPG based	Laboratory: An incremental cycling protocol (20 min) - moderate to vigorous intensity Free-living - ADL (waking hours) (10 min) - moderate to vigorous intensity	- Laboratory: ICC = 0.51; MAPE = 13.00 - free-living: ICC = 0.71; MAPE = 10.20 - Laboratory: ICC = 0.73; MAPE = 6.4 - free-living: ICC = 0.83; MAPE = 7.1

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Navalta et al. (2020) ²⁹	Healthy subjects (n = 21) Age 31.0 ± 11.0 yrs	Polar H7	Suunto Oy (Spartan Sport watch) ECG watch Garmin (Fenix 5 wristwatch) PPG based	- Trail run 1.61 km uphill direction 1.61 km downhill direction	$r_c = 0.955$ (overall) $r_c = 0.964$ (uphill) and 0.949 (downhill) - 95% LoA = -62 to 61 bpm
					$r_c = 0.316$ (overall) $r_c = 0.450$ (uphill) and 0.381 (downhill) - 95% LoA = -32 to 162 bpm
			Scosche Industries Inc. (Rhythm+ forearm band) (-)		$r_c = 0.780$ (overall) $r_c = 0.699$ (uphill) and 0.885 (downhill) - 95% LoA = -114 to 120 bpm
O'Driscoll et al. (2019) ³⁰	Healthy subjects (n = 59) Age 44.4 ± 14.14 yrs	Polar H7	Fitbit (Fitbit charge 2) PPG (LED)	- Treadmill walking, incline walking, running, incline running and cycle ergometer (30 min) - Light to moderate intensity	- Exercise $r_c = 0.23$ and MAPE = 25 (walk) $r_c = 0.29$ and MAPE = 17 (walk incline) $r_c = 0.66$ and MAPE = 8 (run) $r_c = 0.81$ and MAPE = 5 (run incline) $r_c = 0.55$ and MAPE = 12 (cyclic-light) $r_c = 0.62$ and MAPE = 16 (cyclic-mod) - Household task $r_c = 0.29$ and MAPE = 15 (folding) $r_c = 0.31$ and MAPE = 16 (sweeping)
Olstad and Zinner (2019) ³¹	Athletes (n = 26) Age 18.1 ± 2.3 yrs	Polar H10	Polar electro (Polar M600) PPG based	- Swimming (70-85 min) - Moderate to vigorous level	ICC = 0.347 - 0.366 (Moderate level) ICC = 0.209 - 0.438 (Vigorous level)
Parak and Korhonen (2014) ³²	Healthy subjects (n = 21) Age 31.3 ± 10.7 yrs	ECG	Mio Tech. (Mio Alpha) PPG (green LED) Scosche (myRhythm) PPG	Aerobic exercise (Walking, Running, Cycling) (51 min with rest period)	MAPE (Mio) = 5.23 (Overall) MAPE (Scosche) = 6.78 (Overall)

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Parak et al. (2017) ³³	Healthy subjects (n = 24) Age 36.2 ± 8.2 yrs	Polar RS800CX	PulseON® (PulseOn) Optical heart rate sensor	- Outdoor running (≥ 20 mins) and self-selected pace - Running on treadmill (25 mins) and vigorous intensity	MAPE = 1.9 (running) MRD = 0.9 bpm - 95% LoA = -9.5 to 11.3 bpm
Pasadyn et al. (2019) ⁵	Athletes (n = 50) Average Age 29.0 yrs	ECG	Apple (Apple watch III) PPG	- Progressed running at the following levels on the treadmill (12 mins) - Light to vigorous level	CCC = 0.96
			Fitbit (Fitbit iconic) PPG		CCC = 0.89
			Garmin (Vivosmart HR) PPG		CCC = 0.89
			TomTom Int. (Spark 3) PPG		CCC = 0.89
Reddy et al. (2018) ³⁴	Healthy subjects (n = 20) Age 27.5 ± 6.0 yrs	Polar H7	Fitbit (Fitbit charge 2) PPG based	- Maximal physical activity (graded maximal aerobic and resistance circuit workout) - Vigorous intensity - ADLs and HIIT-treadmill or HIIT-cycle ergometer (50 min; 2 times) - Vigorous intensity	r = 0.94 and r _c = 0.92 (treadmill) r = 0.46 and r _c = 0.36 (cycle ergo.) r = 0.88 and r _c = 0.72 (resistance ex.) r = 0.73 and r _c = 0.72 (ADLs) r = 0.88 and r _c = 0.87 (HIIT-treadmill) r = 0.56 and r _c = 0.47 (HIIT-cycle ergo.)
			Garmin Vivosmart HR+ - PPG		r = 0.94 and r _c = 0.93 (treadmill) r = 0.71 and r _c = 0.66 (cycle ergo.) r = 0.90 and r _c = 0.54 (resistance ex.) r = 0.69 and r _c = 0.68 (ADLs) r = 0.90 and r _c = 0.90 (HIIT-treadmill) r = 0.47 and r _c = 0.37 (HIIT-cycle ergo.)

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Reece et al. (2021) ³⁵	Healthy subjects (n = 23) Age 23.57 ± 3.22 yrs	Polar H7	Apple (Apple watch 4) PPG based	CTA protocol (sedentary lifestyle activities, lifestyle activities, walking, jogging, running, and cycling)	ICC = 0.963 and MAPE = 2.95 (ADL) 95% LoA = -38.58 to 38.77 bpm ICC = 0.998 and MAPE = 0.62 (cycling) 95% LoA = -13.40 to 13.57 bpm
			Garmin (Forerunner 735 XT) PPG based	(57 min with rest)	ICC = 0.710 and MAPE = 12.09 (ADL) 95% LoA = -129.89 to 123.49 bpm
			Schösche In (Rhythm 24) PPG based (Infrared LED)		ICC = 0.809 and MAPE = 7.56 (cycling) 95% LoA = -226.68 to 232.95 bpm
					ICC = 0.940 and MAPE = 4.12 (ADL) 95% LoA = -49.50 to 50.09 bpm ICC = 0.997 and MAPE = 0.90 (cycling) 95% LoA = -15.88 to 16.22 bpm
Ruiz-Malago'n et al. (2020) ³⁶	Healthy subjects (n = 37) Age 22.0 ± 9.0 yrs	Polar V800-H10	Polar electro (vantage M) PPG based	Running on treadmill (10 mins) & self-selected pace	r = 0.69 - 0.95; ICC = 0.69 - 0.97 MRD = 0.26 - 10.71 bpm
Shcherbina et al. (2017) ³⁷	Healthy subjects (n = 60) Age 38.0 ± 11.0 yrs	ECG	Apple (Apple watch) PPG based	- Circuit Training (Treadmill, cycle ergometer) (30 mins)	MAPE < 5.0 (cycle ergometer) & 2.5 (walk)
			Basis peak (Basis peak) Optical heart rate sensor	- Low to vigorous level	MAPE < 5.0 (cycle ergometer) & 6.5 (walk)
			Fitbit (Fitbit surge) PPG based		MAPE < 5.0 (cycle ergometer) & 8.8 (walk)
			Microsoft Coro. (MB) Optical heart rate		MAPE < 5.0 (cycle ergometer) & 5.6 (walk)
			Mio Tech. (Mio Alpha 2) PPG based		MAPE < 5.0 (cycle ergometer)
			Samsung (Gear S2) heart rate sensor		MAPE = 5.1 (cycle ergometer)

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Shumate et al. (2021) ³⁸	Athletes (n = 30) Age 26.10 ± 3.39 yrs	ECG	Polar electro (Polar Vantage M) PPG based	- Treadmill at 2.7, 4.0, 5.4, 6.7, 8.0 km/h with 10% increment and increasing 2% every speed stage (at least 15 min) - Low to vigorous intensity	r = 0.865 - light intensity MRD = 0.982 bpm ICC = 0.735 - 0.880 95% LoA = -15.54 to 9.9 bpm - moderate intensity ICC = 0.80 - 0.878 - vigorous intensity MRD = 9.97 bpm ICC = 0.141 - 0.623 95% LoA = -11.53 to 31.47 bpm
Stahl et al. (2016) ³⁹	Athletes (n = 50) Average Age 26.28 yrs	Polar RS400	Scosche (Scosche Rhythm) PPG based Mio Technology (Mio Alpha) PPG based Fitbit (Fitbit Charge HR) PPG based BASIS Science (Basis Peak) PPG based Microsoft Corp (Microsoft) PPG based TomTom International BV (TomTom Runner Cardio) PPG based	- Treadmill (walking and running at 3.2, 4.8, 6.4, 8.0, and 9.6 km/h) (30 min) - moderate to high level	r = 0.930 MAPE = 4.0 r = 0.929 MAPE = 4.6 r = 0.933 MAPE = 6.2 r = 0.954 MAPE = 3.6 r = 0.956 MAPE = 4.8 r = 0.959 MAPE = 3.3

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Støve et al. (2020) ⁴⁰	Healthy subjects (n = 30) Age 24.8 ± 6.3 yrs	Polar H10	Apple (Apple Watch 2) PPG based	- Circuit Training (cycling, walking/running and weightlifting with rapid arm movement) (30 min) - submaximal to near-maximal intensities	Cyclic: - MRD (bpm) = -2.1 (L), -2.3 (M), -1.6 (H) - MAPE = 4.9 (L), 3.1 (M), 2.3 (H) Treadmill: - MRD (bpm) = -1.0 (L), -3.6 (M), -5.0 (H) - MAPE = 3.6 (L), 3.4 (M), 3.6 (H) Rapid arm movements: - MRD = 3.0 bpm; MAPE = 3.1 - 95% LoA = -4.20 to 10.20 bpm
			Garmin International Inc (Garmin Forerunner 235) PPG based		Cyclic: - MRD (bpm) = -16.1 (L), -30.0 (M), -42.1 (H) - MAPE = 26.9 (L), 55.7 (M), 78.1 (H) Treadmill: - MRD (bpm) = 1.3 (L), -0.2 (M), -6.2 (H) - MAPE = 9.8 (L), 5.6 (M), 4.6 (H) Rapid arm movements: - MRD = 6.1 bpm; MAPE = 6.5 - 95% LoA = -10.7 to 23.0 bpm
Sun and Liu. (2020) ⁴¹	Athletes (n = 11) Age 22.0 [IQR = 3.75] yrs	Polar Team 2 Pro	Fifth Zone Fitness Laboratory Company (Fizzo) Optical sensors	- Running on a treadmill (self-selected speed, 6 and 12 km/h) (24 min) on moderate and vigorous intensity	Validity ICC = 0.99 MAPE = 1.62 - 1.82 MRD = 0.64-0.69 bpm 95% LoA = -5.96 to 7.24 bpm

Table 1 Continued

Authors (year)	Target populations	Gold standard	Brands (model) & Technology	Exercise protocol	Result
Thiebaud et al. (2018) ⁴²	Athletes (n = 20) Age 22.0 ± 3.0 yrs	ECG	Fitbit (Fitbit Surges) PPG based Microsoft (Microsoft Band) PPG based TomTom (TT Runner Cardio) PPG based	- Walking and running on a treadmill at 3.2, 4.8, 6.4, 8 and 9.7 km/h (15 min) - Light to vigorous level	r = 0.53 - 0.84 (Overall) MAPE = 2.17 to 8.06% (Overall) r = 0.35 - 0.63 (Overall) MAPE = 1.31 to 7.37% r = 0.52 - 0.91 (Overall) MAPE = 1.01 to 7.49%
Thomson et al. (2019) ⁴³	Healthy subjects (n = 30) Age 23.5 ± 3.0 yrs	ECG	Fitbit (Fitbit Charge HR 2) PPG based Apple (Apple Watch) PPG based	- Walking on treadmill with progressed incline (until volitional fatigue) - Very light to very vigorous intensity	CCC = 0.89 (very light intensity), 0.55 (light intensity), 0.25 (moderate), -0.02 (vigorous), -0.11 (very vigorous) CCC = 0.98 (very light intensity), 0.68 (light intensity), 0.40 (moderate), 0.45 (vigorous), 0.31 (very vigorous)
Wallen et al. (2016) ⁴⁴	Healthy subjects (n = 22) Age 24.0 ± 5.6 yrs	ECG	Apple (Apple Watch) PPG based Fitbit (Fitbit Charge HR) PPG based Samsung (Samsung Gear S) heart rate sensor Mio Technology (Mio Alpha) PPG based	- Activities at rest [lying, sitting, standing] and exercise [walking, cycling] (58 min) - Moderate to vigorous level	r = 0.95, ICC = 0.98, MRD = -1.3 bpm - 95% LoA = -9.9 to 7.3 bpm r = 0.81, ICC = 0.78, MRD = -9.3 bpm - 95% LoA = -26.0 to 7.4 bpm r = 0.67, ICC = 0.80, MRD = -7.1 bpm - 95% LoA = -27.3 to 13.1 bpm r = 0.87, ICC = 0.91, MRD = -4.3 bpm - 95% LoA (bpm) = 0.4.avg - 61.2 to -0.44.avg + 52.69

Remark: CCC and r_c : concordance correlation coefficient, LoA; Limit of Agreement, ECG; Electrocardiogram, MAPE; Mean absolute percentage error, PPG; photoplethysmography, MRD; mean relative difference, L; Light, M; Moderate, H; High

อภิปรายผล

การศึกษานี้เป็นการรวบรวมการรายงานค่าคุณสมบัติในการวัดเชิงจิตวิทยาของเครื่องมือในอุปกรณ์การวัดอัตราการเต้นของหัวใจที่แขน โดยใช้วิธีการทบทวนอย่างเป็นระบบ ซึ่งถือเป็นหลักฐานเชิงประจักษ์ที่มีระดับความน่าเชื่อถือสูง ตามการจัดลำดับของสถาบัน Oxford Centre for Evidence-Based Medicine⁴⁵ เพื่อให้เกิดความมั่นใจว่า สามารถนำผลการศึกษาไปใช้ประโยชน์ในการอ้างอิงต่อไปได้ โดยมีวัตถุประสงค์เพื่อให้ได้ข้อมูลเกี่ยวกับค่าความตรงและความแม่นยำของเครื่องมือในการวัดอัตราการเต้นของหัวใจในขณะที่ยกกำลังกายแบบต่างๆ และแนวทางการนำความรู้ไปประยุกต์ใช้ โดยจากการศึกษานี้รวบรวมข้อมูลของอุปกรณ์การวัดอัตราการเต้นของหัวใจ ประกอบไปด้วย อุปกรณ์จำนวน 15 ตรายี่ห้อ และ 41 รุ่น อุปกรณ์ที่ได้นำมาทดสอบมากที่สุดคือ Fitbit charge HR ในขณะที่อุปกรณ์จาก Xiaomi, Tempo HR, Sunto oy และ fifth zone fitness laboratory company ได้รับการทดสอบเพียง 1 ฉบับเท่านั้น โดยเมื่อเปรียบเทียบกับการศึกษาที่ทบทวนวรรณกรรมอย่างเป็นระบบในอดีตพบว่า มีการศึกษาในอุปกรณ์เพียง 9, 7 และ 2 ตรายี่ห้อเท่านั้น^{1,46,47}

เมื่อประเมินคุณภาพของการศึกษาพบว่า ลักษณะงานวิจัยที่ได้ทบทวนมานั้นมีลักษณะเป็นรูปแบบ cross sectional study ซึ่งเป็นธรรมชาติของงานวิจัยที่ต้องการศึกษาหาความตรงตามสภาพและความแม่นยำ ซึ่งส่งผลทำให้การศึกษาในรูปแบบนี้มีคุณภาพไม่สูงมากนักเมื่อเทียบกับงานวิจัยแบบ randomized control trial อย่างไรก็ดีตาม จากผลการประเมินคุณภาพงานวิจัยจำนวน 33 บทความพบว่าบทความวิจัยส่วนใหญ่ (ร้อยละ 81.82 ของจำนวนบทความทั้งหมด) มีคุณภาพอยู่ในระดับ 2 ตามการจัดระดับของ Oxford level of evidence ซึ่งมีระดับความน่าเชื่อถือในระดับดี รองจากการศึกษาแบบ systematic review และ randomized controlled trial ดังนั้น การศึกษาเหล่านี้สามารถ

นำไปปฏิบัติได้ แต่ต้องมีการฝึกทักษะผู้ปฏิบัติเพิ่มเติม และต้องหาการศึกษาอื่นๆ เพิ่มเติมเพื่อสนับสนุน

สำหรับความแม่นยำ และความตรงในการวัดอัตราการเต้นของหัวใจในการศึกษานี้ พบว่ามีอุปกรณ์เป็นส่วนน้อยที่สามารถวัดได้อย่างแม่นยำ โดยพิจารณาจากค่า MAPE ที่พบว่ามีอุปกรณ์ที่ผ่านเกณฑ์มาตรฐานเพียงแค่ 9 รุ่น จากทั้งสิ้น 41 รุ่น คือ Apple watch, Fitbit surge, TomTom runner cardio, Schosche Rhythm, Microsoft band, Mio Alpha, Basis peak, Fizzo, และ Mio alpha 2 โดยมีความแม่นยำ และตรงแตกต่างกันไปในหลายๆ กิจกรรม โดยพบว่า Apple watch เป็นเครื่องมือที่มีความแม่นยำสูงที่สุดทั้งในการออกกำลังกายแบบการเดิน วิ่ง หรือปั่นจักรยาน หรือแม้แต่การวัดอัตราการเต้นของหัวใจในชีวิตประจำวัน โดยคาดว่ามาจากจำนวน และประสิทธิภาพของ sensor และ photodiode sensor ที่ใช้ ในขณะที่ Fitbit charge 2 เป็นเครื่องมือที่วัดอัตราการเต้นของหัวใจได้ไม่แม่นยำในกิจกรรมการเดิน วิ่ง หรือปั่นจักรยาน แต่มีความแม่นยำมากขึ้นเมื่อออกกำลังกายแบบมีแรงต้าน จึงเป็นที่น่าสนใจว่า ทั้งๆ ที่ในอุปกรณ์รุ่นเดียวกัน แต่กลับมีค่าความแม่นยำ และความตรงไม่เหมือนกัน เมื่อเปลี่ยนกิจกรรมการออกกำลังกาย ดังนั้น นอกจากเทคโนโลยีที่ใช้ อาจจะมีปัจจัยอื่น ๆ ซึ่งส่งผลกระทบต่อค่าความแม่นยำ และความตรง ยกตัวอย่างเช่น การศึกษาของ Olstad และ Zinner ในปี 2020³¹ ที่ต้องการวัดอัตราการเต้นของหัวใจในขณะว่ายน้ำ แต่พบว่ามีความแม่นยำที่ต่ำมากอันเนื่องมาจากน้ำที่อยู่ระหว่าง sensor และผิวหนังอาจจะรบกวนต่อค่าที่เกิดขึ้น หรือการศึกษาของ Gillinov และคณะ ในปี 2017²³ ซึ่งพบว่าเมื่อออกกำลังกายโดยมีการเคลื่อนไหวของแขน ค่าความแม่นยำของอุปกรณ์ชนิดต่างๆ มีค่าลดลง

อย่างไรก็ตาม การศึกษานี้มีข้อจำกัดของการศึกษาบางประการคือ 1) รายงานการวิจัยที่ย่อยนอกเหนือจากฐานข้อมูลที่ระบุไว้ข้างต้น หรือใช้ภาษาอื่น ๆ นอกจากภาษาอังกฤษ และภาษาไทยจะไม่ได้ถูกนำมา

พิจารณาในการศึกษานี้ 2) เครื่องมือ หรืออุปกรณ์บางชนิดที่ได้นำเสนอในการศึกษานี้บางชนิดหยุดการจัดจำหน่าย หรือยกเลิกการขายไปแล้ว รวมถึงมีอุปกรณ์ใหม่เป็นจำนวนมากที่ไม่ถูกรวบรวมอยู่ในการศึกษานี้ เนื่องจากยังไม่ได้รับการศึกษา หรืออยู่ในช่วงระหว่างศึกษาหาความตรง และความแม่นยำ และ 3) บางตราสินค้าได้รับการทดสอบจำนวนน้อยจึงอาจส่งผลกระทบต่อความน่าเชื่อถือของเครื่องมือ 4) ไม่สามารถวิเคราะห์ข้อมูลแบบ meta - analysis ได้เนื่องจากมีอุปกรณ์มาตรฐาน และตราสินค้าจำนวนมาก รวมถึงบางรายงานการวิจัยที่รวบรวมมาไม่ได้ให้รายละเอียดอย่างเพียงพอ และ 5) การศึกษานี้พบว่าส่วนใหญ่เป็นอุปกรณ์ที่มีราคาสูงอยู่ในช่วง 5,590 - 31,900 บาท แต่ยังไม่มีการรายงานค่าต่างๆ ในอุปกรณ์ที่มีราคาต่ำ ซึ่งได้รับการผลิตออกมาเป็นจำนวนมากในปัจจุบัน

ดังนั้น การศึกษาในอนาคตควรจะทำการศึกษาโดยมุ่งเน้นไปที่สินค้าที่เป็นปัจจุบัน ยกตัวอย่างเช่น Apple watch series 4 คือรุ่นล่าสุดที่พบในการศึกษานี้ แต่พบว่า Apple watch series 7 เป็นสินค้าปัจจุบันที่จัดจำหน่ายในร้านค้า ส่งผลทำให้ข้อมูลที่ได้ไม่สอดคล้องกับความเป็นจริง รวมถึงศึกษาปัจจัยรบกวนต่อค่าความตรง และความแม่นยำของอุปกรณ์ และอุปกรณ์ในช่วงราคาที่แตกต่างจากปัจจุบัน เพื่อให้ได้ข้อมูลที่ครอบคลุมต่อผู้ใช้งาน

สรุปผล

การประยุกต์ใช้ในการปฏิบัติ (implication for practice)

จากการประเมินความน่าเชื่อถือของหลักฐานงานวิจัยในการศึกษานี้ตามเกณฑ์ OCEBM พบว่ามีความน่าเชื่อถือในระดับ B ดังนั้น อุปกรณ์ต่างๆ ใน การศึกษานี้ สามารถใช้วัดอัตราการเต้นของหัวใจเพื่อใช้ประเมินความหนักขณะออกกำลังกายในผู้ที่มีสุขภาพดี โดย Apple watch เป็นอุปกรณ์ที่มีความแม่นยำ และความตรงในการวัดอัตราการเต้นของหัวใจ ในผู้ที่มีสุขภาพดีขณะทำกิจกรรมต่างๆ มากที่สุด

ข้อเสนอแนะ

การนำอุปกรณ์อื่นๆ ในการศึกษานี้ที่ใช้เทคโนโลยี PPG เช่นเดียวกัน แต่มีค่าความตรง และความแม่นยำต่างๆ ไปใช้ในการวัดอัตราการเต้นของหัวใจควรคำนึงว่าข้อมูลที่ได้จากอุปกรณ์เหล่านั้นอาจรายงานค่าอัตราการเต้นของหัวใจได้ไม่แม่นยำเท่ากับเครื่องมือมาตรฐาน ซึ่งไม่เหมาะสมสำหรับผู้ที่ต้องการทราบอัตราการเต้นของหัวใจที่แม่นยำสูง แต่อาจใช้ได้ในกลุ่มวัยรุ่น หรือนักกีฬาที่ไม่มีข้อห้าม ข้อควรระวัง ต้องทราบเพียงระดับความหนักของการออกกำลังกายเท่านั้น

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณทุนสนับสนุน จาก คณะกายภาพบำบัด มหาวิทยาลัยศรีนครินทรวิโรฒ ที่ให้การสนับสนุนด้านงบประมาณ ขอขอบคุณนางสาว ชญานิศ ตันติเวทเรืองเดช นางสาวธัญญรัตน์ บุญรอดจัญญ นางสาวนวรินทร์ สภาพการ และ นางสาวสุทธิกานต์ ประสิทธิ์ผล ที่เป็นกำลังสำคัญในการรวบรวมข้อมูล สำหรับการศึกษาในครั้งนี้

เอกสารอ้างอิง

1. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring Steps, energy expenditure, and heart rate: Systematic review. JMIR Mhealth Uhealth 2020; 8:e18694.
2. WHO guidelines approved by the guidelines review committee. In: WHO guidelines on physical activity and sedentary behaviour. World Health Organization® World Health Organization 2020.: Geneva, 2020.

3. Lambrick DM, Faulkner JA, Rowlands AV, et al. Prediction of maximal oxygen uptake from submaximal ratings of perceived exertion and heart rate during a continuous exercise test: the efficacy of RPE 13. *Eur J Appl Physiol* 2009;107:1-9.
4. Kritsana B, Ubon P, Peanchai K. Cardiovascular response and energy expenditure of a novel heart rate maximizer test in healthy volunteers. *Bull Chiang Mai Assoc Med Sci* 2016;49:263-75.
5. Pasadyn SR, Soudan M, Gillinov M, et al. Accuracy of commercially available heart rate monitors in athletes: A prospective study. *Cardiovasc Diagn Ther* 2019;9:379-85.
6. Sartor F, Gelissen J, van Dinther R, et al. Wrist-worn optical and chest strap heart rate comparison in A heterogeneous sample of healthy individuals and in coronary artery disease patients. *BMC Sports Sci Med Rehabil* 2018;10:10.
7. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; 4:1.
8. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539-49.
9. Kaizik MA, Garcia AN, Hancock MJ, et al. Measurement properties of quality assessment tools for studies of diagnostic accuracy. *Braz J Phys Ther* 2020;24:177-84.
10. Swinscow T, Campbell M. *Statistics at Square One*, BMJ Books: Oxford, 2009.
11. Altman D. *Practical statistics for medical research.*, Chapman and Hall: London, 1991.
12. Swanson D. On the relationship among values of the same summary measure of error when it is used across multiple characteristics at the same point in time: An examination of MALPE and MAPE 1. *Rev Econ Finan* 2015;5:1-14.
13. Baek S, Ha Y, Park HW. Accuracy of wearable devices for measuring heart rate during conventional and Nordic walking. *Pm r* 2021;13:379-86.
14. Bai Y, Hibbing P, Mantis C, et al. Comparative evaluation of heart rate-based monitors: Apple watch vs Fitbit charge HR. *J Sports Sci* 2018;36:1734-41.
15. Benedetto S, Caldato C, Bazzan E, et al. Assessment of the fitbit charge 2 for monitoring heart rate. *PLoS One* 2018;13: e0192691.
16. Boudreaux BD, Hebert EP, Hollander DB, et al. Validity of wearable activity monitors during cycling and resistance exercise. *Med Sci Sports Exerc* 2018;50:624-33.
17. Powierza CS, Clark MD, Hughes JM, et al. Validation of a self-monitoring tool for use in exercise therapy. *Pm r* 2017;9: 1077-84.
18. Chow HW, Yang CC. Accuracy of optical heart rate sensing technology in wearable fitness trackers for young and older Adults: Validation and comparison study. *JMIR Mhealth Uhealth* 2020;8:e14707.

19. Claes J, Buys R, Avila A, et al. Validity of heart rate measurements by the Garmin Forerunner 225 at different walking intensities. *J Med Eng Technol* 2017;41: 480-5.
20. Damasceno V, Costa A, Campello M, et al. Criterion validity and accuracy of a heart rate monitor. *Hum Mov* 2022;23:60-8.
21. Dooley EE, Golaszewski NM, Bartholomew JB. Estimating accuracy at exercise intensities: A comparative study of self-monitoring heart rate and physical activity wearable devices. *JMIR Mhealth Uhealth* 2017;5:e34.
22. Düking P, Giessing L, Frenkel MO, et al. Wrist-Worn wearables for monitoring heart rate and energy expenditure while Sitting or performing light-to-vigorous physical activity: Validation study. *JMIR Mhealth Uhealth* 2020;8:e16716.
23. Gillinov S, Etiwy M, Wang R, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc* 2017;49:1697-703.
24. Hettiarachchi IT, Hanoun S, Nahavandi D, et al. Validation of polar OH1 optical heart rate sensor for moderate and high intensity physical activities. *PLoS One* 2019;14:e0217288.
25. Horton JF, Stergiou P, Fung TS, et al. Comparison of polar M600 optical heart rate and ECG heart rate during exercise. *Med Sci Sports Exerc* 2017;49:2600-7.
26. Khushhal A, Nichols S, Evans W, et al. Validity and reliability of the Apple watch for measuring heart rate during exercise. *Sports Med Int Open* 2017;1:e206-e11.
27. Muggeridge DJ, Hickson K, Davies AV, et al. Measurement of heart rate using the polar OH1 and fitbit Charge 3 wearable devices in healthy adults during light, moderate, vigorous, and sprint-based exercise: Validation study. *JMIR Mhealth Uhealth* 2021;9:e25313.
28. Müller AM, Wang NX, Yao J, et al. Heart rate measures from wrist-worn activity trackers in a laboratory and free-living setting: Validation study. *JMIR Mhealth Uhealth* 2019;7:e14120.
29. Navalta JW, Montes J, Bodell NG, et al. Concurrent heart rate validity of wearable technology devices during trail running. *PLoS One* 2020;15:e0238569.
30. O'Driscoll R, Turicchi J, Hopkins M, et al. The validity of two widely used commercial and research-grade activity monitors, during resting, household and activity behaviours. *Health Techno* 2020;10:637-48.
31. Olstad BH, Zinner C. Validation of the Polar OH1 and M600 optical heart rate sensors during front crawl swim training. *PLoS One* 2020;15:e0231522.
32. Parak J, Korhonen I. Evaluation of wearable consumer heart rate monitors based on photoplethysmography. *Annu Int Conf IEEE Eng Med Biol Soc* 2014;2014:3670-3.

33. Parak J, Uuskoski M, Macheek J, et al. Estimating heart rate, energy expenditure, and physical performance with a wrist photoplethysmographic device during running. *JMIR Mhealth Uhealth* 2017;5:e97.
34. Reddy RK, Pooni R, Zaharieva DP, et al. Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: evaluation study. *JMIR Mhealth Uhealth* 2018;6:e10338.
35. Reece JD, Bunn JA, Choi M, et al. Assessing heart rate using consumer technology association standards. *Technologies* 2021;9:46.
36. Ruiz-Malagón E, Ruiz-Alias S, García-Pinillos F, et al. Comparison between photoplethysmographic heart rate monitor from Polar vantage M and Polar V800 with H10 chest strap while running on a treadmill: Validation of the polar precision prime TM photoplethysmographic system. *Proc Inst Mech Eng P J Sport Eng Technol* 2021;235:212-8.
37. Shcherbina A, Mattsson CM, Waggott D, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *J Pers Med* 2017;24:3.
38. Shumate T, Link M, Furness J, et al. Validity of the Polar Vantage M watch when measuring heart rate at different exercise intensities. *PeerJ* 2021;9:e10893.
39. Stahl SE, An HS, Dinkel DM, et al. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc Med* 2016;2:e000106.
40. Støve MP, Holm RS, Kjaersgaard AS, et al. Measurement latency significantly contributes to reduced heart rate measurement accuracy in wearable devices. *J Med Eng Technol* 2020;44:125-32.
41. Sun J, Liu Y. Using smart bracelets to assess heart rate among students during physical education lessons: feasibility, reliability, and Validity study. *JMIR Mhealth Uhealth* 2020;8:e17699.
42. Thiebaud RS, Funk MD, Patton JC, et al. Validity of wrist-worn consumer products to measure heart rate and energy expenditure. *Digit Health* 2018;4:2055207618770322.
43. Thomson EA, Nuss K, Comstock A, et al. Heart rate measures from the Apple watch, fitbit charge HR 2, and electrocardiogram across different exercise intensities. *J Sports Sci* 2019;37:1411-9.
44. Wallen MP, Gomersall SR, Keating SE, et al. Accuracy of heart rate watches: Implications for weight management. *PLoS One* 2016;11:e0154420.
45. Winkelmann RR, Kim GK, Del Rosso JQ. Treatment of cutaneous lupus erythematosus: Review and assessment of treatment benefits based on oxford centre for evidence-based medicine criteria. *J Clin Aesthet Dermatol* 2013;6:27-38.

46. Feehan LM, Geldman J, Sayre EC, et al. Accuracy of fitbit devices: systematic Review and narrative Syntheses of quantitative data. *JMIR Mhealth Uhealth* 2018;6:e10527.
47. Fokkema T, Kooiman TJ, Krijnen WP, et al. Reliability and validity of ten consumer activity trackers depend on walking speed. *Med Sci Sports Exerc* 2017;49:793-800.