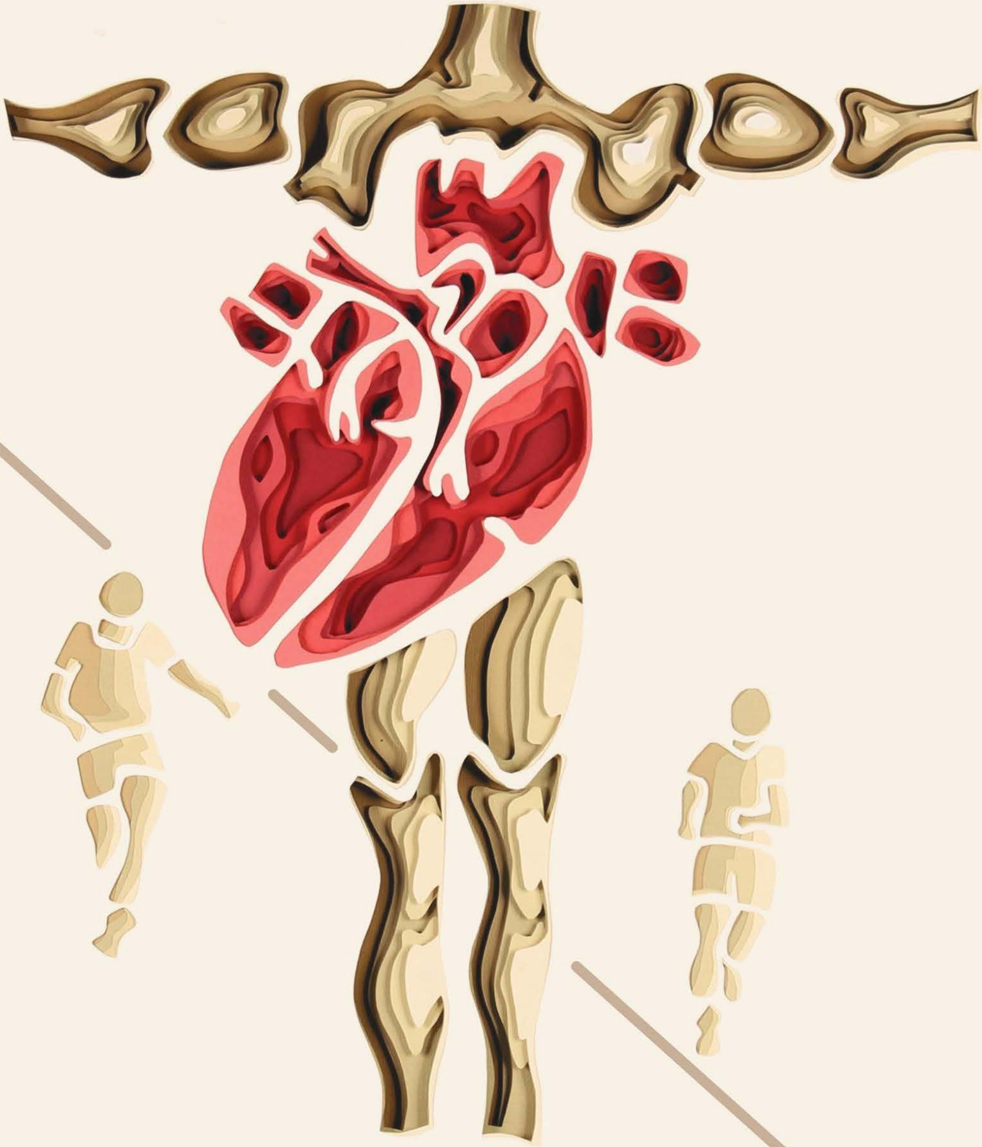
PHYSICAL FITNESS & ACTIVITY

Reference values and a clinical application in major abdominal surgery



Caspar F Mylius

PHYSICAL FITNESS & ACTIVITY

Reference values and a clinical application in major abdominal surgery

Caspar F. Mylius

Financial support for the printing of this thesis by the following sponsors is gratefully acknowledged:

Cover illustration: Nancy Halsema
Layour and printing by Ridderprint | www.ridderprint.nl

rijksuniversiteit
groningen

Physical Fitness & Activity

Reference values and a clinical application in major abdominal surgery

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. ir. J.M.A. Scherpen
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op 28 februari 2024

om 11:00 uur

door

Caspar Mijlius
geboren op 21 januari 1986
te Leeuwarden

Promotores

Prof. dr. C.P. van der Schans
Prof. dr. J.M. Klaase

Copromotores

Dr. W.P. Krijnen
Dr. T. Takken

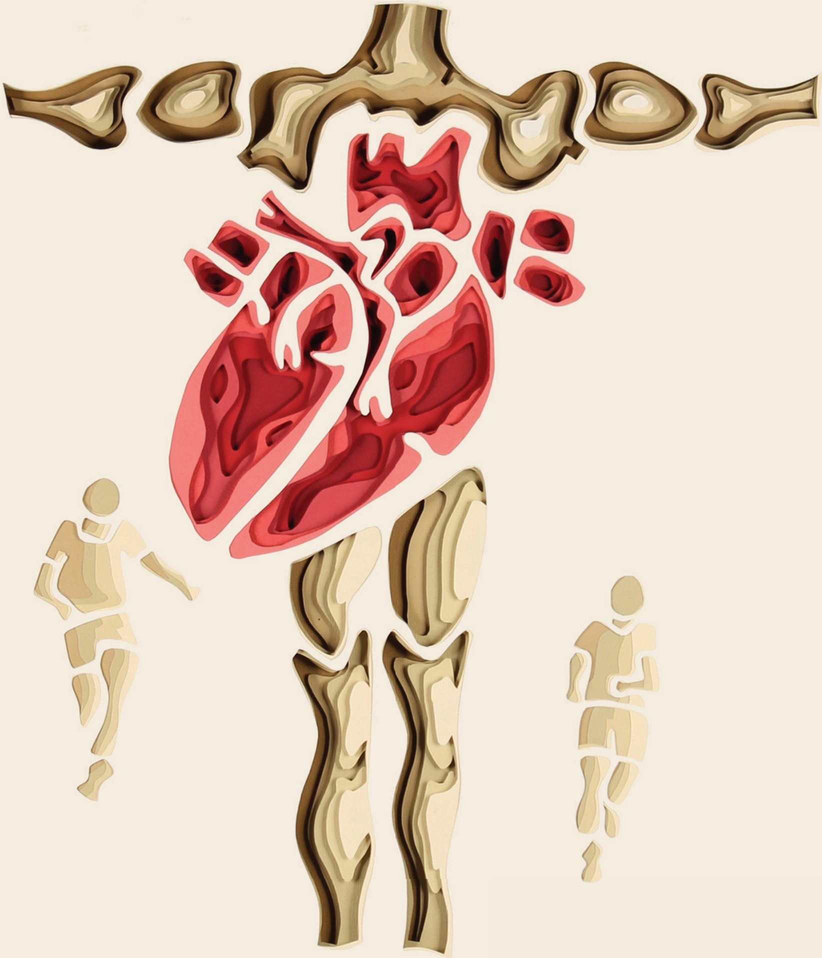
Beoordelingscommissie

Prof. dr. J. Zwerver
Prof. dr. S.W.M. Olde Damink
Prof. dr. R. Dekker

Paranimfen

Harm Brugge
Dr. Mathijs van Ark

TABLE OF CONTENTS



Chapter 1

General Introduction

Physical fitness is defined as a set of attributes that people have or achieve that relates to a person's ability to perform physical activity. (1) Therefore, impairments in physical fitness contribute to limitations in the activity and participation domains of the International Classification of Functioning, Disability and Health (**figure 1**). (2) Additionally, physical fitness is recognized as an important link between diseases or impairments and activity or task limitations in theoretical models of disability onset and progression. (3–5) Namely, reduced or impaired physical fitness, due to sedentary behavior, disease or injury, contributes to activity limitations, and may trigger accommodations to compensate for gaps between fitness and the ability to carry out activities and participation. (6) The physical fitness construct consists of both skill and health-related factors. Skill-related factors pertain the athletic ability of a person like agility, balance, coordination, power and reaction time, whilst health-related factors pertain a state of physical health and well-being. Health-related physical fitness consists of five components, namely cardiorespiratory endurance, flexibility, body composition, muscular strength and muscular endurance. These components are affected by physical training and are associated with important health outcomes. This multidimensional construct definition of health-related fitness is predominantly utilized within health research. (7,8)

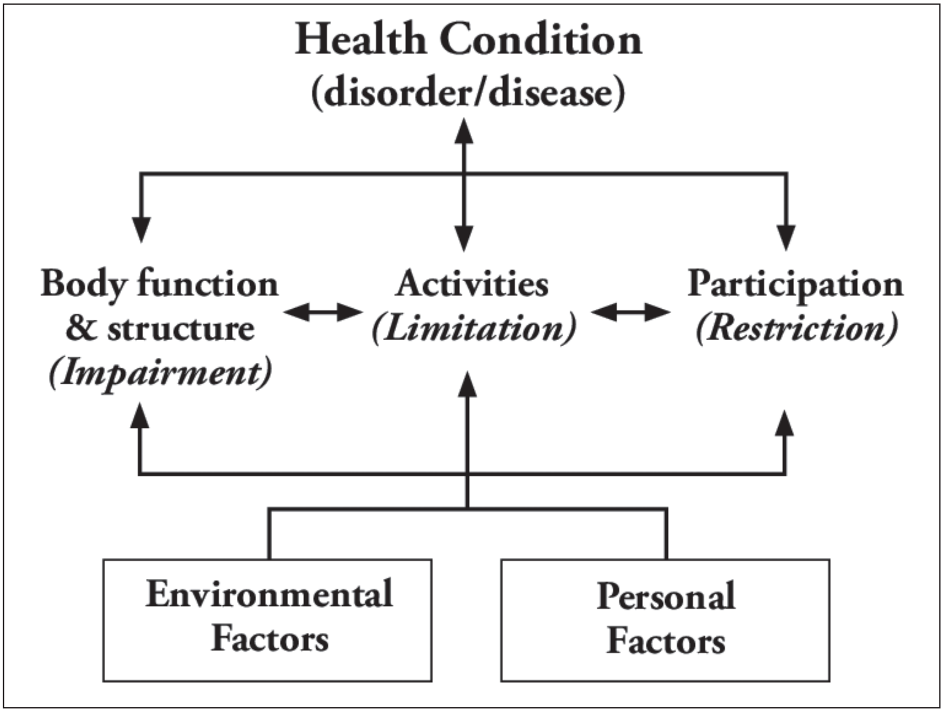


Figure 1. International Classification of Functioning, Disability, and Health (ICF)

A person's physical fitness is commonly determined by measuring the cardiorespiratory endurance. (9) Cardiorespiratory endurance refers to the capacity of the respiratory and cardiovascular systems to work together to provide the needed oxygen and fuel to the body during sustained workloads. (10) Flexibility is the range of motion at a joint. (11) Body composition is the physical make-up of the body, often described as the percentage of muscle, fat, bone and water within the body. (1) Muscle strength is the ability of the muscular system to produce force against a resistance in one maximal effort. (12) Muscle endurance is the ability of the muscular system to produce force over a prolonged period. (10,12) Muscle endurance and muscle strength together contribute to muscular fitness, a term that refers to the ability to do work against a resistance either maximally, explosively, or repeatedly. (10,12)

Although there is limited evidence to suggest an association between health and flexibility, (11) there is ample research providing evidence for the associations between health and cardiorespiratory endurance, body composition and muscular fitness. (10,12–15) The relationship between cardiorespiratory endurance and health is well established. (10,11,16,17) Robust epidemiological evidence has shown a strong association between low cardiorespiratory capacity, and a higher incidence of disease risk including some cancers, cardiovascular disease, and metabolic syndrome, amongst other conditions. (18–20) Adults with low cardiorespiratory endurance have a greatly increased risk of premature all-cause and cardiovascular mortality compared to individuals with the highest levels of cardiorespiratory capacity. (21,22) It has been argued that cardiorespiratory endurance should be regarded as one of the clinical vital signs, and it is expected that assessing cardiorespiratory endurance in clinical practice can improve patient management. (22)

Measurement of cardiorespiratory capacity

As displayed in **figure 2**, cardiorespiratory endurance is dependent on a linked chain of processes that include pulmonary ventilation and diffusion, circulation via the right and left ventricular function, ventricular-arterial coupling, the ability of the vasculature to accommodate and efficiently transport blood from the heart to precisely match oxygen requirements, and the ability of the muscle cells to receive and consume the oxygen and nutrients delivered by the blood, as well as to communicate these metabolic demands to the cardiovascular control center. Therefore, cardiorespiratory endurance is directly related to the integrated function of numerous systems, and it is thus considered a reflection of total body health. Cardiorespiratory endurance can be measured in both direct via gas exchange analysis and indirect via duration or distance. For this purpose, a variety of laboratory and field tests have been developed. Examples of laboratory tests are treadmill or cycle ergometer tests and field-tests encompass the 6-minute walk test.

(22) The 6-minute walk test is an inexpensive submaximal exercise test which measures the distance a participant can walk within a period of 6-minutes. (23) Because of the submaximal nature of the test, and since it closely reflects the activities in daily life, it is used in the assessment of subjects with pulmonary-, cardiovascular-, neurological-, and muscular skeletal pathologies, amongst others.

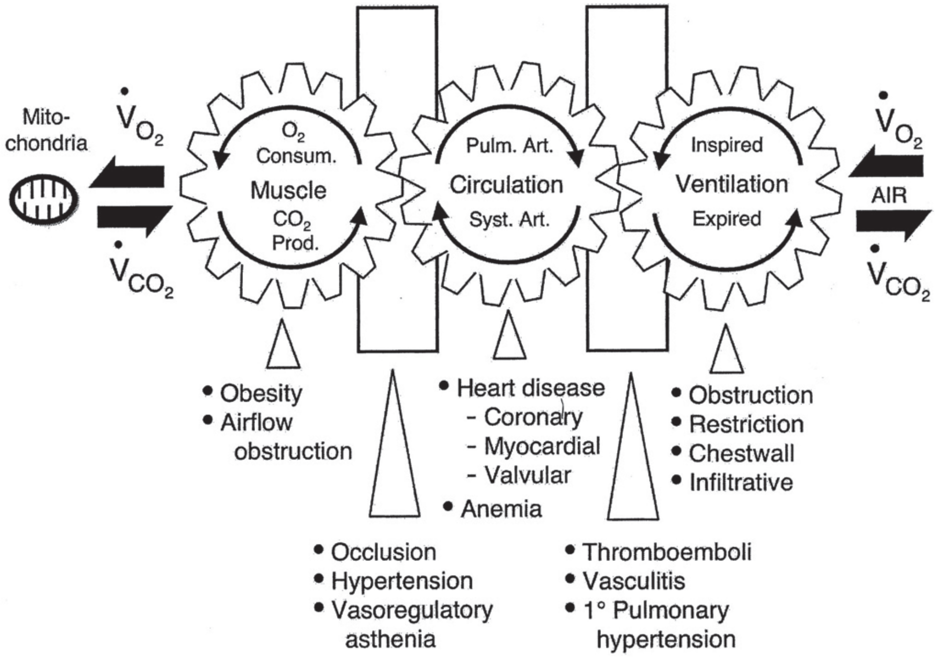


Figure 2. Derangements of gas exchange. The gears represent the functional interdependence of the physiological components of the system. Image from Wasserman, K. (2012). Exercise testing and interpretation [Cover image]. Lippincott Williams & Wilkins.

Cardiorespiratory reference values

Cardiorespiratory endurance values are influenced by underlying biological ageing processes. In most clinical studies, reference values are stratified per paediatric and adult population, resulting in a discontinuity at the transition point between prediction equations. Given its importance, testing, and interpreting cardiorespiratory fitness should be done via a reliable and valid method. Reference values provide the comparative basis for answering questions concerning the normalcy of exercise responses in test subjects, significantly impacting the clinical decision-making process. Reference values obtained from peers can be considered as relevant data to aid the interpretation of test results. (24) The utilization of reference values obtained from peers is critical because cardiorespiratory endurance decreases with age, and higher values are generally observed in men. Therefore, valid reference values can only be established based on data obtained from tests performed in a comparable setting, population and used protocol. (9,23)

Physical activity

As described, physical fitness influences a person's ability to perform physical activity (**figure 1**). Furthermore, health-related fitness is affected positively or negatively by one's habitual physical activity habits. Since benefits of physical activity exhibit a dose-response relationship, the higher the amount of physical activity, the greater the health benefits. Physical activity is defined as any bodily movement produced by skeletal muscles that requires energy expenditure. (1) As in cardiorespiratory fitness, the benefits of higher levels of physical activity behavior are plentiful and significant. Due to these benefits, the World Health Organization has declared physical inactivity the fourth leading risk factor for global mortality. (25,26) With the elimination of physical inactivity in the general population, the life expectancy of the world's population might be expected to increase with 0.68 years, which is comparable to the elimination of the risk factors of smoking and obesity. (27) Due to its known health benefits, it is of a priority to increase physical activity in the general population as a preventative measure. To promote physical activity behavior, the Dutch Health Council physical activity guidelines recommend adults to be physically active at moderate-to-vigorous intensity for at least two and a half hours every week in sessions of 10 minutes or more. (28) The council also recommends engaging in muscle and bone-strengthening activities. People who meet the physical activity guideline recommendations have an estimated 33% lower risk of all-cause mortality compared to those who are not physically active. However, benefits of reducing all-cause mortality, cardiovascular disease, and metabolic disease are seen with any amount of moderate-intensity physical activity. (29–31)

Cardiorespiratory endurance and physical activity in the preoperative phase

Lower cardiorespiratory fitness is a risk factor for adverse post-operative events. (22,32,33) Therefore, increasing cardiorespiratory endurance in the preoperative phase helps to reduce adverse events. Cardiorespiratory endurance is mainly increased by aerobic endurance exercise in healthy subjects. Nonetheless, achieving measurable improvements in cardiorespiratory endurance can be challenging and time consuming. Physical activity behavior, as a proxy of cardiorespiratory endurance, can be altered as a secondary prevention measure. Therefore, the most unfit or diseased individuals have the potential for the greatest reduction in risk, even with small increases in physical activity. By improving the physical activity levels before surgery, the impact of the procedure is reduced, speeding up the time to functional recovery. However, insight into the level of pre-operative physical activity and subsequent postoperative outcomes in patients scheduled for Hepato-Pancreato-Biliary cancer surgery is needed.

AIMS AND OUTLINE OF THIS THESIS

As indicated, there is a need for the information of the reference values for cardiorespiratory fitness measures in the general Dutch population. Additionally, the insight into the physical (in)activity levels of major abdominal surgery patients and the effect of physical activity interventions are limited. This dissertation aims to develop reference values for cardiorespiratory fitness and aims to increase the knowledge on the physical activity level in patients awaiting major abdominal surgery.

A frequently utilized cardiorespiratory fitness test is the 6-minute walk test, this is an inexpensive submaximal exercise test used to quantify the functional exercise capacity in clinical populations. To interpret the results, reference values obtained from peers provide a comparative basis for answering questions concerning the normality of health status, exercise responses and functional exercise capacity. **Chapter 2** is aimed to provide an overview of reference values and reference value prediction equations for the 6-minute walk test in the paediatric and adolescent population.

There are currently no age-related reference values available for the lifespan of individuals in the Dutch population. **Chapter 3** aims to determine the best-fitting regression model for maximal oxygen uptake in the Dutch paediatric and adult populations in relation to age. Furthermore, the gold standard for objectively assessing cardiorespiratory fitness is a cardiopulmonary exercise test (CPET) during which respiratory gas exchange, ventilatory, and heart rhythm measurements are continuously performed throughout an incremental exercise intensity until voluntary exhaustion. **Chapter 4** aims to provide an updated systematic review of the literature on reference values for CPET parameters in healthy subjects across the life span.

As indicated, increasing cardiorespiratory fitness in the preoperative phase helps to reduce adverse events. Therefore, low preoperative cardiorespiratory fitness level has been identified as a modifiable risk factor associated with complications after major abdominal surgery. Additionally, health-related fitness is affected by one’s habitual physical activity habits like performing moderate to vigorous physical activity. To determine the effect of moderate to vigorous physical activity levels in patients awaiting surgical resection for hepato-pancreato-biliary (HPB) cancer, **chapter 5** aims to determine the objectively measured levels of moderate to vigorous physical activity among patients on the waiting list for HPB cancer surgery and their association with postoperative outcomes.

Higher patients' physical activity levels during the preoperative phase, self-reported or objectively measured, improves outcomes in abdominal resection surgery. However, it is not clear to what extend interventions can increase physical activity levels in the preoperative phase, and whether the change measured differs between self-reported and objectively measured outcome measures. Additionally, most previous studies into the physical activity level in major abdominal surgery patients suffer from limited sample size and homogeneity to establish whether physical activity interventions findings are consistent and can be generalized across patients, and treatment variations, or whether findings vary significantly. Therefore, the aim of the meta-analysis described in **chapter 6** is to determine the effect of interventions on physical activity levels of patients awaiting abdominal resection surgery using self-reported as well as objectively measured outcome measures.

Finally, the general discussion of the primary findings of all the chapters taken together is described in **chapter 7** where methodological considerations and further directions for healthcare and research are also discussed.

REFERENCES

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep [Internet]. 1985;100(2):126–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3920711%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1424733
2. Caron A, Ayala A, Damián J, Rodriguez-Blazquez C, Almazán J, Castellote JM, et al. Physical activity, body functions and disability among middle-aged and older Spanish adults. BMC Geriatr. 2017;17(1):1–11.
3. Nagi SZ. An epidemiology of disability among adults in the United States. Milbank Mem Fund Q Health Soc. 1976;54(4):439–67.
4. WHO. Towards a Common Language for Functioning , Disability and Health ICF. Int Classif [Internet]. 2002;1149:1–22. Available from: http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf
5. Verbrugge LM, Jette AM. The disablement process. Soc Sci Med. 1994;38(1):1–14.
6. Freedman VA. Adopting the ICF language for studying late-life disability: A field of dreams? Journals Gerontol - Ser A Biol Sci Med Sci. 2009;64(11):1172–4.
7. Blair S, Durstine L, Eddy D, Hanson P, Painter P, Smith K, et al. Guidelines for Exercise Testing and Prescription, Fourth Edition. Vol. 23, Medicine & Science in Sports & Exercise. 1991. 1215 p.
8. Payne V, Gregory, Isaacs, Larry D. Human Motor Development: A Lifespan Approach. 2016.
9. Ross R, Beck K, Casaburi R, Johnson B, Marciniuk D, Wagner P, et al. ATS/ACCP Statement on Cardiopulmonary Exercise Testing. Am J Respir Crit Care Med. 2003;167(10):1451.
10. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: A powerful marker of health. Int J Obes. 2008;32(1):1–11.
11. Pate RR, Daniels S. Institute of medicine report on fitness measures and health outcomes in youth. JAMA Pediatr. 2013;167(3):221–2.
12. Smith JJ, Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. Sport Med. 2014;44(9):1209–23.
13. Padwal R, Leslie WD, Lix LM, Majumdar SR. Relationship among body fat percentage, body mass index, and all-cause mortality: A Cohort Study. Ann Intern Med. 2016;164(8):532–41.
14. L.A. K, R. A, Ø. E, M.P. H, J. M, C. O, et al. Cardiorespiratory fitness and cardiovascular disease - The past, present, and future. Prog Cardiovasc Dis [Internet]. 2019; Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L2001460054
15. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, et al. The importance of cardiorespiratory fitness in the United States: The need for a national registry: A policy statement from the american heart association. Circulation. 2013;127(5):652–62.
16. Corbin CB, Welk GJ, Richardson C, Vowell C, Lambdin D, Wikgren S. Youth Physical Fitness: Ten Key Concepts. J Phys Educ Recreat Danc. 2014;85(2):24–31.
17. Ross R, Blair SN, Arena R, Church TS, Després JP, Franklin BA, et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement from the American Heart Association. Vol. 134, Circulation. 2016. 653–699 p.

18. Robsahm TE, Falk RS, Heir T, Sandvik L, Vos L, Erikssen J, et al. Cardiorespiratory fitness and risk of site-specific cancers: a long-term prospective cohort study. Cancer Med. 2017;6(4):865–73.

19. Ingle L, Swainson M, Brodie D. Characterization of the Metabolically Healthy Phenotype in Overweight and Obese British Men. Prev Med (Baltim). 2016;94:7–11.

20. Shah R, Murthy V, Colangelo L. Association of Fitness in Young Adulthood with Survival and Cardiovascular Risk: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. Physiol Behav. 2017;176(12):139–48.

21. Kodama S. Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events. J Am Med Assoc. 2009;301(19):2024–35.

22. Ross R, Blair SN, Arena R, Church TS, Després JP, Franklin BA, et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement from the American Heart Association. Vol. 134, Circulation. 2016. 653–699 p.

23. Crapo RO, Casaburi R, Coates AL, Enright PL, MacIntyre NR, McKay RT, et al. ATS statement: Guidelines for the six-minute walk test. Am Thorac Soc. 2002;166(1):111–7.

24. Gräsbeck R. The evolution of the reference value concept. Clin Chem Lab Med [Internet]. 2004;42(7):692–7. Available from: http://nc.ibk.liu.se/nfkk/NORDKEM.

25. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451–62.

26. WHO. Global health risks – Mortality and burden of disease attributable to selected major risks. [Internet]. 2009. Available from: http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf

27. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet. 2012;380(9838):219–29.

28. Weggemans RM, Backx FJG, Borghouts L, Chinapaw M, Hopman MTE, Koster A, et al. The 2017 Dutch Physical Activity Guidelines. Int J Behav Nutr Phys Act. 2018;15(1):1–12.

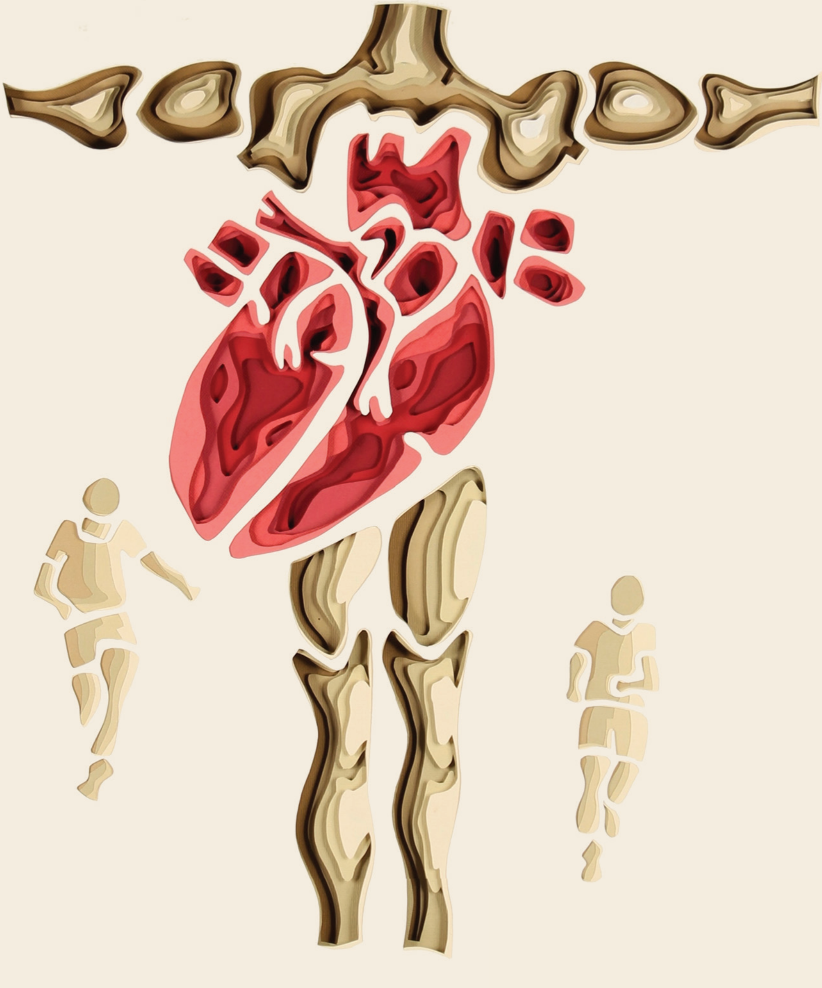
29. Blair SN, Kohl III HW, Barlow CE, Paffenbarger RS, Gibbons LW, Macera CA, et al. Changes in Physical Fitness and All-Cause Mortality A Prospective Study of Healthy and Unhealthy Men From the Cooper Institute for Aerobics Research. Jama [Internet]. 1995;273(Cvd):1093–8. Available from: http://jama.jamanetwork.com/

30. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood J. Exercise capacity and mortality among men referred for exercise testing. ACC Curr J Rev. 2002;346(11):793–801.

31. Blair SN, Iii HWK, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical Fitness and All-Cause Mortality A Prospective Study of Healthy Men and Women. 1989;

32. McCullough PA, Gallagher MJ, Dejong AT, Sandberg KR, Trivax JE, Alexander D, et al. Cardiorespiratory fitness and short-term complications after bariatric surgery. Chest. 2006 Aug;130(2):517–25.

33. Snowden CP, Prentis J, Jacques B, Anderson H, Manas D, Jones D, et al. Cardiorespiratory fitness predicts mortality and hospital length of stay after major elective surgery in older people. Ann Surg. 2013;257(6):999–1004.



Chapter 2

Reference value for the 6-minute walk test in children and adolescents: A systematic review

CF. Mylius, D. Paap, T. Takken

Mylius CF, Paap D, Takken T. 2016. Reference value for the 6-minute walk test in children and adolescents: a systematic review. Expert Rev Respir Med. 10(12):1335–1352. doi:10.1080/17476348.2016.1258305.

ABSTRACT

Background: The 6-minute walk test is a submaximal exercise test used to quantify the functional exercise capacity in clinical populations. It measures the distance walked within a period of 6-minutes. Since the publication of the "American Thoracic Society statement: guidelines for the six-minute walk test", several studies have reported reference values for the pediatric population. The definition of reference values in the pediatric population is especially demanding since not only parameters like height, weight and ethnic background influence the measurement, but maybe as crucial as age. Until now, there is no systematic review on the reference values for the 6-minute walk test in healthy children and adolescents.

Aim: To provide an overview of reference values and reference value prediction equations for the 6-minute walk test in the pediatric population and of the methodology used to obtain them.

Methods: The protocol for this systematic review was based on the PRISMA statement. MEDLINE, EMBASE and Cinahl were searched for eligible articles. Articles were included if the 6-minute walk test was used, were published after 2002 and consists of healthy participants aged ≤18 years.

Results: A total of 22 studies are included. Reported reference values ranged from 383m±41m to 799m±54m. The prediction equation 6MWD=(4.63 * height(cm))–(3.53 * weight)+(10.42 * age)+56.32 yields the highest R^2 value (0.6).

Conclusions: Due to the heterogeneities in the study characteristics and study quality no meta-analysis was performed. It is impossible to present a single best reference value. A flow-chart is presented to aid the selection of reference values or reference value prediction equations.

Implications of key findings: It is recommended that each research department obtained its own reference values and update these regularly because reference values may change over time. Until then, reference values can be selected by using the flow-chart presented.

Keywords: Reference values - prediction equation - six-minute walk test – healthy – pediatric

INTRODUCTION

The 6-minute walk test (6MWT) is an inexpensive submaximal exercise test generally used to quantify the functional exercise capacity in the clinical populations. (1) The test measures the distance a participant can walk within a period of 6-minutes. (1) Because of the submaximal nature of the test, it closely reflects the activities in daily life. (2) The test is frequently used in adults, (3) and is increasingly being utilized in the pediatric populations; it has been used in the assessment of subjects with pulmonary-, (4-6) cardiovascular-, (7) neurological-, (8-15) and muscular skeletal pathologies (16-18) amongst others. (19-23) The test-retest reliability in the healthy pediatric population is high, varying between ICC 0.74 in 6–12-year-olds, (24) ICC 0.80 in 5–6-year-olds to ICC 0.94 in 11-12, (25) and 12–16-year-olds. (26)

In 2002 the American Thoracic Society (ATS) published a statement containing guidelines for the 6MWT in a clinical and research setting. (1) By standardizing the protocol, the aim was to encourage further application of the test and create the possibility to compare achieved values between different studies and populations. The guideline includes facility and procedure related aspects like the track location, lay-out and length plus standardized instructions, encouragements, and preparation procedures.

In order to compare achieved values and establish reference values (RV) the ATS encourages investigators to publish RV for healthy persons using the standardized procedures. (1) RV obtained in healthy subjects provide a comparative basis for answering questions concerning the normality of health status and exercise responses in patients. (27) These questions can significantly impact the clinical decision-making process. (27) In the pediatric population, the definition of RV is especially demanding since not only parameters like height, weight and ethnic background influence the measurement, but maybe as crucial as the development stage and age. For the 6MWT, both the mean walked distance and a prediction equation used to predict the mean walked distance can function as a RV.

In 2014, the European Respiratory Society (ERS) collaborated with the ATS to publish a descriptive review and technical standard regarding the measurement properties of field walk tests in chronic respiratory disease in adults. (28) These documents suggest limits and modifications for the application of the 6MWT regarding the track distance and pretest instructions and present an overview of RV prediction equations for this population. (28)

Since the publication of the ATS statement, several trails investigated RV in the healthy adult and pediatric population. (23-26,29-46) In 2015, Salbach et al. published a descriptive systematic review on RV in healthy adults on multiple walking test including the 6MWT. (48) In this review the RV is reported by median per age decade and RV prediction equation. The RV ranges between 621m for 20-30-years-old males to 350m for 70-79-years-old females, the RV prediction equation R^2 ranges between 0.09 and 0.78.

Despite multiple publications of studies reporting RV or RV prediction equations in the pediatric population, there is no systematic review of these values for the 6MWT. A systematic overview of the available RV in children and adolescents which takes these factors into account can aid a clinician in choosing the set of RV that best reflect the characteristics of the person tested.

Aim

Therefore, by performing a systematic review of the literature on RV for the 6MWT in healthy children and adolescents that were published between 2002 and 2016, the objective of the current study was to provide an overview of RV and RV prediction equations for 6MWT in the pediatric population and the methodology used to obtain them to aid the clinical discussion making process.

METHODS

Design

The protocol for this systematic review was based on the PRISMA statement, (48) the protocol has not been registered.

Information sources & search strategy

The search strategy was created by the first author (CM) and reviewed by an experienced exercise physiologist (TT). The search-string for children and adolescents is based upon the publication of Boluyt from 2008. (49) The used systematic search strategy is revisable in **supplementary material A**. MEDLINE, EMBASE and Cinahl were searched for eligible articles up to March the 21st 2016. Additional records were obtained by screening references from included articles and systematic reviews on related subjects.

Study selection

After combining the results of the electronic searches, duplicates were removed. All unique records were screened by title and abstract for relevance by two reviewers (CM and TT). The first author judged all remaining records based on full text.

Inclusion criteria

Studies were eligible for inclusion if 1) they included healthy subjects with a maximum age of 18 and 2) the study establish a mean walked distance and/or prediction equations for the 6MWT and 3) the study used procedures similar to the ATS guideline either with and without the ERS modifications or another published protocol. Because of the introduction of the ATS guideline in 2002, 1) articles older than 2002 were excluded. Studies that were 2) not published in English, 3) included adults in the sample or were 4) unavailable in full text were also excluded.

Data extraction & synthesis

Data extraction was performed by the first author through standardized extraction forms and consisted of several steps. First, the testing procedures of the 6MWT were evaluated and compared with the existing guidelines. (1,28) The standardized extraction form used to extract the procedure information is displayed in **table 3a** and **3b.** This form includes the track length and layout, instructions prior to the test, encouragement during exercise, and inconsistencies with the ATS/ERS guidelines. Second, the methodological quality of the studies was assessed as described below. Finally, the RV was extracted through the mean walked distance reported per age group, gender and/or overall mean of the study, this is displayed in **table 2.** The RV prediction equations, displayed in **table 4**, were extracted in combination with the fit of the equation (R^2) and the standard error of the estimate (SEE). Any alternative reporting methods are displayed as reported in the original publication.

Methodological quality

In order to assess the methodological quality modifications were made to the assessment list used by Paap et al. (50) to assess the methodological quality in cardiopulmonary exercise testing. The original quality assessment list is based upon study requirements for an optimal set of normal RV as described by the ATS/ACCP guideline. (51) Modifications were made based upon the requirements for the 6MWT as described by the ATS and the modifications by the ATS/ERS review. (1,28) This modified methodological quality assessment list can be found in **supplementary material B**. Each criterion was scored as ‘yes’, ‘no’ or ‘don’t know’ with points only given to ‘yes’. No points are given if the criterion is judged ‘no’ or ‘don’t know’. Studies which only included children younger than 13 years old scored ‘yes’ on the exclusion of smokers’ criterion. This limit is based on an American study by Johnston et al. (52) from 2013 which states that the peak of first-time smokers lies between the age of 11 and 13 years. A study was considered of ‘low quality’ between 0 and 5 points, ‘moderated quality’ if 6-8 points were obtained, and ‘high quality’ if a score of ≥9 was reached.

Quality assessment was independently performed by two reviewers (DP and CM). Afterwards, scores were compared, and disagreements was resolved by consensus. If disagreement persisted, the third reviewer (TT) was consulted for the final rating.

RESULTS

Study selection

The search strategy identified 685 potential studies; five potential studies were identified through other sources. After removal of duplicates and initial screening, 44 studies were regarded potentially eligible. A total of 22 studies were eligible for inclusion after reading the full-text. (23-26,29-46) A flowchart of the study selection procedure is depicted in **figure 1.** Of the excluded studies there were six studies with adults in the domain, two studies did not provide a RV or RV prediction equation, four studies were unavailable in full-text after multiple attempts to obtain the complete article, one potential study was only available in the Spanish language and the last nine potential studies were abstracts about poster presentations, meeting records or commentary letters.

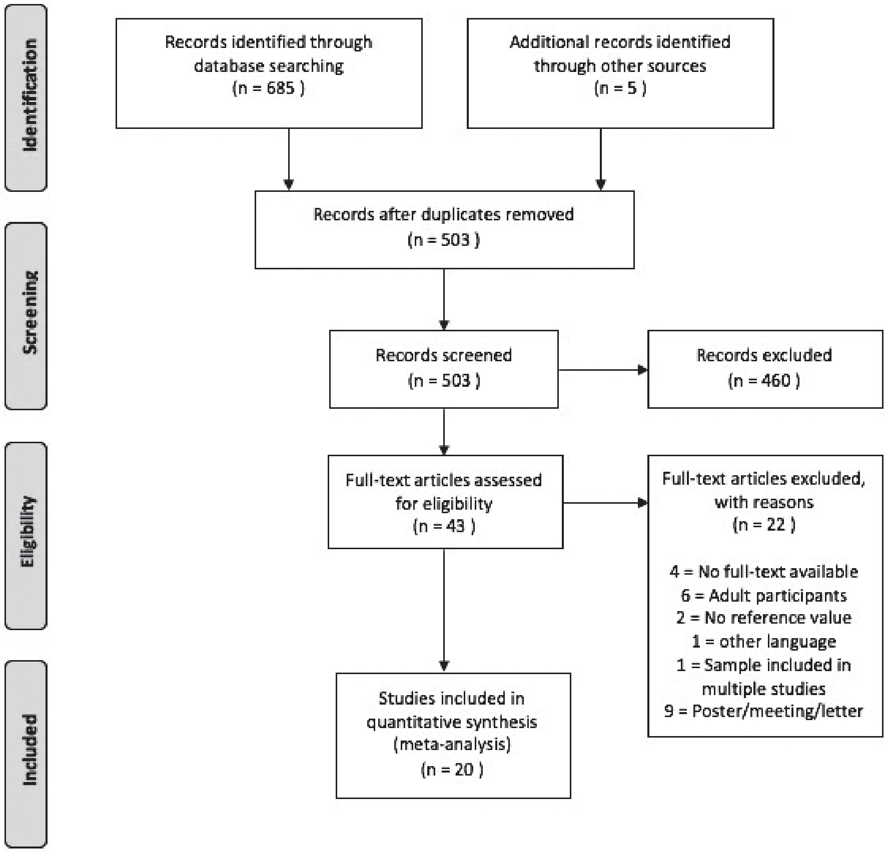


Figure 1. Flow-scheme of selection of studies.

Methodological quality

The quality of the included articles varied between 3 to 10 points out of the 13 criteria points. After a consensus meeting between the two quality assessors', agreement was reached on every criterion except one. The third assessor gave the final verdict for this criteria. Six studies fulfilled five or less criteria and thus received a "low quality" rating. "Moderate quality" ratings were awarded to 11 studies. Five studies scored ≥9 and thus obtained the label "high quality". None of the articles fulfilled all the 13 criteria. Most frequent observed weakness was the lack of cross-validation in a population other than those used to generate the existing data articles, excluding of smokers or lack of report of excluding smokers and measuring activity levels of subjects. The criteria most often met were the prospective and community-based nature of a study. The conclusion of the methodological quality assessment is displayed in **table 1**.

Table 1. Methodological quality assessment.

Study (ref)	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Kanburoglu et al. (45)	1	1	0	1	0	1	1	1	1	1	1	0	1	10
Chen et al. (46)	1	0	1	1	0	1	1	1	1	1	0	1	1	10
Saad et al. (35)	0	1	0	1	1	1	1	0	1	1	1	1	1	10
Goemans et al. (42)	1	0	0	1	0	1	1	1	1	0	1	1	1	9
Li et al. (32)	1	0	0	1	1	1	1	1	1	1	1	0	0	9
D'Silva et al. (40)	1	0	1	1	0	1	1	1	1	1	0	0	0	8
Lammers et al. (32)	1	0	0	1	0	1	1	0	1	1	1	0	1	8
Roush et al. (31)	1	0	0	1	1	1	1	0	1	0	1	0	1	8
Ulrich et al. (27)	1	1	0	1	0	1	1	0	1	1	1	0	0	8
Oliveira et al. (43)	1	0	0	1	0	1	1	0	1	1	1	0	0	7
Rahman et al. (43)	1	0	0	0	0	1	1	1	1	1	1	0	0	7
Tonklang et al. (39)	1	1	1	1	0	0	1	1	1	0	0	0	0	7
Priesnitz et al. (26)	1	0	0	1	0	1	1	0	1	1	1	0	0	7
Geiger et al. (33)	1	0	1	1	0	1	1	0	1	0	1	0	0	7
Li et al. (28)	1	0	0	1	1	0	0	1	1	1	0	0	0	6
Limsuwan et al. (36)	0	0	1	0	0	0	1	0	1	1	1	0	0	5
Fitzgerald et al. (47)	1	0	0	1	0	0	0	0	1	0	1	0	1	5
Klepper et al. (38)	1	0	0	1	0	1	0	0	1	0	1	0	0	4
Morinder et al. (25)	1	0	0	0	0	1	1	0	1	0	0	0	0	4
Pathare et al. (41)	1	0	0	0	1	0	0	0	1	1	0	0	0	4
Basso et al. (37)	0	1	0	1	0	0	0	0	1	0	0	0	0	3

Study characteristics

The 22 studies assessed 7549 subjects in total. The majority of the subjects were male, namely 4093 versus 3456 females. Every article reported the height and weight of the subjects, twenty studies reported the Body Mass index (BMI) of the subjects although two studies used the percentile, and one study used the Z-score for documentation. Eight studies were located in Asia, seven in Europe, three in north America and three south America. One study was located in Africa. Individual study characteristics can be found in **table 2.**

Table 2. Sample characteristics.

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Kanburoglu et al. (45)	Turkey	Yes	High and primary schools		P 949				
					F 482				
				12	109	152±7	46±10	19.92±3.6	604±77
				13	54	158±6	51±10	20.39±3.5	537±79
				14	63	161±6	57±11	21.83±3.8	502±88
				15	95	162±6	55±10	20.82±3.2	508±99
				16	90	164±6	58±8	21.56±2.9	516±92
				17	48	162±6	56±9	21.40±3.3	541±103
				18	9	162±7	54±10	20.67±3.1	561±92
					M 467				
				12	112	151±7	48±11	19.92±3.6	608±95
				13	42	157±9	50±12	20.39±3.5	586±89
				14	43	167±7	63±12	21.83±3.8	528±89
				15	123	172±8	61±12	20.82±3.2	542±87
				16	77	176±7	65±10	21.56±2.9	545±112
				17	55	178±7	71±13	21.40±3.3	543±124
				18	30	177±8	70±11	20.67±3.1	541±109
Chen et al. (46)	Taiwan	Yes	Local elementary, junior high and senior high schools		P 762	148.5±15.3	33.2±14.7	19.5±3.8	513±64
				7	F 380/ M 382				
				8	35/32	124.6±5.5	26.4±5.8	16.9±2.9	473±62
				9	39/38	131.6±5.6	31.9±6.9	18.3±3.1	477±68
				10	36/40	135.8±6.3	33.8±9.1	18.1±3.6	498±57
				11	39/41	141.7±7.5	36.7±8.8	18.1±3.5	503±57
				12	47/41	147.2±6.7	43.1±9.7	19.8±3.9	509±65
				13	40/45	153.7±7.9	48.0±10.3	20.2±3.4	527±71
				14	32/26	156.6±7.3	50.6±9.4	20.6±3.3	527±52
				15	34/31	161.2±8.4	54.8±13.5	21.0±4.1	530±48
				16	26/36	165.1±7.2	57.4±12.2	21.0±3.8	542±54
				17	25/27	165.6±8.0	59.7±12.1	21.6±3.4	543±61
					27/25	164.9±8.6	56.7±12.8	20.7±3.4	545±57

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Saad et al. (35)	Tunisia	No	Offspring of local hospital and medical school workers, athletes excluded		P 200				
					F 100				
				6-7	16	123±7	24±4	16±2	616±53
				8-9	16	132±6	30±5	17±2	648±65
				10-11	21	140±5	32±6	17±2	693±61
				12-13	21	155±6	46±9	19±3	757±51
				14-15	19	159±7	54±10	21±3	718±41
				16	7	164±8	58±5	22±2	730±43
					M 100				
				6-7	16	115±4	22±2	17±2	543±33
				8-9	16	130±6	29±4	17±2	667±55
				10-11	16	142±5	32±4	16±1	715±31
				12-13	21	150±11	44±6	20±3	725±68
				14-15	21	162±8	54±9	20±2	793±84
				16	10	169±7	57±4	20±2	799±54
Goemans et al. (42)	Belgium	Yes	Local primary schools		M 442	135±14.16	31.5±9.63		582.2±88.2
				5	57	115±5.69	20.7±2.5		478.0±44.1
				6	52	121±6.13	23.6±3.2		516.1±61.8
				7	56	127±4.41	25.8±3.4		559.2±65.4
				8	55	133±5.93	29.0±4.2		604.3±72.0
				9	60	139±5.61	32.8±6.0		595.7±69.0
				10	53	144±7.28	35.5±6.2		633.1±70.0
				11	61	148±6.73	39.5±7.8		625.9±83.0
				12	48	154±6.74	46.7±7.3		650.0±76.8

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Li et al. (32)	China	Yes	Primary and secondary school		P 1445	148.7±14.8	41.9±13.6	18.4±3.4	664.0±65.3
					F 640	147.3±13.3	40.4±11.7	18.2±3.3	642.7±58.9
				7	52	124.9±5.4	23.7±4.1	15.1±1.9	
				8	67	128.8±6.3	26.9±7.2	16.0±3.0	
				9	54	135.3±6.1	31.1±6.2	16.9±2.7	
				10	63	143.7±8.0	36.7±8.1	17.6±2.9	
				11	56	148.2±6.9	40.9±11.1	18.4±3.9	
				12	86	152.6±6.5	43.4±8.3	18.6±3.0	
				13	82	155.9±5.8	47.3±8.1	19.4±2.9	
				14	74	156.8±5.6	47.1±7.3	19.2±3.0	
				15	65	159.7±5.0	50.0±5.7	19.6±2.3	
				16	41	158.6±6.0	51.3±8.2	20.4±3.3	
					M 805	149.7±15.8	43.1±14.9	18.6±3.6	680.9±65.3
				7	49	126.7±5.7	25.8±5.7	15.9±2.7	
				8	89	131.4±6.8	29.3±7.0	16.8±2.6	
				9	86	136.3±7.8	32.1±7.6	17.2±3.0	
				10	106	141.1±6.6	36.3±9.0	18.1±3.3	
				11	91	146.2±7.3	39.5±8.5	18.4±3.0	
				12	101	153.5±7.3	44.0±9.3	18.5±3.1	
				13	105	160.6±7.7	50.5±10.1	19.5±3.1	
				14	65	168.5±5.1	57.8±11.9	20.4±4.1	
				15	57	169.0±7.4	63.4±15.2	22.0±4.2	
				16	56	170.8±5.2	60.3±8.5	20.7±3.0	

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
D'Silva et al. (40)	India	Yes	Primary schools		P 400	130.82±10.45	25.30±4.38	14.77±0.75	608±166
					F 198				548.93±44.78
					M 202				670.74±86.21
				7	33/32	118.53±12.19	19.45±3.28	14.01±2.33	625.4±120.22
				8	34/32	121.33±53.44	21.83±3.17	14.79±1.64	526.4
				9	34/34	127.49±6.44	24.74±5.05	15.11±2.24	586.4
				10	34/34	131.71±7.07	25.81±4.72	14.82±2.01	596.2
				11	34/33	142.79±5.30	28.45±4.94	13.93±2.00	658.2
				12	33/33	143.09±6.99	31.56±6.50	15.96±2.54	667.45±181.73
Lammers et al. (34)	United Kingdom	No	Primary schools, healthy siblings and relatives of children attending local hospital		P 328	130±15	29±9	16.9±2.6	470±59
					F 150/ M 178	107±4	18±2	16.0±1.3	383±41
				4	18/18	115±6	21±4	15.7±1.7	420±39
				5	21/19	121±6	24±5	16.5±2.5	463±40
				6	21/19	128±5	26±4	16.0±1.8	488±35
				7	18/22	133±8	30±6	16.8±2.6	483±40
				8	18/27	139±7	34±7	17.8±2.8	496±53
				9	22/27	145±8	38±0	17.8±2.8	506±45
				10	18/30	149±7	41±9	18.4±3.4	512±41
				11	15/15				
Roush et al. (31)	United States	No	Third grade elementary school	7,5-9	P 76				
					F 38	1.32±0.06	28.76±7.32	16.3±2.9	532.2±52.6
					M 38	1.34±0.06	28.11±5.09	15.5±2.0	581.7±58.1

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Ulrich et al. (27)	Switzer-land	No	Schools		P 496	147±2	41±17	17.9±3.3	618±79
					F 252				
				5	19	115±5	19±3	14.5±1.5	506±39
				6	21	118±6	21±4	14.7±1.6	546±51
				7	19	128±6	26±5	16±2.4	586±59
				8	21	133±7	29±6	16.5±1.9	612±40
				9	18	138±5	32±4	16.4±1.8	606±52
				10	22	146±7	37±8	17.5±3.0	638±63
				11	20	150±8	38±10	16.7±2.6	636±54
				12	20	154±8	46±9	19.5±2.7	672±55
				13	27	163±7	52±10	19.6±3.1	622±76
				14	23	165±7	56±8	20.8±2.4	622±64
				15	22	163±6	55±10	20.5±2.9	626±49
				16	20	168±7	62 ±9	21.8±2.4	629±52
					M 244				
				5	19	113±6	20±3	15.5±2.3	494±60
				6	22	123±5	23±3	14.8±1.6	535±73
				7	19	128±6	26±4	15.8±1.7	603±51
				8	22	133±7	29±4	16.1±1.1	596±59
				9	18	138±6	32±4	16.6±1.6	627±70
				10	19	144±7	38±8	18.2±2.5	655±53
				11	23	148±6	40±10	18.1±3.8	624±87
				12	20	156±9	46±10	18.4±2.6	685±74
				13	21	162±9	51±10	19.5±2.9	639±49
				14	20	169±6	56±10	19.7±2.8	684±81
				15	20	176±7	65±13	20.9±3.8	690±71
				16	21	180±9	68 ±13	20.9±3.2	680±55

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Oliveira et al. (43)	Brazil	No	Schools		P 161				
					F 86				
				6-7	19	125.9±6.5	27.3±5.9	0.65±1.16****	608.3±75.6
				8-9	27	136.4±11.3	30.9±7.9	-0.07±1.07****	698.5±50.2
				10-11	18	148.4±9.0	42.0±9.0	0.49±1.16****	701.9±44.0
				12-13	21	159.5±5.0	51.8±5.9	0.48±0.75****	709.7±51.7
					M 75				
				6-7	20	124.8±4.8	25.0±5.0	0.13±1.65****	622.2±60.0
				8-9	18	137.2±7.2	32.8±5.1	0.59±0.96****	688.4±44.4
				10-11	19	147.2±10.2	41.8±9.7	0.83±0.82****	747.2±59.3
				12-13	19	159.5±8.3	49.5±7.5	0.39±0.89****	764.7±54.9
Rahman et al. (44)	Saudi Arabia	Yes	Local primary schools,		F 136	130.60±12.71	28.83±7.53	16.65±1.75***	595.77±61.35
				6	22	115.95±5.15	21.38±2.69	15.86±1.08***	543.68±44.77
				7	23	122.43±4.64	23.80±3.39	15.74±1.21***	564.26±51.30
				8	29	126.10±5.27	26.29±3.76	16.45±1.43***	586.03±41.42
				9	21	133.05±8.59	30.09±5.59	16.95±1.96***	600.86±57.12
				10	20	141.40±8.08	34.63±5.86	17.55±1.88***	647.95±53.56
				11	21	148.38±8.20	38.89±6.19	17.62±2.06***	643.52±51.10
Tonklang et al. (39)	Thailand	Yes	4 to 6 grades of primary schools	9-12	P 739	141.3±8.7	35.3±9.7		677.0±62.2
					F 336	142.4±8.3	36.0±9.9		657.1±51.1
					M 403	140.3±8.9	34.8±9.5		693.5±65.7
Priesnitz et al. (26)	Brazil	No	Elementary schools		P 188	140±10	36.3±11.2	18.5±3.0	579.4±68.1 - 571.3±75.4*
					F 96/ M 92				
				6	12/12	120±6	25.7±4.7	16.9±1.9	508.3±54.0 - 501.7±67.7*
				7	12/13	120±10	27.6±5.2	172.±2.2	550.2±61.6 - 517.7±84.7*
				8	12/12	130±10	33.03±6	18.3±2.4	556.7±67.2 - 570.3±64.1*
				9	15/13	130±10	33.6±6.8	17.3±2.2	594.2±60.6 - 578.3±68.5*
				10	19/13	140±10	38.6±8.7	18.8±3.1	602.4±61.1 - 596.6±59.5*
				11	14/16	150±10	45.7±10.6	20.1±3.3	608.0±54.3 - 610.2±55.7*
				12	12/13	150±10	47.2±12.0	20.3±3.5	618.1±51.4 - 603.1±59.1*

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Geiger et al. (33)	Austria, Caucasian	No	Local schools and kindergartens		P 528				
					F 248				
				3-5	25	113(105-130)	19.0(16.3-28.7)	14.8(13.5-18.9)	501.9±90.2
				6-8	46	128(120-139)	25.2(20.0-36.1)	15.1(13.5-20.1)	573.2±69.2
				9-11	62	145(132-161)	36.7(26.2-57.8)	16.9(14.1-23.4)	661.9±56.7
				12-15	71	164(149-177)	54.0(38.9-72.1)	20.5(16.2-25.5)	663.0±50.8
				≥16	55	170(160-179)	56.9(46.5-73.4)	19.6(17.4-28.2)	664.3±49.5
					M 280				
				3-5	22	114(106-122)	18.1(16.0-23.9)	14.2(13.0-17.0)	536.5±95.6
				6-8	66	130(120-141)	25.8(20.9-36.4)	15.3(13.4-20.9)	577.8±56.1
				9-11	57	147(136-157)	36.9(27.3-52.0)	17.3(13.7-23.4)	672.8±61.6
				12-15	80	166(150-188)	55.7(39.6-77.9)	19.2(16.1-26.9)	697.8±74.7
				≥16	44	182(171-193)	69.4(50.6-91.6)	20.7(16.8-26.2)	725.8±61.2
Li et al. (28)	China	Yes	Secondary schools	12-15	P 74	159.4±7.2	48.6±8.5	19.1±2.6	659.8±58.1
					F 43	156.7±4.7	46.1±5.8	18.7±2.0	637.4±38.6
					M 31	163.2±8.3	52.1±10.4	19.5±3.3	691.0±66.3
Limsuwan et al. (36)	Thailand	No	Hospital	9-12	P 100	131±7.6	37.7±11.1	17.8±4.5	586.1±44.0
					F 47	140.5±8.1	35.7±11.6	19.5±4.0	580.4±47.6
					M 53	141.5±7.3	19.5±10.4	18.7±4.3	591.1±40.4
Fitzgerald et al. (47)	Ireland	No	Local schools, through advertisement in rehabilitation center	4-17	P 137 F 66/ M 71	138±2	36±14.8		528.42±67.77

Study (ref)	Country Ethnicity	Randomized	Recruitment	Age (y)	Sample size & sex distribution	Height (cm)	Weight (kg)	BMI (kg/m²)	Distance (m) mean ± SD
Klepper et al. (38)	United States	No	Public and private schools, after school program, through family and friends of the investigator		P 80	139±9	35.23**	18.52**	518.50±72.56
				7-8	20	132±8	31.52±8.99	18.15±4.02	527.09±64.2
				9	20	135±6	34.14±8.5	18.73±3.86	531.66±80.27
				10	32	144±8	45.1±9.88	21.6±3.5	497.15±66.81
				11	8	150±3	42.1±7.46	18.81±3.45	533±63±85.42
					F 51	138±9	34.77**	19.56±3.92	518.32±73.16
				7-8	14	130±7	29.77**	17.48**	519.64±69.31
				9	12	135±7	30.91**	18.2±3.6	542.54±80.25
				10	21	144±8	42.91±8.97	21.6±3.5	496.69±63.98
				11	4	150±4	43±24±10.52	19.16±4.9	532.33±92.25
					M 29	140±10	39.84±11.33	18.65**	518.73±72.61
				7-8	6	131±8	29.24±4.7	16.65±1.48	543.54±60.3
				9	8	136±5	35.94±9.24	19.39±4.35	515.83±81.4
				10	11	145±10	50**	22.52±3.93	497.94±74.03
				11	4	150±3	40.94±3.95	18.46±1.84	534.93±88.90
Morinder et al. (25)	Sweden	No	Schools		P 97	156.2±15.1	46.0±12.9	18.4±2.2	662.6±61.1
					F 49	154.5±12.3	44.3±11.3	18.2±2.4	655.6±58.5
				8-10	16	140.2±4.4	30.9±3.2	15.7±1.4	665.8±22.7
				11-13	16	157.5±7.6	47.1±6.6	19.0±1.9	681.6±43.7
				14-16	17	165.3±6.6	54.3±6.2	19.7±1.9	619.8±76.6
					M 48	158.0±17.5	47.8±14.3	18.7±1.9	670.4±63.3
				8-10	15	140.9±5.9	34.8±3.7	17.5±1.3	646.0±35.0
				11-13	15	151.0±6.6	41.8±5.5	18.3±1.6	693.6±63.8
				14-16	18	178.0±7.8	63.7±9.1	20.0±1.9	671.4±75.6
Pathare et al. (41)	United States	No	Elementary schools		P 41	124.0±8.8	24.7±4.8	55.5±21.7***	
					F 26	123.0±10.0	25.1±4.0	57.2±20.5***	529.8±54.8
					M 15	125.5±8.8	24.6±6.4	51.7±24.8***	516.7±65.6
				5-6					536.7±61.9
				7-8					539.9±57.8
				9-10					541.7±73.6
Basso et al. (37)	Brazil	No	Public state school and University hospital	11-15	P 19 F 6/ M 13	*150±90*	46±9.8	21±4.8	622±50.8

Abbreviations: P Pooled, F Female, M Male, * Second test, ** Median presented if skewed data, *** percentile, **** Z-score

Results of individual studies

Procedures

Test procedures showed large variation. For example, the two studies by Goemans et al. (40,41) performed the 6MWT protocol as described by McDonald, (10) 12 studies state the usage of the ATS guideline either with or without the ATS/ERS modifications. (24-26,30,32-34,38,39,42-45) The remaining publications do not mention the used guideline but describe comparable procedures. Three studies performed the 6MWT either in groups or in overlapping fashion, (29,37,42) the other studies performed individual tests in accordance with the ATS guideline. The track length ranged between 15m (36) and 70m. (46) Only Fitzgerald et al. (46) used a track length inconsistent with the recommended 15m to 50m of the ERS/ATS review. Most tests were performed in a straight indoor corridor, although three studies were conducted outside (29,35,37) and two studies do not explicitly report the course location. (36,44) In contrast with the ATS guideline, an instructor walked with the participant in four studies for additional measurements or safety. (32,40,41,46) **Table 3a** shows the modifications from the ATS/ERS guideline and additionally reported measurements per study. The procedural factors of influence on the results of the 6MWT, as reported in the respective studies, are displayed in **table 3a.**

Preparations

Although 12 studies reported following the ATS guidelines, (24-26,30,32-34,38,39,42-45) only five studies explicitly report the participants wearing comfortable clothing and shoes during the test (24,33,37,42,43) and an equal number report not using a warm-up period. (26,30,33,37,38) In accordance with the ATS guideline, six studies report not allowing vigorous activities in the 2 hours prior to the test, although a light meal was allowed. (24,26,30,33,37,38) The most reported preparation procedure is the resting period of 10 minute whilst checking either heart rate and blood pressure, (25,26,30,31,33,37,38,42,43) or only heart rate. (23,34,45) In only a few studies measured the dyspnea and fatigue (24,37,44) or only fatigue (23,42) prior to the 6MWT. The measurement of oxygen saturation, which is an optional measurement in the ATS guideline, was measured by nine studies. (24-26,31-33,38,42,43,45) Multiple studies performed the post-test measurements; dyspnea, (24,31,37,38) fatigue (23,24,31,37), oxygen saturation, blood pressure and heart rate. (23-26,30,31,33, 37,38,43)

Instructions

The instructions described by the ATS guideline include the sentence "*remember that the object is to walk as far as possible for 6 minutes, but don't run or jog*". The ERS/ATS systematic review note an average increase in walked distance of 52.7m in interstitial lung disease and pulmonary atrial hypertension patients if they were asked to walk as "*fast as possible*". In seven of the included studies, the instructor used the word "*far*" and three studies used the word "*fast*". Klepper et al. (36) did not instruct the participants they were allowed to stop and rest. Two studies stated that they used the ATS guideline instructions but did not report the given instructions. (33,34) The remaining studies either did not mention the used instructions or used instructions different from those in the ATS guideline or ERS/ATS review, like "*as much ground as possible*". A practice run was made by both the instructor and the participant in three studies. (36,40,41)

Encouragements

The encouragement phrases described in the ATS guideline consist of standardized sentences which should be announced with one minute intervals. However, both the studies performed by Goemans et al. (40,41) used constant encouragements, the encouragement used by Chen et al. (45) were given at random moments and the study conducted by Saad et al. (33) used no encouragements at all. The studies performed by both Morinder et al. (23) and Ulrich et al. (25) only used announcements related to the remaining time. The instructions and encouragements as reported in the respective studies are displayed in **table 3b.**

Synthesis of results

Meta-analysis

Due to the heterogeneities in the study characteristics, testing procedures, reporting method and methodological quality, no meta-analysis is performed. Each of the included studies has various numbers of shortcomings and limitations that are noted through the methodological quality assessment displayed in **table 1**.

Reference value

Table 2 shows the RV of all the studies for the whole sample (23-26,30,32,34-38,44-46), separate gender and separate age groups if reported. (23-25,31-33,36,38-45) The RV ranges from 513m±64m (45) to 677.0m±62.2m (37) for the pooled data. The RV only related to females' ranges between 518.32m±73.16m (36) and 657.1m±51.1m (37) and

516.7m±65.6m (39) till 693.5m±65.7m (37) for all the males. The lowest RV is 383m±41, recorded by a sample of 4-year-olds in the study conducted by Lammers et al. (32) The highest RV is 799m±54m recorded by 16-year-old males in the study performed by Saad et al. (33) Geiger et al. included the youngest sample with an age of 3 till 5-years-old. (31) In this study, the RV for females is 501.9m±90.2m and for males 536.5m±95.6m. (31) The oldest group tested is 18-year-old. In this study, the RV for females is 561m±92m and 541m±109m for males. (44)

Reference value prediction equation

Table 4 lists the 32 RV prediction equations from eight different studies. (24,25,30,31,41,42,45) Of the 32 equations, 14 are developed for females (25,30,31,42) and 12 for males. (25,30,31,41,42) The remaining six are not gender specific. (24,25,33,45) Six studies (24,30,31,33,41,42) reported the R^2 and three studies additionally reported the SEE, (24,31,41) the remaining studies did either; not use the SEE or R^2 or in the case of Ulrich et al. (25) used the Durbin-Watson tests (DW) to detect autocorrelation in multiple linear regression models. The equation yielding the maximum R^2 of 0.6 includes height, weight, and age. Across the 32 RV prediction equations, age was included in the most often (64%) followed by height which was included in 19 equations (61%).

Table 3a Factors of influence on the results of 6MWT - a) Track length & layout; Inconsistency with ATS procedure and additional measurements.

Study (ref)	Track length & layout	Modifications with ATS procedure	Additional measurements
Kanburoglu et al. (45)	30m - -	-	Age*, sex*, height*, weight*, BMI*, pre-HR*, post HR*, BP, Activity levels*, Dyspnea, fatigue
Chen et al. (46)	30m - Hallway	-	Age*, sex*, height*, weight*, BMI*, Rest SaO2*, lowest Sa02, HR
Saad et al. (35)	40m - Seldom traveled flat corridor	-	Age, sex, height, weight, BMI, BSA, FVC, FEV1, FEF, PEF, HR, BP, activity levels*, socioeconomic levels, pubertal status*
Goemans et al. (42)	25m - Flat straight corridor	Orientation video prior to testing, Continuous verbal encouragement, investigator walked behind participant	Age*, weight*, height*, BMI, Knee flexion* and extension* strength
Li et al. (32)	100ft - Internal hallway	-	Age*, Height*, weight*, BMI*, SaO2*, FEV_1*, FVC*, HR*, BP
D'Silva et al. (40)	100ft - Internal hallway	-	Age*, sex*, height*, weight*, BMI, HR, SaO2, BP, Dyspnea*
Lammers et al. (34)	30m & 50m - Flat hard ground	Instructor walked behind child to obtain continuous HR/SaO2 measures	Age*, weight*, height*, BMI, SaO2, HR, ethnicity
Roush et al. (31)	200ft - Playing field	Participants were given straws for every completed lap. Test conducted in groups of 16 participants.	Height, weight, BMI
Ulrich et al. (27)	30m - Flat ground	-	Age*, sex*, height, weight, BP, HR, activity levels*, BMI, SaO2
Oliveira et al. (43)	22m - Corridor	-	Age*, sex*, height*, weight*, BMI, leg length*
Rahman et al. (44)	30m - Seldom traveled indoor Corridor	-	Age*, height*, weight BMI, SaO2, BP, HR
Tonklang et al. (39)	30m - School ground	Technicians arranged the participants in groups with five subjects or less per group on the starting line. Rubber band was given to each participant after completing each lap as a counting token.	Age*, sex*, weight*, height, leg length, Respiratory rate, HR, BP, activity levels, TV time*
Priesnitz et al. (26)	30m - Flat corridor	-	Age*, height*, weight*, SaO2, Δ. HR*, BP, Dyspnea
Geiger et al. (33)	20m - Straight course	Participant used one handed wheel to measure distance, 3- to 4-year-olds allowed to walk/run/jog	Height*, weight*, BMI*, TLL*, SaO2, BP*, HR*, FEV_1, PEF*, Fatigue
Li et al. (28)	- - Corridor	-	Height*, weight, BMI, SaO2, FEV_1*, FVC, HR, HR end*, HR Δ.* BP

Study (ref)	Track length & layout	Modifications with ATS procedure	Additional measurements
Limsuwan et al. (36)	100ft - Internal hallway	-	Age, sex, height*, weight, BMI, leg length*, Start HR, end HR*, Δ HR*, BP
Fitzgerald et al. (47)	70m - Straight corridor	Length of the course was 70 m, turnaround points indicated by signs on the wall, no demonstration lap	Age*, height, BMI, weight
Klepper et al. (38)	15m & 25m	Turnaround point was a line to tip toe.	Age, height*, weight*, BMI*, Leg length*, track length*
Morinder et al. (25)	30m - Indoor corridor	-	Height, Weight, BMI, HR, Exertion, Age
Pathare et al. (41)	18m & 20m - Corridor	-	Height, Weight, BMI*, HR, resting BP*, post BP*, SaO2, exhaustion
Basso et al. (37)	30m - Outdoor area	Participants wore a nose clip during test.	Height, Weight, BMI, Spirometry: FVC, SVC, MVV

*Correlates significantly with walked distance, -: Not/non stated, BMI: body mass index, SaO2: oxygen saturation, FEV1: forced expiratory volume in 1 second, FVC: Forced vital capacity, HR: heart rate, Δ delta, BP: blood pressure, TLL: true leg length, FEF: Forced mid expiratory flow, PEF: Peak expiratory flow, SVC: Slow vital capacity, MVV: maximum voluntary ventilation,

2

Table 3b Factors of influence on the results of 6MWT - b) Verbal instructions and encouragements.

Study (ref)	Reported verbal instructions	Reported verbal encouragements
Kanburoglu et al. (45)	The students were informed about the study and how to perform the 6MWT 1 d before the test.	Observers measured the test time using stopwatches and said the same encouraging sentences as stated by the American Thoracic Society to the students every minute: "you are doing well; you have only 1 minute to go."
Chen et al. (46)	In a period of 6 minutes, the participants were asked to walk back and forth along this hallway as far as possible, at their own best pace but not to run or race.	Standardized phases of encourage or announcement of time remaining, such as "You are doing well," and "You have 3 minutes to go" were often given to the participants, although the frequency of such encouragement varied across studies from providing encouragement every 30 seconds to every 2 minutes. No comments were made with the intention of speeding up or slowing down the participant.
Saad et al. (36)	Not stated	No encouragements.
Goemans et al. (42)	Not stated	The 6MWT as described by McDonald (10).
Li et al. (32)	The subjects were told that the purpose of the test was to see how far they could walk in 6 minutes. They were then instructed to walk up and down the hallway covering as much ground as they could during the 6 minutes. The test was self-paced and the subject could rest if he or she so wished.	The words of encouragement during the testing were standardized ("Keep going," "You are doing fine," "Everything is going well") and were given by the same person at set times during the test.
D'Silva et al. (40)	The subjects were told; the purpose of the test is to see how far they can walk in 6 min. They were instructed to walk up and down the hallway covering as much ground as they can during 6 min. The test was self-paced and the subject could rest if he or she so wished.	The words of encouragement during the testing were standardized ("Keep going," "You are doing fine," "Everything is going well") and was given by the same person at set times during the test.
Lammers et al. (36)	The children were asked to walk up and down the measured lap at their best pace but not to run or race.	Encouragement (eg, ''Keep going'', ''You are doing well'') and announcement of time remaining were given to the children. No comments were made regarding the child's performance, such as, ''You could go faster'' or ''Slow down''.
Roush et al. (31)	The subjects were instructed to "walk, don't run, skip or hop, and stay around the outside of the cones".	During the six minutes the investigators gave words of encouragement such as "good job" and "keep up the good work"
Ulrich et al. (27)	Subjects were instructed to walk as fast as possible (without running) at a steady pace for 6 minutes.	After 5 minutes' time left had been advised to the participant. No other commandos or verbal feedback was given.

Study (ref)	Reported verbal instructions	Reported verbal encouragements
Oliveira et al. (43)	All participants were instructed before the tests that they should walk as fast as possible during 6 min, going back and forth in a demarcated corridor but not being allowed to run or jog. Children were told that in case of fatigue, dyspnea, or abdominal or leg pain they could stop, lean against a wall, or walk slower, although the chronometer would not be stopped.	One of the examiners pronounced standardized phrases of encouragement every minute.
Rahman et al. (44)	"The object of this test is to walk as far as possible for 6 minutes. You will walk back and forth in this hallway. Six minutes is a long time to walk, so you will be exerting yourself. You will probably get out of breath or become exhausted. You are permitted to slow down, to stop and to rest as necessary. You may lean against the wall while resting, but resume walking as soon as you are able. You will be walking back and forth around the cones. You should pivot briskly around the cones and continue back the other way without hesitation. Now I'm going to show you. Please watch the way I turn without hesitation". "Are you ready to do that? I am going to use this counter to keep track of the number of laps you complete. I will click it each time you turn around at this starting line. Remember that the object is to walk as far as possible for 6 minutes, but don't run or jog. Start now or whenever you are ready"	After each minute of the test, the examiner told the student in an even tone "You are doing well" and informed her about the remaining time. The examiner did not use other words of encouragement (or body language to speed up). If the student stopped walking during the test and need a rest, the examiner said this: "You can lean against the wall if you would like; then continue walking whenever you feel able." When the timer was 15 seconds from completion, the examiner said this: "In a moment I'm going to tell you to stop. When I do, just stop right where you are and I will come to you."
Tonklang et al. (39)	Not stated	During the testing, the technicians used standard phrases of encouragement to encourage the participants in the same way.
Priesnitz et al. (25)	The test and its purpose were explained in detail so that the participants would understand what to do. The instructions consisted of walking fast but not running during the test, and going slower if they were very tired, or even stopping if the fatigue were unbearable.	At the end of every minute, they were told motivation sentences like: ''you are doing well, you have five minutes left''; ''good job, there are four minutes left.'' No other type of stimulation was given to avoid interfering in each individual's performance.

2

Study (ref)	Reported verbal instructions	Reported verbal encouragements
Geiger et al. (33)	"The object of this test is to walk as far as possible in 6 minutes, which means to score as many meters on the scale as possible. You will be walking back and forth around the poles. You are permitted to slow down, to stop, and to rest as necessary. You may lean against the wall while resting, but resume walking as soon as you are able to. Are you ready to do that? Remember that the object is to score as many meters as possible in 6 minutes, but without jogging or running. Start now"	After each minute, "You are doing well. You have 5 minutes to go." After the second minute: "Keep up the good work. You have 4 minutes to go." After three minutes: "You are doing well. You are halfway done." After four minutes: "Keep up the good work. You have only 2 minutes left." After five minutes: "You are doing well. You have only 1 minute to go." No other words of encouragement or body language were used to speed up the participant.
Li et al. (28)	The subject was instructed to walk up and down a measured corridor, covering as much ground as possible over a 6-min period.	The wording of encouragement during the testing was standardized (''keep going'', ''you are doing fine'', ''everything is going well'') and given by the same person at set times during the test.
Limsuwan et al. (36)	Not stated	Each child had a personal instructor during the test with a standardized word of encouragement ('you are doing well', keep up with your good work').
Fitzgerald et al. (47)	"the objective of the test is to walk as far as possible in 6 minutes."	No additional verbal comment or encouragement was given other than those recommended by the ATS guide- lines, which allows 1 standardized comment every minute such as "you are doing well, you have only 1 minute to go."
Klepper et al. (38)	Each child walked along the length of the track with the tester and was shown the beginning and end of the course. Demonstration and instructions were given at both ends of the course to "touch the long strip of tape with your foot, turn around and walk back to the other end." Participants were informed that the purpose was to find out how far children walk in 6 minutes. They were told to walk like they were trying to get somewhere they really wanted to go, but hopping, skipping, running, and jumping were not allowed. Subjects were not told they could stop and rest during the test.	Standard phrases of encouragement ("You are doing well," "Things are going well," and "Keep going") were given at 30-second intervals, and participants were informed of the remaining time at each minute mark, for example "You have 5 minutes left."
Morinder et al. (25)	The instructions were to walk as many lengths as possible in six minutes, without running or jogging. To clarify the instructions, the children were also told to walk as fast as possible.	Information was given during the test by telling the children how many minutes they had walked or minutes remaining.

Study (ref)	Reported verbal instructions	Reported verbal encouragements
Pathare et al. (41)	Participants were informed that the purpose of the test was to find out how far they could walk in 6 minutes and were instructed to walk the longest distance possible at their own pace during the allotted time. Hopping, skipping, running, and jumping were not allowed during the test.	Only the standardized phrases for encouragement (eg, 'keep going', 'you are doing well') and announcement of time remaining were given to the participants.
Basso et al. (37)	Not stated	Not stated

Table 4. Prediction equations for the 6-minute walk test reference values

Study (ref)	sex	Prediction equation	R^2	SEE
Chen et al. (46)	Both	$\frac{(6MWD) - B1 - B2xIn(heig}{\sqrt{MSE}}$ Z score F: 4.3204 - 0.3813 MSE = 0.0138 M: 3.5247 - 0.5443 MSE = 0.0132		*
Saad et al. (35)	Pooled	6MWD = (4.63 x height(cm)) – (3.53 x weight) + (10.42 x age) + 56.32	0.6	NS
Goemans et al. (42)	Male	6MWD = 86.795 + (74.547 x age) +(23.0186 x age^2) + (63.2046 x height).	0.41	*
Li et al. (32)	Female Male	6MWD = 526.79 + (ΔHR) x 1.66) + (height(cm) x 0.62) 6MWD = 554.16 + (ΔHR) x 1.76) + (height(cm) x 1.23)	0.373 0.435	*
Rich et al. (27)	Pooled Female Male	6MWD = (11.89 × age) + 486.1 + Height & weight adjusted 6MWD = (391.9 × height) – (2.41 × weight) + 140.2 + HR & PAS adjusted 6MWD = (192.69 x height) + (1.27 x post HR) + 161.55 6MWD = (8.623 x age) + 513.7 + Height & weight adjusted 6MWD = (372.3 x height) – (2.635 x weight) + 172.05 + HR & PAS adjusted 6MWD = (152.58 x height) + (1.38 x post HR) + 197.97 <12y 6MWD = (20.83 x age) + 413.94 + Height & weight adjusted 6MWD = (330.3 x height) + 153.3 + HR & PAS adjusted 6MWD = (279.5 x height) + (.87 x post HR) + 102.45 ≥12y = (-8.66 x age) + 757.42 + Height & weight adjusted 6MWD = (-1.867 x weight) + 734.29 + HR & PAS adjusted 6MWD = (1.79 x post HR) – (1.28 x pre HR) – (2.55 x weight) + (203.3 x height) + (7.83 x PAS) + 298.6 6MWD = (15.36 x age) + 456.92 + Height & weight adjusted 6MWD = (13.40 x age) – (2.16 x weight) + (196.53 x height) + HR & PAS adjusted 6MWD = (14.38 x age) + (1.21 x post HR) – (2.12 x weight) + (166.66 x height) + 146.56 <13y 6MWD = (24.18 x age) + 385.18 + HR & PAS adjusted 6MWD = (28.62 x age) + (1.26 x post HR) – (2.034 x weight) + 239.23≥13y 6MWD = (13.08 x age) + 4.76.69 + HR & PAS adjusted 6MWD = (1.01 x post HR) + (13.3 x age) + 338.25		*

Study (ref)	sex	Prediction equation	R^2	SEE
Oliveira et al. (43)	Female	6MWD = 333.05 + (3.86 x TLL) + (12.93 x age) – (2.1 x weight)	0.47	54.27
	Male	Without TLL: 6MWD = 441.60 + (22.23 x age (y)) + (0.47 x height(cm)) – (0.4 x weight)	0.49	56.55
		6MWD = 351.60 + (17.82 x age) + (3.16 x TLL) – (1.65 x weight)	0.54	53.94
		Without TLL: 6WMD = 287.00 + (2.7 x height(cm)) + (10.04 x age) – (2.26 x weight)	0.33	56.47
Priesnitz et al. (26)	Pooled	6MWD = 145.343 + (11.78 x age) + (292.22 x height) + (0.611 x ΔHR) – (2.684 x weight)	0.366	54.81
Geiger et al. (33)	Female	6MWD = 188.61 + (51.50 x age) – (1.86 x age^2) + (86.10 x height)	0.50	57.52
	Male	6MWD = 196.72 + (39.81 x age) – (1.36 x age^2) + (132.28 x height)	0.49	66.72

Units are as follows (unless stated otherwise): Heart rate (HR): beats per minute, Height: meters, Age: years, Weight: kilogram, PAS - physical activity score, True leg length (TLL): centimeter, * LLN displayed in graph, NS: not stated

2

DISCUSSION

The aim of this study was to review the existing RV and to assess the methodological quality. Twenty-two eligible articles with a RV walked distance or RV prediction equations were evaluated on both methodological quality and testing procedures. It is impossible to determine a single best RV or RV prediction equation due to the heterogeneities in the study characteristics, testing procedures, reporting method and methodological quality. In contrast to Salbach et al., (47) no median per age group is presented because this method does not take into account the heterogeneity in sample characteristics and testing procedures.

The variety in RV is large with a range between 383m±41 (32) and 799m±54m. (33) Multiple explanations can be given for this wide range like methodological details, age, ethnical and cultural differences. The 6MWT was assessed in 13 different counties, representing the continents Asia, Europe, North America, South America, and Africa. Of the countries that are represented multiple times, the United States and Brazil are most frequently studied followed by Belgium, Thailand, and China. Two studies focused on a single racial groups (31,40) and four studies stated only including participants from a single nationality. (34,37,38,45) Ideally, studies report separate results for ethnic groups.

The wide range in RV can partly be explained by increasing age. (24,25,30,32,37,38,41-46) Nonetheless, also within an age group a large variety is shown. For example, in the study by Chen et al. (45) the mean distance of the sample of 7-year-olds scored 463m±62m in comparison to 625.4m±120.22m in the same age group in D'Silva et al. (38) Comparing within age groups is mainly limited by the diversity in reporting method, only nine studies gave age-by-age RV and six studies used age groups larger than two years.

However, the comparison of the RV within an age group in a single country show a smaller range compared to the range between countries. For example, within the United States the sample of Klepper et al. (36) and Roush et al. (29) both included females in the age group of 7 to 9-year-olds. Roush et al. (29) reports a RV of 532.2m±52.6m for the whole group while Klepper et al. (36) reports a RV of 519.64m±69.31m for 7–8-year-olds and 542.54m±80.25m for 9 year olds. Rahman et al. (43) tested the same gender and age group in Saudi Arabia and reports a RV of 564.26m±51.30m in 7-year-olds, 586.03m±41.42m in 8-year-olds and 600.86m±57.12m in 9-year-olds. These differences underline the conclusion by Klepper et al. stating the belief that RV or RV prediction equations for the 6MWT developed for children living in one country may not be applicable to those in other countries. (36) Similar conclusions have been drawn in the adult population.

(54) The differences in reported RV emphasize the recommendations of the ATS/ACCP guidelines that each research department and/or country should have its own RV. It is recommended that these RV are updated regularly because population characteristics may change over time. (51)

The variation in methodological details prior and during the performance of the 6MWT may also be a major factor that affects the results of the 6MWT. Most studies state using the ATS guideline (24-26,30,3-34,38,39,42-45) or comparable protocol. (40,41) nonetheless, these studies often refrained from or only partly describe the actual used method thereby making it impossible to verify the statement and reproduce the used methodology.

Studies using the instruction "*as fast as you can*" (23-25,42) generally scored higher RV compared to the same gender and age group using the instruction "*as far as possible*". (28,31,36,38,39,45,46) The RV in the study by Saad et al., (33) using no encouragement, is higher compared to the RV in the study by Goemans et al. which used constant encouragement. (40,41) This is in conflict with the findings of Guyatt et al. which found a higher walked distance in groups with encouragement compared to no encouragements. (53) This might suggest there is no influence of encouragement during the 6MWT.

Multiple studies aimed to construct a RV prediction equation in order to predict the 6MWD. Of the studies aiming to do so, only the study by Kenan Kanburoglu et al. (44) was unable to compose a RV prediction equation due to a decreasing 6MWD between the ages of 12 and 14 years old. Ulrich et al. (25) also reported a trend reversal around this age and composed different prediction equations for males younger and older than 13-year-old, and females younger and older than 12-year-old. The feasibility of several reported equations is debatable due to the use of variables which are difficult to obtain. Oliveira et al. (42) included the true leg length (TLL) in the equation. Because the measurement of the TLL is not included and the measurement of the pulse oximetry is optional in the ATS guideline, (1) it is not advised to include either in a RV prediction equation.

The RV prediction equation presented by Chen et al. (45) provides a Z-score from which the predicted RV can be derived. In the clinical practice this method is cumbersome and therefore unlikely to be applied. To increase the feasibility of the usage of RV prediction equations it is advisable to present a walked distance as outcome.

Strengths & limitations

This is the first systematic review to describe RV for the 6MWT in the healthy pediatric population. It provides a comprehensive overview of reported RV, methodical differences, and quality assessment.

In accordance with Bartels et al. (3) findings, this study is limited by the absence of an existing quality assessment tool. Consequently, a modified assessment tool is developed. Almost all studies, except one, lack the exclusion of smokers. This criterion might be too strict considering the submaximal nature of the test and age of the sample. However, in the authors opinion this is of importance to assure healthy participants. Furthermore, this criterion is consistent with the ATS cardiopulmonary exercise testing statement (ATS/ACCP) guidelines. (51)

The sample size in five of the included studies was small. A small sample size leads to a reduction of power and limit the ability to generalize the results to the reference population. (55) Salbach et al. (47) recommends a minimum of 15 participants per gender and age decade in order to make the study sufficiently precise. Because of the influence of age in the pediatric population, it is recommended that studies include 10 healthy males and 10 healthy females of similar age. (55)

The youngest group of participants is aged 3-year-old. In this study the 3-year-old participants were allowed to walk or run and jog. Although the study (31) states that reported RV is not affected much by this methodological adjustment, it does disturb the submaximal nature of the 6MWT. Also, it is questionable whether children of this age are able to concentrate and perform the required task during 6-minutes in order to create comparable values. Therefore, it is inadvisable to conduct the 6MWT in participants younger than 5-years-old.

Relevance for clinical practice

This review might help the clinician in choosing the best suited RV in order to make a comparative basis for answering questions concerning the normality of the 6MWD and exercise responses if obtaining own RV is not an option. For this purpose, we present a flow-chart in **figure 2** aimed to aid the process of choosing a suitable RV**.** The flow-chart is a modified version of the flow-chart presented by Paap et al. for cardiopulmonary exercise tests. (50) Modifications were made to suit the 6MWT and to emphasize the influence of age in the pediatric population. The study that best characterizes the sample of healthy volunteers tested should be selected by choosing matching age, gender, and geographical representation. Hereafter the best suited protocol should be chosen and if possible, methodological quality should be taken into account.

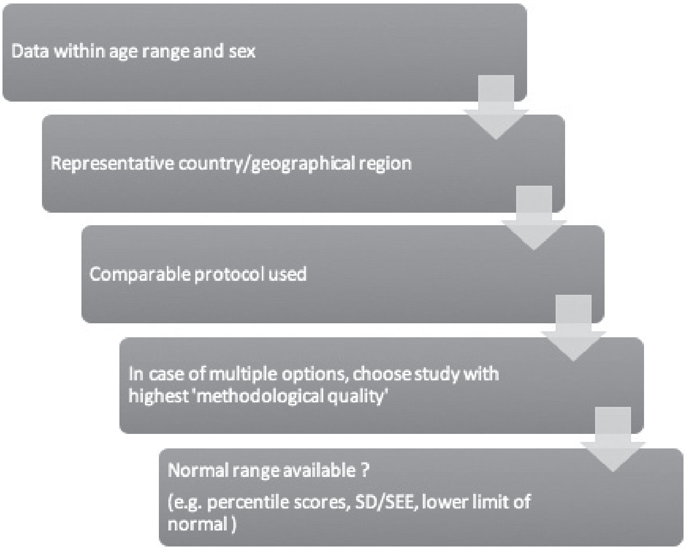


Figure 2. Flow-chart for the selection of '6-minute walk test' reference value in the pediatric population

Recommendation for future research

To increase usability of the 6MWT RV in the pediatric population a stricter appliance of the 6MWT protocol is advised. Furthermore, both a uniform age-by-age reporting method and a more thorough description of used method is needed. The application of the 6MWT in the age group below 5-years is not recommended because it is questionable whether children of this age are able to concentrate and perform the required task during 6-minutes.

Conclusion

The large variation in sample characteristics, applied methodology and quality assessment makes it impossible to present a single best RV for the 6MWT. Further research is needed to obtain RV for every world region and ethnicity. Until then, RV can be selected by using the flow-chart presented in **figure 2**. The flow-chart can aid the process of selecting RV for a research department and clinical practices if obtaining own RV is no possibility.

REFERENCES

1. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111-117.
2. Solway S, Brooks D, Lacasse Y, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. Chest 2001;119(1):256-270.
3. Bartels B, de Groot J, Terwee C. The six-minute walk test in chronic pediatric conditions: a systematic review of measurement properties. Phys Ther 2013;93(4):529-541.
4. Mandrusiak A. Functional capacity tests in young people with cystic fibrosis. New Zealand journal of physiotherapy 2009;37(1):13
5. Cunha M, Rozov T, de Oliveira R, Jardim J. Six-minute walk test in children and adolescents with cystic fibrosis. Pediatr Pulmonol 2006;41(7):618-622.
6. Lammers A, Diller G, Odendaal D, Tailor S, Derrick G, Haworth S. Comparison of 6-min walk test distance and cardiopulmonary exercise test performance in children with pulmonary hypertension. Arch Dis Child 2011;96(2):141-147.
7. Moalla W, Gauthier R, Maingourd Y, Ahmaidi S. Six-minute walking test to assess exercise tolerance and cardiorespiratory responses during training program in children with congenital heart disease. Int J Sports Med 2005;26(9):756-762.
8. Mazzone E, Vasco G, Sormani MP, Torrente Y, Berardinelli A, Messina S, et al. Functional changes in Duchenne muscular dystrophy: a 12-month longitudinal cohort study. Neurology 2011;77(3):250-256.
9. McDonald C, Henricson E, Han J, Abresch RT, Nicorici A, Atkinson L, et al. The 6-minute walk test in Duchenne/Becker muscular dystrophy: longitudinal observations. Muscle Nerve 2010;42(6):966-974.
10. McDonald C, Henricson E, Han J, Abresch RT, Nicorici A, Elfring G, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle Nerve 2010;41(4):500-510.
11. Maher C, Williams M, Olds T. The six-minute walk test for children with cerebral palsy. Int J Rehabil Res 2008;31(2):185-188.
12. Thompson P, Beath T, Bell J, Jacobson G, Phair T, Salbach N, et al. Test-retest reliability of the 10-metre fast walk test and 6-minute walk test in ambulatory school-aged children with cerebral palsy. Dev Med Child Neurol 2008;50(5):370-376.
13. Chong J, Mackey A, Broadbent E, Stott NS. Relationship between walk tests and parental reports of walking abilities in children with cerebral palsy. Arch Phys Med Rehabil 2011;92(2):265-270.
14. de Groot J, Takken T, Gooskens, Rob H J M, Schoenmakers, Marja A G C, Wubbels M, Vanhees L, et al. Reproducibility of maximal and submaximal exercise testing in "normal ambulatory" and "community ambulatory" children and adolescents with spina bifida: which is best for the evaluation and application of exercise training? Phys Ther 2011;91(2):267-276.
15. Montes J, McDermott MP, Martens WB, Dunaway S, Glanzman AM, Riley S, et al. Six-Minute Walk Test demonstrates motor fatigue in spinal muscular atrophy. Neurology 2010;74(10):833-838.
16. Lelieveld, Otto T H M, Takken T, van der Net J, van Weert E. Validity of the 6-minute walking test in juvenile idiopathic arthritis. Arthritis Rheum 2005;53(2):304-307.

17. Paap E, van der Net J, Helders PJM, Takken T. Physiologic response of the six-minute walk test in children with juvenile idiopathic arthritis. Arthritis Rheum 2005;53(3):351-356.

18. Alves VLdS, Avanzi O. Objective assessment of the cardiorespiratory function of adolescents with idiopathic scoliosis through the six-minute walk test. Spine 2009;34(25):E926-E929.

19. Elloumi M, Makni E, Ben Ounis O, Moalla W, Zbidi A, Zaoueli M, et al. Six-minute walking test and the assessment of cardiorespiratory responses during weight-loss programmes in obese children. Physiother Res Int 2011;16(1):32-42.

20. Makni E, Moalla W, Trabelsi Y, Lac G, Brun JF, Tabka Z, et al. Six-minute walking test predicts maximal fat oxidation in obese children. Int J Obes (Lond) 2012;36(7):908-913.

21. Guinhouya B. Outcomes and cardiac response of overweight prepubescent to the 6 minutes walk test. Minerva Pediatr 2011;63(5):375-384.

22. Takken T, Engelbert R, van Bergen M, Groothoff J, Nauta J, van Hoeck K, et al. Six-minute walking test in children with ESRD: discrimination validity and construct validity. Pediatr Nephrol 2009;24(11):2217-2223.

23. Morinder G, Mattsson E, Sollander C, Marcus C, Larsson U. Six-minute walk test in obese children and adolescents: reproducibility and validity. Physiother Res Int 2009;14(2):91-104.

24. Priesnitz C, Rodrigues G, Stumpf CdS, Viapiana G, Cabral C, Stein R, et al. Reference values for the 6-min walk test in healthy children aged 6-12 years. Pediatr Pulmonol 2009;44(12):1174-1179.

25. Ulrich S, Hildenbrand F, Treder U, Fischler M, Keusch S, Speich R, et al. Reference values for the 6-minute walk test in healthy children and adolescents in Switzerland. BMC Pulm Med 2013;13:49-49.

26. Li AM, Yin J, Yu CCW, Tsang T, So HK, Wong E, et al. The six-minute walk test in healthy children: reliability and validity. Eur Respir J 2005;25(6):1057-1060.

27. Singh S, Puhan M, Andrianopoulos V, Hernandes N, Mitchell K, Hill C, et al. An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease. Eur Respir J 2014;44(6):1447-1478.

28. Wasserman K, Hansen JE, Sue DY, et al. Principles of exercise testing and interpretation. J Cardiopulm Rehabil Prev 1987;7(4):189

29. Roush J. Reference values and relationship of the Six Minute Walk Test and Body Mass Index in healthy third grade school children. The Internet journal of allied health sciences and practice 2006;4(3):6p.

30. Li A, Yin J, Au J, So H, Tsang T, Wong E, et al. Standard reference for the six-minute-walk test in healthy children aged 7 to 16 years. Am J Respir Crit Care Med 2007;176(2):174-180.

31. Geiger R, Strasak A, Treml B, Gasser K, Kleinsasser A, Fischer V, et al. Six-minute walk test in children and adolescents. J Pediatr 2007;150(4):395-9, 399.e1.

32. Lammers AE, Hislop AA, Flynn Y, Haworth SG. The 6-minute walk test: normal values for children of 4-11 years of age. Arch Dis Child 2008;93(6):464-468.

33. Ben Saad H, Prefaut C, Missaoui R, Mohamed I, Tabka Z, Hayot M. Reference equation for 6-min walk distance in healthy North African children 6-16 years old. Pediatr Pulmonol 2009;44(4):316-324.

34. Limsuwan A, Wongwandee R, Khowsathit P. Correlation between 6-min walk test and exercise stress test in healthy children. Acta Paediatr 2010;99(3):438-441.

35. Basso R, Jamami M, Pessoa B, Labadessa I, Regueiro EMG, Di Lorenzo VAP. Assessment of exercise capacity among asthmatic and healthy adolescents. Rev Bras Fisioter 2010;14(3):252-258.

36. Klepper S, Muir N. Reference values on the 6-minute walk test for children living in the United States. Pediatr Phys Ther 2011;23(1):32-40.

37. Tonklang N, Roymanee S, Sopontammarak S. Developing standard reference data for Thai children from a six-minute walk test. J Med Assoc Thai 2011;94(4):470-475.

38. D'Silva C, Vaishali K P, Venkatesan. Six-minute walk test-normal values of school children aged 7-12 y in India: a cross-sectional study. Indian J Pediatr 2012;79(5):597-601.

39. Pathare N, Haskvitz E, Selleck M. 6-Minute Walk Test Performance in Young Children who are Normal Weight and Overweight. Cardiopulm Phys Ther J 2012;23(4):12-18.

40. Goemans N, Klingels K, van den Hauwe M, Van Orshoven A, Vanpraet S, Feys H, et al. Test-retest reliability and developmental evolution of the 6-min walk test in Caucasian boys aged 5-12 years. Neuromuscul Disord 2013;23(1):19-24.

41. Goemans N, Klingels K, van den Hauwe M, Boons S, Verstraete L, Peeters C, et al. Six-minute walk test: reference values and prediction equation in healthy boys aged 5 to 12 years. PLoS ONE 2013;8(12):e84120-e84120.

42. Oliveira A, Rodrigues C, Rolim D, Souza AAL, Nascimento O, Jardim J, et al. Six-minute walk test in healthy children: is the leg length important? Pediatr Pulmonol 2013;48(9):921-926.

43. Rahman SAA, Alnegimschi AA. Normative values of six-minute walk distance for healthy Saudi girls. World Appl Sci J, 2014 32 (8): 1721-1730

44. Kanburoglu M, Ozdemir F, Ozkan S, Tunaoglu F. Reference values of the 6-minute walk test in healthy Turkish children and adolescents between 11 and 18 years of age. Respir Care 2014;59(9):1369-1375.

45. Chen C, Chang C, Lin M, Hua Y, Fang W, Wu M, et al. Six-Minute Walking Test: Normal Reference Values for Taiwanese Children and Adolescents. Zhonghua Minguo Xin Zang Xue Hui Za Zhi 2015;31(3):193-201.

46. Fitzgerald D, Hickey C, Delahunt E, Walsh M, O'Brien T. Six-Minute Walk Test in Children With Spastic Cerebral Palsy and Children Developing Typically. Pediatr Phys Ther 2016;28(2):192-199.

47. Salbach N, O'Brien K, Brooks D, Irvin E, Martino R, Takhar P, et al. Reference values for standardized tests of walking speed and distance: a systematic review. Gait Posture 2015;41(2):341-360.

48. Moher D, Liberati A, Tetzlaff J, Altman D. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341.

49. Boluyt N, Tjosvold L, Lefebvre C, Klassen T, Offringa M. Usefulness of systematic review search strategies in finding child health systematic reviews in MEDLINE. Arch Pediatr Adolesc Med 2008;162(2):111-116.

50. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. Expert Rev Cardiovasc Ther 2014;12(12):1439-1453.

51. ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003;167(2):211-277.

52. Johnston LD, Malley PM, Bachman JG, Schulenberg JE. Monitoring the Future national survey results on drug use, 1975-2012. Volume I, Secondary school students, 2013 http://www.monitoringthefuture.org/pubs/monographs/mtf-vol1_2012.pdf.

53. Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman L, Jones NL, et al. Effect of encouragement on walking test performance. Thorax 1984;39(11):818-822.

54. Poh H, Eastwood P, Cecins N, Ho K, Jenkins S. Six-minute walk distance in healthy Singaporean adults cannot be predicted using reference equations derived from Caucasian populations. Respirology 2006;11(2):211-216.

55. Portnet LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd ed. Upper sadle river, NJ: pearson prentice hall 2008

Supplementary material A

Pubmed: 202 hits

Hits 7784

(((((((SMWT[Title/Abstract]) OR SMW[Title/Abstract]) OR SMWD[Title/Abstract]) OR 6MWT[Title/Abstract]) OR 6MWD[Title/Abstract]) OR 6MW[Title/Abstract]) OR 6 min* Walk*[Title/Abstract]) OR Six Min* Walk*[Title/Abstract]

AND

Hits 3680188

(((Highschool*[Title/Abstract]) OR High school*[Title/Abstract]) OR Elementary school*[Title/Abstract]) OR Secondary school*[Title/Abstract]) OR Primary school*[Title/Abstract]) OR Schools[MeSH Terms]) OR Nursery school*[Title/Abstract]) OR Kindergar*[Title/Abstract]) OR Paediatric[Title/Abstract]) OR Paediatrics[MeSH Terms]) OR Prepubescen*[Title/Abstract]) OR Pubescen*[Title/Abstract]) OR Pubert*[Title/Abstract]) OR Puberty[MeSH Terms]) OR Minors*[Title/Abstract]) OR Minors[MeSH Terms]) OR Girl*[Title/Abstract]) OR Boy*[Title/Abstract]) OR Teen[Title/Abstract]) OR Adoles*[Title/Abstract]) OR Adolescent[MeSH Terms]) OR Toddle*[Title/Abstract]) OR kids[Title/Abstract]) OR Kid[Title/Abstract]) OR Preschoo*[Title/Abstract]) OR Schoolchil*[Title/Abstract]) OR School ag*[Title/Abstract]) OR Schoolchild*[Title/Abstract]) OR Child*[Title/Abstract]) OR Child[MeSH Terms]) OR Postmatur*[Title/Abstract]) OR Prematur*[Title/Abstract]) OR Preterm*[Title/Abstract]) OR Neonat*[Title/Abstract]) OR Newbor*[Title/Abstract]) OR Baby*[Title/Abstract]) OR Newborn*[Title/Abstract]) OR infancy[Title/Abstract]) OR Infant*[Title/Abstract]) OR Infant[MeSH Terms]))

AND

Hits 2369762

((((((((((((Standard[Title/Abstract]) OR Criterion[Title/Abstract]) OR reference ranges[Title/Abstract]) OR Reference range[Title/Abstract]) OR Normal values[Title/Abstract]) OR Normal value[Title/Abstract]) OR Normative value[Title/Abstract]) OR Norma*[Title/Abstract]) OR Normal ranges[Title/Abstract]) OR reference values[MeSH Terms]) OR Reference values[Title/Abstract]) OR (Normal ranges))

Embase: 429 hits

Hits 15158

SMWT:ab,ti OR SMW:ab,ti OR SMWD:ab,ti OR 6MWT:ab,ti OR 6MWD:ab,ti OR 6MW:ab,ti OR ‘6 min* Walk*’:ab,ti OR ‘Six Min* Walk*’:ab,ti

AND

Hits 4114688

'highschool*':ab,ti OR 'high school*':ab,ti OR 'elementary school*':ab,ti OR 'secondary school*':ab,ti OR 'primary school*':ab,ti OR 'schools'/exp OR 'nursery school*':ab,ti OR 'kindergar*':ab,ti OR 'paediatric':ab,ti OR 'paediatric' OR 'prepubescen*':ab,ti OR

'pubescen*':ab,ti OR 'pubert*':ab,ti OR 'puberty'/exp OR 'minors*':ab,ti OR 'minors'/exp OR 'girl*':ab,ti OR 'boy':ab,ti OR 'teen':ab,ti OR 'adoles*':ab,ti OR 'adolescent'/exp OR 'toddle*':ab,ti OR 'kids':ab,ti OR 'kid':ab,ti OR 'preschoo*':ab,ti OR 'schoolchil*':ab,ti OR 'school ag*':ab,ti OR 'schoolchild*':ab,ti OR 'child*':ab,ti OR 'child'/exp OR 'postmatur*':ab,ti OR 'prematur*':ab,ti OR 'preterm*':ab,ti OR 'neonat*':ab,ti OR 'newbor*':ab,ti OR 'baby*':ab,ti OR 'newborn*':ab,ti OR 'infancy':ab,ti OR 'infant*':ab,ti OR 'infant'/exp
AND
Hits 2944199
'standard':ab,ti OR 'criterion':ab,ti OR 'reference ranges':ab,ti OR 'reference range':ab,ti OR 'normal values':ab,ti OR 'normal value':ab,ti OR 'normative value':ab,ti OR 'norma*':ab,ti OR 'reference values'/exp OR 'reference values':ab,ti OR 'normal ranges':ab,ti

---

Cinahl: 54 hits
Hits 251557
(Standard OR Criterion OR (reference ranges) OR (Reference range) OR (Normal values) OR (Normal value) OR (Normative value) OR (Norma*) OR (Normal ranges) OR (Reference values) OR (Normal ranges))
AND
Hits 616312
(Highschool* OR (High school*) OR (Elementary school*) OR (Secondary school*) OR (Primary school*) OR Schools OR (Nursery school*) OR Kindergar* OR Paediatric OR Paediatrics OR Prepubescen* OR Pubescen* OR Pubert* OR Puberty OR Minors* OR Minors OR Girl* OR Boy* OR Teen OR Adoles* OR Adolescent OR Toddle* OR kids OR Kid OR Preschoo* OR Schoolchil* OR (School ag*) OR Schoolchild* OR Child* OR Child OR Postmatur* OR Prematur* OR Preterm* OR Neonat* OR Newbor* OR Baby* OR Newborn* OR infancy OR Infant* OR Infant)
AND
Hits 2257
(SMWT OR SMWD OR SMW OR 6MWT OR 6MWD OR 6MW OR (6 min* Walk*) OR (Six Min* Walk*))

Supplementary material B

Modified methodological quality list according the ATS/ERS guidelines.

Population characteristics:

Subjects are community based. *(The subjects studied preferably be community bases rather than hospital based).*

Level of physical activity are reported.

Exclusion of different racial groups.

Exclusion criteria for unhealthy subjects is described.

If relevant (≥12y), exclusion of smokers in the sample studied. *(If the participants were 11 years or younger, a 1 was given)*

Specified characteristics are sub categorized by age group. *(Include gender, and anthropomorphic considerations. Grouped in maximum of 2year).*

Sample size:

The number of subjects tested is sufficiently equal or larger than the appropriately powered sample size, with a uniform distribution of subjects for sex and groups. *(20 per age year, equally distributed amongst gender (min 10))*

Randomization:

Subjects are randomly selected from a larger population *(The study design include a randomized selection process to avoid the potential bias seen when more or less physically active subjects volunteer for the study).*

Design:

A prospective study design

Quality assurance of equipment and methodologies:

There is no lack of quality control. *(Quality was achieved using recommendations contained in the ATS guidelines and the 6MWT protocol in accordance with recommendations specified in the ERS/ATS guidelines*).

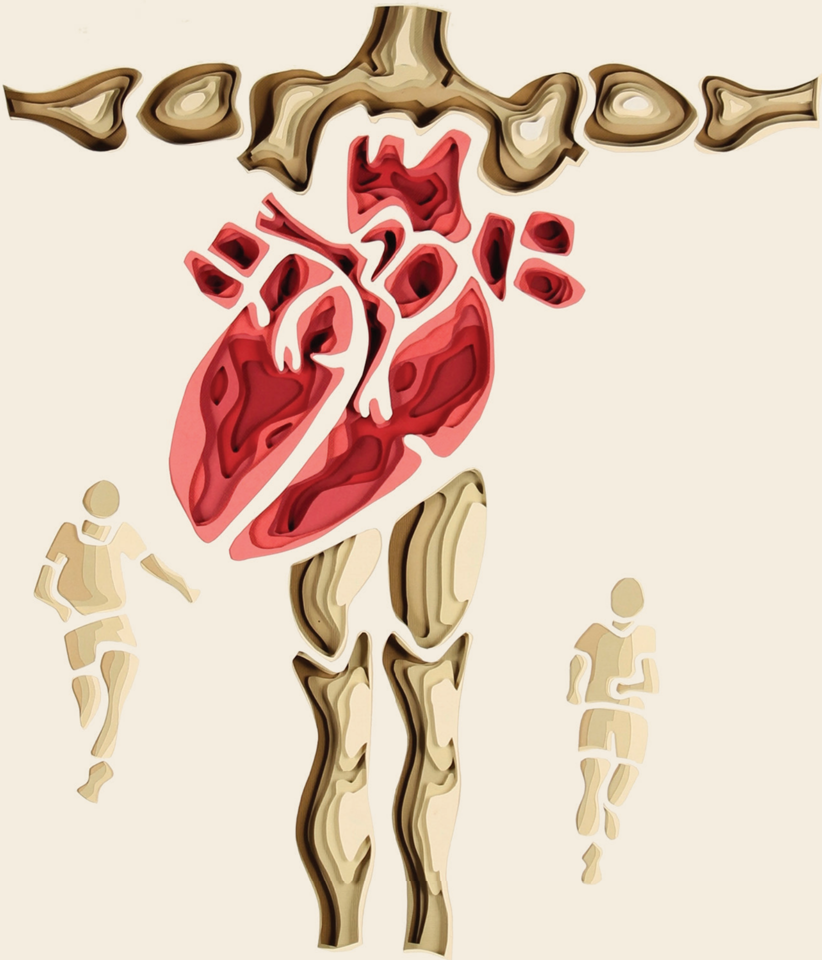
Testing protocol and procedures are described. *(Including track length, instructions and encouragements given prior to test)*

Validation:

Reference equations are cross-validated in population other than those used to generate the existing data.

Statistical validation:

The function that most accurately describes the distribution of the data are used. For example, curvilinear (power) functions may more accurately describe the distribution of the data. Furthermore, the precision of the individual and population predicted values are reported.



Chapter 3

Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the lowLands Fitness registry

CF Mylius, WP Krijnen, CP van der Schans, T Takken

Mylius CF, Krijnen WP, van der Schans CP, et al. Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry. ERJ Open Res 2019; 5: 00056-2018 https://doi.org/10.1183/23120541.00056-2018].

ABSTRACT

Peak oxygen uptake (VO_{2peak}) is recognized as the best expression of aerobic fitness. Therefore, it is essential that VO_{2peak} reference values are accurate for interpreting a cardiopulmonary exercise test. These values are country specific and influenced by underlying biological aging processes. They are normally stratified per pediatric and adult population resulting in discontinuity at the transition point between prediction equations. There are currently no age-related reference values available for the lifespan of individuals in the Dutch population. The aim of this study is to determine the best fitting regression model for the VO_{2peak} in the healthy Dutch pediatric and adult population in relationship to age.

In this retrospective study, cardiopulmonary exercise test cycle ergometry results of 4477 subjects without reported somatic diseases were included (907 female, 8-78 years). Generalized Additive Models were employed to determine the best fitting regression model. Cross-validation was performed against an independent dataset consisting of 3518 subjects. (170 female, 7-59 years).

An additive model was the best fitting with the largest predictive accuracy in both the primary (adj. R^2=0.57, SEE=556.50) and cross-validation (adj. R^2=0.57, SEE=473.15) dataset.

This study provides a robust additive regression model for peak oxygen uptake in the Dutch population.

Keywords: Cardiopulmonary exercise test – CPET – oxygen uptake – exercise capacity – reference value

INTRODUCTION

Peak oxygen uptake (VO_{2peak}) represents the functional limit of the body's ability to deliver and extract oxygen in muscles in order to satisfy the metabolic demands of vigorous exercise; it is recognized as the best expression of aerobic fitness. (1) VO_{2peak} is increasingly utilized to optimize risk stratification and to facilitate clinical decision-making because it reflects therapeutic response and predicts adverse events such as postoperative complications and mortality after abdominal and thoracic surgery. (2–4)

For the interpretation of a cardiopulmonary exercise test (CPET) using cycle ergometry, it is essential to have accurate VO_{2peak} reference values. (4–7). These values are region or country specific and change over time due to cultural differences and evolving population characteristics. (6,7) Therefore, each country must have specific updated VO_{2peak} reference values that optimally reflect the characteristics of the current population tested, the equipment, and the methodology utilized. (6–8). Although multiple countries provided up to date VO_{2peak} reference values derived from large cohorts, (9–11) 30-year-old VO_{2peak} reference values are the most commonly used in clinical settings in the Netherlands as there are none available derived for the Dutch adult population. (12). Thereby, these commonly used VO_{2peak} reference value prediction equations are obtained from a relatively small sample from the North American population. (7,12)

The VO_{2peak} is highly influenced by underlying biological aging processes such as physical development, pubertal status, age induced neuromuscular deterioration, sarcopenia, and cardiopulmonary decline. (5,13,14) It has been hypothesized in both the pediatric and adult population that VO_{2paek} develops in a non-linear and inter-related manner with the progression of age. (5,15–18) Linear regression models are predominantly used to determine VO_{2peak} reference value prediction equations depending upon gender, age, height, and weight. (7,12,14,19)

The frequently used age stratification between the pediatric and adult population is somewhat arbitrary and it introduces a discontinuity at the transition point between the two equations, which leads to a reference value shift from the pediatric to the adult population. Additionally, such a age stratification implies more prediction uncertainty since accuracy is lowest at the boundary of the sample age scale. Estimation of an up-to-date general prediction model across the pediatric as well as the adult population would facilitate a smooth transition into adult care. Therefore, the aim of this study is to determine the best fitting regression model for the VO_{2peak} in the healthy Dutch pediatric and adult population in relationship to age.

METHODS

This retrospective multi-center study is conducted using the Low-lands Fitness Registry, a primary dataset of 8900 subjects from 11 healthcare centers in the Netherlands that was aggregated with the aim of establishing CPET reference values for the Dutch population. Additionally, to determine the external and predictive validity of the reference value prediction model, a cross-validation procedure was performed on an independent sample as recommended by the American Thoracic Society/ American College of Chest Physicians (ATS/ACCP). (6) Specifically, the cross-validation in this study was performed against an additional dataset obtained from the Diving Medical Center from Den Helder, the Netherlands, and the Wilhelmina Children's Hospital from Utrecht, the Netherlands. The cross-validation dataset contained 4536 subjects that were not included in the primary dataset. Both datasets contain incremental CPET measurements collected between January 2010 and December 2016. Institutes that were included satisfied the following criteria: 1) To meet the ATS/ACCP-statement equipment requirements to perform an incremental CPET using an electromagnetically braked cycle ergometry test utilizing gas exchange analysis by bag collection, mixing chamber, or breath by breath analysis based upon averaging the values measured during last 30 to 60 seconds of the test (6); and 2) to perform equipment quality control in accordance with the ATS/ACCP-statement. (6)

Subjects included in both of the datasets underwent an individualized incremental CPET cycle ergometry test for multiple reasons including initiated by a healthcare professional; work and sports related (mandatory) annual health checks; participation in scientific studies; or based on personal motivation such as an exercise response evaluation for the aid of a trainings scheme. Every institute provided anonymized, coded patient information to the data coordinator at the University Medical Center Utrecht, the Netherlands. All records were previously screened for measurement failures. If there were any uncertainties, the testing institute was contacted to ensure the communication of correct data. It has been confirmed by the medical ethical research committee of the University Medical Center Utrecht that the WMO act does not apply to the current study.

STUDY SAMPLE

All of the subjects included in the study were habitants from the Dutch country, aged 65 years or less. The status "healthy" was defined as the absence of any reported somatic signs of disease and the exclusion of registrated available risk factors. (20) Therefore, subjects were excluded if (s)he; [1] reported somatic diseases at the time of testing or; [2] showed irregularities on the electrocardiogram (ECG) prior to testing. Additionally, subjects were excluded from further analysis if the subject; [3] included a missing predictor or outcome values. To ensure subjects reached their maximal VO_2 measurement, subjects were excluded if (s)he did [4] not reach a respiratory exchange peak ratio (RER_{peak}) of at least 1.0 (21) or did not reach a minimum of 85% of the age-predicted maximum heart rate (HR_{peak}) determined as 208-(0.7 * age). (22) Furthermore, due to the abnormal working capacity and cardiovascular responses to exercise in patients affected by underweight and the recognition of obesity as a disease by the World Health Organization, subjects who [5] had a body mass index (BMI) value ≥30 (23) or, in adult subjects, ≤18.5 (24) were excluded. Due to the decrease in VO_{2peak} associated with smoking, [6] subjects who actively smoked at the time of the test were excluded (25); and [7] lastly, professional athletes are excluded because they were considered not representative for the average Dutch population due to the positive effects of exercise training on VO_{2peak}. (26) The exclusion criteria were applied in both the primary and cross-validation datasets.

Statistical analyses

Statistical analyses were performed using R version 3.4.4, released in 2018. (27) Throughout, a p-value ≤.05 was considered significant. Continuous data were summarized as mean (SD) and categorical data as frequencies (percentage). The variables gender, age, weight, and height were included in the analyses as these are commonly used as a basis for VO_{2peak} reference value prediction equations. (7)

Generalized Additive Models (GAM) were utilized to semi-parametrically find the most appropriate fitting regression model. (28–30) To determine the model best fitting the data, criteria such as the adjusted R^2, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) were used. (31,32) A higher adjusted R^2 and a lower AIC and BIC was considered as improving the fit. In cases of inconsistency between these, the BIC criterion was taken as the most decisive. The interpretation of the BIC score was 0 to 2 as “minimal” improvement, 2 to 6 as “positive” improvement, 6 to 10 as “strong” improvement, and a >10 score as a “very strong” improvement. (33)

All models fitted to the data included an interaction age by gender term to account for the different VO_{2peak} levels between male and female subjects. (10) In order to compare with a best performing polynomial regression model, each predictive variable was modeled using linear, quadratic, and cubic effects by stepwise minimum BIC procedures. (34) The resulting model was employed to represent the polynomial model type in the model fitting procedures. Additionally, based upon the hypothesized non-linear age dynamics for VO_{2peak}, an additive model with a smooth spline type of transformation of age was included. (15–18,29)

To determine the fit of the models in the separated pediatric and adult population, the predictive accuracy of the models was measured using stratified age groups by comparing the residual standard error of the estimate (SEE). The groups were stratified by ≤20 and >20 years of age. The better the predictive fit of either of the three types of models, the less variability there is and the smaller the SEE. (35)

Models are of little clinical value unless these have predictive accuracy for independent samples. A cross-validation procedure was performed using each identified model per type (linear, polynomial, GAM) against a cross-validation dataset. Similar to criteria for the primary analysis, the model performance was evaluated by a larger adjusted R^2 and a smaller SEE.

For purpose of illustration, examples of VO_{2peak} predictions are reported using the best performing regression model. For these examples, cases with an increase of 5 years per pediatric case and 10 age years per adult case is used, corresponding average height and weight is used determined by data provided by the Statistics Netherlands (CBS). The 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97.5th prediction intervals are reported.

RESULTS

The complete registry consisted of 8900 cases (1641 female), after applying the exclusion criteria for missing values (N=2674), non-maximal tests (N=480), a BMI exceeding 30 or below 18.5 in adults (N=324), smokers (N=881) and, professional athletes (N=64), a sample of 4477 cases labelled as healthy remained (907 female) with age ranging from 7.9 to 65.0 years. The cross-validation sample contained 4536 subjects, after applying the exclusion criteria for missing values (N=0), non-maximal tests (N=64), a BMI exceeding 30 or below 18.5 in adults (N=260), smoking (N=694) and, professional athletes (N=0) 3518 subjects (170 female) with an age range from 6.8 to 59.0 years remained. **Table 1** summarizes the samples characteristics of both samples. **Figure 1** displays the age distribution of the primary sample.

Table 1. Sample characteristics

	Primary sample					Cross-validation sample				
Gender	**Subjects**	**Age**	**Weight**	**Height**	**BMI**	**Subjects**	**Age**	**Weight**	**Height**	**BMI**
Female	907 (20.4%)	32.24 ±12.76	64.32 ±11.78	168.61 ±9.33	22.46 ±3.04	170 (4.8%)	23.00 ±9.81	61.50 ±14.49	166.62 ±12.49	21.77 ±3.38
Male	3.570 (79.6%)	34.61 ±11.46	81.63 ±11.64	181.75 ±8.11	24.62 ±2.64	3.348 (95.2%)	33.89 ±9.98	84.30 ±11.56	182.75 ±8.19	25.13 ±2.45
All	4.477 (100%)	34.13 ±11.77	78.13 ±13.59	179.09 ±9.90	24.18 ±2.86	3.518 (100%)	33.36 ±10.24	83.20 ±12.72	181.97 ±9.13	24.97 ±2.60

Data are presented as n (%) or mean ±sd; Subject: total per gender, Age: years per decimal, Height: centimeters, Weight: kilograms, Body mass index (BMI): kilogram/meter^2.

3

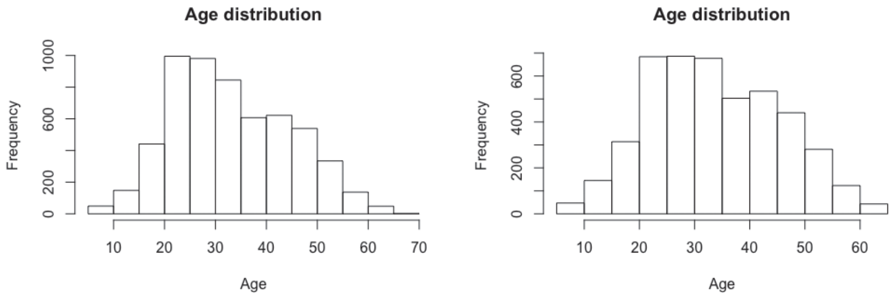


Figure 1. Sample age distribution

The best performing polynomial regression model that was found via stepwise minimum BIC was: VO_{2peak} (mL * min^{-1}) = -1469 + (673.00 * gender) + (16.87 * age) + (-0.47 * age^2) + (0.07 * height^2) + (39.70 * weight) + (-0.16 * weight^2). (adj. R^2= 0.56, AIC= 69480.15, BIC= 69531.40). The male gender is labelled as 1 and female as 0, age is presented in years, height in centimeters, and weight in kilograms.

Table 2 summarizes various estimated models and their fit measures. The best fitting model to the dataset was the additive model that includes a smooth spline transformation for age and an interaction term between age and gender plus linear terms for weight and height. The fit of the model yields an adjusted R^2= 0.57, AIC= 69342.81 and, BIC= 69449.50. This additive model demonstrate "very strong" improvements (33) compared to both the linear model (BIC diff. = 170.34) and the polynomial model (BIC diff. = 81.9). The age depended transformations of VO_{2peak} are displayed in **figure 2** including pointwise 95%-confidence intervals shown by the shaded bands. Additionally, the linear dependencies of weight and length are displayed in **figure 3** and **4**.

Table 2. Regression model type fitting comparison.

	Estimate	Std. error	T-value	P-value	Adj. R^2	AIC	BIC
Linear model							
Intercept	-3039.01	206.02	-14.75	<0.001	0.55	69581.40	69619.84
Gender	634.32	25.75	24.63	<0.001			
Age	-16.50	0.79	-20.66	<0.001			
Height	29.22	1.46	19.95	<0.001			
Weight	16.17	1.11	14.48	<0.001			
Polynomial model							
Intercept	-1469.00	158.80	-9.25	<0.001	0.56	69480.15	69531.40
Gender	673.00	25.89	25.99	<0.001			
Age	16.87	4.81	3.50	<0.001			
Age^2	-0.47	0.06	-7.31	<0.001			
$Height^2$	0.07	<0.01	16.52	<0.001			
Weight	39.70	5.17	7.67	<0.001			
$Weight^2$	-0.16	0.03	-5.05	<0.001			
Additive model							
Intercept	-2537.29	224.98	-11.28	<0.001	0.57	69342.81	69449.50
Gender	743.35	26.30	28.26	<0.001			
Height	24.30	1.52	15.91	<0.001			
Weight	12.57	1.12	11.21	<0.001			
S(age): male	edf*:	Ref.df**:	F-value:	<0.001			
S(age): female	4.263	5.260	22.59	<0.001			
	7.391	8.288	70.38				

Gender: 0= female, 1= male, Age= years, height= centimeters, weight= kilograms, * edf= effective degrees of freedom, ** Ref.df= reference number of degrees of freedom

The fit of the models compared to the separate pediatric and adult population are displayed in **table 3**. The additive model provides the largest predictive accuracy overall with an adjusted R^2= 0.57 and a SEE= 556.50 in the entire primary sample; the polynomial and the additive model performed equal against the cross-validation sample, specifically, R^2=0.57 compared to the linear model with R^2=0.55. Additionally, the additive model also provided the smallest SEE in the stratified age groups in both samples, namely, SEE= 495.18 and 420.72 in ≤20-years-old and SEE= 563.82 and 476.92 in the >20 years old. The largest improvement between models in both samples occurred in the ≤20-years-old age group. In this age group, the additive model has a better fit than both the linear and polynomial model with an equal adjusted R^2 difference of 0.05. Similar improvements are discerned in the SEE value between the additive and the linear and polynomial model of 65.47 and 53.62 ($mL * min^{-1}$) in the primary sample, and 108.14 and 35.81 ($mL * min^{-1}$) in the cross-validation sample, respectively.

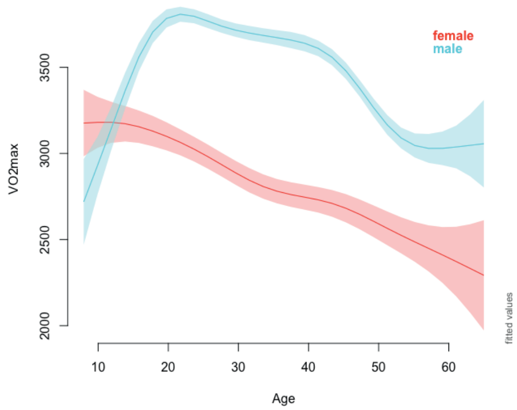


Figure 2. Age depended on transformation of VO2peak. Shaded bands represent pointwise 95%-confidence interval. Red = female, Blue = male

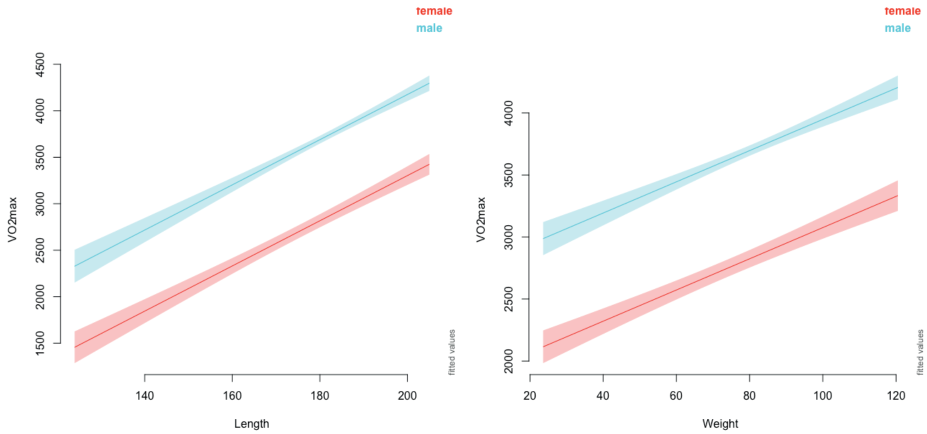


Figure 3. Relation between VO2peak and Weight. Including pointwise 95%-confidence interval. Red = female, Blue = male

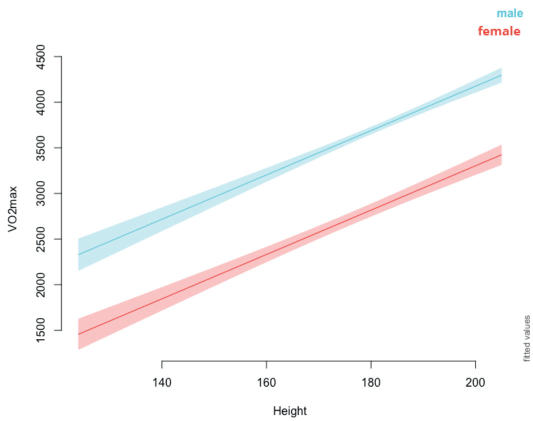


Figure 4. Relation between VO2peak and Height. Including pointwise 95%-confidence interval. Red = female, Blue = male

Table 3. Fit of model type per age group and sample set.

		Primary sample		Cross-validation sample	
	Age group	***Adj. R^2***	***SEE***	***Adj. R^2***	***SEE***
Linear model	All	0.55	572.89	0.54	487.84
	≤ 20 y	0.71	560.65	0.81	528.86
	> 20 y	0.50	574.43	0.37	484.56
Polynomial model	All	0.56	566.20	0.57	476.72
	≤ 20 y	0.71	548.80	0.82	456.03
	> 20 y	0.51	568.37	0.38	478.26
Additive model	All	0.57	556.50	0.57	473.15
	≤ 20 y	0.76	495.18	0.84	420.72
	> 20 y	0.52	563.82	0.38	476.92

Reference values with corresponding prediction intervals are constructed using average weight and height per gender and age provided by the CBS. **Table 4** shows predictions for the female cases, **table 5** shows the predictions for the male cases. In both genders, the 2.5th and 97.5th prediction interval in the 60-year-old cases is largest with 213 mL * min^{-1} for the female case and 352 mL * min^{-1} for the male case. The prediction interval of the 20-year-old cases are smallest with 78- and 131-mL * min^{-1} respectively. Both genders have increasing VO_{2peak} prediction until the age of 20 followed by a decline.

Table 4. Additive model VO2peak (mL * min-1) predictions percentiles per female case

Age	Height	Weight	2.5%	5%	10%	25%	Prediction	75%	90%	95%	97.5%
10	143.0	33.5	1581	1601	1624	1662	1704	1746	1784	1807	1826
15	164.0	52.0	2345	2359	2375	2401	2429	2458	2484	2500	2513
20	168.8	63.2	2543	2556	2570	2593	2619	2645	2668	2682	2694
30	169.3	68.5	2415	2426	2438	2458	2481	2503	2523	2536	2546
40	169.3	70.3	2298	2309	2322	2343	2367	2391	2412	2425	2436
50	167.7	70.5	2089	2104	2120	2148	2180	2211	2239	2255	2270
60	166.6	71.6	1793	1821	1854	1908	1969	2029	2084	2117	2145

Abbreviation: Age= years, Height= centimeters, Weight= kilograms

Table 5. Additive model VO2peak (mL * min-1) predictions percentiles per male case

Age (year)	Height (cm)	Weight (kg)	2.5%	5%	10%	25%	Prediction	75%	90%	95%	97.5%
10	143.0	34.0	1329	1351	1377	1421	1469	1518	1561	1587	1610
15	168.0	53.0	2775	2788	2804	2831	2860	2889	2916	2932	2945
20	183.5	78.1	3808	3816	3825	3841	3858	3875	3891	3900	3908
30	183.7	83.3	3818	3825	3832	3844	3857	3870	3882	3889	3896
40	182.4	85.1	3718	3725	3733	3747	3763	3778	3792	3800	3808
50	181.3	86.4	3292	3301	3311	3327	3346	3364	3381	3391	3399
60	179.2	84.4	2969	2986	3006	3039	3076	3112	3145	3165	3182

Abbreviation: Age= years, Height= centimeters, Weight= kilograms

DISCUSSION

The aim of this study was to determine reference values for VO_{2peak} based upon an optimal regression model in healthy Dutch pediatric and adult population. Based on adjusted R^2, AIC, BIC, and SEE, the additive model was the best fitting with the largest predictive accuracy. From the model, it can be concluded that the VO_{2peak} is gender-specific and depends non-linearly on years of age.

We determined that the additive model results in smaller SEE especially in the ≤20-year-old subjects because, in contrast to the linear model, the additive model is able to adjust for the age-related transformations like the increase in VO_{2peak} associated with the growth-related weight and height gain during childhood and adolescence. The increase in skeletal muscle mass during this life phase accounts for the majority of gained weight. (36) Because skeletal muscle mass is responsible for the majority of utilized oxygen during exercise, the increase in skeletal muscle mass associated with increasing age in ≤20-year-old subjects partially explains the increase in VO_{2peak} during this life phase. (37) During adulthood, the increase in skeletal muscle mass and height are limited. VO_{2peak} decreases during adulthood because of a decrease in muscle mass and a loss of chronotropic competence. (22,38)

Our additive regression model differs from previously utilized linear and polynomial regression models. (7,9,39) The use of the advanced statistical analysis method, GAM, in the current study makes it possible to determine the best fitting regression model for the combined pediatric and adult population. This method fits the data through cubic type of splines with the degree of smoothness determined by generalized cross-validation which facilitates combining the previously hypothesized nonlinear and inter-related fashion of more than one independent variable in the pediatric and adolescent population and the curvilinear decline with age in the adult population. (5,15–18) This method results in an improved fit across the entire population. (15,16,18) Therefore, using the additive model prediction of VO_{2peak} in the transition group between adolescents and adulthood is more precise.

In comparison with the prediction models currently utilized in the Dutch clinical settings, the additive model improves the fit in both the adult and pediatric population. The linear prediction model for adults provided by Jones et al. (12) yields a R^2=0.41 to the primary sample and R^2=0.33 to the cross-validation sample compared to respectively R^2=0.52 and R^2=0.38 in the additive model. The linear prediction equation provided by Ten Harkel et al. (39) is most frequently used in the Dutch pediatric population, this equation yields a R^2= 0.58, and R^2=0.73 compared to R^2=0.76 and R^2=0.84 in the primary and cross-validation sample, respectively. These improved fits make the additive model provided by the current study a more adequate reference prediction equation to utilize in both the pediatric and adult population.

Primary sample analysis and cross-validation showed consistent results, specifically, a stronger predictive accuracy in subjects aged ≤20-years-old and accuracy improvement in >20-year-old subjects and the entire sample. This consistent increase in predictive accuracy indicates a good generalizability to the Dutch population. This is underlined by the fit of R^2=0.54 (SEE=556.55) of the additive model to the whole sample including smokers, all BMI values, and athletes. The somewhat lower obtained adjusted R^2 of the additive model in the cross-validation subgroup >20-year-olds suggests a difference from the primary sample analysis. This is possibly caused using a variety of more institutions providing >20-year-old subjects in the cross-validation sample. Every >20-year-old in this sample was tested at a single institute aimed at test indications such as sports and work-related (mandatory) annual health checks. The underrepresentation of tests initiated by a healthcare professionals result in a cross-validation sample with higher aerobic fitness compared to the more heterogeneous primary sample (healthy workers effect).

The strength of our study is the wide age range of 7.9 to 65 years. The Low-lands Fitness Registry that we used in our study is a reasonable representation of the Dutch population. Additionally, the utilization of a diverse variety of healthcare centers including hospitals, sports medicine clinics, and occupational medicine clinics ensure representation of every conditioning status. The familiarity of the Dutch population with cycling and the low-risk of injury during testing ensures this method of measurement is fitting for the population and participants of all ages. (6)

Study results are limited by the retrospective and institution-based nature of the study. Preferably, VO_{2peak} reference value research should be performed in a prospective community based method (6) as a retrospective study design has potential data quality issues. (40) Although every utilized institution used measurement methods and equipment described by the ACCP/ATS-statement, the exclusion of 4364 subjects emphasizes the variety of data quality in the primary sample. (6) The majority of excluded subjects are due to missing values accounting for 2674 excluded subjects. It is of primary importance that CPET instructors increase skills and knowledge and stringently apply the test guidelines provided by ATS/ACCP-statement in order to facilitate data harmonization. (6)

Representative reference VO_{2peak} values are genuinely needed because of the current lack of reference data in the Dutch population. The currently employed North American reference values from 1985 may plausibly underestimate the aerobic fitness for the Dutch population, hence, subjects are misclassified as having normal aerobic fitness. The additive regression equation presented in the current study can be used to determine a reference value for the Dutch population. In future research aimed at determining reference value prediction equations, the type of regression model fitted to the data may conveniently be modeled by semi-parametric regression. This research can best be performed in a prospective, community-based setting with emphasis on the inclusion of sufficient numbers of female participants.

CONCLUSION

In conclusion, this study has provided a robust additive regression model for peak oxygen uptake in the Dutch population. Peak oxygen uptake is gender-specific and has a non-linear relationship with age. Publicly usable reference values can conveniently be obtained by suitable software implementation.

Acknowledgements

We would like to thank the collaborating centers for providing the anonymous exercise testing data. The Low-Lands Fitness Registry Study Group consists of: Harriet Wittink PhD, Marcel Schmitz MSc, Pieter-Jan van Ooi MD, Geert van Beek MSc, Leendert van Galen MSc, Jeroen Molinger MSc, Robert Rozenberg MD, Marieke van den Oord PhD, Yvonne Hartman PhD, Nicolle Verbaarschot MSc, Aernout Snoek MD, Jaap Stomphorst MD, Joep van Kesteren MSc.

REFERENCES

1. Myers J, McAuley P, Lavie C, Despres J, Arena R, Kokkinos P. Physical Activity and Cardiorespiratory Fitness as Major Markers of Cardiovascular Risk: Their Independent and Interwoven Importance to Health Status. Prog Cardiovasc Dis. 2015;57(4):306–14.
2. Wasserman K, Hansen JE, Sue DY SW and WB. Principles of Exercise Testing and Interpretation: Pathophysiology and Clinical Applications. Vol. 14, Shock. 2005. 1-180 p.
3. Older P, Levett D. Cardiopulmonary exercise testing and surgery. Ann Am Thorac Soc. 2017;14(July):S74–83.
4. Puente-Maestu L, Palange P, Casaburi R, Laveneziana P, Maltais F, Neder JA, et al. Use of exercise testing in the evaluation of interventional efficacy: An official ERS statement. Eur Respir J [Internet]. 2016;47(2):429–60. Available from: http://dx.doi.org/10.1183/13993003.00745-2015
5. Fleg JL, Morrell CH, Bos AG, Brant LJ, Talbot LA, Wright JG, et al. Accelerated longitudinal decline of aerobic capacity in healthy older adults. Circulation. 2005;112(5):674–82.
6. Ross R, Beck K, Casaburi R, Johnson B, Marciniuk D, Wagner P, et al. ATS/ACCP Statement on Cardiopulmonary Exercise Testing. Am J Respir Crit Care Med. 2003;167(10):1451.
7. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. Expert Rev Cardiovasc Ther. 2014;12(12):1439–53.
8. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, et al. The importance of cardiorespiratory fitness in the United States: The need for a national registry: A policy statement from the american heart association. Circulation. 2013;127(5):652–62.
9. Rapp D, Scharhag J, Wagenpfeil S, Scholl J. Reference values for peak oxygen uptake: Cross-sectional analysis of cycle ergometry-based cardiopulmonary exercise tests of 10 090 adult German volunteers from the Prevention First Registry. BMJ Open. 2018;8(3):1–12.
10. Kokkinos P, Kaminsky LA, Arena R, Zhang J, Myers J. A new generalized cycle ergometry equation for predicting maximal oxygen uptake: The Fitness Registry and the Importance of Exercise National Database (FRIEND). Eur J Prev Cardiol. 2018;25(10):1077–82.
11. Kaminsky L, Arena R, Myers J. Reference Standards for Cardiorespiratory Fitness Measured With Cardiopulmonary Exercise Testing. Mayo Clin Proc. 2015;90(11):1515–23.
12. Jones N, Makrides L, Hitchkock C, Chypchar T, McCartney N. Normal Stardards for an Incremental Progressive Cycle Ergometer Test. Am REV Respir Dis. 1985;131:700–8.
13. de Oliveira RJ, Bottaro M, Mota AM, Pitanga F, Guido M, Leite TKM, et al. Association between sarcopenia-related phenotypes and aerobic capacity indexes of older women. J Sport Sci Med. 2009;8(3):337–43.
14. Graves R, Mahnken J, Perea R, Billinger S, Vidoni E. Modeling Percentile Rank of Cardiorespiratory Fitness Across the Lifespan. Cardiopulm Phys Ther J. 2015;December(26(4)):108–13.
15. Hawkins SA, Wiswell RA. Rate and Mechanism of Maximal Oxygen Consumption Decline with Aging: Implications for Exercise Training. Sport Med. 2003;33(12):877–88.
16. Blais S, Berbari J, Counil F-P, Dallaire F. A Systematic Review of Reference Values in Pediatric Cardiopulmonary Exercise Testing. Pediatr Cardiol. 2015;36(8):1553–64.
17. Bouaziz W, Kanagaratnam L, Vogel T, Schmitt E, Dramé M, Kaltenbach G, et al. Effect of Aerobic Training on Peak Oxygen Uptake Among Seniors Aged 70 or Older: A Meta-Analysis of Randomized Controlled Trials. Rejuvenation Res. 2018;rej.2017.1988.

18. Nevill A, Cooke C. The Dangers of Estimating VO2max Using Linear, Nonexercise Prediction Models. Med Sci Sports Exerc. 2017;49(5):1036–42.

19. Graves L, Batterham A, Foweather L, Mcwhannell N, Hopkins N, Boddy L, et al. Scaling of peak oxygen uptake in children: A comparison of three body size index models. Med Sci Sports Exerc. 2013;45(12):2341–5.

20. Gräsbeck R. The evolution of the reference value concept. Clin Chem Lab Med [Internet]. 2004;42(7):692–7. Available from: http://nc.ibk.liu.se/nfkk/NORDKEM.

21. Guazzi M, Adams V, Conraads V, Halle M, Mezzani A, Vanhees L, et al. Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. Eur Heart J. 2012;33(23):2917–27.

22. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol. 2001;37(1):153–6.

23. Allison DB, Downey M, Atkinson RL, Billington CJ, Bray GA, Eckel RH, et al. Obesity as a disease: A white paper on evidence and arguments commissioned by the council of the obesity society. Obesity. 2008;16(6):1161–77.

24. Biadi O, Rossini R, Musumeci G, Frediani L, Masullo M, Ramacciotti C, et al. Cardiopulmonary exercise test in young women affected by anorexia nervosa. Ital Hear J. 2001;2(6):462–7.

25. Suminski R, Wier L, Poston W, Randles A, Arenare B, Jackson A. The Effect Of Habitual Smoking On Measured And Predicted VO2max. J Phys Act Heal. 2009;6:667–73.

26. Bovens AMPM, Van Baak MA, Vrencken JGPM, Wijnen JAG, Saris WHM, Verstappen FTJ. Maximal aerobic power in cycle ergometry in middle-aged men and women, active in sports, in relation to age and physical activity. Int J Sports Med. 1993;14(2):66–71.

27. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Internet]. Vienna, Austra; 2008. Available from: http://www.r-project.org

28. Wood SN. Generalized Additive Models: An Introduction with R. J R Stat Soc A. 2006;170(1):388.

29. Wood SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J Am Stat Assoc. 2004;99:673–86.

30. Wahba G. Spline Models for Observational Data. SIAM Rev. 1991;33(3):502–502.

31. Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control. 1974;19(6):716–23.

32. Schwarz G. Estimating the Dimension of a Model. Ann Stat. 1978;6(2):461–4.

33. Kass R, Raftery A. Bayes Factors. Vol. 90, Journal of the American Statistical Association. 1995. p. 773–95.

34. Zhang Z. Variable selection with stepwise and best subset approaches. Ann Transl Med. 2016;4(7):136–136.

35. Portney G, Watkins P. Foundations of Clinical Research: Applications to Practice. New Jersey Pearson Educ. 2009;892.

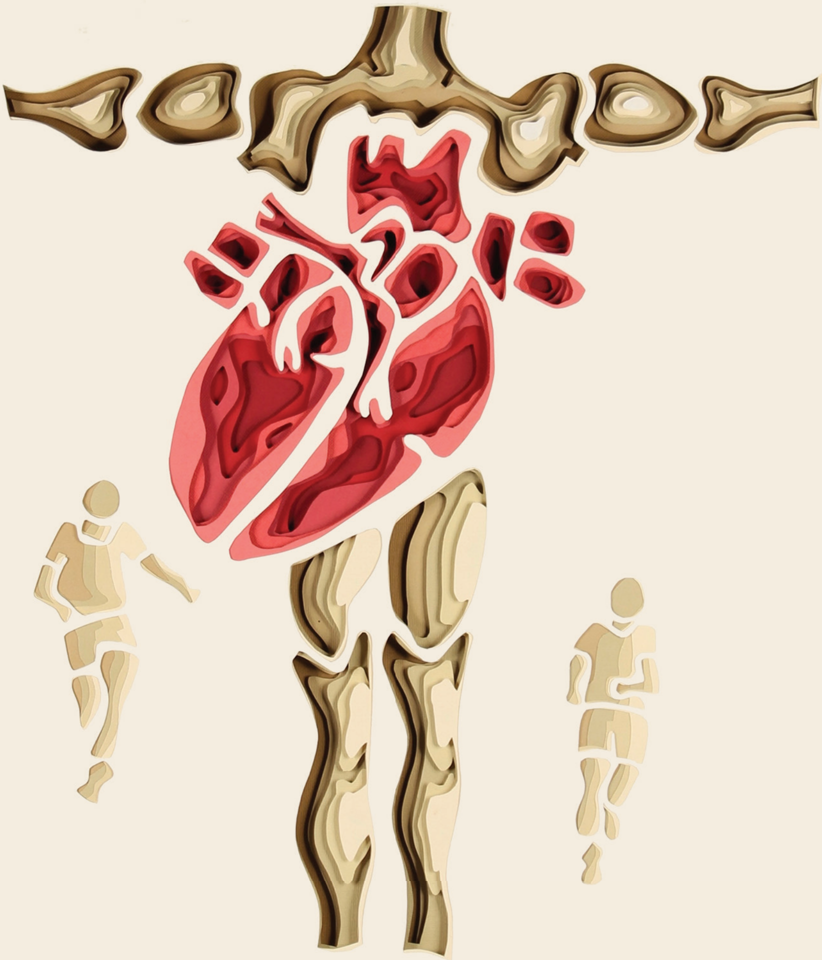
36. Weber DR, Leonard MB, Zemel BS. Body composition analysis in the pediatric population. Pediatr Endocrinol Rev. 2012;10(1):130–9.

37. Loomba-Albrecht L, Styne D. Effect of Puberty on Body Composition. Curr Opin Endocrinol Diabetes Obes. 2009;Feb(16(1)):10–5.

38. Köhler A, King R, Bahls M, Groß S, Steveling A, Gärtner S, et al. Cardiopulmonary fitness is strongly associated with body cell mass and fat-free mass. Scand J Med Sci Sport. 2018;1(Jan):1600-0838.

39. Harkel ADJ Ten, Takken T, Van Osch-Gevers M, Helbing W a. Normal values for cardiopulmonary exercise testing in children. Eur J Prev Cardiol. 2011;18(1):48–54.

40. Parker R, Berman N. Planning Clinical Research. Cambridge: Cambridge University Press; 2016.



Chapter 4

Reference values for cardiopulmonary exercise testing in healthy subjects – an updated systematic review

T Takken, CF Mylius, D Paap, W Broeders, HJ Hulzebos, M Van Brussel, BC Bongers

T. Takken, C.F. Mylius, D. Paap, W. Broeders, H.J. Hulzebos, M. Van Brussel & B.C. Bongers (2019) Reference values for cardiopulmonary exercise testing in healthy subjects – an updated systematic review, Expert Review of Cardiovascular Therapy, 17:6, 413-426, DOI: 10.1080/14779072.2019.1627874

ABSTRACT

Introduction: Reference values for cardiopulmonary exercise testing (CPET) parameters provide the comparative basis for answering important questions concerning the normalcy of exercise responses in patients, and significantly impacts the clinical decision-making process.

Areas covered: The aim of this study was to provide an updated systematic review of the literature on reference values for CPET parameters in healthy subjects across the life span.

A systematic search in MEDLINE, Embase, and PEDro databases was performed for articles describing reference values for CPET published between March 2014 and February 2019.

Expert Opinion/Commentary: Compared to the review published in 2014, more data have been published in the last five years compared to the 35 years before. However, there is still a lot of progress to be made. Quality can be further improved by performing a power analysis, a good quality assurance of equipment and methodologies, and by validating the developed reference equation in an independent (sub)sample. Methodological quality of future studies can be further improved by measuring and reporting the level of physical activity, by reporting values for different racial groups within a cohort as well as by the exclusion of smokers in the sample studied. Normal reference ranges should be well-defined in consensus statements.

Keywords: Cardiopulmonary exercise testing; Healthy adults; Healthy children; Exercise physiology; Reference values; Maximal oxygen uptake; Aerobic capacity; VO_{2max}.

INTRODUCTION

Cardiopulmonary exercise testing (CPET) is an important diagnostic tool for assessing aerobic fitness of individuals. (1) Although many different exercise testing protocols are employed to estimate aerobic fitness, (2) the gold standard for objectively assessing aerobic fitness remains cardiopulmonary exercise testing (CPET) during which respiratory gas exchange, ventilatory, and heart rhythm measurements are continuously performed throughout an incremental exercise intensity until voluntary exhaustion. (3) As such, CPET provides an evaluation of the integrative exercise response of the cardiovascular, respiratory, and metabolic systems to an incremental work rate. (4) This relatively non-invasive, dynamic physiologic test permits the evaluation of resting, submaximal, and peak exercise responses, as well as recovery responses, providing the clinician relevant information for clinical decision-making. (4) Examples concerning the usefulness of CPET for clinical decisions are the evaluation of exercise intolerance, (4) eligibility for organ transplantation, and preoperative risk stratification. (5)

Adequate reference values provide the comparative basis for answering important questions concerning the normality of exercise responses, and can significantly impact the clinical decision-making process. (6,7) As recommended by the American Thoracic Society/American College of Chest Physicians (ATS/ACCP) guideline, each exercise laboratory must select an appropriate set of reference values that best reflects the characteristics of the population tested, and the equipment, protocol, and methodology utilized to collect the reference values. (4) Many reference values for different CPET parameters obtained in different populations are available in the literature. We have previously published a systematic review of reference values for CPET parameters published up to 2014. (8) The current article is an update of our previous publication, including recent papers, as well as an extension towards the pediatric population. Reference values for pediatric CPET published up to 2014 were previously reviewed by Blais et al. (9) The aim of this study was to provide an updated systematic review of the literature on reference values for CPET parameters in healthy subjects across the life span.

METHODS

This systematic review of the literature followed the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. (10)

2.1 Data sources and search strategy

A search strategy was created and critically reviewed and approved by experienced exercise physiologists with the support of a medical librarian. After approval, published articles in the electronic databases MEDLINE, Embase, and PEDro were searched up to February 2019 (articles published from March 2014). We used the systematic search strategy as described in **Appendix A**. The search strategy did not have any limitations on ethnicity and language. Relevant reference lists were hand-searched as a method to supplement electronic searching.

2.2 Selection of studies

Results of the searches in different electronic databases were combined, where after duplicates were removed by two reviewers (CM and DP). The same two reviewers screened all unique records for potential relevance using the title, abstract or descriptors, or both. Hereafter, remaining articles were screened by the two reviewers on compliance with the eligibility criteria based on the full text of the articles. Reasons for possible article exclusion based on its full text were recorded.

2.3 Eligibility criteria

Studies with the objective to evaluate reference values for maximal CPET were included. Furthermore, inclusion criteria were: studies that included healthy subjects (no age restriction), studies using cycle or treadmill ergometry for CPET, cross-sectional studies or cohort studies, and studies that reported CPET parameters. Exclusion criteria were: studies published before March 2014, studies of which the full-text was not available, intervention studies, studies in which no maximal exercise protocol was used, and studies that exclusively included elite athletes.

2.4 Data extraction

All authors extracted data using a standard data extraction form. Data extraction was performed in pairs of reviewers (TT and MB, CM and DP, EH and WB), and discrepancies in extracted data were discussed with an independent reviewer (BB) till consensus was reached. If data were missing or further information was required, serious attempts were made to contact the corresponding authors to request for further information.

2.5 Methodological quality

Methodological quality of the selected studies was assessed using a quality list as provided in the ATS/ACCP guideline (see **appendix B**) 4. This list is a combination of study requirements to obtain an optimal set of reference values as described in the ATS/ACCP guideline and the code number scheme of shortcomings and limitations. Each criterion was scored as 'yes', 'no', or 'don't know', with one point for each 'yes'. A study was considered to be of high quality when it scored ≥10 points (≥75% of the maximum score of 14), of moderate quality when it scored 7 to 9 points, and of low quality when it scored ≤6 points. Quality assessment of all studies was performed in pairs of reviewers as well, and discrepancies in the scoring of criterions were discussed till consensus was reached. There was no blinding on authors or journal.

RESULTS

3.1. Selected studies

We identified 578 potential studies published between March 2014 and February 2019. After initial screening, 125 studies were regarded potentially eligible. After reading the full-text, 29 studies were considered eligible for inclusion. A flowchart displaying exact details of the selection process, including the reasons for exclusion, is presented in **Figure 1**.

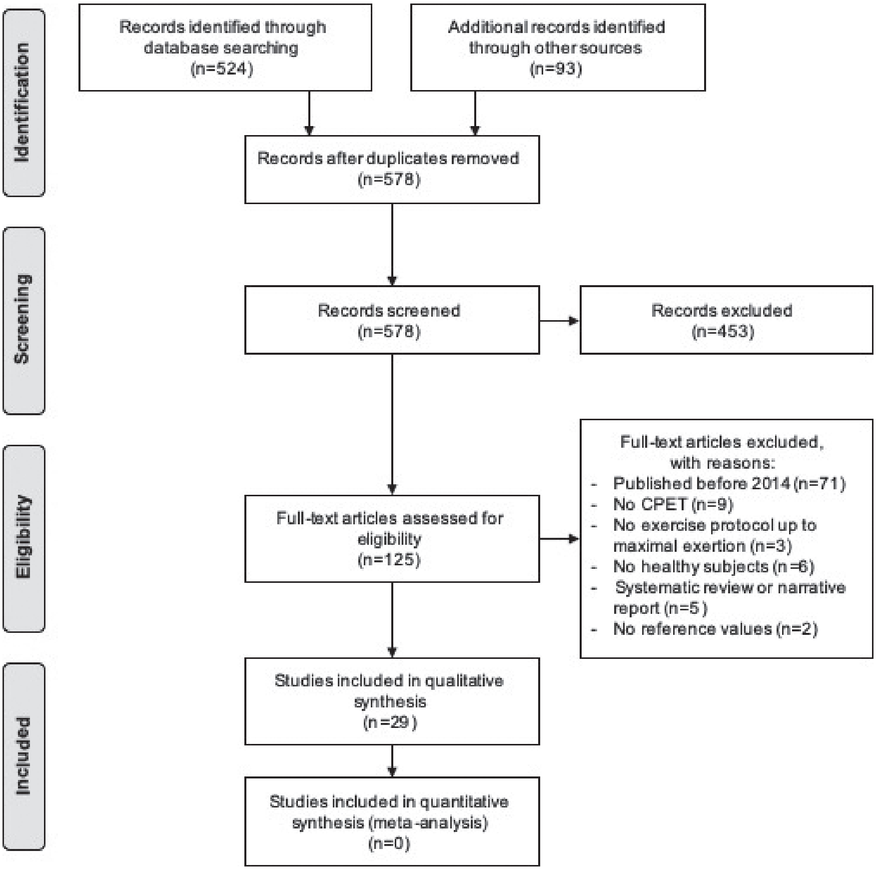


Figure 1. The PRISMA flow diagram displaying the selection of studies and reasons for exclusion.

Table 1. Overall study characteristics

Reference	Sample size (males/females)	Age (years)	Sample characteristics	Country	Smokers included	Treadmill or cycle ergometry	Protocol	Primary parameters measured	Methodology	Time averaging (s)
Aadland 2016	765 (402/363)	20-85	Population-based, retrospective	Norwegian	Yes	TM	Modified Balke protocol	VO_2, HR, RER	Gas analyzer	30 s
Abella 2016	215 (138/77)	6-17	Hospital-based, retrospective	Argentina	?	TM	Bruce protocol	VO_2[a], HR, RER, O_2-pulse, VE/VCO_2-slope, SpO_2	B×B	10-60 s
Agostini 2017	500 (260/240)	18-77	Population-based, prospective	Italy	Yes	CY	Personalized incremental ramp protocol	VO_2, CO, arteriovenous oxygen difference, HR, SV, CI	B×B	20 s
Almeida 2014	3119 (1624/1495)	8-90	Hospital-based, retrospective	Brazil	Yes	TM	Personalized incremental ramp protocol	HR, SBP, DBP, RER, VE, VO_2	B×B	20 s
Blanchard 2018	228 (112/116)	12-17	Population-based, prospective	Canada	No	CY	Personalized incremental ramp protocol	VO_2, O_2-pulse, WR[a], VE, HR, RER, OUES, OUES-slope below VAT, VE/VCO_2-slope, VE/VCO_2-slope below VAT, VE/VCO_2 at VAT, VO_2/WR-slope, O_2-pulse/WR-slope, HRR[c]	B×B	?
Bongers 2015	214 (114/100)	8-19	Population-based, prospective	The Netherlands	?	CY	Godfrey protocol (10, 15, or 20 W/min)	WR, HR, RER, VO_2[a], VE, VE/VCO_2-slope, OUES, OUEP, OUE at VAT	B×B	30 s
Buys 2014	1411 (877/534)	20-60	Population-based, prospective	Belgium	?	CY	Incremental protocol (20 W/min)	VO_2, WR, HR, RER, OUES	B×B	30 s

4

Reference	Sample size (males/ females)	Age (years)	Sample character-istics	Country	Smokers included	Treadmill or cycle ergometry	Protocol	Primary parameters measured	Metho-dology	Time averaging (s)
Dilber 2015	164 (99/65)	11-17	Hospital-based	Croatia	?	TM	Bruce protocol	WR[1], HR[a,b], RER, VO_2[a], O_2-pulse[a,b], $\Delta VO_2/\Delta WR$, SBP[a,b], BF[a,b], VT[a,b], VE[a], VE/VO_2[a], VE/VCO_2[a], VD/VT[a], $PETCO_2$[a]	B×B	15 s
Duff 2017	70 (33/37)	10-18	Population-based, prospective	Canada	?	TM	Incremental TM protocol (start at 2.0 mph, 1%, increase of 0.5 mile/hr/min)	VO_2, VE, HR, RER	B×B	15 s
Genberg 2016	181 (90/91)	50	Population-based, prospective	Sweden	Yes	CY	Incremental protocol (10 W/min, with initial work rate of 30 W (women) and 50 W (men)	WR, VO_2[a], VE/VCO_2 at VAT	B×B	?
Herdy 2015	3922 (2388/1534)	15-74	Hospital-based, prospective	Brazil	No	TM	Personalized incremental ramp protocol	VO_2	Mixing chamber gas analyzer	10 s
Hossri 2018	217 (69/148)	4-21	Hospital-based, retrospective	Brazil	?	TM	Personalized incremental ramp protocol	OUES, $PETCO_2$ at rest, VE/VCO_2-slope, VAT, O_2-pulse, RER, SpO_2[2]	Gas analyzer	30 s
Kaafarani 2017	184 (113/71)	6-18	Hospital-based, retrospective	The Netherlands	No	CY	Godfrey protocol (10, 15, or 20 W/min)	VO_2, WR, RER, SBP	B×B	30 s
Kaminsky 2015[d]	7783 (4611/3172)	20-79	Population-based, retrospective	United States	?	TM	Personalized incremental ramp protocol	VO_2, HR, RER	Gas analyzer	20-30 s

Reference	Sample size (males/ females)	Age (years)	Sample character-istics	Country	Smokers included	Treadmill or cycle ergometry	Protocol	Primary parameters measured	Metho-dology	Time averaging (s)
Kaminsky 2017[d]	4494 (1717/2777)	20-79	Population-based, retrospective	United States	?	CY	Personalized incremental ramp protocol	WR, HR, RER	Gas analyzer	20-30 s
Kaminsky 2018[d]	5232 (3043/2189)	20-79	Population-based, retrospective	United States	?	TM	Personalized incremental ramp protocol	VE, VO_2, HR[2], SBP at rest, DBP at rest	Gas analyzer	20-30 s
Kokkinos 2018[d]	5100 (3378/1722)	20-79	Random, population-based, retrospective	United States	Yes	CY	?	VO_2	Open circuit spirometry	30-60 s
Lintu 2014	140 (71/69)	9-11	Hospital-based, retrospective	Finland	?	CY	Incremental (1 W/6 s, with initial work rate of 20 W)	WR, VO_2, VE, RER, VE/ VCO_2 (lowest), O_2-pulse, HR[b,c], SBP[b,c]	B×B	15 s
Loe 2014	3512 (1758/1754)	20-90	Random, population-based, prospective	Norway	Yes	TM	Incremental (0.5-1.0 km/h/min or 1-2% incline)	WR[a], HR[a], VO_2[a], VE[a], BF[a], VT[a], VCO_2[a], RER[a]	B×B	?
Meyers 2017[d]	7759 (4601/3158)	20-79	Population-based, retrospective	United States	?	TM	Personalized incremental ramp protocol	VO_2, HR, RER, SBP, DBP	?	20-30 s
Mylius 2019	4477 (3570/907)	7-65	Population-based, retrospective	The Netherlands	No	CY	Personalized incremental ramp protocol	VO_2	B×B	30-60 s
Neto 2019	18189 (12555/5634)	13-69	Population-based, retrospective	Brazil	?	TM	Personalized incremental protocol	VO_2	B×B	30 s

Reference	Sample size (males/ females)	Age (years)	Sample character-istics	Country	Smokers included	Treadmill or cycle ergometry	Protocol	Primary parameters measured	Metho-dology	Time averaging (s)
Ozemek 2017	2644 (1510/1134)	18-76	Population-based, retrospective	United States	No	TM	Bruce, modified Bruce, BSU Bruce ramp, Balke, modified Balke, and personalized incremental protocol	VO_2, HR[b]	Open circuit spirometry	?
Pistea 2016	99 (58/41)	>70	Population-based, prospective	France	Yes	CY	Incremental 10, 15, 20, 25, or 30 W/min (depending on subjects age, body mass, and physical fitness level)	VO_2, HR[b], WR, VE[b], VE/VCO_2, VE/VO_2, RER	B×B	20 s
Rapp 2018	10090 (6462/3628)	21-83	Population-based, retrospective	Germany	Yes	CY	Ramp protocol + multistage protocols	VO_2, SBP, DBP	B×B	10 s
Sabbahi 2018[d]	2736 (1525/1211)	20-79	Random, population-based, retrospective	United States	?	TM	?	SBP[b], DBP[b], HR	NA	NA
Stensvold 2017	310 (150/160)	70-77	Random, population-based, prospective	Norway	Yes	CY/TM	10 W/30 s, on CY, or incremental protocol on TM	HR[c], VO_2[a], RER[a], VCO_2[a], BF[a], VE[a], BR, VT, O_2-pulse, VE/VO_2[a], VE/VCO_2[a], SBP, DBP	B×B	Average of three highest consecu-tive values

Reference	Sample size (males/ females)	Age (years)	Sample character-istics	Country	Smokers included	Treadmill or cycle ergometry	Protocol	Primary parameters measured	Metho-dology	Time averaging (s)
Tompouri 2017	38 (18/20)	9-11	Hospital-based, prospective	Finland	?	CY	Incremental 1 W/6 s, with initial work rate of 20 W	WR[a], VO_2[a,b], RER	B×B	15 s
van de Poppe 2018	3463 (2868/595)	20-60	Population-based, retrospective	The Netherlands, Belgium	No	CY	Personalized incremental ramp protocol	WR, VO_2, HR, RER	B×B	30 s

If not explicitly stated, a variable was obtained at peak exercise.
Abbreviations: B×B=breath-by-breath; BF=breathing frequency; BR=breathing reserve; CI=cardiac index; CO=cardiac output; CY=cycle ergometry; DBP=diastolic blood pressure; HR=heart rate; HRR=heart rate reserve; NA=not applicable; O_2-pulse=oxygen-pulse; O_2-pulse/WR-slope=relation between oxygen-pulse and work rate; OUE=oxygen uptake efficiency; OUEP=oxygen uptake efficiency plateau; OUES=oxygen uptake efficiency slope; $PETCO_2$=end tidal carbon dioxide pressure; RER=respiratory exchange ratio; s=seconds; SBP=systolic blood pressure; SpO_2=peripherally measured oxygen saturation; SV=stroke volume; TM=treadmill ergometry; VAT=oxygen uptake at the ventilatory anaerobic threshold; VCO_2=carbon-dioxide production; VD/VT=physiologic dead space to tidal volume ratio; VE=minute ventilation; VE/VCO_2=minute ventilation to carbon dioxide production ratio; VE/VCO_2-slope=relationship between minute ventilation to carbon dioxide production; VE/VO_2=minute ventilation to oxygen uptake ratio; VO_2=oxygen uptake; VO_2/WR-slope=relation between oxygen uptake and work rate; $\Delta VO_2/\Delta WR$=delta oxygen uptake to delta work rate ratio (oxygen cost of work); VT=tidal volume; WR=work rate; ?=unknown.
[a]: Variable(s) also obtained at the VAT; [b]: Variable(s) also obtained at rest; [c]: Variable(s) also obtained during recovery; [d]: data from the FRIEND registry.

3.2 Study characteristics

Table 1 depicts overall study characteristics. The 29 included studies assessed 87.256 subjects in total, of which were 54.214 males and 33.042 females. Age of included subjects ranged between 6 and 90 years. CPET was performed using a cycle ergometer in fourteen studies (48.3%) and using a treadmill in fourteen studies (48.3%), whereas one study (3.4%) used both modalities. There was a wide variety in the used CPET protocols, in which all studies used a continuous stepwise or ramp incremental protocol. Included studies included data from three different continents, of which most represented countries were European (n=16), North-American (n=9), and South-American (n=5). Sample size ranged from 38 to 18.189 subjects. Sixteen studies (55.2%) were performed in adults, eight studies (27.6%) in children, and five studies (17.2%) in a combined sample. Some of the publications included CPET data from the same core database (e.g., FRIEND database, LowLands Fitness Registry).

3.3 Methodological quality assessment

Quality of the included studies varied, and none of the studies fulfilled all 14 quality criteria. A 'quality score' ≥10 was seen in four studies, fifteen studies received a score of 7 to 9, and eleven studies received a score of ≤6. Frequently observed weaknesses were a lack of power analysis, quality assurance of equipment and methodologies, and reference equation validation. **Table 2** provides a detailed overview of the methodological score of the included studies on the ATS/ACCP quality list. (4)

3.4 Meta-analysis

Each of the included studies has various numbers of shortcomings and limitations, which are noted in **Table 2**. Meta-analysis of the data was not meaningful, as a large heterogeneity of methods and subjects (including sampling bias, uneven quality of primary data, and inadequate statistical treatment of the data) was observed.

Table 2. Methodological quality of the included studies list based on the ATS/ACCP guideline

Reference	A/P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total score
Aadland, 2016	A	0	1	0	0	1	0	0	0	1	1	1	1	1	1	9
Abella, 2016	P	0	0	0	0	1	0	0	0	0	1	1	0	0	0	3
Agostini, 2017	A	1	0	0	0	1	0	0	1	0	1	1	1	0	1	7
Almeida, 2014	A+P	0	1	0	0	1	1	0	0	0	1	1	0	1	0	6
Blanchard, 2018	P	1	1	0	1	1	0	0	1	0	1	1	1	0	1	9
Bongers, 2016	P	1	0	0	0	1	0	0	1	1	1	1	1	0	1	8
Buys, 2014	A	1	0	1	0	1	1	0	1	1	1	1	1	1	1	11
Dilber, 2015	P	0	0	0	0	1	0	0	?	0	1	1	0	0	0	3
Duff, 2017	P	1	0	0	0	0	1	0	1	1	1	1	0	0	0	6
Genberg, 2016	A	1	1	0	0	1	1	1	1	1	1	1	0	0	0	9
Herdy, 2016	A+P	1	1	0	1	1	1	0	1	0	1	1	0	1	1	10
Hossri, 2018	A+P	0	0	0	0	1	1	0	0	0	0	1	1	0	1	5
Kaafarani, 2017	P	1	0	0	1	1	0	0	0	1	1	1	1	0	1	8
Kaminsky, 2015	A	1	0	0	0	1	1	1	0	1	0	1	1	0	1	8
Kaminsky, 2017	A	1	0	0	0	1	1	1	0	1	0	1	1	0	1	8
Kaminsky, 2018	A	1	0	0	0	1	1	1	0	1	0	1	1	1	1	9
Kokkinos, 2018	A	1	0	0	0	1	1	1	0	1	1	0	1	1	1	9
Lintu, 2015	P	0	0	0	1	1	1	0	0	0	1	1	0	0	1	6
Loe, 2014	A	1	0	0	0	1	1	1	1	1	1	1	0	0	1	9
Myers, 2017	A	1	0	0	0	1	1	0	0	1	1	1	1	1	1	9
Mylius, 2019	A+P	1	0	0	1	1	1	0	0	1	1	1	1	1	1	10
Neto, 2019	A+P	1	0	0	0	1	1	0	0	1	1	1	1	0	1	8
Ozemek, 2017	A	1	0	0	0	1	1	0	0	0	1	1	0	0	1	6
Pistea, 2016	A	0	1	0	0	1	0	0	1	0	1	1	0	0	0	5
Rapp, 2018	A	1	0	0	1	1	1	0	0	1	1	1	1	0	1	9
Sabbahi, 2017	A	1	0	0	0	1	1	0	0	1	1	1	1	0	0	7
Stensvold, 2017	A	1	1	0	0	1	1	0	1	1	1	1	1	0	1	10
Tompuri, 2017	P	1	1	0	1	1	0	0	1	0	1	1	0	0	1	8
van de Poppe, 2018	A	1	0	0	1	1	1	0	0	1	1	1	1	1	1	10

Abbreviations: A=adult subjects; P=pediatric subjects, 0= criterion is not met, 1= criterion is not met.

3.5 Results of individual studies

Table 3 shows reference values for cardiovascular, ventilatory, and ventilatory efficiency parameters. Studies differed in the way of reporting reference values. Studies that did report reference values using regression equations are included in **Table 3**. Several studies reported their reference values in tables. We refer to these specific tables of the respective study for further details.

3.6 Cardiovascular parameters

3.6.1 Oxygen uptake at peak exercise

Twenty-six studies reported oxygen uptake at peak exercise (VO_{2peak}) in L/min, mL/min, or in mL/kg/min, (11-28), but not all studies provided reference values. Several different parameters were used to predict VO_{2peak}. Body height, body mass, age, and sex were often included in prediction equations. VO_{2peak} (absolute values) increased with body height and body mass, was lower in females, decreased with age during adulthood, but increased with age during childhood.

3.6.2 Ventilatory anaerobic threshold

Only one study in children reported ventilatory anaerobic threshold (VAT) values 29, no study reported VAT values in adult subjects. Reference values for VAT (mL/min) increased with body height and body mass in children and were provided for male and female subjects separately.

3.6.3 Heart rate at peak exercise

One study in children 29 and one study performed in adults 30 provided prediction equations for heart rate at peak exercise (HR_{peak}). The pediatric study reported four different equations, two for males, and two for females. Body height, body mass, and age were predictors of $HR_{peak.}$ (29) Six prediction equations for HR_{peak} in adults were reported using both cross-sectional and longitudinal data. Males had a higher HR_{peak} during young adulthood compared to females; however, males showed a somewhat faster decline in HR_{peak} values with age compared to females. (30)

3.6.4 Oxygen pulse

One study 29 performed in children provided four different equations for peak oxygen pulse (O_2-pulse), two for males, and two for females. No study reported O_2-pulse reference values in adults.

3.6.5 Blood pressure

One study (31) performed in children provided two prediction equations for systolic blood pressure at peak exercise. Systolic blood pressure increased with attained work rate at peak exercise (WR_{peak}), and the increment in systolic blood pressure was independent of age and sex. There was no study that provided reference values in adults for systolic blood pressure at peak exercise.

3.6.6 Work rate at peak exercise

Two studies 29,32 reported equations for the attained WR_{peak} during CPET. These studies reported 18 different equations for the prediction of WR_{peak}. In adults, WR_{peak} increased with body height, body mass, and was significantly higher in male subjects. In children, WR_{peak} increased with the development of body height and body mass (**Table 3**).

3.7 Ventilatory parameters

3.7.1 Minute ventilation at peak exercise

Ten studies (29,33-41) reported data for minute ventilation at peak exercise (VE_{peak}). Almost all studies reported VE_{peak} data using tabulated data. Two sex-specific prediction equations were provided for children. (29) One prediction equation was provided for adults, (37) in which VE_{peak} values were lower in females and declined with age throughout adulthood.

3.7.2 Tidal volume at peak exercise

Four studies 29,35,39,41 reported reference values for tidal volume at peak exercise (TV_{peak}). Two studies were performed in children [29,35] and two in adults [39,41]. One study [29], performed in children, provided a prediction equation for TV, the other studies provided tabulated data.

3.7.4 Breathing frequency at peak exercise

Two studies 35,41 reported breathing frequency at peak exercise (BF_{peak}). One study 35 was performed in children and one in older adults (70-77 years of age) 35. Results were only provided in tabulated data.

3.8 Ventilatory efficiency parameters

3.8.1. Oxygen uptake efficiency plateau and oxygen uptake efficiency slope

One study (34) in children reported a reference equation for oxygen uptake efficiency plateau (OUEP). No results in adults were found. Five studies reported oxygen uptake efficiency slope (OUES) values, two in adults, (42,43) two in a pediatric population, (29,34) and one study reporting up to young adulthood (21 years of age). (44) Results were reported for males and females separately. Other commonly used predictors were age, body height, body mass, or body surface area. OUES values were determined using data from 10 to 100% of the exercise test and normalized for body surface area or body mass.

3.8.2. Minute ventilation to carbon dioxide production

Minute ventilation (VE) to carbon dioxide production (VCO_2) coupling was reported in eight studies, of which four studies were performed in children (29,35,38,45) and four studies in adults. (39-41,46) VE to VCO_2 coupling was expressed in many different ways: VE/VCO_2-slope, VE/VCO_2 ratio at the VAT, the lowest VE/VCO_2 ratio during the test, or VE/VCO_2 ratio at peak exercise (see **Table 3**).

Table 3. Reference values of the included studies for cardiovascular parameters, ventilatory parameters, and ventilatory efficiency parameters.

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
Cardiovascular parameters					
VO_{2max}/VO_{2peak} (mL/kg/min)	Aadland 2016	<55	M	VO_{2max}/VO_{2peak} = 22.04 + (-0.18 × age (years)) + (-0.13 × body mass (kg)) + (2.61 × time to exhaustion (min))	?,4.46
VO_{2max}/VO_{2peak} (mL/kg/min)	Aadland 2016	>55	M	VO_{2max}/VO_{2peak} = 40.05 + (-0.27 × age (years)) + (-0.13 × body mass (kg)) + (1.49 × time to exhaustion (min))	?,3.91
VO_{2max}/VO_{2peak} (mL/kg/min)	Aadland 2016	<55	F	VO_{2max}/VO_{2peak} = 23.03 + (-0.15 × age (years)) + (0.10 × body mass (kg)) + (1.95 × time to exhaustion (min))	?,3.87
VO_{2max}/VO_{2peak} (mL/kg/min)	Aadland 2016	>55	F	VO_{2max}/VO_{2peak} = 39.67 + (-0.25 × age (years)) + (0.13 × body mass (kg)) + (1.07 × time to exhaustion (min))	?,3.19
VO_{2max}/VO_{2peak} (mL/kg/min)	Almeida 2014	8-90	M/F	VO_{2max}/VO_{2peak} = 53.478 + (-7.518 × sex) + (-0.254 × age) + (0.430 × BMI) + (6.132 × physical activity)	0.679,?
VO_{2max}/VO_{2peak} (mL/kg/min)	Kokkinos 2018	20-79	M/F	VO_{2max}/VO_{2peak} = 1.74 × (WR_{peak} × 6.12 / body mass (kg)) + 3.5	
VO_{2max}/VO_{2peak} (mL/kg/min)	Kokkinos 2018	20-79	M	VO_{2max}/VO_{2peak} = 1.76 × (WR_{peak} × 6.12 / body mass (kg)) + 3.5	
VO_{2max}/VO_{2peak} (mL/kg/min)	Kokkinos 2018	20-79	F	VO_{2max}/VO_{2peak} = 1.65 × (WR_{peak} × 6.12 / body mass (kg)) + 3.5	
VO_{2max}/VO_{2peak} (mL/kg/min)	Myers 2017	20-79	M/F	VO_{2max}/VO_{2peak} = 79.9 - (0.39 × age) - (13.7 × sex (0=male; 1=female) - (0.127 × body mass (lbs))	0.62, 7.2
VO_{2max}/VO_{2peak} (mL/min)	Blanchard 2018	12-17	M	VO_{2max}/VO_{2peak} = (-0.297 × body height^2) + (105.9 × body height) + (36.6 × corrected body mass) + (0 × age) + -8660	
VO_{2max}/VO_{2peak} (mL/min)	Blanchard 2018	12-17	F	VO_{2max}/VO_{2peak} = (-0.24 × body height^2) + (86.8 × body height) + (14.7 × corrected body mass) + (0 × age) + -6424	
VO_{2max}/VO_{2peak} (mL/min)	Blanchard 2018	12-17	M	Z-score = VO_{2peak} – [(-0.3 × body height^2) + (105.88 × body height) + (36.59 × body mass) + (-8660.14)] / (6.35 × body height) + (-717.05)	

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
VO_{2max}/VO_{2peak} (mL/min)	Blanchard 2018	12-17	F	Z-score = VO_{2peak} – [(-0.24 × body height^2) + (86.856 × body height) + (14.7 × body mass) + (-6424.42)] / (2.12 × body height) + (-45.9)	
VO_{2max}/VO_{2peak} (mL/min)	Mylius 2019	7.9-65	M	VO_{2max}/VO_{2peak} = -2537.29 + 743.35 + (24.3 × body height) + (12.57 × body mass) + (spline function for age: estimate degrees of freedom: 4.263, reference degrees of freedom 5.260)	0.57, 556.5
VO_{2max}/VO_{2peak} (mL/min)	Mylius 2019	7.9-65	F	VO_{2max}/VO_{2peak} = -2537.29 + (24.3 × body height) + (12.57 × body mass) + (spline function for age: estimate degrees of freedom: 7.391, reference degrees of freedom 8.288)	0.57, 556.5
VAT (mL/min)	Blanchard 2018	12-17	M	VAT = (-0.146 × body height^2) + (56.3 × body height) + (18.0 × corrected body mass) + (-48.3 × age) + -3898	
VAT (mL/min)	Blanchard 2018	12-17	F	VAT = (-0.00407 × body height^2) + (-2.14 × body height) + (15.9 × corrected body mass) + (-26.7 × age) + 1282	
VAT (mL/min)	Blanchard 2018	12-17	M	Z-score = VAT – [(-0.13 × body height^2) + (52.37 × body height) + (17.21 × body mass) + (-51.9 × age) + (-3565.48)] / (3.24 × body height) + (-109.49)	
VAT (mL/min)	Blanchard 2018	12-17	F	Z-score = VAT – [(-0.004 × body height^2) + (-2.14 × body height) + (15.91 × body mass) + (-26.72 × age) + (1281.8)] / (0.45 × body height) + (215.33)	
HR_{peak} (beats/min)	Ozemek 2017	18-76	M	HR_{peak} = (-0.005 × age^2) – (0.33 × age) + 205 (cross-sectional)	0.386
HR_{peak} (beats/min)	Ozemek 2017	18-76	F	HR_{peak} = (0.0002 × age^3) – (0.02 × age^2) + (0.44 × age) + 191 (cross-sectional)	0.358
HR_{peak} (beats/min)	Ozemek 2017	18-76	M/F	HR_{peak} = (0.0002 × age^3) – (0.02 × age^2) + (0.44 × age) + 211 (cross-sectional)	0.369
HR_{peak} (beats/min)	Ozemek 2017	18-76	M	HR_{peak} = -0.83 × age + 215 (longitudinal)	BIC provided
HR_{peak} (beats/min)	Ozemek 2017	18-76	F	HR_{peak} = -0.74 × age + 211 (longitudinal)	BIC provided
HR_{peak} (beats/min)	Ozemek 2017	18-76	M/F	HR_{peak} = (0.0002 × age^2) – (0.03 × age^2) + 0.84 + 185 (longitudinal)	BIC provided
HR_{peak} (beats/min)	Blanchard 2018	12-17	M	HR_{peak} = (-0.000532 × body height^2) + (0.313 × body height) + (-0.259 × corrected body mass) + (0 × age) + 169.5	

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
HR_{peak} (beats/min)	Blanchard 2018	12-17	F	HR_{peak} = (-0.0213 × body height^2) + (7.198 × body height) + (-0.193 × corrected body mass) + (-0.809 × age) + -391.1	
HR_{peak} (beats/min)	Blanchard 2018	12-17	M	Z-score = HR_{peak} – [(-0.0005 × body height^2) + (0.31 × body height) + (-0.26 × body mass) + (169.45)] / (0.1 × body height) + (-7.47)	
HR_{peak} (beats/min)	Blanchard 2018	12-17	F	Z-score = HR_{peak} – [(-0.02 × body height^2) + (7.2 × body height) + (-0.19 × body mass) + (-0.81 × age) + (-391.11)] / (-0.12 × body height) + (28.41)	
O_2-pulse$_{peak}$ (mL/beat)	Blanchard 2018	12-17	M	O_2-pulse$_{peak}$ = (-0.00131 × body height^2) + (0.459 × body height) + (0.214 × corrected body mass) + (0 × age) + -37.48	
O_2-pulse$_{peak}$ (mL/beat)	Blanchard 2018	12-17	F	O_2-pulse$_{peak}$ = (-0.00019 × body height^2) + (0.075 × body height) + (0.1007 × corrected body mass) + (0 × age) + -1.83	
O_2-pulse$_{peak}$ (mL/beat)	Blanchard 2018	12-17	M	Z-score = O_2-pulse$_{peak}$ – [(-0.001 × body height^2) + (0.41 × body height) + (0.2 × body mass) + (-0.2 × age) + (-35.14)] / (0.03 × body height) + (-2.69)	
O_2-pulse$_{peak}$ (mL/beat)	Blanchard 2018	12-17	F	Z-score = O_2-pulse$_{peak}$ – [(-0.0002 × body height^2) + (0.07 × body height) + (0.1 × body mass) + (-1.83)] / (-0.003 × body height) + (2.17)	
Blood pressure (mm Hg)	Kaafarani 2017	6.2-18.6	M/F	Normality SBP = 0.00004 × (WR_{peak}^2) – 0.00526 × (WR_{peak}) + 0.46541 Mean SBP = 0.2853 × (WR_{peak}) + 111.46	
WR_{peak} (W)	Blanchard 2018	12-17	M	WR_{peak} = (-0.0182 × body height^2) + (-5.324 × body height) + (2.824 × corrected body mass) + (4.170 × age) + 378.9	
WR_{peak} (W)	Blanchard 2018	12-17	F	WR_{peak} = (-0.06025 × body height^2) + (20.57 × body height) + (0.741 × corrected body mass) + (0 × age) + -1622	
WR_{peak} (W)	Blanchard 2018	12-17	M	Z-score = WR_{peak} – [(-0.02 × body height^2) + (-5.32 × body height) + (2.82 × body mass) + (-4.17 × age) + (378.86)] / (0.22 × body height) + (-7.62)	
WR_{peak} (W)	Blanchard 2018	12-17	F	Z-score = WR_{peak} – [(-0.06 × body height^2) + (20.57 × body height) + (0.74 × body mass) + (-1622.29)] / (-0.28 × body height) + (-24.41)	
WR_{peak} (W)	Poppe 2018	20-60	M/F	WR_{peak} = -102 + (1.5 × body mass (kg)) + (1.9 × body height (cm)) – (2.0 × age) – (sex × 60 (M:1; F:0))	0.57, 44.2
WR_{peak} (W)	Poppe 2018	20-60	M	WR_{peak} = (-0.967 × age^2) + (5.2057 × age) + 257.12	0.99

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
WR_{peak} (W)	Poppe 2018	20-60	M	$WR_{peak} = (-0.0372 \times \text{body mass}^2) + (8.0074 \times \text{body mass}) - 92.929$	0.99
WR_{peak} (W)	Poppe 2018	20-60	M	$WR_{peak} = (0.0162 \times \text{body height}) - (2.4774 \times \text{body height}) + 227$	0.99
WR_{peak} (W)	Poppe 2018	20-60	F	$WR_{peak} = (-0.0012 \times \text{age}^3) + (0.1147 \times \text{age}^2) - (4.7471 \times \text{age}) + 278.7$	0.99
WR_{peak} (W)	Poppe 2018	20-60	F	$WR_{peak} = (0.002 \times \text{body mass}^3) - (0.4715 \times \text{body mass}^2) + (38.12 \times \text{body mass}) - 818.6$	0.99
WR_{peak} (W)	Poppe 2018	20-60	F	$WR_{peak} = (-0.0642 \times \text{body height}^2) + (24.481 \times \text{body height}) - 2101.7$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	M/F	$WR_{peak} = 2.45 - (0.026 \times \text{body mass (kg)}) + (0.024 \times \text{body height (cm)}) - (0.024 \times \text{age}) - (\text{sex} \times 0.84\ (\text{M: 1; F: 0}))$	0.4,0.54
WR_{peak} (W/kg)	Poppe 2018	20-60	M	$WR_{peak} = (-0.0008 \times \text{age}^2) + (0.0247 \times \text{age}) + 3.9059$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	M	$WR_{peak} = (7E\text{-}06 \times \text{body mass}^3) + (0.0016 \times \text{body mass}^2) + (0.109 \times \text{body mass}) + 2.022$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	M	$WR_{peak} = (-4E\text{-}07 \times \text{body height}^4) + (0.0003 \times \text{body height}^3) - (0.083 \times \text{body height}^2) + (9.8777 \times \text{body height}) - 435.9$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	F	$WR_{peak} = (-0.0005 \times \text{age}^2) + (0.0139 \times \text{age}) + 3.2404$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	F	$WR_{peak} = (-0.0004 \times \text{body mass}^2) + (0.029 \times \text{body mass}) + 2.8378$	0.99
WR_{peak} (W/kg)	Poppe 2018	20-60	F	$WR_{peak} = (-0.0009 \times \text{body height}^2) + 0.31 \times \text{body height}) - 24.466$	0.99
Ventilatory parameters					

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
VE_{peak} (L/min)	Almeida 2014	8-90	M/F	VE_{peak} = 75.32 ± 15.78 (range 33.10-121.9) Tabulated data (n=2495)	SD provided
VE_{peak} (L/min)	Blanchard 2018	12-17	M	Z-score = VE_{peak} – [(-0.002 × body height²) + (-0.42 × body height) + (0.98 × body mass) + (3.17 × age) + (2.7)] / [(0.4 × body height) + (-52.54)]	
VE_{peak} (L/min)	Blanchard 2018	12-17	F	Z-score = VE_{peak} – [(-0.007 × body height²) + (2.56 × body height) + (0.53 × body mass) + (1.13 × age) + (-202.86)] / [(0.07 × body height) + (3.72)]	
VE_{peak} (L/min)	Bongers 2016	8-18	M	VE_{peak} = 80 ± 25 (range 42-157) Tabulated data (n=114)	SD provided
VE_{peak} (L/min)	Bongers 2016	8-18	F	VE_{peak} = 71 ± 21 (34-152) Tabulated data (n=100)	SD provided
VE_{peak} (L/kg/min)	Bongers 2016	8-18	M	VE_{peak} = 1.7 ± 0.3 (0.9-2.5) Tabulated data (n=114)	SD provided
VE_{peak} (L/kg/min)	Bongers 2016	8-18	F	VE_{peak} = 1.5 ± 0.3 (0.8-2.1) Tabulated data (n=100)	SD provided
VE_{peak} (L/min)	Dilber 2015	11-17	M	VE_{peak} = 89.09 ± 30.1 Tabulated data (n=99)	SD provided
VE_{peak} (L/min)	Dilber 2015	11-17	F	VE_{peak} = 67.29 ± 19.6 Tabulated data (n=65)	SD provided
VE_{peak} (L/min)	Duff 2017	10-18	M/F	VE_{peak} = 99.2 (75.6-120.0) (median + IQR) Tabulated data (n=70)	
VE_{peak} (L/min)	Kaminsky 2018	20-79	M/F	VE_{peak} = 17.32 – (28.33 × sex (M=0; F=1)) – (0.79 × age (years)) – (1.85 × body height (inches))	21.7
VE_{peak} (L/min)	Lintu 2015	9-11	M	VE_{peak} = 69.0 ± 20.0 Tabulated data (n=71)	SD provided
VE_{peak} (L/min)	Lintu 2015	9-11	F	VE_{peak} = 63.0 ± 18.0 Tabulated data (n=69)	SD provided

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
VE_{peak} (L/min)	Loe 2014	20-90	M	$VE_{peak} = 123.7 \pm 25.7$ Tabulated data per age group	SD provided
VE_{peak} (L/min)	Loe 2014	20-90	F	$VE_{peak} = 81.8 \pm 17.6$ Tabulated data per age group	SD provided
VE_{peak} (L/min)	Pistea 2016	>70	M	$VE_{peak} = 72.77 \pm 18.31$ Tabulated data (n=58)	SD provided
VE_{peak} (L/min)	Pistea 2016	>70	F	$VE_{peak} = 49.50 \pm 13.22$ Tabulated data (n=41)	SD provided
VE_{peak} (L/min)	Stensvold 2017	70-77	M	$VE_{peak} = 96.2 \pm 21.7$ Tabulated data (n=768)	SD provided
VE_{peak} (L/min)	Stensvold 2017	70-77	F	$VE_{peak} = 61.1 \pm 21.6$ Tabulated data (n=769)	SD provided
VT_{peak} (L)	Blanchard 2018	12-17	M	Z-score = VT_{peak} – [(0.00002 × body height^2) + (0.002 × body height) + (0.02 × body mass) + (0.09 × age) + (-1.22)] / [(0.004 × body height) + (-0.46)]	
VT_{peak} (L)	Blanchard 2018	12-17	F	Z-score = VT_{peak} – [(0.00005 × body height^2) + (-0.009 × body height) + (0.01 × body mass) + (0.06 × age) + (0.35)] / [(0.0008 × body height) + (0.17)]	
VT_{peak} (L)	Dilber 2015	11-17	M	$VT_{peak} = 2.22 \pm 0.6$ Tabulated data (n=99)	SD provided
VT_{peak} (L)	Dilber 2015	11-17	F	$VT_{peak} = 1.84 \pm 0.8$ Tabulated data (n=65)	SD provided
VT_{peak} (L)	Loe 2014	20-90	M	$VT_{peak} = 2.83 \pm 0.67$ Tabulated data per age group	SD provided
VT_{peak} (L)	Loe 2014	20-90	F	$VT_{peak} = 1.90 \pm 0.43$ Tabulated data per age group	SD provided
VT_{peak} (L)	Stensvold 2017	70-77	M	$VT_{peak} = 2.3 \pm 0.5$ Tabulated data (n=768)	SD provided
VT_{peak} (L)	Stensvold 2017	70-77	F	$VT_{peak} = 1.6 \pm 0.3$ Tabulated data (n=769)	SD provided

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
BF_{peak} (breaths/min)	Dilber 2015	11-17	M	$BF_{peak} = 49.64 \pm 11.7$ Tabulated data (n=99)	SD provided
BF_{peak} (breaths/min)	Dilber 2015	11-17	F	$BF_{peak} = 49.49 \pm 9.1$ Tabulated data (n=65)	SD provided
BF_{peak} (breaths/min)	Stensvold 2017	70-77	M	$BF_{peak} = 41.8 \pm 8.0$ Tabulated data (n=768)	SD provided
BF_{peak} (breaths/min)	Stensvold 2017	70-77	F	$BF_{peak} = 39.7 \pm 7.1$ Tabulated data (n=769)	SD provided
Ventilatory efficiency parameters					
OUEP	Bongers 2016	8-18	M	$OUEP = 26.34 – (0.029 \times age^2) + (1.641 \times age)$	0.9998
OUEP	Bongers 2016	8-18	F	$OUEP = 28.437 – (0.00363 \times age^2) + (1.1409 \times age)$	0.9999
OUES	Barron 2015	25-84	M	$OUES = 0.7 – (11.51 \times age) + (5.67 \times body\ height) + (8.62 \times body\ mass) – (49.99 \times beta\ blocker) – (214.53 \times current\ smoker) + (172.97 \times FEV_1)$	P5 and P95 provided
OUES	Barron 2015	25-80	F	$OUES = -182.4 – (8.89 \times age) + (10.12 \times body\ height) + (10.51 \times body\ mass) – (117.65 \times beta\ blocker) – (21.45 \times current\ smoker) + (40.31 \times FEV_1)$	P5 and P95 provided
OUES	Buys 2014	20-60	M	$OUES = 3930 – (12.5 \times age)$ $OUES = 1093 – (18.5 \times age) + (1479 \times BSA)$	
OUES	Buys 2014	20-60	F	$OUES = 3013 – (15 \times age)$ $OUES = 842 – (18.5 \times age) + (1280 \times BSA)$	
OUES	Bongers 2016	8-18	M	$OUES = 577.2 + 6.2 \times age^2 + 52 \times Age$	0.997
OUES	Bongers 2016	8-18	F	$OUES = 342.4 - 2.589 \times Age^2 \times 214.6 \times age$	0.9993

4

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
OUES (10-100)	Blanchard 2018	12-17	M	Z-score = $OUES_{10\text{-}100}$ – [(-0.24 × body height^2) + (81.44 × body height) + (38.25 × body mass) + (-6176.58)] / [(9.29 × body height) + (-1137.43)]	
OUES (10-100)	Blanchard 2018	12-17	F	Z-score = $OUES_{10\text{-}100}$ – [(-0.37 × body height^2) + (130.32 × body height) + (15.27 × body mass) + (-19.1 × age) + (-9 721.78)] / [(4.91 × body height) + (-474.83)]	
OUES/BSA	Hossri 2018	4-21	M/F	OUES/BSA LLN: 1200	
OUES/kg	Hossri 2018	4-21	M/F	OUES/kg ULN: 34.63	
OUES/kg	Bongers 2016	8-18	M	OUES/kg = 21.757 – (0.0011 × age^4) + (0.0562 × age^3) – (1.0675 × age^2) + (8.8991 × age)	0.9063
OUES/kg	Bongers 2016	8-18	F	OUES/kg = 41.3 + (0.0006 × age^4) + (0.0045 × age^3) + (0.3241 × age^2) + (1.4446 × age)	0.991
VE/VCO_2 at the VAT	Loe 2014	20-90	M	VE/VCO_2 = 26.7 ± 2.4 Tabulated data per age group	SD provided
VE/VCO_2 at the VAT	Loe 2014	20-90	F	VE/VCO_2 = 28.5 ± 3.6 Tabulated data per age group	SD provided
VE/VCO_2 at the VAT	Genberg 2016	50	M	VE/VCO_2 = 27.5 ± 2.70	SD provided
VE/VCO_2 at the VAT	Genberg 2016	50	F	VE/VCO_2 = 27.9 ± 3.24	SD provided
VE/VCO_2 minimum	Lintu 2015	9-11	M	VE/VCO_2 normal range: 24-32.9	SD provided
VE/VCO_2 minimum	Lintu 2015	9-11	F	VE/VCO_2 normal range: 25-33.8	
VE/VCO_2 peak	Pistea 2016	>70	F	VE/VCO_2 = 34.83 ± 5.66	SD provided

Variable	Reference	Age-range	Sex	Prediction equation or reference data	R^2, SEE
VE/VCO_2 peak	Pistea 2016	>70	M	$VE/VCO_2 = 34.19 \pm 4.63$	SD provided
VE/VCO_2 peak	Loe 2014	20-90	M	$VE/VCO_2 = 29 \pm 3.3$ Tabulated data per age group	SD provided
VE/VCO_2 peak	Loe 2014	20-90	F	$VE/VCO_2 = 29.3 \pm 4$ Tabulated data per age group	SD provided
VE/VCO_2 peak	Stensvold 2017	70-77	M	$VE/VCO_2 = 32.6 \pm 4.4$ (26.6 - 28.7)	P5 and P95 provided
VE/VCO_2 peak	Stensvold 2017	70-77	F	$VE/VCO_2 = 31.8 \pm 4.1$ (26.3 - 38.3)	P5 and P95 provided
VE/VCO_2-slope	Abella 2016	6-17	M/F	Data shown in graph only, no equation provided	R^2=0.336
VE/VCO_2-slope (up to the VAT)	Dilber 2015	11-17	M	VE/VCO_2-slope = 27 ± 2.9	SD provided
VE/VCO_2-slope (up to the VAT)	Dilber 2015	11-17	F	VE/VCO_2-slope = 28.16 ± 2.8	SD provided
VE/VCO_2-slope (up to the VAT)	Blanchard 2018	12-17	M	Z-score = VE/VCO_2-slope – [(-0.0004 × body height^2) + (0.24 × body height) + (-0.1 × body mass) + (-1.01 × age) + (15.1)] / [(-0.03 × body height) + (8.71)]	
VE/VCO_2-slope (up to the VAT)	Blanchard 2018	12-17	F	Z-score = VE/VCO_2-slope – [(-0.002 × body height^2) + (0.63 × body height) + (0.06 × body mass) + (-0.31 × age) + (-24.88)] / [(-0.02 × body height) + (5.8)]	

Abbreviations: BF_{peak}=breathing frequency at peak exercise; BMI-body mass index; BSA=body surface area; F=women; HR_{peak}=heart rate at peak exercise; IQR=interquartile range; LLN=lower limit of normal; M=men; O_2-$pulse_{peak}$=oxygen-pulse at peak exercise; OUEP=oxygen uptake efficiency plateau; OUES=oxygen uptake efficiency slope; SBP=systolic blood pressure; SD=standard deviation; SEE=standard error of the estimate; ULN=upper limit of normal; VAT=oxygen uptake at the ventilatory anaerobic threshold; VE_{peak}=minute ventilation at peak exercise; VE/VCO_2=minute ventilation to carbon dioxide production ratio; VE/VCO_2-slope=relation between minute ventilation and carbon dioxide production; VO_{2max}=maximal oxygen uptake; VO_{2peak}=oxygen uptake at peak exercise; VT_{peak}=tidal volume at peak exercise; WR_{peak}=work rate at peak exercise

DISCUSSION

The aim of our study was to review recently published studies in the last five years on reference values for CPET parameters in healthy children and adults. In this update of the literature, 29 studies with reference values for CPET parameters were included, in which data of 87.256 subjects (54.214 males and 33.042 females) were reported. This number is more than three times the number of subjects included in our original systematic review of the literature (25.826 subjects). (8) This increase in number shows that the sample size of the studies is increasing over time. For an adequate interpretation of CPET, the normal range of a variety of CPET parameters (e.g. VO_{2peak}, VAT, HR_{peak}, VE/VCO_2-slope) is essential. In many studies however, only the mean or median value for the population is provided. We recommend that studies should also report the lower and upper limit of normal. As shown in the study of Blanchard et al., (29) the use of the 80% of predicted as lower limit of normal should be abandoned. Instead a Z-score should be used with a lower and upper limit of normal of -1.96 SD and +1.96 SD, respectively. Moreover, authors should try to statistically model their data instead of merely providing tabulated data. In addition, authors are encouraged to publish multiple different CPET parameters in one publication, such as for example in Bongers et al. (47) This will help clinicians to select the optimal set of reference values for their tests. The use of reference values from different sources to interpretation one CPET will provide additional noise in its interpretation.

4.1 Comparison with previous review

Compared to our original review, more data from South-America are available. In the original protocol, one study in 120 adult subjects from Brazil was available. In the last five years, four new studies from Brazil and one from Argentina were added to the literature, including the study by Neto et al. (48) among 18.189 healthy subjects between 13 and 69 years of age. These studies significantly added to the available reference values for CPET in this geographic region.

Cycle ergometry was still more commonly employed as CPET method compared to treadmill ergometry. The large variety in CPET protocols, equipment, study methodology, and parameters reported indicates the need for standardization of CPET as a clinical outcome tool. Without a robust standardization of the CPET methodology, data pooling and multi-center studies are very hard to perform.

4.2 Conclusion

In the last five years, 29 studies with CPET reference values of 87.256 subjects were published. We found no single set of ideal reference values, as characteristics of each population are too diverse to pool data in a single equation for each CPET parameter. Harmonization of CPET data is still urgently needed to facilitate pooling of data from difference sources.

4.3 Expert commentary

Strength of this updated review is the inclusion of many studies from around the world with large databases. However, harmonization for CPET data is still urgently needed. Without harmonization, pooling of CPET data from difference sources is hardly possible. This is well illustrated by the various parameters used for the coupling of VE and VCO_2. Many different metrics such as the ratio of the two at the VAT, at peak, or the slope are used to describe this relationship. These different metrics give all different values and thus cannot be used interchangeably

Another limitation identified in the current review is that only a limited amount of CPET parameters are reported in the literature. An international database like the FRIEND database (49) with raw breath-by-breath data will help to report reference values for a large number of CPET parameters in a standardized manner. Using novel big data analytic methods, this database enables the continuous generation of up-to-date reference values.

The reporting of CPET reference values is still in its infancy. For instance, we recommend that in the future researchers are not only reporting the mean or median value of a population or tabulated data, but obtained data should be modeled and reference ranges including upper and lower limits of normal should be provided.

Compared to the review published in 2014, more data have been published in the last five years compared to the 35 years before. However, there is still a lot of progress to be made. Quality can be further improved by performing a power analysis, a good quality assurance of equipment and methodologies, and by validating the developed reference equation in an independent (sub)sample. Methodological quality of future studies can be further improved by measuring and reporting the level of physical activity, by reporting values for different racial groups within a cohort as well as by the exclusion of smokers in the sample studied. Normal reference ranges should be well-defined in consensus statements. For example, should we use the 5^{th} to 95^{th} percentile or the 2.5^{th} to 97.5^{th} percentile as normative range? Moreover, advanced data modeling techniques should be used. Tabulated data and simple linear regression techniques should be abandoned, since they have quite large prediction errors. For example Z-scores will provide a more qualitative analysis of the performance of a CPET parameter instead of a binary normal/abnormal.

We expect that in the near future more CPET data harmonization initiatives are undertaken to establish robust reference values for CPET. Researchers, end-users, and industry should collaborate to establish a continuous development and update of adequate reference values using an open-source database technology. This database should also include longitudinal data. Using big data techniques such as curve matching, a prediction for the future development of CPET outcomes in a subject can be made. Furthermore, we expect that open-source platforms for the interpretation and reporting of CPET data are developed for the harmonization of interpretation and reporting of CPET results.

4.4 Key issues

- There is no single set of ideal reference values; population characteristics of each population are too diverse to pool data in a single equation.
- Each exercise laboratory must select an appropriate set of reference values that best reflect the characteristics of the (patient) population tested, and equipment and methodology utilized.
- Adequate reference values provide the comparative basis for answering important questions concerning the normalcy of exercise responses in patients, and can significantly impact the clinical decision-making process.
- Researchers, end-users, and industry should collaborate to establish a continuous development and update of reference values for CPET parameters using an open source database technology. There is a growing number of geographic regions in which reference values are established: Europe, Japan, South-America, and Scandinavia were most frequently studied regions. Data from other regions such as other Asian countries, Middle East, and Africa are needed.
- Reference values for CPET parameters may change over time and should be regularly updated and/or validated.
- Standardization of the methodology to generate reference values, reporting of CPET parameters, reporting on specific software and hardware settings of the equipment, and data harmonization are necessary to facilitate interpretation and to optimize the clinical applications of CPET.

REFERENCES

1. Guide to Physical Therapist Practice. Second Edition. American Physical Therapy Association. Phys Ther. 2001 Jan;81(1):9-746.
2. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Physical therapy. 2000 Aug;80(8):782-807.
3. Mezzani A, Agostoni P, Cohen-Solal A, et al. Standards for the use of cardiopulmonary exercise testing for the functional evaluation of cardiac patients: a report from the Exercise Physiology Section of the European Association for Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil. 2009 Jun;16(3):249-67.
4. ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med. 2003 Jan 15;167(2):211-77.
5. Levett DZH, Jack S, Swart M, et al. Perioperative cardiopulmonary exercise testing (CPET): consensus clinical guidelines on indications, organization, conduct, and physiological interpretation. British journal of anaesthesia. 2018 Mar;120(3):484-500.
   * *Outstanding resource for the conduction and interpretation of CPET.*
6. Van Brussel M, Bongers BC, Hulzebos EHJ, et al. A Systematic Approach to Interpreting the Cardiopulmonary Exercise Test in Pediatrics. Pediatric exercise science. 2019 Jan 28:1-10.
7. Wasserman K, Hansen JE, Sue DY, et al. Principles of Exercise Testing and Interpretation. J Cardiopulm Rehabil Prev. 1987;7(4):189.
8. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. Expert review of cardiovascular therapy. 2014 Dec;12(12):1439-53.
9. Blais S, Berbari J, Counil FP, et al. A Systematic Review of Reference Values in Pediatric Cardiopulmonary Exercise Testing. Pediatric cardiology. 2015 Dec;36(8):1553-64.
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8(5):336-41.
11. Inbar O, Oren A, Scheinowitz M, et al. Normal cardiopulmonary responses during incremental exercise in 20- to 70-yr-old men. Med Sci Sports Exerc. 1994 May;26(5):538-46.
12. Storer TW, Davis JA, Caiozzo VJ. Accurate prediction of VO2max in cycle ergometry. Med Sci Sports Exerc. 1990;22(5):704-12.
13. Koch B, Schaper C, Ittermann T, et al. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. The European respiratory journal. 2009 Feb;33(2):389-97.
14. Ong KC, Loo CM, Ong YY, et al. Predictive values for cardiopulmonary exercise testing in sedentary Chinese adults. Respirology. 2002;7(3):225-231.
15. Akinola AB, Land JM, Mathias CJ, et al. Contribution of nitric oxide to exercise-induced hypotension in human sympathetic denervation. Clin Auton Res. 1999 Oct;9(5):263-9.
16. Jones NL, Makrides L, Hitchcock C, et al. Normal standards for an incremental progressive cycle ergometer test. Am Rev Respir Dis. 1985 May;131(5):700-8.
17. Edvardsen E, Scient C, Hansen BH, et al. Reference values for cardiorespiratory response and fitness on the treadmill in a 20- to 85-year-old population. Chest. 2013 Jul;144(1):241-8.
18. Magrani P, Pompeu FA. [Equations for predicting aerobic power (VO(2)) of young Brazilian adults]. Arq Bras Cardiol. 2010 Jun;94(6):763-70.

19. Singh R, Singh HJ, Sirisinghe RG. Cardiopulmonary fitness in a sample of Malaysian population. Jpn J Physiol. 1989;39(4):475-85.

20. Blackie SP, Fairbarn MS, McElvaney GN, et al. Prediction of maximal oxygen uptake and power during cycle ergometry in subjects older than 55 years of age. Am Rev Respir Dis. 1989;139(6):1424-9.

21. Fairbarn MS, Blackie SP, McElvaney NG, et al. Prediction of heart rate and oxygen uptake during incremental and maximal exercise in healthy adults. CHEST Journal. 1994;105(5):1365-1369.

22. Nelson MD, Petersen SR, Dlin RA. Effects of age and counseling on the cardiorespiratory response to graded exercise. Med Sci Sports Exerc. 2010 Feb;42(2):255-64.

23. Habedank D, Reindl I, Vietzke G, et al. Ventilatory efficiency and exercise tolerance in 101 healthy volunteers. Eur J Appl Physiol Occup Physiol. 1998 Apr;77(5):421-6.

24. Hollenberg M, Ngo LH, Turner D, et al. Treadmill exercise testing in an epidemiologic study of elderly subjects. J Gerontol A Biol Sci Med Sci. 1998 Jul;53(4):B259-67.

25. Itoh H, Ajisaka R, Koike A, et al. Heart rate and blood pressure response to ramp exercise and exercise capacity in relation to age, gender, and mode of exercise in a healthy population. J Cardiol. 2013 Jan;61(1):71-8.

26. John N, Thangakunam B, Devasahayam AJ, et al. Maximal oxygen uptake is lower for a healthy Indian population compared to white populations. J Cardiopulm Rehabil Prev. 2011 Sep-Oct;31(5):322-7.

27. Tammelin T, Nayha S, Rintamaki H. Cardiorespiratory fitness of males and females of northern Finland birth cohort of 1966 at age 31. Int J Sports Med. 2004 Oct;25(7):547-52.

28. Mylius CF, Krijnen WP, van der Schans CP, et al. Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry. ERJ Open Res. 2019 Apr;5(2).

29. Blanchard J, Blais S, Chetaille P, et al. New Reference Values for Cardiopulmonary Exercise Testing in Children. Medicine and science in sports and exercise. 2018 Jun;50(6):1125-1133.
** Article reports referenc values for multiple CPET paramters in children using statistical modelling.*

30. Ozemek C, Whaley MH, Finch WH, et al. Maximal heart rate declines linearly with age independent of cardiorespiratory fitness levels. European journal of sport science. 2017 Jun;17(5):563-570.

31. Kaafarani M, Schroer C, Takken T. Reference values for blood pressure response to cycle ergometry in the first two decades of life: comparison with patients with a repaired coarctation of the aorta. Expert review of cardiovascular therapy. 2017 Dec;15(12):945-951.

32. Van de Poppe DJ, Hulzebos E, Takken T, et al. Reference values for maximum work rate in apparently healthy Dutch/Flemish adults: data from the LowLands fitness registry. Acta cardiologica. 2018 Jun 22:1-8.

33. Almeida AE, Stefani Cde M, Nascimento JA, et al. An equation for the prediction of oxygen consumption in a Brazilian population. Arquivos brasileiros de cardiologia. 2014 Oct;103(4):299-307.

34. Bongers BC, Hulzebos EH, Helbing WA, et al. Response profiles of oxygen uptake efficiency during exercise in healthy children. European journal of preventive cardiology. 2016 May;23(8):865-73.

35. Dilber D, Malcić I, Čaleta T, et al. Reference values for cardiopulmonary exercise testing in children and adolescents in nortwest Croatia. Paediatria Croatica. 2015;59:195-201.

36. Duff DK, De Souza AM, Human DG, et al. A novel treadmill protocol for exercise testing in children: the British Columbia Children's Hospital protocol. BMJ open sport & exercise medicine. 2017;3(1):e000197.

37. Kaminsky LA, Harber MP, Imboden MT, et al. Peak Ventilation Reference Standards from Exercise Testing: From the FRIEND Registry. Medicine and science in sports and exercise. 2018 Dec;50(12):2603-2608.

38. Lintu N, Viitasalo A, Tompuri T, et al. Cardiorespiratory fitness, respiratory function and hemodynamic responses to maximal cycle ergometer exercise test in girls and boys aged 9-11 years: the PANIC Study. European journal of applied physiology. 2015 Feb;115(2):235-43.

39. Loe H, Steinshamn S, Wisloff U. Cardio-respiratory reference data in 4631 healthy men and women 20-90 years: the HUNT 3 fitness study. PloS one. 2014;9(11):e113884.

40. Pistea C, Lonsdorfer E, Doutreleau S, et al. Maximal aerobic capacity in ageing subjects: actual measurements versus predicted values. ERJ open research. 2016 Jan;2(1).

41. Stensvold D, Bucher Sandbakk S, Viken H, et al. Cardiorespiratory Reference Data in Older Adults: The Generation 100 Study. Medicine and science in sports and exercise. 2017 Nov;49(11):2206-2215.

42. Barron AJ, Dhutia NM, Glaser S, et al. Physiology of oxygen uptake kinetics: Insights from incremental cardiopulmonary exercise testing in the Study of Health in Pomerania. IJC metabolic & endocrine. 2015 Jun;7:3-9.

43. Buys R, Coeckelberghs E, Vanhees L, et al. The oxygen uptake efficiency slope in 1411 Caucasian healthy men and women aged 20-60 years: reference values. European journal of preventive cardiology. 2015 Mar;22(3):356-63.

44. Hossri CA, Souza IPA, de Oliveira JS, et al. Assessment of oxygen-uptake efficiency slope in healthy children and children with heart disease: Generation of appropriate reference values for the OUES variable. European journal of preventive cardiology. 2019 Jan;26(2):177-184.

45. Abella IT, Tocci AC, Iglesias DE, et al. Cardiopulmonary Exercise Testing in Healthy Children. Rev Argent Cardiol. 2016;84(5):412-417.

46. Genberg M, Andren B, Lind L, et al. Commonly used reference values underestimate oxygen uptake in healthy, 50-year-old Swedish women. Clinical physiology and functional imaging. 2018 Jan;38(1):25-33.

47. Bongers BC, Hulzebos HJ, van Brussel M, et al. Pediatric norms for cardiopulmonary exercise testing. 2nd edition ed. 's Hertogenbosch, the Netherlands: Uitgeverij BOXPress; 2014.
*** In this book, reference values for 32 different CPET parameters are provided for children and adolescents: a good example for reporting of CPET reference data.*

48. Rossi Neto JM, Tebexreni AS, Alves ANF, et al. Cardiorespiratory fitness data from 18,189 participants who underwent treadmill cardiopulmonary exercise testing in a Brazilian population. PloS one. 2019;14(1):e0209897.
*** Article reports reference values for CPET in largest population to date (18189 subjects)*

49. Kaminsky LA, Imboden MT, Arena R, et al. Reference Standards for Cardiorespiratory Fitness Measured With Cardiopulmonary Exercise Testing Using Cycle Ergometry: Data From the Fitness Registry and the Importance of Exercise National Database (FRIEND) Registry. Mayo Clinic proceedings. 2017 Feb;92(2):228-233.
** The Fitness Registry and the Importance of Exercise National Database (FRIEND) Registry is an excellent source for CPET reference values.*

APPENDIX A

Search strategy

MEDLINE: ((((((((exercise test[MeSH Terms]) OR exercise test[Title/Abstract]) OR ergometry test[Title/Abstract]) OR ergometry tests[Title/Abstract]) OR Treadmill test[Title/Abstract]) OR Treadmill tests[Title/Abstract]) OR bicycle test[Title/Abstract]) OR bicycle tests[Title/Abstract])) AND (((((((((reference values[MeSH Terms]) OR reference values[Title/Abstract]) OR normal range[Title/Abstract]) OR normal ranges[Title/Abstract]) OR norms[Title/Abstract]) OR normative value[Title/Abstract]) OR normal value[Title/Abstract]) OR normal values[Title/Abstract]) OR reference ranges[Title/Abstract]) OR reference range[Title/Abstract]).

Embase: ('exercise test':ab,ti OR 'ergometry':ab,ti OR 'exercise tests':ab,ti OR 'cardiopulmonary exercise test':ab,ti OR 'cardiopulmonary exercise tests':ab,ti OR 'cardiopulmonary exercise testing':ab,ti OR 'cycle ergometry':ab,ti OR 'incremental exercise':ab,ti) AND ('values, reference':ab,ti OR 'normal range':ab,ti OR 'normal ranges':ab,ti OR 'reference values':ab,ti OR 'reference ranges':ab,ti OR 'reference range':ab,ti OR 'normal responses':ab,ti).

PEDro: "cardiopulmonary exercise test" AND "reference values".

Appendix B

Modified methodological quality list according the ATS/ACCP guidelines

Population characteristics:

Subjects are community based. *(The subjects studied preferably be community bases rather than hospital based).*

Level of physical activity are reported.

Exclusion of different racial groups.

Exclusion of smokers in the sample studied.

No lack of definition of de confidence limits for individual or specified characteristics. *(Include age, sex, and anthropomorphic considerations).*

Sample size:

The number of subjects tested is sufficiently equal or larger than the appropriately powered sample size, with a uniform distribution of subjects for sex and groups. *(Specific attention is given to include women and older individuals, given the changing demographics and paucity of reliable population- based CPET data for these groups).*

Randomization:

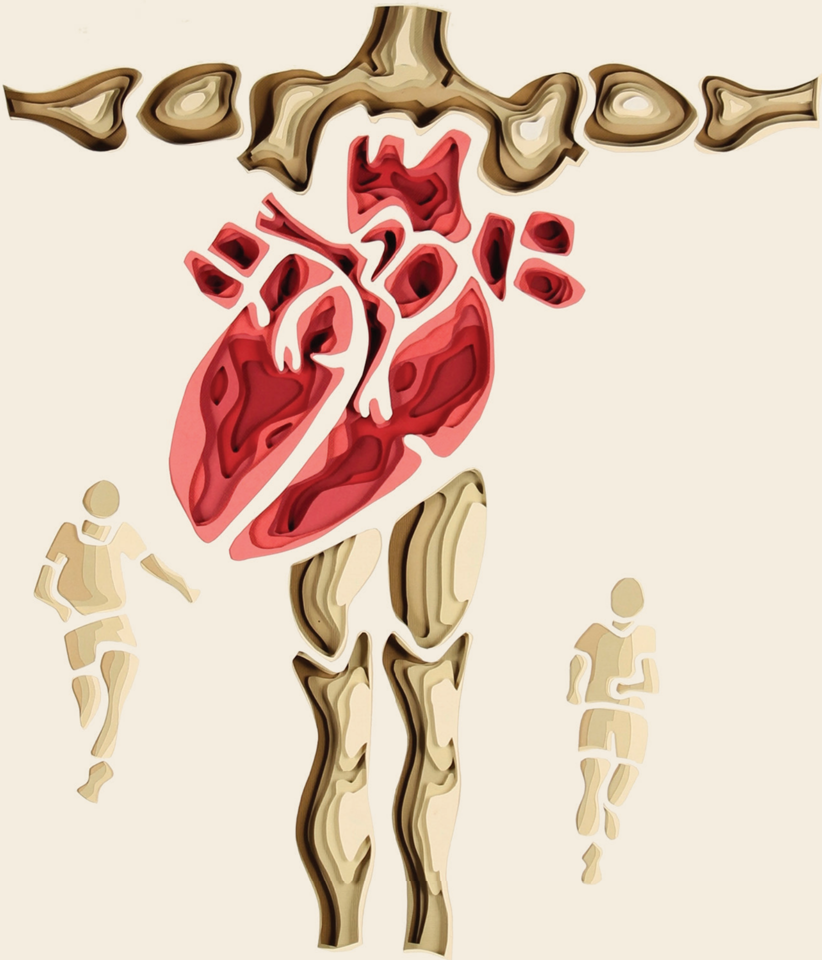
Randomization was applied.

(The study design includes a randomization process to avoid the potential bias seen when more physically active subjects volunteer for the study).

Design:

A prospective study design

Quality assurance of equipment and methodologies:
Quality control was applied.
(Quality was achieved using recommendations contained in the ATS/ACCP guidelines and the CPET protocols in accordance with recommendations specified in the ATS/ACCP guidelines).
Exercise testing protocol and procedures are described.
Results are obtained by either breath- by – breath analysis or mixing chamber treated in accordance with recommendation contained in the ATS/ACCP guidelines.
Treatment of data:
CPET result in interval averaged, preferably every 30-60 seconds (to avoid the noise of shorter interval), and the peak value reported represent the mean of the last completed stage or of all the data collected during the final stage, but preferably for no less than 30 seconds.
Validation:
Reference equations are validated in population other than those used to generate the existing data.
Statistical treatment of data:
The function that most accurately describes the distribution of the data are used. For example, curvilinear (power) functions may more accurately describe the distribution of the data. Furthermore, the precision of the individual and population predicted values are reported.



Chapter 5

Objectively measured preoperative physical activity is associated with time to functional recovery after Hepato-Pancreato-Biliary cancer surgery: a pilot study

CF Mylius, WP Krijnen, T Takken, DJ Lips, H Eker, CP van der Schans, JM Klaase

Mylius CF, Krijnen WP, Takken T, Lips DJ, Eker H, van der Schans CP, Klaase JM. Objectively measured preoperative physical activity is associated with time to functional recovery after hepato-pancreato-biliary cancer surgery: a pilot study. Perioper Med (Lond). 2021 Oct 4;10(1):33. doi: 10.1186/s13741-021-00202-7. PMID: 34602089; PMCID: PMC8489102.

ABSTRACT

Surgical resection is currently the cornerstone of hepato-pancreato-biliary (HPB) cancer treatment. A low preoperative aerobic fitness level has been identified as a modifiable risk factor associated with complications after major abdominal surgery. A person's aerobic fitness is influenced by performing moderate to vigorous physical activity (MVPA). This study aims to determine the activity monitor measured levels of MVPA performed among patients on the waiting list for HPB cancer surgery and their association with postoperative outcomes.

Methods: A prospective, observational multi-center cohort pilot study was conducted. Patients enlisted for resection surgery on suspicion of HPB (pre)malignancy were enrolled. Performed MVPA was measured by an Actigraph wGT3X-BT. Additionally, aerobic fitness was measured via the Incremental Shuttle Walk Test, and (post)operative variables were collected from the electronic patient files. The association between MVPA and the pre and postoperative variables was determined by univariate and multivariable (logistic) robust regression.

Results: A total of 38 participants, median age 66.0 (IQR 58.25 – 74.75) years, were enrolled. The median daily MVPA was 10.7 (IQR 6.9 – 18.0) minutes, only 8 participants met Dutch MVPA guidelines. Participant's age and aerobic fitness were associated with MVPA by multivariable statistical analysis. Time to functional recovery was 8 (IQR 5 - 12) days and was associated with MVPA and type of surgery (major/minor) in multivariable analysis.

Conclusion: 76% of patients enlisted for resection of HPB (pre)malignancy performed insufficient MVPA. A higher level of MVPA was associated with a shorter time to functional recovery.

Keywords: Hepato-pancreato-biliary cancer, perioperative, preoperative, physical activity, time to functional recovery

INTRODUCTION

Hepato-Pancreato-Biliary (HPB) cancer is a frequently diagnosed disease with an incidence of 248.800 patients diagnosed with HPB cancer in Europe in 2018, of which pancreatic cancer constituted the majority with 132.600 diagnoses. (1) Since advancing age of the population is the most important factor contributing to the incidence of pancreatic cancer, the incidence, and the average age of HPB cancer patients is set to increase in the coming years due to increasing life expectancy. (2,3) Surgical resection and adjuvant therapy are currently the cornerstone of treatment for HPB cancer. (4) Currently, approximately 20-30% of patients develop major postoperative complications which lead to increased length of hospital stay (LOS), decreased postoperative quality of life and delay to chemotherapy. (5–7) Since complications and mortality rates following pancreatic and liver surgery increase with advancing age, (8) identifying modifiable risk factors in HPB cancer patients may help to reduce postoperative complications, LOS, and hospital costs. (9)

Preoperative aerobic fitness level has been identified as a modifiable risk factor in a variety of patients who need surgery. (10–12) A person's aerobic fitness reflects the physiological reserve available to endure the physical stress of surgery and postoperative recovery. (13) Low preoperative aerobic fitness is associated with negative postoperative outcomes such as prolonged LOS and increase in incidence of unplanned readmissions, morbidity, and mortality after major intraabdominal surgery. (14,15) A person's aerobic fitness is influenced by his or her physical activity (PA) level. (16,17) Consequently, current (inter)national guidelines for PA advocate to spend at least 150 minutes per week in activities with a moderate to vigorous intensity (MVPA). (18,19)

Multiple studies investigated the relation between preoperative (self) reported PA levels and outcome after surgery concluding that a higher preoperative level of PA is not significantly associated with the presence of postoperative complications (OR=2.60; 95%CI=0.59 to 11.37). However, it has been previously reported that PA is significantly associated with shorter LOS following abdominal surgery (OR=3.66; 95%CI= 1.38 to 9.6). (20) Nevertheless, correlations between self reported PA and actual PA are generally low-to-moderate and ranging from R= 0.2 to 0.96. (21,22) Furthermore, previous studies have demonstrated that cancer patients overestimate their self-reported PA level when compared to objective measures. (23)

Therefore, insight into the level of actual, objectively measured, PA and subsequent postoperative outcomes in patients scheduled for HPB cancer surgery is needed. This study aims to determine the activity monitor measured levels of MVPA performed among patients on the waiting list for HPB cancer surgery. Additionally, the secondary aim of the study is to determine the association between preoperative MVPA and the association with postoperative outcomes.

METHODS

Study design and study population

This prospective, observational multi-center cohort pilot study was performed at the University Medical Centre Groningen (UMCG), the Medical Center Leeuwarden (MCL), and the Medical Spectrum Twente (MST) in the Netherlands. All centers are connected via a Managed Clinical Network HPB surgery. Ethical approval was obtained from the Central Ethics Review Committee of the UMCG under registration number 201800539, and all participants provided written informed consent. The primary objective of the study was the total of activity monitor measured MVPA performed by subjects in one week whilst awaiting HPB cancer surgery. The secondary objective was 1) the association between the subject characteristics and the performed MVPA, and 2) the associations between these parameters and the surgery outcome.

The research population consisted of adult (18 years and older) patients scheduled for resection of HPB (pre)malignancy between October 2018 and September 2019. Exclusion criteria were 1) receiving an intervention aimed at influencing PA in the pre-operative period. Performing health enhancing physical activity (e.g., fitness, jogging) on own initiative was allowed since this is part of the participants normal PA behavior; 2) receiving neo-adjuvant chemo(radio)therapy during the measurement period.

Potential participants who met the inclusion criteria were identified by the responsible surgeon directly after surgery enlistment and were invited to participate immediately after being informed about their pending surgical procedure. If eligible, potential participants received instructions on the purpose of the study and were provided an information letter. After giving informed consent, participants were visited at home to perform measurements and provide the activity monitor. After surgery, participants were treated by the Enhanced Recovery After Surgery protocol as part of the care as usual.

Data collection

The primary outcome of the study was the total of activity monitor measured MVPA performed by subjects in one week whilst awaiting HPB cancer surgery. The secondary outcomes were 1) subject characteristics, 2) Aerobic fitness, and 3) the Functional recovery.

After informed consent baseline characteristics were collected: age, height and weight, BMI (formula: weight / height^2), smoking behavior- (yes/no), occupation (work/volunteer, yes/no), living- (alone/together), education- (lower/ higher), and alcohol consumption status. Alcohol consumption was coded as above the norm or equal to/ below norm of a maximum of one consumption per day as defined by the Dutch health council. (24) Lower education was defined as (preparatory) vocational or primary education and higher education as (preparatory) academic or higher education.

Aerobic fitness was measured directly after providing informed consent. This was measured using the Incremental Shuttle Walk Test (ISWT) to determine the influence of aerobic fitness on the PA level. As an externally paced walk test, the ISWT yields greater physiological responses in comparison with self-paced walk tests. (31) The test was performed once, in accordance with the Singh protocol. (31) The maximum walking distance expressed in meters and the percentage of the predicted distance based on Probst et al., was used to determine a participant's aerobic fitness level. (32) Conventionally, the variability between healthy subjects is taken to be a standard deviation of 10%, the normal predicted range would be from 80% to 120%. Therefore, participants reaching a distance below 80% of the predicted distance, were labelled unfit. (33)

MVPA level was measured using a hip worn activity monitor, the Actigraph wGT3X- BT+ (Actigraph, Pensacola, FL, USA) was provided. (25–27) The measuring period started the day after baseline characteristics were collected and lasted 7 consecutive days. Instructions for use included performing regular PA as they were used to and wearing the device during waking hours to minimize influencing sleep quality. The used cut-off counts per activity intensity level were sedentary time (<100 counts/min), moderate- (2020–5999 counts/min) and vigorous intensity PA (≥5999 counts/min) with 100 Hz measurement epoch. (28) The total amount of MVPA is determined both as the daily median of total accumulated minutes and as the daily median minutes accumulated in at least 10-minute bouts, (19) where the latter is generally defined as a 10-minute period with an interruption of no more than 2 minutes below the threshold of 2020 counts per minute. (28) MVPA measured in 10-minute bouts was used for further analyses. To identify non-wear time, the algorithm of Choi et al. (2011) was used. (29) This algorithm defines non-wear times as periods of consecutive 0-counts for the duration of 90 minutes. A minimum of 6 measurement days or more had to be completed to be included in the analysis. Participants who wore the activity monitor for less than 6 days or did not undergo resection were excluded from analysis.

After completion of the activity tracker measurement week, the symptom burden of the past 24 hours was determined by completing a translated version of the “MD Anderson Symptom Inventory” (MDASI) questionnaire. The MDASI median scores, and the sub-domain “symptom burden” and “activity interference” scores were used to determine the participants symptom burden. (30) Median scores are used per sub-domain.

After surgery characteristics data of the surgery and outcome were collected from the electronic patient files. These included the surgery type (target organ, major/minor surgery, open/laparoscopic surgery). Major surgery was defined as any pancreatic, or liver resection of at least three liver segments. (33) Mortality was defined as in-hospital all-cause mortality or within 30 days after discharge. Overall complications consisted of all surgical and non-surgical complications within 30 days of surgery. Major complication was defined as any Clavien–Dindo grade ≥III complication. (34)

In the post-surgery phase Functional recovery was determined as the number of days between surgery and the day that adequate pain control requiring oral analgesia only was reached without signs of active (wound)infection, tolerance of solid foods and independent mobility sufficient to perform activities of daily living at the preoperative level. (35) LOS was determined at discharge and expressed in days between surgery and hospital discharge.

Statistical analyses

Statistical analyses were performed using R software version 3.6.1. (36) A p-value ≤ 0.05 was considered significant. Continuous data were summarized by median and interquartile range (IQR), categorical data by frequency and percentage. Range was reported if deemed relevant.

MVPA data are frequently non-normally distributed due to outlying observations for a few persons having PA levels away from the bulk of the data. Therefore, a robust regression approach was undertaken throughout this study. Robust regression is a regression method suitable for non-normally distributed data with outliers, this method prevents a large influence on the association coefficients by outlying observations. (37–39) All enlisted patients, recording 6 or more measurement days, were used in the MVPA analyses, participants only receiving an exploratory laparotomy or laparoscopy without resection and those who were eventually not operated upon were excluded from the complication's analyses. Furthermore, all participants that reached discharge from hospital were included in the time to FR analysis.

The association between the level of MVPA in 10-minute bouts and the preoperative variables, and time to FR and pre- and peri-operative variables was determined by univariate and multivariable robust regression. (40) Furthermore, univariate, and multivariable robust logistic regression was used to determine the Odds Ratio (OR) of the occurrence of complications based on the preoperative and per- operative variables. (38–40) All multivariable analyses were performed using the measured independent explanatory variables identified to potentially have a significant association with the dependent variable from univariate regression analysis. Lastly, a subset analysis was performed to determine the association between MVPA in 10-minute bouts and time to FR within the major complications group via univariate robust regression. LOS analysis is reported in the supplementary material.

RESULTS

A total of 154 patients who met the inclusion criteria were approached for participation, 40 patients (26%) consented to participate in the study. Two participants were excluded from PA analysis due to not meeting wear-time criteria, the measurements from the remaining 38 participants were used for further analysis. Five participants had either no surgery procedure (one participant) or received a procedure without resection (exploratory laparotomy only, four participants). These patients were excluded from complications analyses. Furthermore, two participants were excluded from the time to FR analyses due to postoperative mortality. **Figure 1** displays the flowchart of participant inclusion. Median time between placement on surgery awaiting list until baseline measurements and between baseline measurements, including the start of the activity monitor period, until surgery day was 0 days (IQR 0 – 0.75, range 0 - 14) and 31.5 days (IQR 22.25 - 45, range 9 -171) respectively.

Characteristics

Of the 38 participants, 22 participants were male, and the mean age of participants in both the PA and surgery outcome group were 65.8 years (±9.4) and 65.5 years (±9.8), respectively. The label unfit was given to 22 participants, with a median 69% (±31%) distance covered of the predicted ISWT distance in the PA group and 65% (±28%) in the surgery outcome group. Of the 33 participants that underwent the surgery procedure, 10 developed major complications. Participant characteristics and perioperative data are presented in **Table 1**.

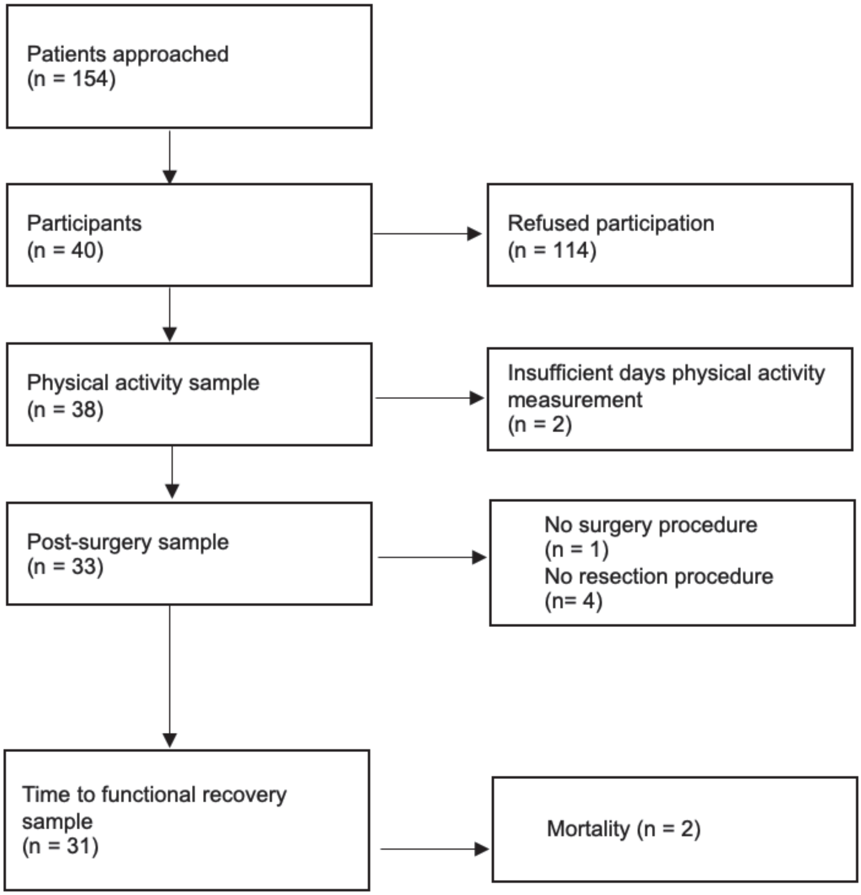


Figure 1. Inclusion flowchart

Physical activity

The participants median level of MVPA was 10.7 minutes per day, wearing the activity monitor 66% (±29%) of waking hours per day. The MVPA variability between participants was large, ranging from zero to 60.1 minutes per day. Eight participants (21%) met the PA guideline of 150 minutes MVPA per week. The level of MVPA reduced with 0.52 minutes per advancing age year, (R^2=.31, p=.001), and increased by .02 minutes per meter covered during the ISWT (R^2 = .35, p = .008), and subjects labeled as fit (7.90 minutes more in fit subjects, R^2= .20, p=.023) were identified as correlating with MVPA via univariate robust regression. Since the aerobic fitness level was derived from the ISWT distance covered, this variable was omitted from multivariable regression. The multivariable regression model for performed MVPA determined by multivariable robust regression was 29.05 + (ISWT $^{(meters)}$ * 0.01) + (Age $^{(years)}$ * -0.35) (adj. R^2= .41). The association between MVPA and preoperative variables via univariate and multivariable robust regression is displayed in **Table 2**.

Table 1 Patient characteristics

Variable	Physical activity sample (N=38) N (%) or Median (IQR)	Surgery outcome sample (N=33) N (%) or Median (IQR)
Gender		
Female	16 (42%)	14 (42%)
Male	22 (58%)	19 (58%)
Age (years)	66 (58.25 – 74.75)	66 (56 – 74)
Height (cm)	173 (167.8 – 182.8)	173 (167 – 182)
Weight (kg)	77 (70.1 – 87.6)	77 (70.5 - 88)
BMI (kg / cm^2)	24.9 (22.8 – 28.2)	24.8 (22.7 – 28.3)
Living situation		
Living alone	8 (21%)	7 (21%)
Living together	30 (79%)	26 (79%)
Education level		
Lower education	23 (61%)	19 (58%)
Higher education	15 (39%)	14 (42%)
Work		
Employed	14 (37%)	12 (36%)
Unemployed	24 (63%)	21 (64%)
Alcohol consumption		
Above norm	10 (26%)	9 (27%)
Equal to or below norm	28 (74%)	24 (73%)
Smoker	5 (13%)	5 (15%)
Non smoker	33 (87%)	28 (85%)
MDASI total (sum score)	1.87 (±1.75) (N=32)	1.86 (±1.85) (N=28)
Symptoms	1.81 (±1.68) (N=32)	1.84 (±1.81) (N=28)
Activity	1.91 (±2.29) (N=32)	1.92 (±2.30) (N=28)
ISWT (meters)	430 (310 – 473.1)	430 (280 - 620)
Percentage of predicted (%)	69 (±31)	65 (±28)
Labelled fit / unfit	16 / 22 (42% / 58%)	14 / 19 (42% / 58%)

Variable	Physical activity sample (N=38) N (%) or Median (IQR)	Surgery outcome sample (N=33) N (%) or Median (IQR)
Physical activity (minutes)		
Time spend sedentary (per day)	564.4.5 (310.4 – 662.4)	580.4 (417.0 – 668.3)
Time spend sedentary (per week)	3951.1 (2172.8 – 4636.6)	4062.8 (2918.9 – 4678.5)
MVPA (total accumulated per day)	26.4 (16.8 – 43.8)	24.4 (16.6 – 34.5)
MVPA (total accumulated per week)	184.8 (117.7 – 306.9)	170.6 (116.1 – 241.2)
Adherence to guideline (yes / no) *	21 / 17 (55% / 45%)	17 / 16 (48% / 52%)
MVPA (total 10-minute bouts per day)	10.7 (6.9 – 18.0)	11.6 (7.5 – 21.3)
MVPA (total 10-minute bouts per week)	74.7 (48.3 – 125.94)	81 (52.6 – 148.9)
Adherence to guideline (yes / no) **	8 / 30 (21% / 79%)	8 / 25 (24% / 76%)
Wear time in percentage	66.6% (±27.6)	69.2% (±26.7)
Highest Clavien-Dindo rating		
No complication		16 (49%)
Minor / Major		7/10 (21%/30%)
Grade I		2 (6%)
Grade II		5 (15%)
Grade IIIa		3 (9%)
Grade IIIb		5 (15%)
Grade IV		0 (0%)
Grade V		2 (6%)
Target organ		
Pancreas		13 (39%)
Liver		20 (61%)
Open procedure		24 (72%)
Laparoscopic procedure		9 (28%)
Major procedure		18 (54%)
Minor procedure		15 (46%)
Length of hospital stay (days)		9 (7 - 15)
Time to functional recovery (days)		8 (5 - 12)
Mortality		2 (6%)

Abbreviations: BMI = Body Mass Index, MDASI = MD Anderson Symptom Inventory, ISWT = Incremental Shuttle Walk Test, PA = Physical Activity, MVPA = Moderate to Vigorous Physical Activity, avg. = average, * = 150 minutes per week accumulated total bouts, ** = 150 minutes per week accumulated 10-minute bouts

Complications

Seventeen participants (51%) had complications of which ten (30%) were major. The association found between MVPA and the presence of major complications (OR = 0.99, 95%CI= 0.95 – 1.04, p= .703) was not statistically significant. A statistically significant association was found between the presence of major complications and BMI (OR = .71, 95%CI= 0.52 – 0.98, p= .036), % of predicted ISWT (OR= .98, 95%CI .97 – .99, p=.008) and surgery type (OR = .24, 95%CI = 0.06 – 0.95, p= .043). The odds of major complications decrease with increasing BMI, more distance covered on the ISWT compared to the predicated distance and a minor surgery procedure. The OR from multivariable robust logistic regression including surgery type and ISWT $^{(\% \text{ of predicted})}$ was found to be: (surgery type $^{(minor)}$ * 0.144) + (ISWT $^{(\% \text{ of predicted})}$ * 0.948). The OR from robust univariate and multivariable logistic regression for the occurrence of major complications are displayed in **Table 3**.

Table 2. Uni- and multivariable robust regression association between preoperative variables and moderate to vigorous physical activity (10-minute bouts)

Variable	Estimate	Std. error	R^2	t-value	p-value	Adj. R^2
Gender (female)	-2.140	2.867	0.016	-0.746	.460	
Age (years)	-0.521	0.146	0.309	-3.565	.001*	
BMI (kg/cm^2)	0.358	0.642	0.0332	0.557	.581	
Living situation (together)	-0.999	4.335	0.002	-0.230	.819	
Work status (employed)	5.491	3.686	0.092	1.490	.145	
Education (high)	5.045	3.153	0.083	1.600	.118	
Smoking status (no)	1.263	2.771	0.002	0.456	.651	
Alcohol norm (above)	5.935	3.415	0.101	1.738	.091	
MDASI total (avg, n=32)	-1.058	0.848	0.045	-1.247	.222	
MDASI symptoms (avg, n=32)	-0.884	0.823	0.031	-1.074	.291	
MDASI activities (avg, n=32)	-0.865	0.579	0.050	-1.496	.145	
ISWT (meters)	0.020	0.007	0.346	2.796	.008*	
ISWT (% of predicted)	0.091	0.050	0.112	1.811	.078	
Aerobic fitness (labelled fit)	7.905	3.325	0.204	2.378	.023*	
Multivariable						
Constant	29.048	11.107		2.615	.013*	.414
Age (years)	-0.348	0.135		-2.576	.014*	
ISWT (meters)	0.013	0.006		2.238	.031*	

Abbreviations: BMI = Body Mass Index, ISWT = Incremental Shuttle Walk Test, MDASI = MD Anderson Symptom Inventory, * = P ≤.05, avg. = average

5

Table 3. Uni- and multivariable logistic robust regression association between pre- and per- operative variables and complications Clavien-dindo grade ≥III

Variable	OR	95% CI	p-value
Gender (female)	0.952	0.226 – 4.006	.947
Age (years)	1.017	0.944 – 1.096	.650
BMI (kg/cm^2)	0.714	0.521 – 0.979	.036*
Living situation (living together)	0.706	0.129 – 3.868	.688
Working status (works)	1.250	0.507 – 3.081	.627
Education level (high)	1.900	0.759 – 4.759	.171
Alcohol consumption (above norm)	0.533	0.142 – 1.996	.351
Smoking status (no)	0.893	0.273 – 2.914	.851
MDASI total (avg, n=28)	0.980	0.728 – 1.319	.893
MDASI symptom (avg, n=28)	0.937	0.680 – 1.290	.688
MDASI activity (avg, n=28)	1.016	0.827 – 1.249	.879
ISWT (meters)	0.998	0.996 – 1.001	.066
ISWT (% of predicted)	0.981	0.967 – 0.995	.008*
Aerobic fitness norm (labelled fit)	0.875	0.332 – 2.304	.787
Time spend sedentary (minutes)	0.999	0.997 – 1.001	.321
Daily MVPA - total accumulated (minutes)	0.996	0.976 – 1.017	.745
Daily MVPA - 10-minute bouts (minutes)	0.991	0.949 – 1.036	.703
Laparoscopic / closed surgery (Laparoscopic)	0.242	0.034 – 1.720	.156
Major / minor surgery (minor)	0.240	0.060 – 0.953	.043*
Multivariate			
Major / minor surgery (minor)	0.144	0.024 – 0.858	.033*
ISWT (% of predicted)	0.984	0.954 – 1.015	.309

Abbreviations: OR = Odds Ratio, 95%CI = 95% Confidence Interval, BMI = Body Mass Index, avg. = average, ISWT= Incremental Shuttle Walk Test, MDASI = MD Anderson Symptom Inventory, MVPA = Moderate to Vigorous Physical Activity, * = P≤.05

Time to functional recovery

The median time to FR was 8 (IQR 5 - 12) days, ranging from 2 till 56 days. Higher MVPA in both total accumulated bouts (-0.07 less days per minute increase, p= .009), and 10-minute bouts (-0.14 less days per minute increase, p=.007), a minor surgery procedure (-6.39 less days, p=<.001), and a higher BMI (-0.46 less days per kg/cm^2 increase, p=.006) resulted in less time to FR. The multivariable model yields an adj. R^2 .43, the model is as follows 12.54 + (MVPA $^{(minutes)}$ * -.08) + (surgery size $^{(1\ if\ minor,\ 0\ if\ major)}$ * -5.64). The association between MVPA in 10-minute bouts and time to FR in the subset where major complications occurred was -0.352 less days to FR per minute increase (R^2=.460, p=.023). Time to FR analysis is displayed in **Table 4**.

Table 4 Uni- and multivariable robust Cox proportional hazard association with time to functional recovery

Variable	Estimate	Std. error	R^2	t-value	p-value	Adj. R^2
Gender (female)	0.754	2.206	0.006	0.342	.735	
Age (years)	0.140	0.097	0.097	1.449	.158	
BMI (kg/cm²)	-0.463	0.156	0.206	-2.974	.006*	
Living situation (together)	0.902	1.979	0.005	0.456	.652	
Working status (employed)	-0.655	1.903	0.004	-0.344	.733	
Education level (high)	0.365	2.343	0.001	0.156	.877	
Alcohol norm (above)	-1.255	2.061	0.014	-0.609	.547	
Smoking status (no)	2.942	2.882	0.044	1.021	.316	
MDASI total (avg., n=27)	0.269	0.390	0.011	0.691	.496	
MDASI symptoms (avg, n=27)	0.166	0.399	0.004	0.417	.680	
MDASI activities (avg, n=27)	0.425	0.379	0.031	1.124	.272	
ISWT (meters)	-0.007	0.004	0.098	-1.679	.104	
ISWT (% of predicted)	-0.054	0.036	0.076	-1.508	.142	
Aerobic fitness norm (labelled fit)	-2,172	1.862	0.044	-1.167	.253	
Time spend sedentary (minutes)	-0.006	0.004	0.056	-1.345	.189	
Daily MVPA - total accumulated (minutes)	-0.068	0.024	0.157	-2.793	.009*	
Daily MVPA - 10-minute bouts (minutes)	-0.145	0.050	0.174	-2.905	.007*	
Laparoscopic / open surgery (Laparoscopic)	-1.555	2.045	0.021	-0.761	.453	
Major / minor surgery (minor)	-6.392	1.638	0.426	-3.902	<.001*	
Multivariate						
Constant	12.545	2.062		6.083	<.001*	.432
Major / minor surgery (minor)	-5.643	1.803		-3.130	.004*	
Daily MVPA - 10-minute bouts (minutes)	-0.079	0.031		-2.573	.016*	
Major complications subset analysis						
Daily MVPA – 10-minute bouts (minutes)	-0.352	0.126	0.460	-2.798	.023*	

Abbreviations: BMI = Body Mass Index, ISWT = Incremental Shuttle Walk Test, MDASI = MD Anderson Symptom Inventory, MVPA = Moderate to Vigorous Physical Activity, * = P≤.05

DISCUSSION

To our knowledge, this is the first study investigating device measured MVPA levels in HPB resection candidates not receiving PA interventions. Patients scheduled for HPB surgery engage in low daily MVPA at baseline whilst waiting for surgery. Furthermore, a relation was found between the level of MVPA and time to FR after HPB surgery for (pre) malignancy; patients with higher levels of PA require less time to FR. The current findings suggest that increasing a patient's pre-operative MVPA level might be an intervention to improve the postsurgical outcome.

Physical activity

The median MVPA level measured in the current study was low but comparable to other preoperative activity monitor measured MVPA studies, e.g., gastric bypass and lumbar fusion surgery. (41,42) However, this comparison is somewhat arbitrary due to the influence of age, and the variety in symptom burden experienced amongst different pathologies. Furthermore, the variety in activity monitor device configuration like MVPA cut-off point and wear-time validation highly influences the results. (43) Nevertheless, this study demonstrates that the majority (79%) of the patients, enlisted for HPB surgery did not perform sufficient MVPA to meet the guideline of 150 minutes MVPA per week. (18,19)

These findings might be explained by the psychological impact of being enlisted for surgery because of malignancy. Namely, being informed about the presence of a tumor can result in changes in PA behavior. (44) Participants were recruited directly after being enlisted and measurements were performed during the first week after enlistment. Due to the design of the study, it remains unclear whether this effect is temporarily, the minutes of MVPA, and the performed distance covered on the ISWT might reach higher levels over time. Previous studies have reported an increase in PA during the waiting period. (45) The observed increase might have been caused by an increased awareness or social desirability of the participant, as they had to fill out PA questionnaires, wore a PA monitor, or had to perform physical fitness measures during this study. Furthermore, it seems likely that patients perform less MVPA due to the interference of tumor related symptoms. However, there was no evidence for an association between the experienced symptom burden like pain and fatigue, measured with the MDASI, and the level of performed MVPA. Notably, participants experienced fairly low symptom interference in our study, 1.87 points on mean out of 10. It therefore seems probable that subjects with high symptom interference were more likely to reject study participation. Due to the small sample size, no subcategory analysis with subjects experiencing high levels of symptom burden could be performed.

Post-operative outcomes

A significant association was found between MVPA and time to FR (R^2= 0.17, p= .006) but no significant association was found between MVPA and the occurrence of postoperative complications (OR = 0.99, 95%CI= 0.95 – 1.03, p= .67). These findings are in accordance with the systematic review and meta-analysis in preoperative cancer patients by Steffens et al., who found an association between higher levels of preoperative MVPA and a shorter absolute LOS (OR=3.66; 95%CI= 1.38 to 9.6), but not with postoperative complications (OR=2.60; 95%CI=0.59 to 11.37). The majority of studies in this meta-analysis used self-reported MVPA and participants undergoing neo-adjuvant (physical) therapy. (20) However, the meta-analysis as well as the current study consistently indicate that higher levels of MVPA positively influences a patient's capacity to endure the demands of surgery. (20)

A subject's level of preoperative MVPA was associated with reduced time to FR, 43% of the time variance to FR could be explained via multivariable robust regression including surgery size and MVPA levels. This reduction might be explained by the lower relative capacity needed to perform activities in daily living by patients with higher levels of aerobic fitness. FR is determined by both functional and physiological criteria, that is, higher levels of aerobic fitness increase a patient's functional capacity to perform activities of daily living. (46) However, caution is needed when interpreting these results since we did not directly measure aerobic reserves at the moment of FR.

Furthermore, a higher percentage of predicted distance covered on the ISWT was associated with reduced OR for the occurrence of major postoperative complications found by univariate robust regression. Similar reductions have been reported in multiple studies amongst a large variety of surgical procedures. (7,10–13) These reductions might be explained by the higher aerobic reserves enhancing the bodies capacity to cope with the responses to the surgical procedure. Nevertheless, a higher percentage of predicted distance covered on the ISWT was not found to have a significant association in multivariable robust regression including surgery size. Notably, the current study found lower OR for the occurrence of major complications in subjects with a higher BMI. This result is inconsistent to previous studies showing increased OR for the development of major complications in obese and overweight subjects undergoing pancreatectomy procedures. (47) This difference might be explained by the overrepresentation of subjects with high BMI scores undergoing a major surgery procedure in the present study (Wilcoxon rank sum test, W = 74, p= .02). Major surgery has a higher risk of resulting in major complications. Therefore, BMI was removed from multivariable regression analysis in the current study. Additionally, we found a reduction in time to FR after major complications in subjects performing higher levels of PA. Therefore, it could be concluded that because subjects with a higher level of MVPA have more capacity to cope with the demands endured by complications, the impact of complications is less. Nevertheless, these results should be interpreted with some caution since only nine subjects reached FR after major complications.

Treatment opportunity

This study identifies preoperative MVPA as a modifiable patient factor to reduce time to FR. Multiple associations between performed MVPA and preoperative variables were found, namely MVPA decreased with advancing age with 0.52 minutes per age year (p=<.001), and increased in participants with higher aerobic fitness, covering more distance during the ISWT (0.02 minutes per meter, p=.008). Since both PA and aerobic fitness declines with age, (48) these findings underpin the hypothesis that unfit and older patients could benefit most from interventions aimed to improve aerobic fitness and to increase MVPA levels, especially in the waiting time before surgery. Furthermore, although this study does not include a detailed cost analysis, increasing the level of preoperative MVPA via relative low-cost treatment modalities as education, wearables, and physiotherapy, may be of particular relevance for the reduction of hospital costs due to the shorter hospital stay (see supplementary data). (9)

Limitations

There are some limitations to this observational pilot study. The first is that no pre-trail sample-size calculation was performed. This can impact the results with a higher risk of type II errors. Since a limited number of HPB resections are yearly performed, participants were included via convenience sampling in a multi-center design. However, the final sample size obtained is comparable with several other studies aimed at measuring PA via activity monitor devices in major abdominal surgery. (49,50) Unfortunately, only 26% of the approached subjects provided consent to participate in the study. A reason for this low participation rate might have been the moment of inclusion, namely directly after being enlisted for surgery. Frequently mentioned reasons for declining participation were the feeling of being emotionally overwhelmed and currently not having the energy to endorse participation. These reasons might have induced a sample slightly biased in the direction of somewhat fitter patients. Larger sample sizes and less strenuous PA measurements can be more easily acquired via questionnaires. Nevertheless, activity monitor measured PA is a feasible and more reliable method of determining PA and is therefore recommended. (20,41)

Conclusion

This study demonstrates that 79% of the patients, enlisted for resection of HPB (pre) malignancy performed insufficient MVPA. A higher level of MVPA, objectively measured with an activity monitor was independently associated with a shorter time to FR. However, levels of MVPA were not associated with postoperative complications. Stimulating MVPA in the waiting time for surgery might help to reduce the LOS. These findings add to a growing body of evidence suggesting that higher levels of MVPA positively influence a patient's capacity to endure the demands of surgery and improve the outcome of surgery.

Acknowledgements

The authors would like to acknowledge the invaluable contribution of the research nurses at the University Medical Hospital Groningen, Medical Centre Leeuwarden, the Medical Spectrum Twente, Rebekka van der Laan, Haike Verbree and Sander Koops.

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer [Internet]. 2018;103:356–87. Available from: https://doi.org/10.1016/j.ejca.2018.07.005
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
3. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. Lancet [Internet]. 2017;389(10076):1323–35. Available from: http://dx.doi.org/10.1016/S0140-6736(16)32381-9
4. Kommalapati A, Tella SH, Goyal G, Ma WW, Mahipal A. Contemporary management of localized resectable pancreatic cancer. Cancers (Basel). 2018;10(1):1–15.
5. Pinto A, Faiz O, Davis R, Almoudaris A, Vincent C. Surgical complications and their impact on patients' psychosocial well-being: A systematic review and meta-analysis. BMJ Open. 2016;6(2).
6. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, et al. Mortality after surgery in Europe: A 7 day cohort study. Lancet [Internet]. 2012;380(9847):1059–65. Available from: http://dx.doi.org/10.1016/S0140-6736(12)61148-9
7. Kumar R, Garcea G. Cardiopulmonary exercise testing in hepato-biliary & pancreas cancer surgery – A systematic review: Are we any further than walking up a flight of stairs? Int J Surg [Internet]. 2018;52:201–7. Available from: https://doi.org/10.1016/j.ijsu.2018.02.019
8. Raill T. What Is the Effect of Age on Pancreatic Resection? Adv Surg [Internet]. 2009;43:233–49. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf
9. Straatman J, Cuesta MA, De Lange-De Klerk ESM, Van Der Peet DL. Hospital cost-analysis of complications after major abdominal surgery. Dig Surg. 2015;32(2):150–6.
10. Snowden CP, Prentis JM, Anderson HL, Roberts DR, Randles D, Renton M, et al. Submaximal cardiopulmonary exercise testing predicts complications and hospital length of stay in patients undergoing major elective surgery. Ann Surg. 2010;251(3):535–41.
11. Simões CM, Carmona MJC, Hajjar LA, Vincent JL, Landoni G, Belletti A, et al. Predictors of major complications after elective abdominal surgery in cancer patients. BMC Anesthesiol. 2018;18(1):1–8.
12. Van Beijsterveld CA, Bongers BC, Den Dulk M, Van Kuijk SMJ, Dejong KCH, Van Meeteren NLU. The association between preoperative physical functioning and short-term postoperative outcomes: a cohort study of patients undergoing elective hepatic resection. HPB [Internet]. 2019;21(10):1362–70. Available from: https://doi.org/10.1016/j.hpb.2019.02.009
13. Levett DZH, Grocott MPW. Cardiopulmonary exercise testing , prehabilitation , and Enhanced Recovery After Surgery (ERAS). Can J Anaesth. 2015;62(2):131–42.
14. Moran J, Wilson F, Guinan E, McCormick P, Hussey J, Moriarty J. Role of cardiopulmonary exercise testing as a risk-assessment method in patients undergoing intra-abdominal surgery: A systematic review. Br J Anaesth. 2016;116(2):177–91.
15. Chandrabalan V V., McMillan DC, Carter R, Kinsella J, McKay CJ, Carter CR, et al. Pre-operative cardiopulmonary exercise testing predicts adverse post-operative events and non-progression to adjuvant therapy after major pancreatic surgery. HPB. 2013;15(11):899–907.

16. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, et al. Global physical activity levels: Surveillance progress, pitfalls, and prospects. Lancet [Internet]. 2012;380(9838):247–57. Available from: http://dx.doi.org/10.1016/S0140-6736(12)60646-1

17. Chastin SFM, De Craemer M, De Cocker K, Powell L, Van Cauwenberg J, Dall P, et al. How does light-intensity physical activity associate with adult cardiometabolic health and mortality? Systematic review with meta-analysis of experimental and observational studies. Br J Sports Med. 2018;bjsports-2017-097563.

18. Weggemans RM, Backx FJG, Borghouts L, Chinapaw M, Hopman MTE, Koster A, et al. The 2017 Dutch Physical Activity Guidelines. Int J Behav Nutr Phys Act. 2018;15(1):1–12.

19. WHO. Global health risks – Mortality and burden of disease attributable to selected major risks. [Internet]. 2009. Available from: http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf

20. Steffens D, Beckenkamp PR, Young J, Solomon M, da Silva TM, Hancock MJ. Is preoperative physical activity level of patients undergoing cancer surgery associated with postoperative outcomes? A systematic review and meta-analysis. Eur J Surg Oncol [Internet]. 2019;45(4):510–8. Available from: https://doi.org/10.1016/j.ejso.2018.10.063

21. Colley RC, Butler G, Garriguet D, Prince SA, Roberts KC. Comparison of self-reported and accelerometer-measured physical activity among Canadian youth. Heal Reports. 2019;30(7):3–12.

22. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. Int J Behav Nutr Phys Act. 2008;5.

23. Vassbakk-Brovold K, Kersten C, Fegran L, Mjåland O, Mjåland S, Seiler S, et al. Cancer patients participating in a lifestyle intervention during chemotherapy greatly over-report their physical activity level: A validation study. BMC Sports Sci Med Rehabil [Internet]. 2016;8(1):1–9. Available from: http://dx.doi.org/10.1186/s13102-016-0035-z

24. Health Council of the Netherlands. Richtlijnen goede voeding 2015. Gezondheidsraad [Internet]. 2015;24:1–95. Available from: https://www.gezondheidsraad.nl/sites/default/files/201524_richtlijnen_goede_voeding_2015.pdf

25. Knaier R, Höchsmann C, Infanger D, Hinrichs T, Schmidt-Trucksäss A. Validation of automatic wear-time detection algorithms in a free-living setting of wrist-worn and hip-worn ActiGraph GT3X+. BMC Public Health. 2019;19(1):1–6.

26. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport [Internet]. 2011;14(5):411–6. Available from: http://dx.doi.org/10.1016/j.jsams.2011.04.003

27. Moran J, Wilson F, Guinan E, McCormick P, Hussey J, Moriarty J. The preoperative use of field tests of exercise tolerance to predict postoperative outcome in intra-abdominal surgery: a systematic review. J Clin Anesth [Internet]. 2016;35:446–55. Available from: http://dx.doi.org/10.1016/j.jclinane.2016.09.019

28. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, Mcdowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40(1):181–8.

29. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time. Med Sci Sport Exerc. 2011;43(2):357–64.

30. Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle MT, Morrissey M, et al. Assessing symptom distress in cancer patients: The M.D. Anderson Symptom Inventory. Cancer [Internet]. 2000;89(7):1634–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11013380

31. Singh S, Morgan M, Scott S, Walters D, Hardman A. Development of a shuttle walking test of disability in patients with chronic airflow obstruction. Thorax, 47. 1992;1019–24.

32. Probst VS, Hernandes NA, Teixeira DC, Felcar JM, Mesquita RB, Gonalves CG, et al. Reference values for the incremental shuttle walking test. Respir Med. 2012;106(2):243–8.

33. Stanojevic S, Wade A, Stocks J. Reference values for lung function: Past, present and future. Eur Respir J. 2010;36(1):12–9.

34. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240(2):205–13.

35. Wong-Lun-Hing EM, van Dam RM, van Breukelen GJP, Tanis PJ, Ratti F, van Hillegersberg R, et al. Randomized clinical trial of open versus laparoscopic left lateral hepatic sectionectomy within an enhanced recovery after surgery programme (ORANGE II study). Br J Surg. 2017;104(5):525–35.

36. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Internet]. Vienna, Austra; 2013. Available from: http://www.r-project.org

37. Yohai V, Stabel W, Zamar R. A Procedure for Robust Estimation and Inference in Linear Regression. Directions in RObust Statistics and Diagnostics Part II. 1991. 365–374 p.

38. Koller M, Stabel W. Sharpening Wald-type Inference in RObust Regression for Small Samples. Comput Stat Data Anal. 2011;55(8):2504–25115.

39. Hubert M, Rousseeuw PJ. Robust regression with both continuous and binary regressors. J Stat Plan Inference. 1997;57(1):153–63.

40. Portney G, Watkins P. Foundations of Clinical Research: Applications to Practice. New Jersey Pearson Educ. 2009;892.

41. Van der Meij E, Van der Ploeg HP, Van Den Heuvel B, Dwars BJ, Meijerink WJHJ, Bonjer HJ, et al. Assessing pre- and postoperative activity levels with an accelerometer: a proof of concept study. BMC Surg. 2017;17(1):1–10.

42. Lotzke H, Jakobsson M, Gutke A, Hagströmer M, Brisby H, Hägg O, et al. Patients with severe low back pain exhibit a low level of physical activity before lumbar fusion surgery: A cross-sectional study. BMC Musculoskelet Disord. 2018;19(1):1–9.

43. Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: A systematic review and data analysis. Eur Rev Aging Phys Act. 2014;11(1):35–49.

44. Allender S, Hutchinson L, Foster C. Life-change events and participation in physical activity: A systematic review. Health Promot Int. 2008;23(2):160–72.

45. Kim DJ, Mayo NE, Carli F, Montgomery DL, Zavorsky GS. Responsive measures to prehabilitation in patients undergoing bowel resection surgery. Tohoku J Exp Med. 2009;217(2):109–15.

46. Jackson AS, Sui X, Hébert JR, Church TS, Blair SN. Role of lifestyle and aging on the longitudinal change in cardiorespiratory fitness. Arch Intern Med. 2009;169(19):1781–7.

47. Lovasik BP, Kron P, Clavien PA, Petrowsky H, Kooby DA. Pancreatectomy and body mass index: an international evaluation of cumulative postoperative complications using the comprehensive complications index. Hpb [Internet]. 2019;21(12):1761–72. Available from: https://doi.org/10.1016/j.hpb.2019.04.006

48. Milanović Z, Pantelić S, Trajković N, Sporiš G, Kostić R, James N. Age-related decrease in physical activity and functional fitness among elderly men and women. Clin Interv Aging. 2013;8:549–56.

49. Mungovan SF, Huijbers BP, Hirschhorn AD, Patel MI. Relationships between perioperative physical activity and urinary incontinence after radical prostatectomy: An observational study. BMC Urol. 2013;13:1–9.

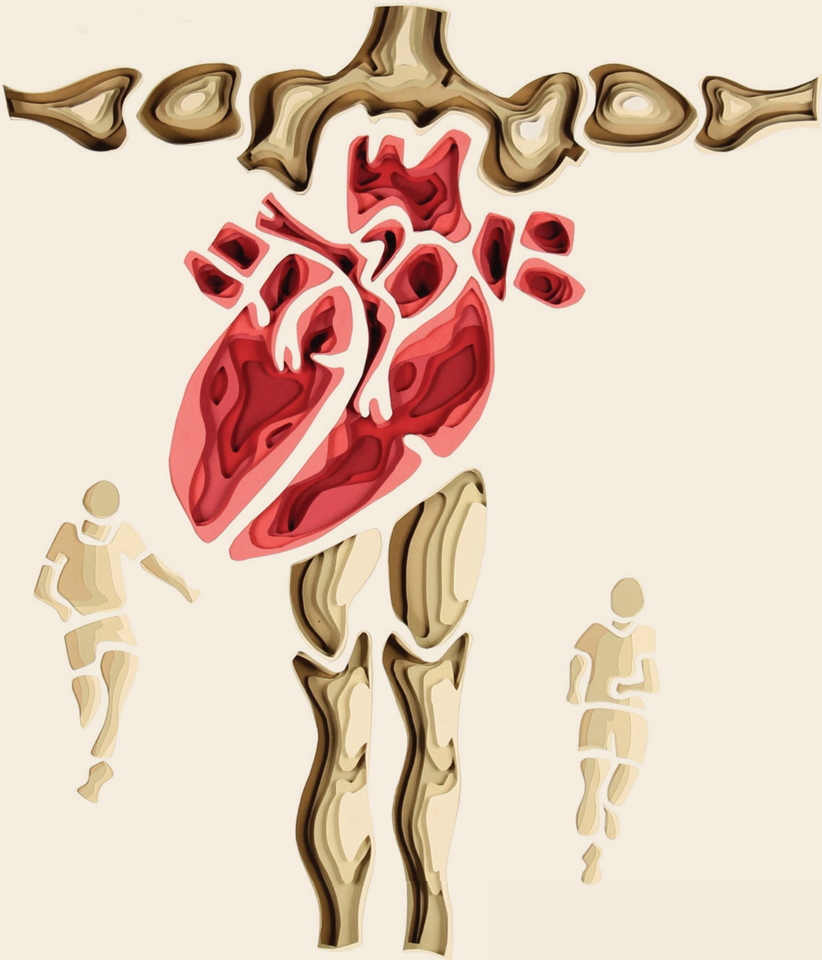
50. Dronkers JJ, Chorus AMJ, Van Meeteren NLU, Hopman-Rock M. The association of pre-operative physical fitness and physical activity with outcome after scheduled major abdominal surgery. Anaesthesia. 2013;68(1):67–73.

Supplementary table

Uni- and multivariate robust regression association with absolute length of hospital stay

Variable	Estimate	Std. error	R^2	t-value	p-value	Adj. R^2
Gender (female)	-0.464	2.146	0.002	-0.216	.830	
Age (years)	0.146	0.114	0.070	1.288	.208	
BMI (kg/cm^2)	-0.394	0.214	0.120	-1.845	.075	
Living situation (together)	1.522	2.056	0.013	0.740	.465	
Working status (employed)	-1.158	2.087	0.011	-0.555	.583	
Education level (high)	0.692	2.612	0.004	0.265	.793	
Alcohol Norm (above)	-1.473	2.273	0.017	-0.648	.522	
Smoking status (no)	2.313	3.275	0.022	0.706	.486	
ISWT (meters)	-0.007	0.004	0.100	-1.709	.098	
ISWT (% of predicted)	-0.076	0.035	0.132	-2.177	.038*	
MDASI total (avg., n=27)	-0.047	0.455	0.000	-0.104	.918	
MDASI symptoms (avg, n=27)	-0.174	0.449	0.004	-0.388	.701	
MDASI activities (avg, n=27)	0.118	0.368	0.025	0.321	.751	
Time spend sedentary (minutes)	-0.008	0.005	0.089	-1.760	.089	
Daily MVPA - total accumulated (minutes)	-0.070	0.027	0.132	-2.561	.016*	
Daily MVPA - 10-minute bouts (minutes)	-0.150	0.058	0.149	-2.574	.015*	
Laparoscopic / open surgery (Laparoscopic)	-2.426	2.086	0.043	-1.163	.254	
Major / minor surgery (minor)	-7.063	1.688	0.489	-4.185	<.001*	
Multivariate						
Constant	17.184	2.359		7.285	<.001*	.533
ISWT (% of predicted)	-0.051	0.024		-2.095	.045*	
Major / minor surgery (minor)	-6.609	1.349		-4.897	<.001*	
Daily MVPA - 10-minute bouts (minutes)	-0.035	0.031		-1.147	.261	

Abbreviations: BMI = Body Mass Index, ISWT = Incremental Shuttle Walk Test, MDASI = MD Anderson Symptom Inventory, MVPA = Moderate to Vigorous Physical Activity, * = P ≤.05, avg. = average



Chapter 6

Changes in self-reported and device-measured physical activity before abdominal resection surgery: A meta-analysis

CF Mylius, Y Mooiweer, WP Krijnen, T Takken, BC van Munster, CP van der Schans, JM Klaase

Mylius CF, Mooiweer Y, Krijnen WP, Takken T, van Munster BC, Van der Schans CP, Klaase JM.. Changes in Self-Reported and Device-Measured Physical Activity Before Abdominal Resection Surgery: A Meta-Analysis. Clinical Rehabilitation. 2023;0(0). doi:10.1177/02692155231202215

ABSTRACT

Objective: To determine the effect of interventions on physical activity levels of patients awaiting abdominal resection surgery using self-reported as well as device-measured outcome measures.

Data source: PubMed and EMBASE databases were searched on the 18th of april-2023 up to April 2023 for studies on interventions to promote physical activity during the preoperative phase.

Review methods: Studies were included if pre- and post-intervention physical activity was measured between diagnosis and abdominal surgery. Risk of bias was assessed by the Physiotherapy Evidence Database (PEDro) assessment tool for trials. Meta-analyses were performed to assess the effect of the pre-surgery activity promoting interventions on self-reported and device-measured physical activity.

Results: Seventeen studies were included in the analysis with 452 subjects in the intervention groups. The random-effect meta-analysis showed a moderate improvement in intervention groups measures in pre-surgery physical activity levels compared to the baseline (SMD= 0.67, [CI= 0.30;1.03], I2= 79%). The self-reported subgroup meta-analysis showed the largest increase in performed physical activity, (SMD= 0.78, [CI= 0.4;1.15], I2= 79%) whilst non-significant increase was shown in the device-measured subgroup (SMD= 0.16, [CI= -0.64;0.97], I2 = 58%).

Conclusion: Increasing physical activity in the preoperative phase is feasible. Self-reported physical activity outcome measures show larger effects compared to device-measured outcome measures. More high-quality research should be performed utilizing objective measures.

Keywords: prehabilitation – physical activity – abdominal surgery

INTRODUCTION

Abdominal resection surgery is indicated for the cure or palliation of numerous cancer types (1). The surgery itself represents a stress event with an increased risk for adverse effects unrelated to the treatment goals. These adverse effects have a profound negative impact on a person's ability to perform daily activities and may negatively impact postoperative quality of life (2–6). To help cope with the stressors involved with surgery, surgeons recommend patients to avoid inactivity and progress towards reaching at least 150 minutes of moderate to vigorous intensity physical activity and 2 to 3 moderate intensity resistance exercise sessions per week (7).

The role of physical activity for health is well recognized, with recent publications by the World Cancer Research Fund/American Institute for Cancer Research and the 2020 guidelines by the World Health Organization (8). These guidelines indicate that more physical activity contributes to optimal health outcomes in both the general population and specific subgroups like elderly and subjects suffering from chronic diseases (8). Due to its positive effect, there is a high interest in the role of preoperative physical activity to improve postoperative recovery. Evidence shows that a patient's physical activity level during the preoperative phase influences surgery outcomes by altering a person's ability to cope with the stressors of surgery (9). Namely, lower levels of preoperative physical activity compromise postoperative recovery, leading to an increase in postoperative complication rates (10–13) and length of hospital stay (9,14,15), and a decrease in health-related quality of life (16).

Although there are many different methods used to assess physical activity, namely self-monitored physical activity by questionnaire or diary log, or via device measured by motion sensors, reliability of the measurement instruments remains an important field of research (17). Because of low concurrent validity between the self-reported and the device measured methods, determining the best method is of importance (18,19). The cheapest way of measuring physical activity is the administration of physical activity questionnaires, which can assess all types of physical activity. However, very few physical activity questionnaires show acceptable reliability and validity across age groups (20,21), partly because answers may be distorted due to social desirability or recall bias (22). Motion sensors, such as pedometers or accelerometers are increasingly implemented as a measure of physical activity in a free-living environment. Accelerometers are small electronic devices that record acceleration associated with bodily movement and provide an objective estimate of duration and intensity of locomotion (23).

An increasing number of preoperative intervention studies, also known as prehabilitation, are performed aiming to increase the physical activity level of patients awaiting an abdominal operation amongst others (24–26). A wide variety of interventions are used, ranging from behavioral strategies to prescribed exercise interventions to increase physical activity. Nevertheless, most studies lack the sample size and homogeneity to establish whether physical activity interventions findings are consistent and can be generalized across patient groups, and treatment variations, or whether findings vary significantly. Therefore, the current meta-analysis aims to determine the effect of physical activity interventions in patients awaiting abdominal resection surgery. Furthermore, our secondary objective is to determine whether self-reported and device measured physical activity outcome measures result in similar outcomes.

METHODS

This systematic review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (27).

The databases PubMed and EMBASE were searched for relevant studies (Supplementary material A). The search strategy used blocks of search terms related to preoperative physical activity amongst major abdominal surgery patients. Terms were added for intervention studies aimed at altering physical activity levels including pre and post intervention measurements. The search was conducted on April 18, 2023 ranging from January 2000 to April 2023. Reference lists of eligible studies were hand-searched as a method to supplement electronic searching.

Results of the searches were combined, whereafter duplicates were removed. All unique records were screened for potential relevance using the title, abstract or descriptors, or both. Hereafter, the corresponding studies were screened on compliance with the eligibility criteria based on the full text of the articles. If not compliant, studies were excluded and reasons for exclusion based on its full text were recorded. Study selection and risk of bias assessment were performed by two reviewers (J.K. and C.M.).

Studies that measured pre- and post-intervention physical activity during the period between diagnosis and abdominal resection surgery were included, all physical activity outcome measures were considered eligible. Furthermore, additional criteria were: written in the English language, clinical trial study designs, and case–control studies. Exclusion criteria were: physical activity requirements in sample selection, conference proceedings, non-peer reviewed papers, opinion pieces, letters to the editor, commentaries, abstracts. Furthermore, studies including cosmetic-, bariatric surgery procedures or nonelective surgery were excluded.

A standardized form was used to extract data from eligible studies for assessment of the study quality and evidence synthesis. Information regarding the following was extracted: participant characteristics (age, gender, body mass index, surgery type or site, American Society of Anaesthesiologists classification, (neo)adjuvant therapy and baseline fitness), and study characteristics (sample size, design). All physical activity measurements were extracted. The method of measuring physical activity was extracted to determine the effect self-monitored *versus* device physical activity measures. Furthermore, measurements during the last week of an intervention were considered post-intervention measurement. To determine the effect of the interventions, pre- and post-intervention physical activity measurements, time between baseline measurement and post intervention measurement, and time until surgery was extracted. The extracted characteristics of the intervention were content, type, and duration. Since supervised interventions are reported to be more effective in persons with low habitual physical activity levels, interventions were categorized into supervised prescribed interventions and/or behavioral interventions aimed to increase daily living physical activity (28).

The Physiotherapy Evidence Database scale was used to determine the risk of bias of all included studies (29). This tool is deemed a valid measure of the methodological quality of (randomized) clinical trials (30). Each study was graded using 11 criteria (a score of one was awarded if the response was 'Yes' and zero if the response was 'No'. A total Physiotherapy Evidence Database scale score of 0-3 is considered 'poor', 4-5 'fair', 6-8 'good', and 9-10 'excellent'.

All included studies were summarized in tables. Normally distributed data were reported by mean and standard deviation (SD). Otherwise, the median and interquartile range or frequency and percentages were provided. Meta-analyses to determine the pooled effect of physical activity interventions were conducted using R-studio (31), package *Meta* for outcomes where mean and SD data were available (32). If multiple physical activity measures were available, total sum of physical activity was preferred over subcategories (e.g., time spend sedentary, moderate to vigorous physical activity or step count). To allow for the comparison between different measures across studies, within group pre-post changes were calculated by standardized mean difference using Hedges' adjusted g, which includes a correction for sample size bias (33). Furthermore, prescribed exercise versus behavioral interventions, and device versus self-monitored measured physical activity subset analyses were performed to increase homogeneity. To determine the effect of methodological quality, similar assessment was performed including studies rated 'fair', 'good' or 'excellent'. A Standardized mean difference = 0.0 to 0.2 is considered as a 'small' effect size, 0.2 to 0.8 represents a 'medium' effect size and >0.8 a 'large' effect size. Statistical heterogeneity was assessed by the I^2 test (34), which is based upon the percentage of variability across studies beyond what would be expected by chance. I^2 values of 25, 50 and 75 serve as limits for low, moderate, and high heterogeneity, respectively.

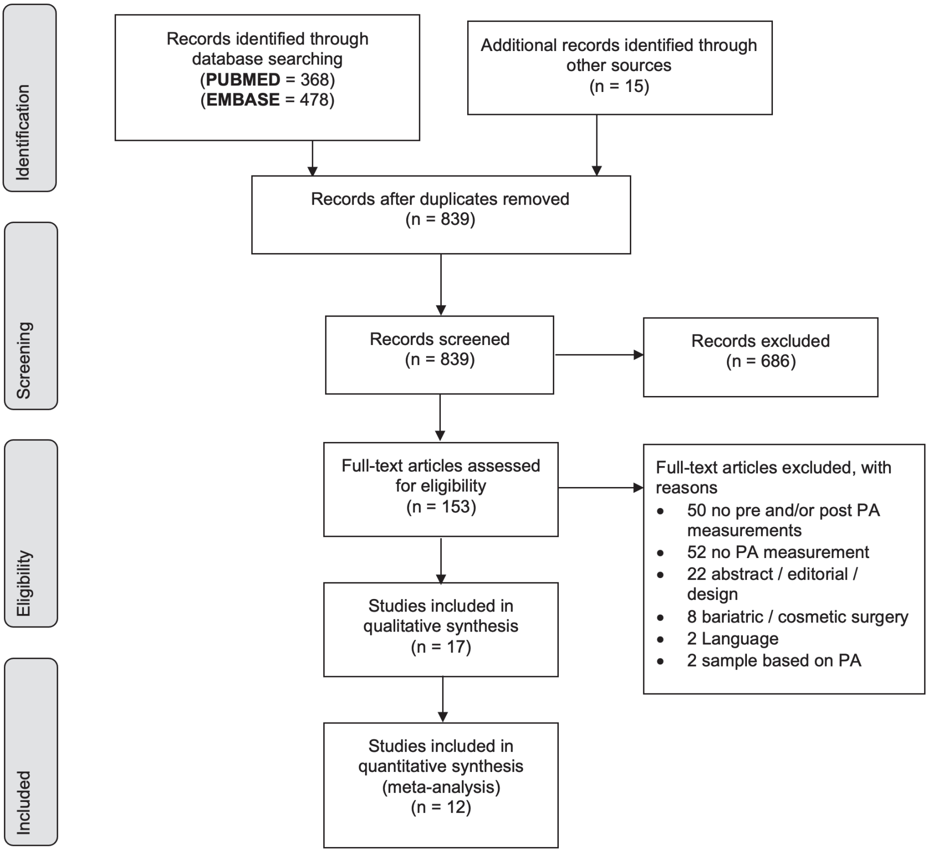


Figure 3. The PRISMA flow diagram displaying the selection of studies and reasons for exclusion.

RESULTS

Selected studies

We identified 861 potential unique studies published between January 2000 and April 2023. After initial screening, 153 studies were regarded potentially eligible. After reading the full-text, 17 studies passed all inclusion criteria. The final selection consisted of two non-randomized controlled trials, seven randomized controlled trials and six single arm intervention trials. Furthermore, 12 studies were considered eligible for inclusion in the meta-analysis. A flowchart displaying details of the selection process, including the reasons for exclusion, is presented in **Figure 1**.

Table 1. Methodological quality of included studies based on PEDro scale.

Author, year	Design	1	2	3	4	5	6	7	8	9	10	11	Total
Alejo, 2019 (36)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
Barberan-Garcia, 2017 (40)	RCT	1	1	0	0	0	0	1	0	1	0	1	5
Bousquet-Dion, 2018 (37)	RCT	1	1	0	1	0	0	1	1	1	1	1	8
Carli, 2010 (47)	RCT	1	1	0	1	0	0	0	0	1	1	1	6
Chmelo, 2022 (44)	CCT	1	0	0	0	0	0	0	1	0	0	1	3
Gillis, 2014 (42)	RCT	1	1	0	1	0	0	1	0	1	1	1	7
Halliday, 2021 (35)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
Li, 2012 (50)	CCT	1	0	0	1	0	0	0	1	1	1	1	6
Loughney, 2021 (49)	RCT	1	1	0	0	0	0	1	1	1	1	1	7
Moug, 2019 (38)	RCT	1	1	0	1	0	1	1	0	1	1	1	8
NGO-Huang, 2017 (51)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
Singh, 2017 (41)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
Steffens, 2021 (39)	RCT	1	1	0	1	0	1	1	1	1	1	1	9
Suen, 2021 (45)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
Timmerman, 2011 (46)	CCT	1	0	0	0	0	0	0	1	1	0	1	4
Waterland, 2022 (43)	CCT	1	0	0	0	0	0	0	0	1	0	1	3
West, 2015 (48)	CCT	1	0	0	1	0	0	0	1	1	1	1	6

1. Eligibility criteria were specified; **2.** Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received); **3.** Allocation was concealed; **4.** The groups were similar at baseline regarding the most important prognostic indicators; **5.** There was blinding of all subjects; **6.** There was blinding of all therapists who administered the therapy; **7.** There was blinding of all assessors who measured at least one key outcome; **8.** Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups; **9.** All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analyzed by "intention to treat"; **10.** The results of between-group statistical comparisons are reported for at least one key outcome; 11. The study provides both point measures and measures of variability for at least one key outcome

Methodological quality assessment

Quality of the included studies varied, and none of the studies fulfilled all eleven quality criteria. An excellent quality score was observed in one study, seven studies received a 'good' score, and four studies received a score 'fair', and five studies received a 'poor' score. On average, randomized controlled trials received higher scores compared to non-randomized trials with 7.1 and 4.4 points, respectively. Due to the nature of the intervention, no study included blinding of participants and allocation concealment for participants. Table 1 provides a detailed overview of the methodological quality of the included studies based on the Physiotherapy Evidence Database scale.

Study characteristics

Table 2 displays the overall study characteristics. The 17 included studies assessed 856, 261 (30%) female and 513 (60%) male subjects in total, of which were 452 (164 females, 302 males) in intervention and 315 (97 females, 211 males) in control groups. One study did not report the sex distribution (35). Mean and median age of included subjects ranged between 54.4 and 74 years. Subjects received neoadjuvant chemo- and/or radiotherapy during the intervention in ten studies (58.8%), before the intervention in two studies (11.8%) and no neoadjuvant therapy in five studies (29.4%).

Interventions

Characteristics of the interventions are reported in **Table 3**. Six studies used behavioral interventions aimed to increase physical activity (36–41). These interventions consist of a pedometer or a time goal to be physical active, a motivational interview to maintain and/or increase physical activity, or education on the importance of physical activity. All studies used a prescribed exercise scheme to be performed either in hospital-based or home-based intervention to increase physical activity (35–47). Four studies performed solely supervised hospital based exercise sessions (46,48–50), most studies combined home-based and hospital-based sessions or did not include a statement about the location. The frequency of the exercise sessions ranged from 0.5 to 5 sessions per week and lasted between 10 and 120 min per session. The duration of the interventions ranged from 2 to 16 weeks.

Table 2. Overall study characteristics

Author	characteristics		Intervention	Control	Time frame
Alejo, 2019 (36)	N: 12, Surgery site: Rectal NARCT: during intervention	Age (year) Sex (male) VO_{2peak} (mL/kg[1]/min[1])	61±7 3(25%) 24.4±9.1		Intervention: 5 weeks Follow-up to surgery: 6-8 weeks
Barberan-Garcia, 2017 (40)	N: 125, IG *n*: 54. CG *n*: 56 Surgery site: Abdominal NARCT: Not reported	Age (year) Sex (male) ASA index II, III, IV 6MWD (meters)	71±11 43(69%) 19(30%), 43(68%), 1(2%) 472±94	71±10 51(81%) 24(38%), 36(56%), 4(6%) 471±95	Intervention: 6±2 weeks Follow-up to surgery: 1 week
Bousquet-Dion, 2018 (37)	N: 63, IG *n*: 37. CG *n*: 26 Surgery site: Colorectal NARCT: during intervention	Age (year) Sex (male) ASA index I, II, III+ 6MWD (% predicted)	74[IQR 67.5 – 78] 30(81%) 1(3%), 23(62%), 13(35%) 69%	71[IQR 54.5 – 74.5] 16(62%) 3(12%), 11(42%), 12(46%) 71%	Intervention: ±4 weeks Baseline to surgery: IG: 32(25-48) days CG: 20.5(15-32) days
Carli, 2010(47)	N: 112, IG *n*: 58. CG *n*: 54 Surgery site: abdominal NARCT: not excluded	Age (year) Sex (male) ASA index I, II, III 6MWD (% predicted) VO_{2peak} (mL/kg[1]/min[1])	61±16 34(58%) 3(5%), 42(72%), 13(22%) 71±15 18±7	60±15 31(57%) 4(7%), 39(72%), 11(20%) 74±15 19±6	Intervention: 59±60.7 days Baseline to surgery: 52±51 days
Chmelo, 2022 (44)	N: 39, Surgery site: Esophageal NARCT: during intervention	Age (year) Sex (male) VO_{2peak} (mL/kg[1]/min[1])	68(51 – 81) 33(85%) 19.4±4.2		Intervention: 91(84-105 days) Follow-up to surgery: 35(31-47 days)
Gillis, 2014 (42)	N: 77, IG *n*: 38. CG *n*: 39 Surgery site: Colorectal NARCT: during intervention	Age (year) Sex (male) ASA index I, II, III+ TNM cancer stage 1-2, 3 6MWD (% predicted)	65.7±13.6 21(55%) 3(11%), 24(63%), 10(26%) 21(55%), 17(45%) 65±17	66±9.1 27(69%) 4(10%), 26(67%), 9(23%) 26(67%), 13(33%) 65±11	Intervention: 24.5 days Baseline to surgery: IG: 24.5(20-35) days CG: 20(11-40) days
Halliday, 2021 (35)	N: 67, Surgery site: Esophageal, NARCT: during intervention	Age (year) Predicted VO_{2max}	66±9.7 23.8±6.4		Intervention: ±16 weeks Follow-up to surgery 1 week
Li, 2012 (50)	N: 87, IG *n*: 42, CG *n*: 45, Surgery site: abdominal NARCT: Not reported	Age (year) Sex (male) ASA index I, II, III 6MWD (% predicted)	67.4±11 22(54%) 3(7%), 31(74%), 8(19%) 66±12	66.4±12 29(64%) 6(13%), 29(65%), 10(22%)	Intervention: 33(21-46) days Follow-up to surgery: ±1 week

Abbreviations: NARCT = Neoadjuvant chemoradiotherapy, TNM = classification of malignant tumors, ASA = American Society of Anesthesiologists classification, VO_{2peak} = Peak oxygen uptake, 6MWD = six-minute walk distance, sec = seconds, min = minutes, IG = intervention group, CG = control group

Table 2. Overall study characteristics (continued)

Author	characteristics		Intervention	Control	Time frame
Loughney, 2021 (49)	N: 33, IG *n*: 17, CG *n*: 16, Surgery site: abdominal, NARCT: before intervention	Age (year) Sex (male) TNM staging T2, T3, T4 VO_{2peak} (mL/kg^{1}/min^{1})	64±14 14(82%) 1(6%), 13(77%), 3(18%) 11.6±3.4	57±10 12(75%) 2(13%), 12(75%), 2(13%) 10.8±2.5	Intervention: 6 – 9 weeks Follow-up to surgery: 3 weeks
Moug, 2019 (38)	N: 48, IG *n*: 24, CG *n*: 24, Surgery site: abdominal, NARCT: during intervention	Age (year) Sex (male) ASA index II, III 6MWD (meters)	65.2±11.4 18(75%) 18(78%), 5(22%) 435.7±91.7	66.5±9.6 13(54%) 14(58%), 10(42%) 436.7±66.4	Intervention: 8 weeks Follow-up to surgery: 1-2 weeks
NGO-Huang, 2017 (51)	N: 15, Surgery: pancreatic, NARCT: During intervention	Age (year) Sex (male) 10-m walk test (sec)	66±6.3 9(60%) 6.61±0.91		Intervention: 17(5-25) weeks
Singh, 2017 (41)	N: 10, Surgery site: Rectal NARCT: During intervention	Age (year) Sex (male)	54.4±12.9 5(50%)		Intervention: 16 weeks Follow-up to surgery: 6±4.1
Steffens, 2021 (39)	N: 22, IG *n*: 11, CG *n*: 11, Surgery site: abdominal, NARCT: during intervention	Age (year) Sex (male) VO_{2max} (mL/kg^{-1}/min^{-1}) 6MWD (meter)	62(48-72) 6(54.5%) 18.8(14 – 23.3) 490(370 – 585)	66(46-70) 6(54.5%) 19.9(14.7 – 23) 525(459 – 585)	Intervention: 2-6 weeks Follow-up to surgery: 1 week
Suen, 2021(45)	N: 22, Surgery site: Colorectal, NARCT: None	Age (year) Sex (male) 6MWD (meter)	72.5(56 – 86) 12(54.6%) 434.6±108.1		Intervention: 2-4 weeks Follow-up to surgery: 1-2 days
Timmerman, 2011 (46)	N: 39, IG *n*: 15, CG *n*: 24 Surgery site: abdominal, NARCT: None	Age (year) Sex (male) VO_{2peak} (mL/kg^{-1}/min^{-1})	59±8 12(80%) 25±0.5	64±13 17(71%)	Intervention: 5 weeks
Waterland, 2022 (43)	N: 50, Surgery site: abdominal NARCT: Not reported	Age (year) Sex (male) VO_{2max} (mL/kg^{-1}/min^{-1})	71(44-84) 26(52%) 14±2.9		Intervention: 6 weeks
West, 2015 (48)	N: 35, IG *n*: 22, CG *n*: 13 Surgery site: rectal, NARCT: Before intervention	Age (year) Sex (male) TNM staging T2, T3, T4 TNM staging N0, N1, N2 VO_{2peak} (mL/kg/min)	62(45-82) 14(64%) 2(9%), 17(77%), 3(14%) 2(9%), 12(55%), 8(36%) 18.9±5.1	72(62-84) 9(69%) 1(8%), 10(77%), 2(15%) 2(15%), 7(54%), 4(31%) 17.9±3.1	Intervention: 6 weeks Follow-up to surgery: 9 weeks

Abbreviations: NARCT= Neoadjuvant chemoradiotherapy, TNM= malignant tumor classification, ASA= American Society of Anesthesiologists classification, VO_{2peak}= Peak oxygen uptake, 6MWD= six-minute walk distance, sec= second, min= minute, IG= intervention group, CG= control group

Table 3. Intervention description

Author	Intervention
Alejo, 2019 (36)	Six educational lectures, supervised hospital and community-based exercise sessions. 45min avg prescribed weekly physical activity.
Barberan-Garcia, 2017 (40)	IG: 1) motivational interview: promote physical activity by increasing daily pedometer steps; optimization of walking intensity; 2) Patients with severely reduced aerobic capacity / physical activity: home-based functional resistance exercises; 3) triweekly supervised high intensity endurance exercise training. 47-141min avg prescribed weekly physical activity. CG: physical activity recommendation, nutritional counseling, advice on intoxications, intravenous iron.
Bousquet-Dion, 2018 (37)	IG: 3-4 weekly 30min personalized homebased aerobic / resistance program. Pedometer to encourage walking and, weekly 60min supervised aerobic training and nutrition intervention. 90min avg prescribed weekly physical activity. CG: care as usual.
Carli, 2010 (47)	IG: Triweekly 10-15min resistance sessions, daily 20-30min aerobic training. 170-255min avg prescribed weekly physical activity. CG: Encouraged to daily walk 30min, 5min breathing- and 5-10min circulation exercise.
Chmelo, 2022 (44)	Homebased exercise consisting of targeted daily step-based aerobic exercise and daily strengthening exercise with motivational weekly phone call supervision.
Gillis, 2014 (42)	IG: Trimodal program of triweekly 50min homebased personalized supervised exercise sessions alternating between aerobic and resistance training. No physical activity instructions. 150min avg prescribed physical activity weekly. CG: care as usual.
Halliday, 2021 (35)	Personalized exercise program aimed at achieving 600 MET weekly. Supervised by weekly phone call from exercise therapist in addition to self-monitored and regulated exercise. Furthermore, motivational interview by clinical nurse specialist if needed. Interview aimed at reinforcing need and facilitators, removing potential barriers. 150-300min avg prescribed physical activity weekly.
Li, 2012 (50)	Trimodal program: individualized exercise, nutrition, and anxiety reduction. Prescribed triweekly 30min aerobic and resistance exercise. No physical activity instructions.
Loughney, 2021 (49)	IG: Triweekly sessions of aerobic exercise program ranging from 40-60min. Consisting of moderate to high intensity interval, high intensity interval and resistance training. No physical activity instruction. CG: care as usual.
Moug, 2019 (38)	IG: Step target based on pre-NACRT pedometer step level. After achieving goal, increase until surgery up to 3000 steps. Weekly diary including target and motivational material. Supervised by biweekly motivational phone call. Achieving the goal accumulates to daily 30min physical activity, performed 5 times weekly. 150min avg prescribed physical activity weekly. CG: care as usual.
NGO-Huang, 2017 (51)	Moderate intensity program consisting of triweekly 20min walking and twice weekly 30min resistance training. Advised to reach 120min of exercise weekly. Supervised by personal trainer, DVD- and dietician instructions. 240min avg prescribed physical activity weekly.
Singh, 2017 (41)	Twice weekly supervised exercise session for 60min. Homebased physical activity log sheets aimed at 150min MVPA weekly. 150min avg prescribed physical activity weekly.
Steffens, 2021 (39)	IG: 60min sessions 4 times weekly physiotherapist individualized homebased aerobic and resistance training and dairy to track exercise. Weekly progressive exercise supervised by physiotherapist for 60min and encouraged to walk 30min daily tracked by Fitbit. 510min avg prescribed physical activity weekly. CG: advised to maintain physically active.

Author	Intervention
Suen, 2021(45)	Twice weekly 60min personalized supervised aerobic and resistance exercise sessions. Triweekly homebased aerobic exercise sessions and encouraged to engage in self-monitoring. Twice weekly contact with supervisor, namely oncology exercise physiologist and written information on nutrition. 120min avg prescribed physical activity weekly.
Timmerman, 2011 (46)	Twice weekly individualized endurance and resistance training for 120min. Supervised by physiotherapist and exercise assistant. 240min avg prescribed physical activity weekly.
Waterland, 2022 (43)	Personalized homebased exercise prescription provided by therapist including strength and aerobic exercises, preoperative education. Avg 180min prescribed physical activity.
West, 2015 (48)	Triweekly supervised in-hospital exercise sessions. 40min interval training including 5min warm-up and cooldown. 120min avg prescribed physical activity weekly.

Abbreviations: MVPA = moderate to vigorous physical activity, NARCT = Neoadjuvant chemoradiotherapy, MET = Metabolic equivalent of Task, IG = intervention group, CG = control group, avg = average

Physical activity

Table 4 shows the baseline and pre-operative physical activity measures. Performed physical activity increased in ten out of the thirteen and four out of seven studies for the intervention and control groups, respectively. The random-effect meta-analysis, depicted in **Figure 2.**, showed a moderate improvement in intervention groups compared to the baseline measures; however, high heterogeneity was detected (SMD = 0.67, [CI = 0.30; 1.03], I^2= 79%). Excluding studies of 'poor' quality showed no significant increase with an SMD of 0.56 ([CI = -0.11; 1.24] I^2: 89%). When comparing baseline and pre-operative physical activity behavior in control groups, see Figure 2., no changes (SMD = 0.03 [CI = -0.18; 0.24] I^2 = 0%) were found without heterogeneity.

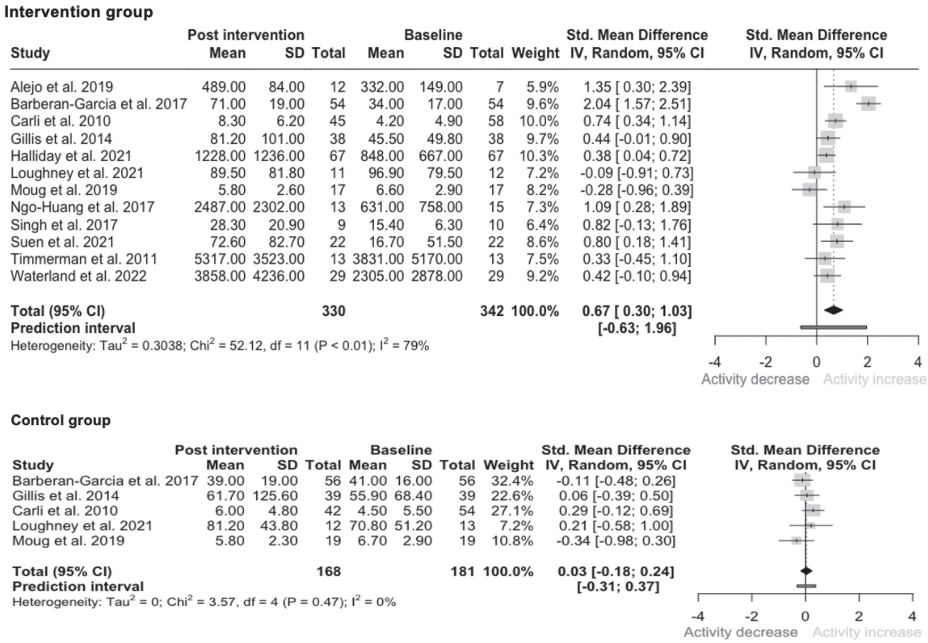


Figure 2. Meta-analysis of baseline and post-intervention physical activity in the intervention and control group- Meta-analysis and forest plot of random effect baseline and post-intervention physical activity in the intervention and control group

Outcomes

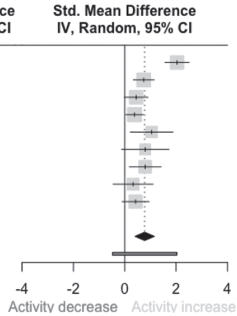
Physical activity performed during the last four weeks (37,42,47,50) or the last week of the intervention was used as the measurement period in the majority of the studies (35,36,39–41,45,46,48,51)**.** Twelve studies employed self-report measures of physical activity; the CHAMPS questionnaire was most frequently used (37,42,47,50). Outcomes were either a sum score (40), hours and/or minutes of physical activity per week (37,41,42,45,47) or metabolic equivalent of task (MET) per week (35,39,43,46,51). The intervention arms of the self-report measurement studies reported an increase in physical activity in ten studies (35,37,40–43,45–47,50,51), while three out of seven studies reported increase in physical activity in control groups (42,47,49). Five studies used device measures of physical activity, namely, one reporting on activity bouts, three reporting step counts, and two studies reporting both step counts and activity bouts. (36,38,44,48,49) The self-reported intervention subgroup meta-analysis showed a large increase in performed physical activity, notably high statistical heterogeneity was found (SMD = 0.78, [CI = 0.4; 1.15], I^2 = 79%), while no significant changes were shown in the device-measured subgroup (SMD = 0.16, [CI = -0.64; 0.97], I^2 = 58%) as seen in Figure 3.a. Analysis in 'fair' to excellent' quality studies show a SMD 0.61 ([0.31; 0.91], I^2: 0%) in self-reported, and a non-significant SMD -0.2 ([-0.73; 0.32] I^2: 0%) in device measured physical activity, respectively.

Although studies including behavioral interventions yielded a large improvement with a standardized effect of SMD 0.99 ([CI = -0.03; 2.01], I^2 = 90%) compared to solely prescribed exercise, namely SMD 0.51 ([CI = 0.33; 0.69], I^2= 1%), these differences in outcome are not statistically significant. Control groups solely receiving advice to maintain physical activity reported conflicting results (39,40,42). Analyses of intervention type are shown in **Figure 3.b**.

Self-reported group

Study	Post intervention Mean	SD	Total	Baseline Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Barberan-Garcia et al. 2017	71.00	19.00	54	34.00	17.00	54	12.3%	2.04 [1.57; 2.51]
Carli et al. 2010	8.30	6.20	45	4.20	4.90	58	13.0%	0.74 [0.34; 1.14]
Gillis et al. 2014	81.20	101.00	38	45.50	49.80	38	12.5%	0.44 [-0.01; 0.90]
Halliday et al. 2021	1228.00	1236.00	67	848.00	667.00	67	13.5%	0.38 [0.04; 0.72]
Ngo-Huang et al. 2017	2487.00	2302.00	13	631.00	758.00	13	8.8%	1.05 [0.22; 1.88]
Singh et al. 2017	28.30	20.90	10	15.40	6.30	10	8.0%	0.80 [-0.12; 1.72]
Suen et al. 2021	72.60	82.70	22	16.70	51.50	22	10.8%	0.80 [0.18; 1.41]
Timmerman et al. 2011	5317.00	3523.00	13	3831.00	5170.00	13	9.3%	0.33 [-0.45; 1.10]
Waterland et al. 2022	3858.00	4236.00	29	2305.00	2878.00	29	11.8%	0.42 [-0.10; 0.94]
Total (95% CI)			**291**			**304**	**100.0%**	**0.78 [0.40; 1.15]**
Prediction interval								**[-0.48; 2.03]**

Heterogeneity: Tau^2 = 0.2441; Chi^2 = 38.55, df = 8 (P < 0.01); I^2 = 79%



Accelerometer group

Study	Post intervention Mean	SD	Total	Baseline Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Alejo et al. 2019	489.00	84.00	7	332.00	149.00	7	25.3%	1.22 [0.04; 2.39]
Loughney et al. 2021	89.50	81.80	11	96.90	79.50	12	35.1%	-0.09 [-0.91; 0.73]
Moug et al. 2019	5.80	2.60	17	6.60	2.90	17	39.7%	-0.28 [-0.96; 0.39]
Total (95% CI)			**35**			**36**	**100.0%**	**0.16 [-0.64; 0.97]**
Prediction interval								**[-8.56; 8.89]**

Heterogeneity: Tau^2 = 0.3041; Chi^2 = 4.82, df = 2 (P = 0.09); I^2 = 58%

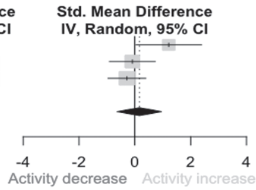
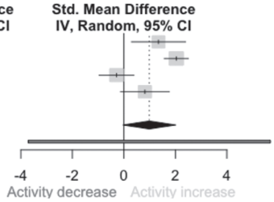


Figure 3.a meta-analysis of effect of measurement method on physical activity - Meta-analysis and forest plot of random effect baseline and post-intervention physical activity in the intervention group measured by self-monitor and accelerometer instruments.

Combined behavioral and prescribed physical activity intervention group

Study	Post intervention Mean	SD	Total	Baseline Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Alejo et al. 2019	489.00	84.00	12	332.00	149.00	7	22.5%	1.35 [0.30; 2.40]
Barberan-Garcia et al. 2017	71.00	19.00	54	34.00	17.00	54	27.8%	2.04 [1.57; 2.51]
Moug et al. 2019	5.80	2.60	17	6.60	2.90	17	26.1%	-0.28 [-0.96; 0.39]
Singh et al. 2017	28.30	20.90	9	15.40	6.30	10	23.6%	0.82 [-0.13; 1.77]
Total (95% CI)			**92**			**88**	**100.0%**	**0.99 [-0.03; 2.01]**
Prediction interval								**[-3.71; 5.69]**

Heterogeneity: $Tau^2 = 0.9226$; $Chi^2 = 31.48$, df = 3 ($P < 0.01$); $I^2 = 90\%$



Prescribed physical activity intervention group

Study	Post intervention Mean	SD	Total	Baseline Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carli et al. 2010	8.30	6.20	45	4.20	4.90	58	20.2%	0.74 [0.34; 1.14]
Gillis et al. 2014	81.20	101.00	38	45.50	49.80	38	15.8%	0.44 [-0.01; 0.90]
Halliday et al. 2021	1228.00	1236.00	67	848.00	667.00	67	28.0%	0.38 [0.04; 0.72]
Loughney et al. 2021	89.50	81.80	11	96.90	79.50	12	4.9%	-0.09 [-0.91; 0.73]
Ngo-Huang et al. 2017	2487.00	2302.00	13	631.00	758.00	15	5.1%	1.09 [0.28; 1.89]
Suen et al. 2021	72.60	82.70	22	16.70	51.50	22	8.6%	0.80 [0.18; 1.41]
Timmerman et al. 2011	5317.00	3523.00	13	3831.00	5170.00	13	5.4%	0.33 [-0.45; 1.10]
Waterland et al. 2022	3858.00	4236.00	29	2305.00	2878.00	29	12.1%	0.42 [-0.10; 0.94]
Total (95% CI)			**238**			**254**	**100.0%**	**0.51 [0.33; 0.69]**
Prediction interval								**[0.29; 0.74]**

Heterogeneity: $Tau^2 < 0.0001$; $Chi^2 = 7.06$, df = 7 ($P = 0.42$); $I^2 = 1\%$

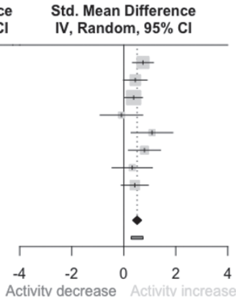


Figure 3.b meta-analysis of effect of intervention type on physical activity – Meta-analysis and forest plot of random effect baseline and post-intervention physical activity in the intervention group receiving behavioral intervention and a prescribed physical activity scheme and the group receiving solely a supervised and/or unsupervised prescribed physical activity scheme.

Table 4. Baseline and pre-operative physical activity measures

Author	Device	Measurement period	Outcome measures	Group	Baseline*	Follow-up*	SMD [95% CI]
Alejo, 2019 (36)	ActiGraph GT3X	T0: Prior to intervention T1: Week before surgery	Mean inactivity min /week Mean MVPA min /week	IG (n = 7) IG (n = 7)	513±111 332±149	510±302 489±84	-0.01[-0.94;0.92] 1.35[0.30;2.40]**
Barberan-Garcia, 2017 (40)	YPAS	T0: start intervention T1: prior to surgery	Mean sum score	IG (n = 54) CG (n = 56)	34±17 41±16	71±19 39±19	2.04[1.57;2.51]** -0.11[-0.48;0.26]
Bousquet-Dion, 2018 (37)	CHAMPS	T0: prior to intervention T1: prior to surgery	Kcal per kg /week total physical activity	IG (n = 37) CG (n = 26)	37.5[15.3;64.5] 38.5[10;58.3]	45[19.1;69.3] 31.1[7.8;70.8]	
Carli, 2010 (47)	CHAMPS	T0: planning surgery T1: prior to surgery	Mean moderate physical activity hours /week	IG (n = 45) CG (n = 42)	4.2±4.9 4.5±5.5	8.3±6.2 6±4.8	0.74[0.34;1.14]** 0.29[-0.12;0.69]
Chmelo, 2022 (44)	Walking Style One	T0: prior to intervention T1: prior to surgery	Daily step count	IG (n = 19)	5528[2303;8515]	5792[2361;9980]	
Gillis, 2014 (42)	CHAMPS	T0: prior to intervention T1: 4 weeks prior to surgery	Mean kcal per kg /week	IG (n = 38) CG (n = 39)	45.5±49.8 55.9±68.4	81.2±101 61.7±125.6	0.44[-0.01;0.90] 0.06[-0.39;0.50]
Halliday, 2021 (35)	Diary log	T0: at completion NARCT T1: week prior to surgery	Mean MET /week	IG (n = 67)	989±805	1228±1236	0.23[-011;0.57]
Li, 2012 (50)	CHAMPS	T0: Prior to intervention T1: week prior to surgery	Median kcal per kg /week	IG (n = 42)	17[10;36]	36[19;74]	
Loughney, 2021 (49)	Sensewear	T0: After completion NARCT T1: prior to surgery	Mean step count / day Mean physical activity duration min / day	IG (n = 11) CG (n = 12) IG (n = 11) CG (n = 12)	7058±4981 6321±3456 96.9±79.5 70.8±51.2	7023±5562 6749±2959 89.5±81.8 81.2±43.8	-0.01[-0.82;0.81] 0.13[-0.66;0.91] -0.09[-0.91;0.73] 0.21[-0.58;1.00]
Moug, 2019 (38)	ActivPAL	T0: prior undergoing NACRT T1: 1–2 weeks pre-surgery	Mean % active / day Mean steps / day	IG (n = 17) CG (n = 19) IG (n = 17) CG (n = 19)	6.6%±2.9 6.7%±2.9 7779±4045 7773±3871	5.8±2.6 5.8±2.3 6675±3100 5920±3152	-0.28[-0.96;0.39] -0.34[-0.98;0.30] -0.30[-098;0.38] -0.51[-1.16;0.13]
NGO-Huang, 2017 (51)	IPAQ-SF	T0: At enrollment T1: week prior to surgery	Mean MET /week	IG (n = 13)	631±758	2487±2302	1.09[0.28;1.89]**
Singh, 2017 (41)	Godin Leisure-time	T0: Prior to intervention T1: Week prior to surgery	Mean leisure time /week	IG (n = 9)	15.4±6.3	28.3±20.9	0.82[-0.13;1.77]

MVPA = moderate to vigorous physical activity, YPAS = Yale Physical Activity Survey, CHAMPS = Community Healthy Activities Model Program for Seniors, NARCT = Neoadjuvant chemoradiotherapy, IPAQ-SF = International Physical Activity Questionnaire Short Form, SQUASH = Short Questionnaire to Assess Health-enhancing physical activity, MET = Metabolic equivalent of Task, min = minutes, Kcal = Kilocalorie, kg = kilogram, SD = standard deviation, IQR = interquartile range, SMD = standardized mean difference, * = Mean±SD or frequency (%) or median [IQR] ,** = P value <.05, T0 = baseline, T1 = Follow-up

Table 4. Baseline and pre-operative physical activity measures (continued)

Author	Device	Measurement period	Outcome measures	Group	Baseline*	Follow-up*	SMD [95% CI]
Steffens, 2021 (39)	IPAQ-SF	T0: prior to intervention T1: day before surgery	Median MET /week Median walking /week	IG (n = 11) CG (n = 11) IG (n = 11) CG (n = 11)	1386[792;3252] 1386[660;41580] 420[210;840] 420[200;1680]	313.5[82.5;1297] 231[66;594] 50[25;175] 95[0;187]	
Suen, 2021(45)	Godin Leisure-time	T0: prior to surgery T1: one/two days before surgery	Mean moderate leisure activity /week	IG (n = 22)	16.7±51.5	72.6±82.7	0.80[0.18;1.41]**
Timmerman, 2011 (46)	SQUASH	T0: at preoperative screening T1: prior to surgery	Total MET /week	IG (n = 13)	3831±5170	5317±3523	0.33[-0.45;1.10]
Waterland, 2022 (43)	IPAQ	T0: prior to undergoing NACRT T1: Post intervention	Total MET /week	IG (n=29)	2305(2878)	3858(4236)	0.42[-0.10;0.94]
West, 2015 (48)	Sensewear	T0: Post-NARCT T1: Post intervention	Median steps / day	IG (n = 22) CG (n = 13)	3399(2009;4867) 2274(1339;5118)	5465(4409;8244) 4792(3206;7411)	

MVPA = moderate to vigorous physical activity, YPAS = Yale Physical Activity Survey, CHAMPS = Community Healthy Activities Model Program for Seniors, NARCT = Neoadjuvant chemoradiotherapy, IPAQ-SF = International Physical Activity Questionnaire Short Form, SQUASH = Short Questionnaire to Assess Health-enhancing physical activity, MET = Metabolic equivalent of Task, min = minutes, Kcal = Kilocalorie, kg = kilogram, SD = standard deviation, IQR = interquartile range, SMD = standardized mean difference, * = ±SD or frequency (%) or median [IQR] ,** = P value <.05, T0 = baseline, T1 = Follow-up

DISCUSSION

The principal finding of this review was that interventions aimed to increase preoperative physical activity in major abdominal surgery appear to be effective. However, self-reported outcome measures seem to report larger increases in physical activity levels compared to device measures. Nonetheless, too few high-quality studies utilized device measures to determine the effect of measurement instruments. Overall, interventions including a combination of prescribed exercise and behavioral strategies seem to be the most effective to increase physical activity.

Moderate-sized improvement in physical activity levels were found in intervention groups compared to the baseline measures. Similarly, interventions aimed at increasing physical activity in the healthy and chronically ill adults' population have been proven to be moderately effective (52–54). In these populations, larger effects are reported when using individually adapted behavioral interventions (52,53). This method, aimed to have participants incorporate physical activity into their daily routines, are more cost-effective when compared with prescribed physical activity interventions in the general population (55,56). However, the brief preoperative window of opportunity to increase physical activity behavior, and the effects of the stress involved with being enlisted for surgery favors an intensive collaboration between patient and healthcare provider (57). This assumption is supported in the current meta-analysis, which shows that the largest increase was achieved through intensive collaboration in the form of combined of prescribed exercise and behavioral intervention. However, behavioral changes are time consuming. Although patients are likely to be open for behavioral changes to increase their ability to cope with the stressors of surgery during the preoperative period, the available preoperative time is limited to achieve long lasting behavioral changes (58).

Our meta-analysis suggests that studies utilizing device measurement showed a lower or non-significant increase compared to self-monitored questionnaires. This difference may be due to overestimation in the self-monitored group resulting from recall bias (59,60). Thereby, owning an over optimistic version of performed physical activity, may represent a distorted version of reality. This effect might be enlarged by knowing the potential effect of inactivity on surgery outcomes. This knowledge likely increases when subjects participate in an intervention aimed at improving the patient's ability to cope with the stressors of surgery via increased physical activity. Furthermore, participation in an intervention built around physical activity can result in an increased awareness on performed physical activity, duration neglect, and socially desirable answers to please researchers (61,62). However, device measurement instruments are subject to the researcher's analysis and configuration choices, which can affect the reliability of the data. Therefore, to gather reliable information into performed physical activity, device measures should be used, utilizing standardized configurations.

Since homogeneous results were found in the studies using device measurement instruments, the high heterogeneity between studies found in the analysis of the overall effect and in analysis of the different intervention types, might be caused by the application of different measurement instruments. However, the number of studies using device measurement instruments were low. Furthermore, in general, studies with a high therapeutic validity show homogeneous results towards better postoperative outcomes (63,64). Peer-reviewed studies on prehabilitation schemes for major abdominal elective surgery candidates show heterogeneous designs in terms of measurement instruments, duration, and modalities of the intervention. The design of these interventions depends in part on organizational aspects of healthcare providers, like the interval of time before surgical date, availability of healthcare providers and identifying eligible patients as well as by type of surgical intervention (63). In the current study, designs varied significantly between studies. Nevertheless, most studies showed an increase in physical activity. Notably, the biggest improvement was observed in the study conducted by Barberan-Garcia et al., this trial performed a personalized program including a motivational interview aimed at increasing physical activity, home-based functional training, and supervised endurance training (40). However, more studies are needed to confirm the effect of this combined intervention.

PubMed and Embase were the chosen database resources for the current study, based on their well-established reputation for providing comprehensive, diverse, and high-quality biomedical literature. Both databases offer broad coverage, unique content, and accessibility, with their use of different indexing systems providing complementary results. The decision to limit the study to these two databases was a deliberate one, made to ensure a rigorous and comprehensive literature review. However, our analyses are limited by the low number of studies included in the selected subgroups. Therefore, the effect of device versus self-reported measurement for each intervention type could not be determined since only a single study performed both a behavioral intervention combined with device measurement. Furthermore, the discrepancy between the device measured, and self-monitored physical activity outcomes limits the generalizability of the overall intervention effects. However, the current meta-analysis provides an up-to-date overview of the effect of prehabilitation literature incorporating interventions aimed to increase physical activity levels in patients awaiting major abdominal surgery. Prehabilitation is a broad term encompassing various interventions with different theoretical underlying mechanisms but have a common goal to improve a person’s ability to cope with the stressors of surgery. The ultimate goal is to improve outcomes after surgery, so this is just the first step of the proof of concept that it is possible. Narrowing the focus to a specific area of prehabilitation, physical activity, provided meaningful analysis of the practicalities of the intervention and a clearer understanding of its effectiveness.

In conclusion, increasing physical activity in the preoperative phase is feasible. Self-reported physical activity outcome measures show larger effects compared to device measured physical activity outcome measures. To gather reliable information into performed physical activity, device measures should be used, utilizing standardized configurations.

REFERENCES

1. Kommalapati A, Tella SH, Goyal G, Ma WW, Mahipal A. Contemporary management of localized resectable pancreatic cancer. Cancers (Basel). 2018;10(1):1–15.
2. Janssen TL, de Vries J, Lodder P, Faes MC, Ho GH, Gobardhan PD, et al. The effects of elective aortic repair, colorectal cancer surgery and subsequent postoperative delirium on long-term quality of life, cognitive functioning and depressive symptoms in older patients. Aging Ment Heal [Internet]. 2021;25(5):896–905. Available from: https://doi.org/10.1080/13607863.2020.1725807
3. Pinto A, Faiz O, Davis R, Almoudaris A, Vincent C. Surgical complications and their impact on patients' psychosocial well-being: A systematic review and meta-analysis. BMJ Open. 2016;6(2).
4. Mina DS, Scheede-Bergdahl C, Gillis C, Carli F. Optimization of surgical outcomes with prehabilitation. Appl Physiol Nutr Metab. 2015;40(9):966–9.
5. Brandeis JM, Litwin MS, Burnison CM, Reiter RE. Quality of life outcomes after brachytherapy for early stage prostate cancer. J Urol. 2000;163(3):851–7.
6. Phillips-Bute B, Mathew JP, Blumenthal JA, Grocott HP, Laskowitz DT, Jones RH, et al. Association of neurocognitive function and quality of life 1 year after Coronary Artery Bypass Graft (CABG) surgery. Psychosom Med. 2006;68(3):369–75.
7. Galvão DA, Newton RU. Review of exercise intervention studies in cancer patients. J Clin Oncol. 2005;23(4):899–909.
8. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451–62.
9. Mylius CF, Krijnen WP, Takken T, Lips DJ, Eker H, van der Schans CP, et al. Objectively measured preoperative physical activity is associated with time to functional recovery after hepato-pancreato-biliary cancer surgery: a pilot study. Perioper Med. 2021;10(1):1–12.
10. West MA, Parry MG, Lythgoe D, Barben CP, Kemp GJ, Grocott MPW, et al. Cardiopulmonary exercise testing for the prediction of morbidity risk after rectal cancer surgery. Br J Surg. 2014 Aug;101(9):1166–72.
11. Lee L, Schwartzman K, Carli F, Zavorsky GS, Li C, Charlebois P, et al. The association of the distance walked in 6 min with pre-operative peak oxygen consumption and complications 1 month after colorectal resection. Anaesthesia. 2013;68(8):811–6.
12. Nutt CL, Russell JC. Use of the pre-operative shuttle walk test to predict morbidity and mortality after elective major colorectal surgery. Anaesthesia. 2012 Aug;67(8):839–49.
13. Saratzis A, Shakespeare J, Jones O, Bown MJ, Mahmood A, Imray CHE. Pre-operative Functional Cardiovascular Reserve Is Associated with Acute Kidney Injury after Intervention. Eur J Vasc Endovasc Surg [Internet]. 2017;53(5):717–24. Available from: http://dx.doi.org/10.1016/j.ejvs.2017.01.014
14. Wilson RJT, Davies S, Yates D, Redman J, Stone M. Impaired functional capacity is associated with all-cause mortality after major elective intra-abdominal surgery. Br J Anaesth. 2010;105(3):297–303.
15. Snowden CP, Prentis J, Jacques B, Anderson H, Manas D, Jones D, et al. Cardiorespiratory fitness predicts mortality and hospital length of stay after major elective surgery in older people. Ann Surg. 2013;257(6):999–1004.

16. Hoogeboom TJ, Dronkers JJ, Hulzebos EHJ, van Meeteren NLU. Merits of exercise therapy before and after major surgery. Curr Opin Anaesthesiol. 2014 Apr;27(2):161–6.

17. Skender S, Ose J, Chang-Claude J, Paskow M, Brühmann B, Siegel EM, et al. Accelerometry and physical activity questionnaires - A systematic review. BMC Public Health [Internet]. 2016;16(1):1–10. Available from: http://dx.doi.org/10.1186/s12889-016-3172-0

18. Seves BL, Hoekstra F, Schoenmakers JWA, Brandenbarg P, Hoekstra T, Hettinga FJ, et al. Test-retest reliability and concurrent validity of the Adapted Short QUestionnaire to ASsess Health-enhancing physical activity (Adapted-SQUASH) in adults with disabilities. J Sports Sci [Internet]. 2021;39(8):875–86. Available from: https://doi.org/10.1080/02640414.2020.1850983

19. Harris TJ, Owen CG, Victor CR, Adams R, Ekelund U, Cook DG. A comparison of questionnaire, accelerometer, and pedometer: Measures in older people. Med Sci Sports Exerc. 2009;41(7):1392–402.

20. Helmerhorst HJF, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act. 2012;9.

21. Sember V, Meh K, Sorić M, Jurak G, Starc G, Rocha P. Validity and reliability of international physical activity questionnaires for adults across eu countries: Systematic review and meta analysis. Int J Environ Res Public Health. 2020;17(19):1–23.

22. Althubaiti A. Information bias in health research: Definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211–7.

23. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. Practical guide to measuring physical activity. J Acad Nutr Diet. 2014;114(2):199–208.

24. Moran J, Guinan E, McCormick P, Larkin J, Mockler D, Hussey J, et al. The ability of prehabilitation to influence postoperative outcome after intra-abdominal operation: A systematic review and meta-analysis. Surgery. 2016 Nov;160(5):1189–201.

25. Luther A, Gabriel J, Watson RP, Francis NK. The Impact of Total Body Prehabilitation on Post-Operative Outcomes After Major Abdominal Surgery: A Systematic Review. World J Surg [Internet]. 2018;42(9):2781–91. Available from: https://doi.org/10.1007/s00268-018-4569-y

26. Hijazi Y, Gondal U, Aziz O. A systematic review of prehabilitation programs in abdominal cancer surgery. Int J Surg [Internet]. 2017;39:156–62. Available from: http://dx.doi.org/10.1016/j.ijsu.2017.01.111

27. Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;6(7).

28. Fennell C. Effects of Supervised Training Compared to Unsupervised Training on Physical Activity, Muscular Endurance, and Cardiovascular Parameters. MOJ Orthop Rheumatol. 2016;5(4):265–70.

29. Cashin AG, McAuley JH. Clinimetrics: Physiotherapy Evidence Database (PEDro) Scale. J Physiother [Internet]. 2020;66(1):59. Available from: https://doi.org/10.1016/j.jphys.2019.08.005

30. de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust J Physiother [Internet]. 2009;55(2):129–33. Available from: http://dx.doi.org/10.1016/S0004-9514(09)70043-1

31. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Internet]. Vienna, Austra; 2013. Available from: http://www.r-project.org

32. Schwarzer G, Mair P, Hatzinger R. meta : An R Package for Meta-Analysis meta : An R Package for Meta-Analysis. CranRstudioOrg [Internet]. 2016;7(January):40–5. Available from: https://cran.rstudio.org/doc/Rnews/Rnews_2007-3.pdf#page=40

33. Hedges L V. Distribution Theory for Glass's Estimator of Effect size and Related Estimators. J Educ Stat. 1981;6(2):107–28.

34. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58.

35. Halliday LJ, Doganay E, Wynter-Blyth V, Osborn H, Buckley J, Moorthy K. Adherence to Pre-operative Exercise and the Response to Prehabilitation in Oesophageal Cancer Patients. J Gastrointest Surg Off J Soc Surg Aliment Tract. 2021 Apr;25(4):890–9.

36. Alejo LB, Pagola-Aldazabal I, Fiuza-Luces C, Huerga D, de Torres MV, Verdugo AS, et al. Exercise prehabilitation program for patients under neoadjuvant treatment for rectal cancer: A pilot study. J Cancer Res Ther [Internet]. 2019;15(1):20–5. Available from: http://www.cancerjournal.net/article.asp?issn=0973-1482;year=2019;volume=15;issue=1;spage=20;epage=25;aulast=Alejo

37. Bousquet-Dion G, Awasthi R, Loiselle SÈ, Minnella EM, Agnihotram R V., Bergdahl A, et al. Evaluation of supervised multimodal prehabilitation programme in cancer patients undergoing colorectal resection: a randomized control trial. Acta Oncol (Madr) [Internet]. 2018;57(6):849–59. Available from: https://doi.org/10.1080/0284186X.2017.1423180

38. Moug SJ, Mutrie N, Barry SJE, Mackay G, Steele RJC, Boachie C, et al. Prehabilitation is feasible in patients with rectal cancer undergoing neoadjuvant chemoradiotherapy and may minimize physical deterioration: results from the REx trial. Colorectal Dis. 2019 May;21(5):548–62.

39. Steffens D, Young J, Beckenkamp PR, Ratcliffe J, Rubie F, Ansari N, et al. Feasibility and acceptability of a preoperative exercise program for patients undergoing major cancer surgery: results from a pilot randomized controlled trial. Pilot feasibility Stud. 2021 Jan;7(1):27.

40. Barberan-Garcia A, Ubré M, Roca J, Lacy AM, Burgos F, Risco R, et al. Personalised Prehabilitation in High-risk Patients Undergoing Elective Major Abdominal Surgery : A Randomized Blinded Controlled Trial. Ann Surg [Internet]. 2018;267(1):50–6. Available from: https://www.embase.com/search/results?subaction=viewrecord&id=L620048007&from=export

41. Singh F, Newton RU, Baker MK, Spry NA, Taaffe DR, Galvão DA. Feasibility and Efficacy of Presurgical Exercise in Survivors of Rectal Cancer Scheduled to Receive Curative Resection. Clin Colorectal Cancer [Internet]. 2017;16(4):358–65. Available from: http://dx.doi.org/10.1016/j.clcc.2017.03.010

42. Gillis C, Li C, Lee L, Awasthi R, Augustin B, Gamsa A, et al. Prehabilitation versus Rehabilitation: A Randomized Control Trial in Patients Undergoing Colorectal Resection for Cancer. Anesthesiology. 2014;121(5):937–47.

43. Waterland JL, Ismail H, Granger CL, Patrick C, Denehy L, et al. Prehabilitation in high-risk patients scheduled for major abdominal cancer surgery: a feasibility study. Perioper Med. 2022;11(1):1–13.

44. Chmelo J, Phillips AW, Greystoke A, Charman SJ, Avery L, Hallsworth K, et al. A feasibility trial of prehabilitation before oesophagogastric cancer surgery using a multi-component home-based exercise programme: the ChemoFit study. Pilot Feasibility Stud [Internet]. 2022;8(1):1–10. Available from: https://doi.org/10.1186/s40814-022-01137-6

45. Suen M, Liew A, Turner JD, Khatri S, Lin Y, Raso KL, et al. Short-term multimodal prehabilitation improves functional capacity for colorectal cancer patients prior to surgery. Asia Pac J Clin Oncol. 2021 Apr;

46. Timmerman H, de Groot JF, Hulzebos HJ, de Knikker R, Kerkkamp HEM, van Meeteren NLU. Feasibility and preliminary effectiveness of preoperative therapeutic exercise in patients with cancer: a pragmatic study. Physiother Theory Pract. 2011 Feb;27(2):117–24.

47. Carli F, Charlebois P, Stein B, Feldman L, Zavorsky G, Kim DJ, et al. Randomized clinical trial of prehabilitation in colorectal surgery. Br J Surg. 2010;97(8):1187–97.

48. West MA, Loughney L, Lythgoe D, Barben CP, Sripadam R, Kemp GJ, et al. Effect of prehabilitation on objectively measured physical fitness after neoadjuvant treatment in preoperative rectal cancer patients: a blinded interventional pilot study. Br J Anaesth. 2015 Feb;114(2):244–51.

49. Loughney L, West MA, Dimitrov BD, Kemp GJ, Grocott MP, Jack S. Physical activity levels in locally advanced rectal cancer patients following neoadjuvant chemoradiotherapy and an exercise training programme before surgery: a pilot study. Perioper Med (London, England). 2017;6:3.

50. Li C, Carli F, Lee L, Charlebois P, Stein B, Liberman AS, et al. Impact of a trimodal prehabilitation program on functional recovery after colorectal cancer surgery: A pilot study. Surg Endosc. 2013;27(4):1072–82.

51. Ngo-Huang A, Parker NH, Wang X, Petzel MQB, Fogelman D, Schadler KL, et al. Home-based exercise during preoperative therapy for pancreatic cancer. Langenbeck's Arch Surg. 2017;402(8):1175–85.

52. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. Am J Public Health. 2011;101(4):751–8.

53. Conn VS, Hafdahl AR, Brown SA, Brown LM. Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. Patient Educ Couns. 2008;70(2):157–72.

54. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med. 2019;9(1):147–57.

55. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The Effectiveness of Interventions to Increase Physical Activity A Systematic Review. Am J Prev Med. 2002;22(4S).

56. Sevick MA, Dunn AL, Morrow MS, Marcus BH, Chen GJ, Blair SN. Cost-effectiveness of lifestyle and structured exercise interventions in sedentary adults. Results of project ACTIVE. Am J Prev Med. 2000;19(1):1–8.

57. Stults-Kolehmainen MA, Sinha R. The effects of stress on physical activity and exercise. Vol. 44, Sports Medicine. 2014. 81–121 p.

58. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: A scoping review of methods used. Transl Behav Med. 2018;8(2):212–24.

59. Colombo D, Suso-Ribera C, Fernández-Álvarez J, Cipresso P, Garcia-Palacios A, Riva G, et al. Affect Recall Bias: Being Resilient by Distorting Reality. Cognit Ther Res [Internet]. 2020;44(5):906–18. Available from: https://doi.org/10.1007/s10608-020-10122-3

60. Douma JAJ, de Beaufort MB, Kampshoff CS, Persoon S, Vermaire JA, Chinapaw MJ, et al. Physical activity in patients with cancer: self-report versus accelerometer assessments. Support Care Cancer. 2020;28(8):3701–9.

61. Grimm P. Social Desirability Bias. 2010;(1999):2010.

62. Redelmeier DA, Dickinson VM. Determining whether a patient is feeling better: Pitfalls from the science of human perception. J Gen Intern Med. 2011;26(8):900–6.

63. Thomas G, Tahir MR, Bongers BC, Kallen VL, Slooter GD, Van Meeteren NL. Prehabilitation before major intra-abdominal cancer surgery: A systematic review of randomised controlled trials. Eur J Anaesthesiol [Internet]. 2019;36(12):933–45. Available from: https://www.embase.com/search/results?subaction=viewrecord&id=L630501968&from=export

64. Herbert RD, Bø K. Analysis of quality of interventions in systematic reviews. Br Med J. 2005;331(7515):507–9.

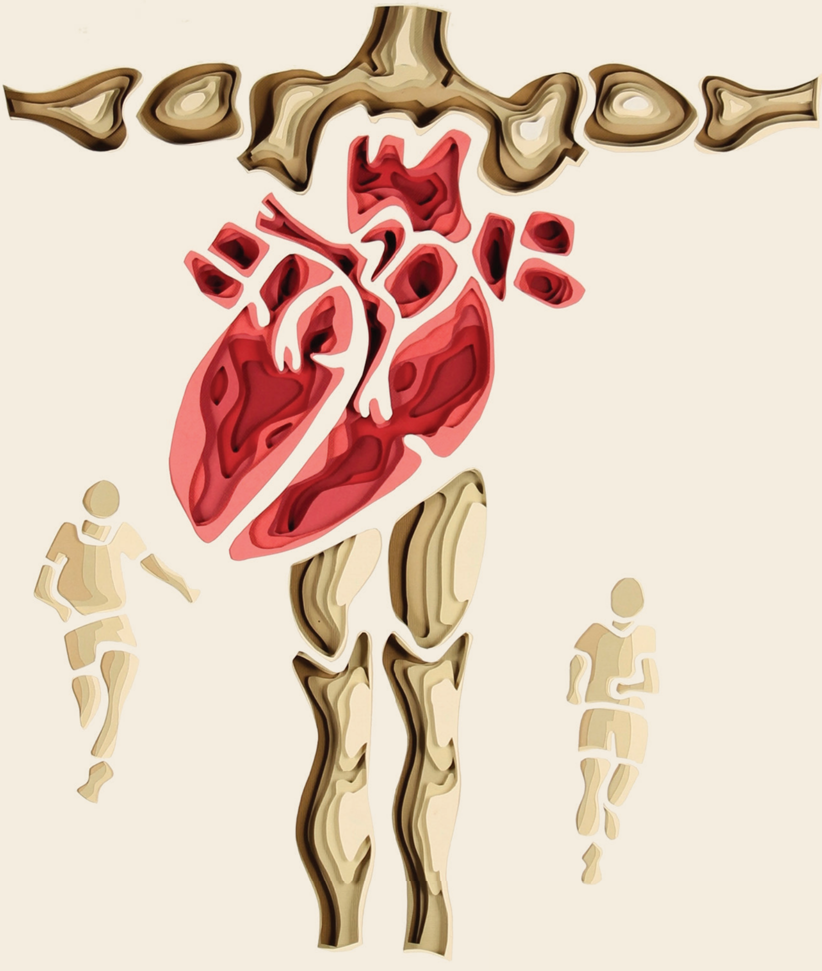
SUPPLEMENTARY MATERIAL A

EMBASE

(((((exercise training program) OR (physical capacity/de)) OR (physical fitness)) OR (physical activities)) OR (physical activity/de)))
AND ((((((((((Intestines) OR (Abdominal)) OR (Colorectal)) OR (Hepatic)) OR (Pancreatic)) OR (Liver)) OR (Gastric)) OR (Esophageal)) OR (Rectal)) OR (Colon)) OR (HIPEC)))
AND ((Surgery/de) OR (Operation)))
AND ((Prehabilitation) OR (preoperative period)))
NOT ((((Antibiotic agent) OR (Protein/de)) OR (anesthesia)) OR (breathing exercise))
AND (((((("2000"[Date - Publication] : "3000"[Date - Publication]))

PUBMED

(((((exercise training program) OR (physical capacity)) OR (physical fitness[MeSH Terms])) OR (physical activities[MeSH Terms])) OR (physical activity[MeSH Terms])))
AND ((((((((((Intestines[MeSH Terms]) OR (Abdominal)) OR (Colorectal)) OR (Hepatic)) OR (Pancreatic)) OR (Liver)) OR (Gastric)) OR (Esophageal)) OR (Rectal)) OR (Colon)) OR (HIPEC)))
AND ((Surgery) OR (Operation)))
AND ((Prehabilitation) OR (Preoperative)))
NOT ((((Antibiotics) OR (Protein)) OR (Anaesthesia)) OR (Breathing therapy))
AND (((((("2000"[Date - Publication] : "3000"[Date - Publication]))



Chapter 7

Summary & General Discussion

This thesis determined the reference values for exercise capacity tests in the general population (**Chapter 2, 3, and 4**) and evaluated the effect of physical activity in patients awaiting major abdominal surgery (**Chapter 5 and 6**). An overview of the main findings per study and concluding remarks is provided below.

One specific objective of this thesis was to determine reference distance values for the six-minute walk test amongst the healthy pediatric and adolescent population. For this purpose, a systematic literature review was conducted in which the results of six-minute walk tests were combined and evaluated (**Chapter 2**). A total of 22 studies were included in the systematic review, of which eight studies were performed in Asia, seven in Europe, three in north America, three south America and one study was situated in Africa. Reference values were reported as either absolute values or by reference equation. The systematic review could not provide a single set of reference values based on a meta-analysis due to variations in applied procedures amongst studies (e.g., given instructions, encouragement, track lay-out and, preparations). The reference equation yielding the highest R^2 was based on both male and female subjects between the age of 6 till 16 years old. The equation was as followed: distance = (4.63 x height (in centimeters)) – (3.53 x weight (in kilograms)) + (10.42 x age (in years)) + 56.32. The systematic review highlighted that further research is needed to present a single set of reference values for the six-minute walk test amongst the pediatric and adolescent population. Furthermore, it presented a flow chart to aid the selection of reference values for the time being.

In **Chapter 3** a reference value model for the peak oxygen uptake based upon a regression model in the healthy Dutch pediatric and adult population was determined. Furthermore, to determine the external and predictive validity of the reference value prediction model, a cross-validation procedure was performed. The database used consisted of cycle ergometry results from several test institutes across the Netherlands. All included institutes used the ATS/ACCP statement equipment requirements and procedures to perform an incremental cardiopulmonary exercise test using an electromagnetically braked cycle ergometry test. Oxygen uptake analyses were performed by utilizing gas exchange analysis by bag collection, mixing chamber or breath-by-breath analysis based upon averaging the values measured during the last 30–60 s of the test. The status *healthy* was defined as the absence of any reported somatic signs of disease and the exclusion of registered available risk factors. Generalized additive models were utilized to semi parametrically find the most appropriate fitting regression model. The additive model results in a smaller standard error of the estimate especially in the   20-year-old subjects, in contrast to the linear model. This is because the additive model can adjust for age-related transformations such as the increase in peak oxygen uptake associated with the growth-related weight and height gain during childhood and adolescence.

In conclusion, the study provided a robust additive regression model for peak oxygen uptake in the Dutch population and concluded that peak oxygen uptake is sex specific and has a nonlinear relationship with age.

Chapter 4 provides an updated systematic review originally published in 2014 on the cardiopulmonary exercise test reference values in healthy subjects. Compared to the original review, a large increase in published reference values is seen on the basis of 29 eligible studies including 87256 subjects in total. As with the 6-minute walk test review, meta-analysis of the data was not meaningful, as a large heterogeneity of methods and subjects was observed. The review recommends abandoning the use of the 80% of predicted as lower limit of normal values. Instead, a Z-score should be used with a lower and upper limit of normal of –1.96 SD and +1.96 SD, respectively. Furthermore, the review concludes that future reference value study quality can be further improved by performing a power analysis, a good quality assurance of equipment and methodologies, and by validating the developed reference equation in an independent (sub)sample.

Preoperative aerobic fitness level has been identified as a modifiable risk factor in a variety of patients who need surgery. Aerobic fitness can be modified by habitual physical activity habits like performing moderate to vigorous physical activity. Therefore, a prospective multi-center study was done (**Chapter 5)**. The study had the aim to get insight into the level of actual, objectively measured physical activity performed by patients awaiting hepato-pancreato-biliary cancer resection surgery. The secondary aim of the study is to determine the association between preoperative moderate to vigorous physical activity with postoperative outcomes. Physical activity level was measured using a hip-worn activity monitor; the Actigraph wGT3X- BT+. The median moderate to vigorous physical activity was 10.7 minutes per day; only eight participants met the Dutch moderate to vigorous physical activity guidelines. Furthermore, the study showed that a higher level of moderate to vigorous physical activity was strongly associated with a shorter time to functional recovery. Namely, a multivariable linear model including surgery size and physical activity level yields an adj. R^2 .43; the model is as follows 12.54 + (moderate to vigorous physical activity (minutes) * –.08) + (surgery size (1 if minor, 0 if major) * –5.64).

Due to the relation between time to functional recovery and the amount of preoperative physical activity, **Chapter 6** describes a meta-analysis of the effect of interventions to increase physical activity levels of patients awaiting abdominal resection surgery. Additionally, a comparison between self-reported and objectively measured outcome measures to determine if the used tools result in similar outcomes. The random-effect meta-analysis showed a moderate improvement in intervention groups compared to the baseline measures (SMD= 0.69, [CI= 0.29;1.09], I2= 80%). Therefore, it was concluded that it is feasible to increase physical activity in the preoperative phase. Self-reported physical activity outcome measures show larger effects compared to objectively measured outcome measures.

Reference values for activities of daily living

Peak oxygen uptake represents the functional limit of the body's ability to deliver and extract oxygen in muscles to satisfy the metabolic demands of vigorous exercise. The oxygen consumption above which aerobic energy production is supplemented by anaerobic mechanisms, causing a sustained increase in lactate and metabolic acidosis, is termed the anaerobic threshold. (1) Both the peak oxygen uptake, as described in **chapter 3 & 4,** and the ventilatory threshold(s) are currently the main parameters utilized in the clinical reasoning process when performing cardiopulmonary exercise testing to evaluate a subjects cardiopulmonary fitness level. (2) As described, these values reflect the functional limit of the body's ability to deliver and extract oxygen in muscles to satisfy the metabolic demands of vigorous exercise at the maximal aerobic and peak intensity. Although it provides insight into a subject's cardiopulmonary fitness, the clinical relevance to perform activities of daily living throughout the day may be limited. In people without cardiopulmonary or musculoskeletal impairments, the cardiopulmonary reserve capacity is thought to be barely tapped during activities of daily living. (3) However, in people with pathology, this reserve can be greatly reduced, and a greater than usual proportion of a person's capacity may be needed to perform these activities. (4) To determine if a subject has sufficient cardiopulmonary capacity to functionally perform activities of daily living, the evaluation of the submaximal exercise capacity compared to the relative physical requirements of these activities might potentially provide a more informative insight.

To be able to functionally perform and maintain activities throughout the day, the majority of these activities should be performed at a relatively low intensity level or submaximal intensity level. A subject performs activities at a submaximal intensity if the intensity stays below the first ventilatory threshold (VT_1). The VT_1 represents a level of intensity at which blood lactate accumulates faster than it can be cleared, which causes the person to breath faster to blow off the extra CO_2 produced by the buffering of acid metabolites. Prior to this threshold, only small amounts of lactate are being produced, within the capacity of the body to clear it, enabling it to maintain an activity for a prolonged period. Past the VT_1 point, ventilation rates begin to increase exponentially as oxygen demands outpace the oxygen-delivery system and lactate begins to accumulate in the blood. This threshold can be determined by analyzing the breath-by-breath carbon dioxide production (VCO_2) and ventilation during incremental exercise testing.

The goal of submaximal exercise testing is to produce a sufficient level of exercise stress without physiologic or biomechanical strain resulting in an intensity higher than the VT_1. This capacity is typically measured through performance-based testing, such as the 6 minute walk test, which has been examined in depth in **Chapter 2.** (5) Additional examples of these are, amongst others, the Shuttle Walk Test, the 12-Minute Walk Test or the Timed Up & Go test. These tests involve measuring the ability to perform standardized physical activities for a preset period or set of repetitions that are typically encountered in everyday life. (5) Although the results are influenced by many components like musculoskeletal-, or neuromuscular impairments, or the use of aids, submaximal exercise testing provides insight into the subject's capacity to perform the specific task. Due to the reflection of daily activities, submaximal exercise testing appears to have greater applicability to physiotherapists in their role as clinical exercise specialists compared with maximal exercise testing. Nevertheless, the intensity of the performance is unclear since breath-by-breath analysis is not included in the protocol and thus not performed in the clinical setting. Therefore, it is possible that a participant performs at a lower or higher intensity than the VT_1.

Physiotherapy is indicated when a subject lacks the capacity to perform activities of daily living at a submaximal level. To identify subjects that require physiotherapy treatment, adequate measurements should be performed. Although there are numerous submaximal tests proposed. These exercise tests and their applications have been less well developed compared to maximal exercise tests. Reports in the literature on the common submaximal tests, vary with respect to the adequacy of establishing validity, reliability, and sensitivity, limiting their interpretation with respect to decision making by the physiotherapist. Hence, physiotherapists should determine what information will be added by performing these tests and how that information will alter clinical decision making. In its recommended applications, (2) incremental cardiopulmonary exercise testing with breath-by-breath analysis consists of applying a gradually increasing intensity exercise from minimal effort until exhaustion or until the appearance of limiting symptoms and/or signs. Therefore, both the VT_1, anaerobic threshold (VT_2) and possibly the peak oxygen uptake can be determined by analyzing gas exchange. Given the availability of this measurement, it is recommended to communicate VT_1 values to treating physiotherapists after completion of the test. These values provide reliable and valid information to determine if physiotherapy is indicated with a tailormade treatment scheme, and these can additionally be used as an evaluative tool to determine progression.

To increase the clinical value of cardiopulmonary exercise testing, the oxygen uptake at VT_1 can be compared with reference values of activities of daily living. If the oxygen consumption during a selected activity surpasses oxygen consumption at VT_1 it is likely the test subject cannot perform the activity for a prolonged period. To make an estimation of a subject's ability to perform prolonged activities, reference values of oxygen uptake during activities are required. An estimation for energy expenditure or intensity of activities of daily living are the "Metabolic equivalent of tasks" (METs). (6,7) The MET concept represents a simple and practical procedure to express the energy cost of physical activities as a multiple of the resting metabolic rate. One MET is defined as 3.5 ml O_2 per kg * min^{-1}, the amount of oxygen consumed while sitting at rest. Therefore, energy cost of an activity can be determined by dividing the relative oxygen cost of the activity by 3.5. Multiple publications have attempted to quantify the requirements of physical activities, like household chores, providing a comparative basis to evaluate the submaximal capacity of a subject. (6–8) However, these values are rough estimations of activities that might have changed through the years and lack adaptions to a subject's personal situation, e.g., the precents of pathologies, differences in body composition and the execution of the task highly influence oxygen consumption. Therefore, to increase the applicability of task specific reference values, research into population specific oxygen uptake measurements during task execution are required to aid clinical decision making to determine a patient's ability to perform activities.

In conclusion, to be able to functionally perform and maintain activities of daily living throughout the day, most of these activities should be performed at a submaximal intensity level. To determine if a subject has sufficient cardiopulmonary capacity to functionally perform these activities, the evaluation of the submaximal exercise capacity and the relative physical requirements of these activities should be compared. In patient categories, actual oxygen uptake measurements during task execution are required for a valuable comparison with oxygen uptake at $VT_{1.}$ Furthermore, additional research is needed to determine the value of VT_1 reference values in relation to population participation rate.

Physical activity measurements

The impact of physical inactivity has been characterized as similar to that of smoking in relation to the burden of noncommunicable diseases. (9) Inactivity is associated with many of the most common chronic diseases and conditions, including heart disease, type 2 diabetes mellitus, hypertension, obesity, osteoporosis, depression, and breast and colorectal cancers. (10) In contrast, higher levels of physical activity have numerous health benefits, including reducing the time required for functional recovery after major surgeries such as hepato-pancreato-biliary resection surgery, as shown in **Chapter 5**. Therefore, adequately identifying inactivity is desirable. At present, there is little information available to guide the selection of a physical activity assessment method that is appropriate for the wide variety of potential applications. There are two broad categories of methods available to assess a person's level of physical activity: subjective methods and objective methods. Subjective methodologies rely on the individuals either to record activities as they occur or to recall previous activities. Objective methodologies, as utilized in **Chapter 5,** include all wearable monitors that directly measure one or more bio signals, such as acceleration, heart rate, or some other indicator of physical activity or energy expenditure, as they occur.

Subjective methods that are commonly used include diaries, logs, interviews, and questionnaires. Participants are asked to keep track of or recall the physical activity they have or are performing. These methods are influenced by socioeconomic and sociodemographic factors, and measurement bias including misinterpretation or deliberate changes (e.g., social desirability) and recall bias. (11,12) Hence, the accuracy, volume, and duration of the memory may be influenced by subsequent events and experiences resulting in over- or underestimation due to interference of other constructs like the experienced effort. (13) Therefore, it is probable that subjects with less physical fitness overestimate performed physical activities. Nevertheless, subjective methods are highly feasible with low cost to use thus can therefore provide a gross estimation of physical activity levels in large groups and can provide immediate insight into activity levels of individuals.

Objective measures like motion sensors (e.g., accelerometers and pedometers), biomarkers and calorimetry are prone to similar pitfalls and benefits. Given the direct measurement of this method, no recall bias or distortion due to intensity or duration is involved. However, measurement accuracy is dependent on consistently wearing the device, the wearing location and the activity performed. For example, not all accelerometers are capable of measuring water-based activities, cycling is inaccurately measured with a hip or wrist worn pedometer, and activities involving upper extremity movements are inaccurately measured with lower extremity placed devices. Furthermore, since the measurement is prospective, it is time and cost intensive. Hence, participants must wear the device for a prolonged period and therefore this method is difficult to administer to a large-scale population. (14–16) Given these characteristics, both objective or subjective measurement methods can be preferred depending on the setting and goal. Since objective measures are more accurate and valid to determine physical activity levels, this method should be recommended to identify at risk patients due to inactivity or to evaluate treatment progression. (17,18) In contrast, to determine the physical activity levels of the general public, subjective methods should be preferred since it is an inexpensive, feasible method that provides a gross estimation.

Objective measurements of performed physical activity do not provide information on the perceived intensity of the physical activity. Perceived exertion has been defined as the subjective rating of the intensity of physical work and can be quantified by Borgs 'Rate of Perceived Exertion' (RPE). (19) The processing of sensory cues related to physical performance enables an individual to perceive general feelings of exertion incorporating more than solely fatigue. The interference of multiple constructs when measuring RPE is apparent given the lower correlation between RPE measurements and biomarkers in individuals with lower physical fitness levels. (20) It has been suggested that the RPE is a merger of many feelings and sensations related to the physical fatigue and activity including effort and exertion. (21,22) The Borg RPE is an affordable, practical, and valid tool for monitoring and prescribing physical activity intensity, independent of gender, age, exercise modality, and physical activity level. (23) Monitoring RPE can be valuable addition to identify changes in patients' perceived effort to perform physical activity. (21) Therefore, given the importance of early detection of complicated care paths, screening of mental wellbeing and minimizing perceived fatigue to maintain physical activity (24), the Borg RPE tool might proof valuable when added to a point of care diagnostic device designed to objectively measure physical activity levels. (25)

In conclusion, adequately identifying inactivity and drops in activity levels is desirable to provide timely interventions. Objective measures are recommended to identify at risk subjects due to timely measurement and lack of recall bias. The addition of RPE measurements provide healthcare providers with insight into the perceived effort and exertion of a patient.

Preoperative physical activity

Performing physical activities can optimize physical fitness, (26) which is beneficial in major (abdominal) surgery patients. Especially in high-risk patients, optimizing physical fitness contributes to post surgery recovery. (27) Increases in, or maintaining sufficient levels of, physical activity should be identified as pursuable treatment outcomes if performed physical activity levels are low or decreasing. Therefore, preoperative physiotherapeutic care should include education on the importance of, address limiting factors to perform, and provide guidance to achieve desired levels of physical activity in high-risk patients. Furthermore, the reported correlation between performed moderate to vigorous physical activity and the duration of time to functional recovery, provides arguments to incorporate remote monitoring of performed physical activity via point of care devices in the care path of patients awaiting major abdominal surgery. (25) Moreover, increased adherence to a prehabilitation program is reported by the utilization of point of care devices in colorectal surgery. (28) Additionally, the integration of RPE measurements via these devices provide an easily accessible and informative addition to physical activity measurements. (21) Therefore, regular physical activity monitoring and RPE measurements should be integrated in the preoperative care-path to early identify at risk patients experiencing problems to perform activities and/or monitor therapy adherence. (28)

Although this dissertation focusses on prehabilitation and physical activity levels in major abdominal surgery patients, similar protective mechanisms are likely to be present in other major surgery patients. Therefore, implementing monitoring of physical activity and perceived exertion levels will probably provide valuable information in high-risk patients scheduled for other major surgery types. Furthermore, this study found moderate improvements in physical activity levels during the preoperative phase are feasibly. However, evidence that improving physical activity levels consequently reduces postoperative complications are inconclusive. Therefore, research into the incidence and impact of postoperative complications, length of hospitalization and quality of life after the completion of a prehabilitation scheme aimed at increasing physical activity will contribute to understanding the influence of physical activity on surgery outcomes.

To summarize, interventions aimed at improving physical fitness and physical activity levels are indicated when a subject does not have the capacity to perform activities of daily living. Moreover, preoperative physical inactivity should be addressed to potentially reduce length of hospital stay in major (abdominal) surgery.

REFERENCES

1. Wasserman K. The anaerobic threshold: definition, physiological significance and identification. Adv Cardiol. 1986;35:1–23.
2. Ross R, Beck K, Casaburi R, Johnson B, Marciniuk D, Wagner P, et al. ATS/ACCP Statement on Cardiopulmonary Exercise Testing. Am J Respir Crit Care Med. 2003;167(10):1451.
3. Fletcher GF, Froelicher V, Hartley H, Haskell WL, Pollock ML. Exercise Standards: A Statement for Health Professionals From the American Heart Association. Circulation. 1990;82:2286–322.
4. Hagberg JM. Exercise assessment of arthritic and elderly individuals. Baillieres Clin Rheumatol. 1994;8(1):29–52.
5. Noonan V, Dean E. Submaximal exercise testing: Clinical application and interpretation. Phys Ther. 2000;80(8):782–807.
6. Weir JB d. V. New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol. 1949;109(1–2):1–9.
7. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol. 1990;13(8):555–65.
8. de Almeida Mendes M, da Silva I, Ramires V, Reichert F, Martins R, Ferreira R, et al. Metabolic equivalent of task (METs) thresholds as an indicator of physical activity intensity. PLoS One. 2018;13(7):1–10.
9. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet. 2012;380(9838):219–29.
10. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451–62.
11. Chinapaw MJM, Mokkink LB, Van Poppel MNM, Van Mechelen W, Terwee CB. Physical activity questionnaires for youth: A systematic review of measurement properties. Sport Med. 2010;40(7):539–63.
12. Sabia S, Van Hees VT, Shipley MJ, Trenell MI, Hagger-Johnson G, Elbaz A, et al. Association between questionnaire-and accelerometer-assessed physical activity: The role of sociodemographic factors. Am J Epidemiol. 2014;179(6):781–90.
13. Lee PH, Macfarlane DJ, Lam T, Stewart SM. Validity of the international physical activity questionnaire short form. Int J Behav Nutr Phys Act. 2011;8(1):1–11.
14. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. Int J Behav Nutr Phys Act. 2008;5.
15. Skender S, Ose J, Chang-Claude J, Paskow M, Brühmann B, Siegel EM, et al. Accelerometry and physical activity questionnaires - A systematic review. BMC Public Health [Internet]. 2016;16(1):1–10. Available from: http://dx.doi.org/10.1186/s12889-016-3172-0
16. Gobbi E, Elliot C, Varnier M, Carraro A. Psychometric properties of the physical activity questionnaire for older children in Italy: Testing the validity among a general and clinical pediatric population. PLoS One. 2016;11(5):1–12.

17. Bowles HR. Measurement of active and sedentary behaviors: closing the gaps in self-report methods. J Phys Act Health. 2012;9 Suppl 1(Suppl 1):7–10.

18. Freedson, P., Bowles, H.R., Troiano, R. & Haskell W. Monitors : Recommendations for Monitor. Med Sci Sports Exerc. 2012;44(1 Suppl 1):S1–4.

19. Sullivan SBO. Perceived Exertion: A Review. Phys Ther. 1984;64(3):343–6.

20. Travlos AK, Marisi DQ. Perceived Exertion During Physical Exercise Among Individuals High and Low in Fitness. Percept Mot Skills. 1996;(82):419–24.

21. Williams N. The Borg Rating of Perceived Exertion (RPE) scale. Occup Med (Chic Ill). 2017;67(5):404–5.

22. Abbiss CR, Peiffer JJ, Meeusen R, Skorski S. Role of Ratings of Perceived Exertion during Self-Paced Exercise: What are We Actually Measuring? Sport Med. 2015;45(9):1235–43.

23. Scherr J, Wolfarth B, Christle JW, Pressler A, Wagenpfeil S, Halle M. Associations between Borg's rating of perceived exertion and physiological measures of exercise intensity. Eur J Appl Physiol. 2013 Jan;113(1):147–55.

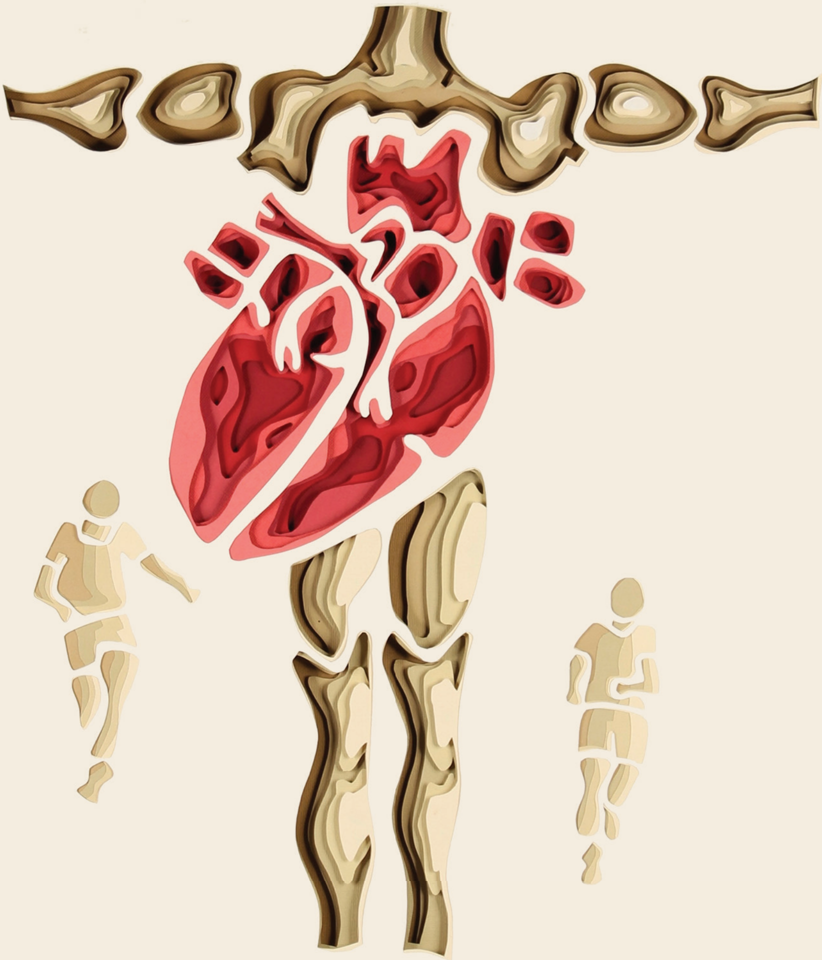
24. Egerton T, Chastin SFM, Stensvold D, Helbostad JL. Fatigue May Contribute to Reduced Physical Activity Among Older People: An Observational Study. Journals Gerontol - Ser A Biol Sci Med Sci. 2016;71(5):670–6.

25. Vashist SK. Point-of-care diagnostics: Recent advances and trends. Biosensors. 2017;7(4):10–3.

26. Vaara JP, Vasankari T, Wyss T, Pihlainen K, Ojanen T, Raitanen J, et al. Device-Based Measures of Sedentary Time and Physical Activity Are Associated With Physical Fitness and Body Fat Content. Front Sport Act Living. 2020;2(December):1–13.

27. Thomas G, Tahir MR, Bongers BC, Kallen VL, Slooter GD, Van Meeteren NL. Prehabilitation before major intra-abdominal cancer surgery: A systematic review of randomised controlled trials. Eur J Anaesthesiol. 2019;36(12):933–45.

28. Kimura CS, Bidwell S, Gurland B, Morris A, Shelton A, Kin C. Association of an Online Home-Based Prehabilitation Program With Outcomes After Colorectal Surgery. JAMA Surg [Internet]. 2022 Nov 2; Available from: https://doi.org/10.1001/jamasurg.2022.4485



Appendices

SAMENVATTING

Voldoende fysieke fitheid is belangrijk omdat dit bepaalt of een persoon dagelijks activiteiten kan uitvoeren en omdat het ziektes en beperkingen kan voorkomen. Fysieke fitheid bestaat uit cardiorespiratoire uithoudingsvermogen, vaardigheden zoals balans en coördinatie evenals gezondheid gerelateerde factoren zoals, lichaamssamenstelling, spierkracht en spieruithoudingsvermogen. Om te bepalen wat de mate van fysieke fitheid van de algehele Nederlandse populatie is, werden in dit proefschrift de referentiewaarden voor inspanningstesten in de algemene gezonde bevolking bepaald. Ook is het effect van fysieke activiteit bij patiënten die wachten op een grote buikoperatie onderzocht om te evalueren wat het effect van bewegen op de operatie uitkomsten zijn.

Fysiotherapeuten gebruiken frequent de zes minuten wandeltest als meetinstrument om de fysieke fitheid te bepalen. In het eerste deel van dit proefschrift, **hoofdstuk twee,** is een literatuuronderzoek beschreven waarin referentiewaarden voor deze zes minuten wandeltest bij kinderen en adolescenten zijn geëvalueerd. In dit literatuuronderzoek werden 22 studies geïncludeerd. De beste referentievergelijking voor zowel jongens als meisjes tussen de 6 en 16 jaar oud was als volgt: afstand = (4,63 * hoogte (in centimeters)) - (3,53 * gewicht (in kilogram)) + (10,42 * leeftijd (in jaren)) + 56,32.

Een andere veel gebruikte maat van fysieke fitheid is het maximale vermogen van mensen om zuurstof te op te nemen tijdens inspanning. Dit is een test waarbij wordt gekeken hoe goed het hart en de longen functioneren tijdens lichamelijke inspanning. In het **derde hoofdstuk** is er een referentiemodel gemaakt voor de maximale hoeveelheid zuurstof te voorspellen die mensen kunnen opnemen tijdens een cardiopulmonaire inspanningstest. Voor de test moesten mensen met behulp van een ergometer fiets zich maximaal inspannen en werd er gemeten hoeveel zuurstof ze opnamen. Op basis van een dataset van 4477 test uitslagen van gezonde Nederlandse kinderen en volwassenen is een model bepaald welke het beste aansluit bij de dataset. Voor de bepaling van gezonde deelnemers, zijn er geen mensen met gediagnostiseerde ziekte of mensen met vooraf bepaalde risicofactoren meegenomen in de test. Om representativiteit van het model te verbeteren, zijn data gebruikt vanuit verschillende testcentra in Nederland. Het model welke het beste bij de data paste was een additief model. Uit de studie bleek dat de maximale zuurstofopname afhankelijk is van geslacht en dat de relatie tussen zuurstofopname en leeftijd niet lineair is.

In aanvulling op het bepalen van referentiemodel in het derde hoofdstuk is in **hoofdstuk vier** een geactualiseerde review gepresenteerd over referentiewaarden voor gezonde mensen tijdens dezelfde cardiopulmonaire inspanningstest. In vergelijking met de oorspronkelijke review uit 2014 zijn er meer studies beschikbaar waaruit referentiewaarden zijn afgeleid. In totaal zijn er 29 studies met 87.256 proefpersonen beoordeeld. Helaas was het ook hier niet mogelijk om de gegevens van deze studies te combineren in een meta-analyse vanwege de grote verschillen in de methoden en de proefpersonen die werden gebruikt. Op basis van de uitkomst van het review wordt geadviseerd om een Z-score te gebruiken om de normale waarden te bepalen. Dit is een statistische maat die aangeeft hoeveel standaarddeviaties een bepaalde waarde afwijkt van het gemiddelde. De review concludeert dat toekomstige studies die referentiewaarden bepalen, verbeterd kunnen worden door onder meer een power-analyse te gebruiken, goede kwaliteitsborging van de apparatuur en methodologieën en door de ontwikkelde referentievergelijking te valideren in een onafhankelijke groep proefpersonen.

Het tweede deel van het proefschrift richt zich op de toepassing van fysieke fitheidsmaten bij mensen die een grote buikoperatie moeten ondergaan. De mate van lichamelijke fitheid die iemand heeft voor de operatie is een factor die mogelijk het risico op complicaties kan beïnvloeden. Algemeen wordt aangenomen dat fitheid kan worden verbeterd door regelmatig fysieke actief te zijn, waardoor dit kansen kan bieden om het risico op postoperatieve complicaties te beïnvloeden. **Hoofdstuk vijf** beschrijft een onderzoek welke is uitgevoerd om te begrijpen hoeveel lichaamsbeweging patiënten uitvoeren voordat ze een resectieoperatie ondergaan voor kanker van de lever, alvleesklier of galblaas. Het secundaire doel van dit onderzoek was om te kijken of er een verband was tussen de hoeveelheid matige tot intensieve lichaamsbeweging die patiënten uitvoerden en de resultaten na de operatie. Het niveau van lichaamsbeweging werd gemeten met een activiteitenmonitor die aan de heup werd gedragen. Het bleek dat het mediane niveau van matige tot intensieve lichaamsbeweging slechts 10,7 minuten per dag was, en slechts acht deelnemers voldeden aan de Nederlandse richtlijnen voor lichaamsbeweging. De studie liet ook zien dat een hoger niveau van matige tot intensieve lichaamsbeweging sterk geassocieerd was met een kortere tijd tot functioneel herstel na de operatie. Dit betekent dat patiënten die voor de operatie meer lichaamsbeweging uitvoerden, over het algemeen sneller herstelden na de operatie.

Aansluitend op deze bevindingen beschrijft het **zesde hoofdstuk** een systematische literatuur review welke gericht was op de vraag of het mogelijk is de hoeveelheid fysieke activiteit van patiënten die een buikoperatie ondergaan, te verhogen. De studie bevat ook een vergelijking tussen zelf-gerapporteerde en objectief gemeten uitkomstmaten om te bepalen of de gebruikte methoden tot vergelijkbare resultaten leiden. De resultaten van de studie laten zien dat interventies om fysieke activiteit te verhogen een matige verbetering in het niveau laten zien in vergelijking met de meting voorafgaande aan de interventie. Het is dus mogelijk om fysieke activiteit in de preoperatieve fase te verhogen.

Het review laat ook zien dat bij zelf gerapporteerde uitkomstmaten voor fysieke activiteit de effecten groter zijn dan bij objectief gemeten uitkomstmaten.

In conclusie, de studies in dit proefschrift bevestigen het belang van fysiek fitheid en het onderhouden van fysieke activiteit. De vastgestelde referentiewaarden en modellen bieden een basis voor het beoordelen van de fysieke fitheid van de Nederlandse bevolking, terwijl het onderzoek naar het effect van fysieke activiteit op operatieresultaten inzichten verschaft om het herstel na een operatie te bevorderen.

ABOUT THE AUTHOR

Caspar Frederik Mylius was born on January 21st, 1986, in Leeuwarden, the Netherlands. After completing a Bachelor in Physiotherapy at the Hanze University of applied sciences in 2012 he started teaching at the same university and worked as a primary care physiotherapist in Groningen. In 2017 he graduated at the Master Clinical Health Sciences at Utrecht University. Subsequently started as a PhD candidate on cardiorespiratory fitness and physical activity at the Hanze University of Applied Sciences, in Groningen.

A

DANKWOORD

Het is nu dan toch echt klaar! Terwijl ik deze woorden op papier zet, realiseer ik me dat niet alleen ikzelf, maar ook velen met mij opgelucht zullen ademhalen dat deze dissertatie nu eindelijk af is. Mijn oprechte dank gaat uit naar degenen die gedurende de afgelopen jaren betrokken zijn geweest bij dit promotietraject. De lijst met mensen die mij in welke vorm dan ook hebben geholpen is lang, maar er zijn een aantal mensen die ik graag specifiek wil bedanken.

In de eerste plaats, en bovenal, wil ik mijn diepe waardering uitspreken voor de deelnemers aan de onderzoeken. Het vertrouwen en de bereidheid om deel te nemen hebben dit werk mogelijk gemaakt. Evenzeer wil ik de toegewijde medewerkers van het Medisch Spectrum Twente, het Medisch Centrum Leeuwarden en het Universitair Medisch Centrum Groningen bedanken voor de onvermoeibare inzet, flexibiliteit en toewijding.

Mijn dank gaat ook uit naar mijn promotieteam, bestaande uit Dr. Tim Takken, Dr. Wim Krijnen, Prof. Dr. Joost Klaase en Prof. Dr. Cees van der Schans. De samenwerking gedurende deze jaren heeft mijn traject zowel inhoudelijk als persoonlijk verrijkt. De begeleiding heeft mij de ruimte gegeven om te groeien en te leren. Zonder deze steun was het niet mogelijk om dit proefschrift tot stand te brengen.

Tim, vanaf het begin van de studie Klinische Gezondheidswetenschappen heeft u mij bijgestaan in de wereld van onderzoek. Ik ben dankbaar voor uw begeleiding en uw doelgerichte, pragmatische inzichten die u altijd bereid was te delen.

Wim, uw intensieve begeleiding, toewijding en betrokkenheid hebben indruk op mij gemaakt. Elke ontmoeting met u bracht nieuwe inzichten en stof tot nadenken. U informeerde altijd belangstellend naar mijn voortgang en gemoedstoestand, en stond klaar om te helpen wanneer ik vastliep.

Joost, uw bevlogenheid en passie om de patiëntenzorg te verbeteren werkt aanstekelijk. Het hielp mij niet alleen om het grotere doel voor ogen te houden, maar heeft mij ook door moeilijke momenten geholpen. U was altijd genegen mij te helpen, uw constructieve feedback en waardevolle adviezen zijn van onschatbare waarde geweest.

Cees, hartelijk dank voor alle ruimte die u mij heeft gegeven om mijn weg te vinden. U had veel vertrouwen in mij, veel geduld en altijd bereid een luisterend oor te bieden. Ook was u, waar nodig altijd bereid mij een subtiel zetje in de goede richting te geven zodat we op de juiste route bleven. Het heeft mij geholpen grond onder mijn voeten te houden. Ook wil ik u bedanken voor uw openheid de afgelopen jaren. Vooraf hadden we nooit kunnen bedenken dat u een Case-studie over uw eigen situatie zou gaan schrijven. Bedankt dat u mij heeft betrokken in dit proces.

Leden van de beoordelingscommissie, Prof. dr. Dekker, Prof. dr. Olde Damink en prof. Dr Zwerver hartelijk dank voor uw bereidheid om plaats te nemen in de leescommissie en voor de tijd die u heeft genomen om dit manuscript te beoordelen. Ik kijk ernaar uit om met u van gedachten te kunnen wisselen.

Mijn collega's bij de Hanzehogeschool en Expertisecentrum Eerstelijnszorg Groningen, hoewel het ondoenlijk is om iedereen bij naam te noemen, wil ik mijn dank uitspreken voor jullie steun en welkome afleiding gedurende dit traject. Ik voel mij gelukkig met collega's die altijd klaarstaan om te ondersteunen.

Mijn paranimfen en goede vrienden, Dr. Mathijs van Ark en Harm Brugge, ik kijk ernaar uit om samen met jullie van de verdediging een geweldige dag te maken. Ik sterk mij met de gedachte dat jullie, waar nodig, een kopbeurt kunnen doen.

Pap en mam, jullie hebben me voortdurend aangemoedigd en ondersteund om mijn educatieve pad voort te zetten. Jullie hebben me altijd de vrijheid gegeven om mijn eigen weg te verkennen en mijn eigen tempo te bepalen. Ik ben jullie zeer erkentelijk voor het sterke fundament dat jullie hebben gelegd. Dit proefschrift is niet alleen mijn prestatie, maar zeker ook van jullie.

Lieve Amarins, als laatste wil ik graag van de gelegenheid gebruik maken om mijn waardering voor jou uit te spreken. Jouw relativeringsvermogen, bemoedigende woorden en onophoudelijke steun hebben mij door de moeilijke momenten geholpen; zonder jou was dit project al meerdere keren gestrand. Wat hebben we het goed en wat zijn we gelukkig samen! Nu kunnen we uitkijken naar nieuwe avonturen, en met jou aan mijn zijde kan het niet anders dan een prachtige reis worden.

RESEARCH INSTITUTE SHARE

This thesis is published within the **Research Institute SHARE** (Science in Healthy Ageing and healthcaRE) of the University Medical Center Groningen / University of Groningen. Further information regarding the institute and its research can be obtained from our internet site: https://umcgresearch.org/w/share
More recent theses can be found in the list below (supervisors are between brackets).

2023
Boersema HJM
The concept of 'Inability to Work Fulltime' in work disability benefit assessment
(Prof S Brouwer, Dr FI Abma, Dr T Hoekstra)

Ots P
The role of individual and contextual factors in paid employment of workers with a chronic
disease
(Prof S Brouwer, Dr SKR van Zon)

Kool E
Untangling the elements of midwives' occupational wellbeing: A study among newly qualified and experienced midwives
(Prof ADC Jaarsma, Prof FG Schellevis, Dr EI Feijen-de Jong)

Jansma FFI
Self-management in rehabilitation practice: On the design and implementation of a serious
theory-based analogue problem-solving game called 'Think Along?'
(Prof R Sanderman, Dr I Wenzler)

Erpecum CPL van
The role of fast-food outlet exposure in Body Mass Index
(Dr N Smidt, Prof U Bültmann, Dr SKR van Zon)

Kerver N
The effectiveness and cost-effectiveness of upper limb prostheses
(Prof CK van der Sluis, Dr RM Bongers, Dr S van Twillert)

Deviandri R
Management of anterior cruciate ligament injury in lower-middle income countries: Focus on outcomes and health economics in Indonesia
(Dr I van den Akker-Scheek, Prof MJ Postma, Dr HC van der Veen, Dr Andri MT Lubis)

Mangot Sala L
Disruptive Life Events and Health: Longitudinal evidence from a large cohort in the Netherlands
(Prof AC Liefbroer, Dr N Smidt)

Wijk DC
From prosperity to parenthood: How employment, income, and perceived economic uncertainty influence family formation
(Prof AC Liefbroer, Prof HAG de Valk)

Dai Y
Effects of exposure to polycyclic aromatic hydrocarbons and heavy metals on placental trophoblasts and childhood inflammation
(Dr MM Faas, Prof X Xu, Prof X Huo)

Menting SGP
Picking up the pace: The development of pacing behaviour during adolescence
(Dr MT Elferink-Gemser, Prof FJ Hettinga)

Vos M
My name is legion for we are many: Lessons learned from linking and splitting psychiatric Disorder
(Dr CA Hartman, Prof NNJ Rommelse)

Haan-Du J De
Cancer risk, stage, and survivorship among patients with type 2 diabetes
(Prof GH de Bock, Dr GWD Landman, Dr N Kleefstra)

Nieboer P
Teaching and learning in the operating room: Navigating treacherous waters
(Prof SK Bulstra, Prof M Huiskes, Dr M Stevens, Dr F Cnossen)

He Z
Risk factors for elevated blood pressure: focus on perimenopausal women and potential causality
(Prof H Snieder, Dr CHL Thio, Prof QYZ Qingying Zhang)

Peeters CMM
Brace therapy and radiographic imaging in adolescent idiopathic scoliosis; where do we stand?
(Prof PC Jutte, Dr C Faber, Dr FH Wapstra, Dr DHR Kempen)

For earlier theses visit the website: Find Research outputs — the University of Groningen research portal (rug.nl)

A

Research Institute
SHARE