RELIABILITY AND VALIDITY OF PHYSICAL FITNESS TESTS IN PEOPLE WITH MENTAL DISORDERS: A SYSTEMATIC REVIEW AND META-ANALYSIS.

Abstract

Background: Several tests are available to assess the different components of physical fitness, including cardiorespiratory fitness, muscular strength, and flexibility. However, the reliability and validity of physical fitness tests in people with mental disorders has not been meta-analysed. **Aims:** To examine the reliability, concurrent and convergent validity of physical fitness tests in people with mental disorders. **Methods:** Studies evaluating the reliability, concurrent and convergent validity of physical fitness tests in people with mental disorders were searched from major databases until January 20th, 2020. Random-effects meta-analyses were performed pooling (1) reliability: test-retest correlations at two-time points; (2) convergent validity between submaximal tests and maximal protocols; or (3) concurrent validity between two submaximal tests. Associations are presented using *r*-values and 95% confidence intervals. Methodological quality was assessed using the Quality Appraisal of Reliability Studies and the Critical Appraisal Tool. **Results:** A total of 11 studies (N=504; 34% females) were included. Reliability of the fitness tests, produced *r*-values ranging from moderate [balance test-EUROFIT; [*r*=0.75 (0.60 – 0.85), *p*=0.0001)] to very strong [explosive leg power EUROFIT; [*r*=0.96 (0.93 – 0.97, *p*=0.0001]. Convergent validity between the six-minute walk test (6MWT) and submaximal cardiorespiratory tests was moderate [0.57 (0.26–0.77); *p*=.0001]. Concurrent validity between the two-minute walk test (2MWT) and 6MWT [r= 0.86 (0.39–0.97); *p*=.0004] was strong. **Conclusion:** The present study demonstrates that physical fitness tests are reliable and valid in people with mental disorders.

Keywords: Exercise capacity; aerobic; mental disorders;

Introduction

The life expectancy of people with mental illness is shortened by more than ten years compared to the general population (Walker et al., 2015). A significant proportion of this gap in life expectancy is due to the increased prevalence of cardiometabolic diseases (Correll et al., 2017; Firth et al., 2019a; Vancampfort et al., 2016a, 2015d). Although genetic factors and shared pathophysiological mechanisms contribute, treatment-related factors and unhealthy lifestyle habits (i.e. low physical activity) also play a prominent role in this increased risk (Firth et al., 2019a).

Physical activity is a key determinant of physical fitness. Physical fitness comprises multiple components, including cardiorespiratory fitness, muscle endurance, muscular strength, body composition, balance and flexibility (Caspersen et al., 1985; Vanhees et al., 2005). In the general population, some components of physical fitness, such as cardiorespiratory capacity and handgrip strength, are independent predictors of all-cause mortality (Firth et al., 2019b; García-Hermoso et al., 2018; Kodama, 2009).

Previous research demonstrated that people with mental disorders have lower physical fitness levels, although it can be improved with physical activity (Vancampfort et al., 2017). Also, there is compelling evidence that, in this vulnerable population, higher levels of physical fitness are associated with a better quality of life and less symptom severity (Vancampfort et al., 2013a, 2013b).

Testing physical fitness is necessary to: (1) determine the exercise workload needed to provide health benefits and minimize the risk of injury, (2) evaluate physical functioning, and (3) monitor the progress of physical activity interventions (Meyer and Broocks, 2000; Stubbs et al., 2018). Previous evidence demonstrates the validity of many physical fitness tests in the general population (Bennett et al., 2016; Scalco et al., 2018). Since people with mental disorders often experience fluctuations in their energy and motivation levels (Taylor et al., 2003), or may avoid or abandon vigorous exercise fearing an exacerbation of somatic and anxiety symptoms (Vliet et al., 1999), it is important to examine the validity and reliability of physical fitness tests in this population.

To the best of our knowledge, no systematic review or meta-analysis has summarized the validity and reliability of physical fitness tests in people with mental disorders. Findings from this meta-analysis could assist exercise professionals in selecting the most reliable and valid tests to apply in this population. This study aimed to review the literature that evaluated the concurrent and convergent validity and reliability of fitness tests, such as cardiorespiratory fitness, exercise functional capacity, muscular strength, balance, flexibility, speed and agility in people with mental disorders.

Methods

Protocol and registration

The present systematic review followed the PRISMA guidelines (Moher et al., 2009). The study protocol is available upon request.

Eligibility criteria

Studies were included if they: 1) included clinically diagnosed patients with mental disorders, such as major depressive disorder, bipolar disorder, anxiety or stress-related disorders, psychotic disorders or substance abuse/dependence, according to the Diagnostic and Statistical Manual for mental Disorders (DSM-V or 5) (Hasin et al., 2013) or the international classification of diseases (ICD-10) (World Health Organization, 2016); (2) assessed the test-retest correlation (reliability), or the correlation between a submaximal cardiorespiratory test against a maximal cardiorespiratory test (convergent validity), or other submaximal validated test (concurrent validity).

Searches

Relevant articles were searched in the following databases: PubMed, PsycINFO, SportDiscus and Web of Science, from inception to January 20th 2020, using the following search terms: (depression OR bipolar and related disorders OR schizophrenia OR psychosis) AND (oxygen consumption OR Vo2 OR oxygen uptake OR aerobic OR cardiovascular OR cardiopulmonary exercise test OR cardiopulmonary fitness OR physical fitness OR fitness OR physical functional performance OR exercise OR leisure activit* OR Physical activity OR "muscle strength" OR "muscle power" OR stretching OR walking OR ability) AND (Valid* OR test-retest OR Reliabilit* OR reproducibility). The searches were slightly adapted depending on the specific requirements of the database. The string search for each database can be found in the supplementary material 1.

The screening for relevant articles was performed in two steps: First, the titles and abstracts of relevant articles were read. Second, potentially relevant articles, on the basis of their title and abstract, were read in full according to the inclusion and exclusion criteria. The two steps were performed by two independent reviewers (FS and DT) until consensus. No third reviewer was required.

Data extraction

All data were extracted by two (FS and DT) independent reviewers. Information about, volunteers; diagnoses, gender and age, measures; test utilized, retest period and author were extracted.

Study quality assessment

The within-study risk of bias was assessed using the Quality Appraisal of Reliability Studies (QAREL) for the convergent validity (Lucas et al., 2010). The QAREL is a checklist comprised of 11 items and has a maximum score of 110%. These questions can be answered as "yes" (good quality), "no" (poor quality) or "unclear" (insufficient information). For reliability risk of bias, we used the Critical Appraisal Tool (CAT) (Brink and Louw, 2011). The CAT scale has 13 items and we eliminated items 4,5, 6 and 8, as these items do not apply to reliability studies (Muñoz-Bermejo et al., 2019). The CAT scales can be scored as "yes", "no", or "not applicable" for research methods to be able to distinguish between high and low-quality studies and have a maximum score of 90%.

In each table, there is a column with the final percentage (%) representing the methodological quality of each study. Studies were considered high quality if they scored above 45% (Muñoz-Bermejo et al., 2019).

Statistical Analysis

Meta-analyses were performed for assessing the reliability, convergent validity and concurrent validity of the physical fitness tests. Correlations (r) and sample size data from each study were pooled together to calculate reliability and validity parameters. To determine test-retest reliability, meta-analyses were performed for: 1) each fitness tests, separately; 2) functional exercise capacity tests (walking tests), and 3) submaximal cardiovascular capacity (bike tests). A convergent validity meta-analysis was performed pooling data of studies associating submaximal tests and maximal protocols for functional exercise capacity tests. To determine concurrent validity, a meta-analysis was performed on the correlation between two submaximal tests (2-min and 6-min walk tests). For all analysis, the correlational values (*r, rho, ICC)* were converted into Fischer z scores and then, Fischer z scores were reconverted into Spearman correlation (*r)*. Outcomes were provided in *r* values, together with the 95% confidence interval (95%CI). Correlations were classified as: 0 to 0.19, very weak; 0.2 to 0.39, weak; 0.40 to 0.59, moderate; 0.6 to 0.79, strong; 0.8 to 0.9, very strong; and 1.0, perfect correlation (Walker & Almond, 2010). The Q and I^2 statistic were used to assess and quantify the heterogeneity, respectively. Scores of <25%, 25-50% and >50% indicated low, moderate and high heterogeneity, respectively (Higgins et al., 2003). Analyses were run with random-effect (DerSimonian and Laird, 1986). All analyses were performed using Comprehensive Meta-Analysis software (version 3) (Borenstein et al., 2013).

Results

Studies selection

The initial search yielded 17,295 results. After removal of duplicates and exclusion at the title plus abstract level, 17,223 abstracts were considered. At the full-text review stage, 72 studies were considered. Finally, a total of 11 studies were included. The flowchart of study selection can be seen in figure 1.

Insert figure 1 here

Studies and participants characteristics

Across eleven studies, a total of 504 (34% women) participants, with a median age of 37.4 (95% CI: 34.1 – 42.2 years) were included. Studies were conducted in several countries, including Belgium (Knapen et al., 2003; Vancampfort et al., 2010, 2011, 2014, 2015a, 2015c, 2016c, 2019b), Uganda (Vancampfort et al., 2019c, 2019a) and Portugal (Gomes et al., 2016). A total of 5 studies assessed the test-retest reliability of functional walk tests, which are a measure-of-proxy for cardiorespiratory fitness (Gomes et al., 2016; Vancampfort et al., 2011, 2016c, 2019c, 2019a). Three studies investigated the test-retest reliability of fitness tests, measured by the EUROFIT test, including flamingo balance (FBA) (balance), plate tapping (PLT) (speed), sit-and-reach (SAR) (flexibility), standing broad jump (SBJ) (explosive strength), handgrip strength (HGR), sit-ups (SUP) (muscular endurance), the 5m shuttle run (SHR) (speed and agility) (Vancampfort et al., 2019b, 2015c, 2011), 3 studies measured the 6MWT (Gomes et al., 2016; Vancampfort et al., 2015a, 2010) and 2 studies measured the 2MWT (Vancampfort et al., 2019a, 2019c). Two studies assessed the test-retest reliability of a submaximal cardiorespiratory fitness tests, one study tested the Franz test (Physical Work Capacity, PWC) (Knapen et al., 2003) and one study the Åstrand-Rhyming test (Vancampfort et al., 2014). Two studies assessed the convergent validity of the 6MWT (Gomes et al., 2016; Vancampfort et al., 2015a) and 2 studies for concurrent validity of the 2MWT against the 6MWT (Vancampfort et al., 2019c, 2019a). The details of all studies are shown in Table 1.

Insert table 1 here

Meta-analysis

Test-Retest Reliability of Fitness Tests

The test-retest reliability of each individual test ranged from moderate [balance test-EUROFIT, r=0.75 (95%CI: 0.60 – 0.85), p>0.0001)], to very strong [explosive leg power test-EUROFIT, r=0.96 (95%CI: 0.93 – 0.97), p>0.0001)]. Results of individual tests are shown in Table 2.

Insert table 2 here

Test-Retest Reliability of Functional Exercise Tests

We found a very strong test-retest reliability for functional exercise capacity tests [r=0.95 (95%CI: 0.94 – 0.97), p<0.0001] with a high heterogeneity (I^2= 57.16%). Data are presented in a forest plot, see Figure 2.

Insert figure 2 here

Test-Retest Reliability of Submaximal Cardiorespiratory Fitness Tests

We found a very strong test-retest reliability of submaximal cardiorespiratory tests [r=0.85 (95%CI: 0.77 – 0.91), p<0.0001]. The analysis presented a high heterogeneity (I^2= 58.35%). Data are presented in a forest plot, see Figure 3.

Insert figure 3 here

Convergent Validity of the 6MWT test

We found a moderate convergent validity of the 6MWT [r=0.57 (95%CI: 0.26 – 0.77), p=0.0001]. There was no heterogeneity (I^2= 0). Data are presented in a forest plot, see Figure 4.

Insert figure 4 here

Concurrent Validity of the 2MWT against the 6MWT

We found a strong correlation between the 6MWT and 2MWT [r=0.86 (95%CI: 0.39 – 0.97) p=0.0004)]. The analysis presented high heterogeneity (I^2= 94.60%). Data are presented in a forest plot, see Figure 5.

Insert figure 5 here

Discussion

To our knowledge, the present meta-analysis is the first to evaluate the reliability and concurrent and convergent validity of physical fitness tests in people with mental disorders. Our main findings are: (1) physical fitness tests can be reliably used in people with a mental disorder; (2) the 6MWT has a moderate convergent validity to estimate cardiorespiratory capacity, and (3) the 2MWT has a strong concurrent validity with the 6MWT.

Mental health disorders are associated with a high prevalence of somatic co-morbidity, particularly cardiovascular disease (Batelaan et al., 2016). It is now recognized that a complete patient assessment should involve screening physical activity habits (Correll et al., 2017), and with this, implementing appropriate tests of physical fitness to develop optimal exercise prescriptions (Hallgren et al., 2017).

Of the included studies, the EUROFIT was the only battery used to assess a range of physical fitness components, such as cardiorespiratory fitness, muscular strength, balance, flexibility, speed and agility in this population. While people with mental disorders present fluctuations in their energy and motivation levels (Firth et al., 2016), fear of somatic symptom exacerbation when performing exercise (Ho et al., 2018), muscular weakness and other abnormal myopathic and electrophysiological features (Reznik et al., 2000), as well as, cognitive symptoms that could potentially affect their performance in such tests (Firth et al., 2018a), the EUROFIT battery has demonstrated value to provide reliable assessments in people with mental disorders. The EUROFIT battery is a safe, and inexpensive protocol that can be implemented in low-resource settings.

We have found evidence that the 6MWT is a reliable test, with a moderate correlation with submaximal or maximal cardiorespiratory tests in people with mental disorders. The 6MWT walk test does not require equipment, such as an ergometer, it is relatively low-cost and easier to administer than cycle or treadmill tests (Solway et al., 2001). Therefore, these tests can be easily be implemented in clinical practice to assess functional exercise capacity levels and to estimate cardiorespiratory fitness. Moreover, the 2MWT showed high convergent validity when compared to the 2MWT. Therefore, in settings with limited staff or time constraints, the 2MWT can be used reliably.

Physical fitness components are associated with better cardiometabolic functioning (Gorczynski and Faulkner, 2010), lesser severity of psychiatric symptoms (Firth et al., 2017), increased quality of life (Vancampfort et al., 2015b), preservation of brain health (Firth et al., 2018b). Given these important health benefits, and the relative ease with which fitness can reliably be assessed by exercise professionals, we argue that physical fitness tests should be incorporated in routine mental health care. Their inclusion may also help to provide a more holistic and comprehensive assessment of the patient and their level of physical functioning.

Some limitations are acknowledged. The number of studies identified was small, therefore, our analyses of test validity may suffer from low statistical power. Also, due the small number of studies, the heterogeneity could not be explored. The EUROFIT battery was the only test examined for most components, and it is not possible to extend these results to other physical fitness tests. Also, worth mentioning that medications may impaired the performance of physical tests regardless of the time of day it is ingested. However, antipsychotic medication can provide unpleasant side-effects, such as feeling of fatigue, and may favour poor performance in the physical test by providing body balance, and lower muscular strength (Perez-Cruzado et al., 2017; Vancampfort et al., 2016b). Lastly, we were not able to explore whether age, sex, body mass index, diagnosis and symptom severity impact the use of such tests.

Conclusion

The present review indicates that physical fitness tests are reliable and valid to assess physical fitness components in people with mental disorders. Maximal aerobic exercise testing is the gold standard to assess a person's cardiorespiratory fitness level, the 6MWT and the 2MWT can be used as a measure-of-proxy for assessing cardiorespiratory fitness in people with mental disorders, especially in low resource settings.

References

Batelaan, N.M., Seldenrijk, A., Bot, M., Van Balkom, A.J.L.M., Penninx, B.W.J.H., 2016. Anxiety and new onset of cardiovascular disease: Critical review and meta-analysis. Br. J. Psychiatry 208, 223–231. https://doi.org/10.1192/bjp.bp.114.156554

Bennett, H., Parfitt, G., Davison, K., Eston, R., 2016. Validity of Submaximal Step Tests to Estimate Maximal Oxygen Uptake in Healthy Adults. Sport. Med. 46, 737–750. https://doi.org/10.1007/s40279-015-0445-1

Borenstein, M., Hedges, L., Higgins, J., Rothstein, H., 2013. Comprehensive Meta-Analysis Version 3. Biostat, Englewood, NJ.

Brink, Y., Louw, Q.A., 2011. Clinical instruments: Reliability and validity critical appraisal. J. Eval. Clin. Pract. 18, 1126–1132. https://doi.org/10.1111/j.1365-2753.2011.01707.x

Caspersen, C.J., Powell, K.E., Christenson, G.M., 1985. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 100, 126–131.

Correll, C.U., Solmi, M., Veronese, N., Bortolato, B., Rosson, S., Santonastaso, P., Thapa-Chhetri, N., Fornaro, M., Gallicchio, D., Collantoni, E., Pigato, G., Favaro, A., Monaco, F., Kohler, C., Vancampfort, D., Ward, P.B., Gaughran, F., Carvalho, A.F., Stubbs, B., 2017. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. World Psychiatry

16, 163–180. https://doi.org/10.1002/wps.20420

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control. Clin. Trials 7, 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Firth, J., Firth, J.A., Stubbs, B., Vancampfort, D., Schuch, F.B., Hallgren, M., Veronese, N., Yung, A.R., Sarris, J., 2018a. Association between muscular strength and cognition in people with major depression or bipolar disorder and healthy controls. JAMA Psychiatry 75, 740–746. https://doi.org/10.1001/jamapsychiatry.2018.0503

Firth, J., Rosenbaum, S., Stubbs, B., Gorczynski, P., Yung, A.R., Vancampfort, D., 2016. Motivating factors and barriers towards exercise in severe mental illness: A systematic review and meta-analysis. Psychol. Med. 46, 2869–2881. https://doi.org/10.1017/S0033291716001732

Firth, J., Siddiqi, N., Koyanagi, A., Siskind, D., Rosenbaum, S., Galletly, C., Allan, S., Caneo, C., Carney, R., Carvalho, A.F., Chatterton, M. Lou, Correll, C.U., Curtis, J., Gaughran, F., Heald, A., Hoare, E., Jackson, S.E., Kisely, S., Lovell, K., Maj, M., McGorry, P.D., Mihalopoulos, C., Myles, H., O'Donoghue, B., Pillinger, T., Sarris, J., Schuch, F.B., Shiers, D., Smith, L., Solmi, M., Suetani, S., Taylor, J., Teasdale, S.B., Thornicroft, G., Torous, J., Usherwood, T., Vancampfort, D., Veronese, N., Ward, P.B., Yung, A.R., Killackey, E., Stubbs, B., 2019a. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. The Lancet Psychiatry 6, 675–712. https://doi.org/10.1016/s2215-0366(19)30132-4

Firth, J., Siddiqi, N., Koyanagi, A., Siskind, D., Rosenbaum, S., Galletly, C., Allan, S., Caneo, C., Carney, R., Carvalho, A.F., Chatterton, M. Lou, Correll, C.U., Curtis, J., Gaughran, F., Heald, A., Hoare, E., Jackson, S.E., Kisely, S., Lovell, K., Maj, M., McGorry, P.D., Mihalopoulos, C., Myles, H., O'Donoghue, B., Pillinger, T., Sarris, J., Schuch, F.B., Shiers, D., Smith, L., Solmi, M., Suetani, S., Taylor, J., Teasdale, S.B., Thornicroft, G., Torous, J., Usherwood, T., Vancampfort, D., Veronese, N., Ward, P.B., Yung, A.R., Killackey, E., Stubbs, B., 2019b. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. The Lancet Psychiatry 6, 675–712. https://doi.org/10.1016/S2215-0366(19)30132-4

Firth, J., Stubbs, B., Rosenbaum, S., Vancampfort, D., Malchow, B., Schuch, F., Elliott, R., Nuechterlein, K.H., Yung, A.R., 2017. Aerobic exercise improves cognitive functioning in people with schizophrenia: A systematic review and meta-analysis.

Schizophr. Bull. 43, 546–556. https://doi.org/10.1093/schbul/sbw115

Firth, J., Stubbs, B., Vancampfort, D., Schuch, F., Lagopoulos, J., Rosenbaum, S., Ward, P.B., 2018b. Effect of aerobic exercise on hippocampal volume in humans: A systematic review and meta-analysis. Neuroimage 166, 230–238. https://doi.org/10.1016/j.neuroimage.2017.11.007

García-Hermoso, A., Cavero-Redondo, I., Ramírez-Vélez, R., Ruiz, J.R., Ortega, F.B., Lee, D.C., Martínez-Vizcaíno, V., 2018. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. Arch. Phys. Med. Rehabil. 99, 2100-2113.e5. https://doi.org/10.1016/j.apmr.2018.01.008

Gomes, E., Bastos, T., Probst, M., Carlos, J., Silva, G., 2016. Reliability and validity of 6MWT for outpatients with schizophrenia : A preliminary study. Psychiatry Res. 237, 37–42. https://doi.org/10.1016/j.psychres.2016.01.066

Gorczynski, P., Faulkner, G., 2010. Exercise therapy for schizophrenia. Cochrane Database Syst. Rev. https://doi.org/10.1002/14651858.cd004412.pub2

Hallgren, M., Stubbs, B., Vancampfort, D., Lundin, A., Jääkallio, P., Forsell, Y., 2017. Treatment guidelines for depression: Greater emphasis on physical activity is needed. Eur. Psychiatry 40, 1–3. https://doi.org/10.1016/j.eurpsy.2016.08.011

Hasin, D.S., O'Brien, C.P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., Compton, W.M., Crowley, T., Ling, W., Petry, N.M., Schuckit, M., Grant, B.F., 2013. DSM-5 criteria for substance use disorders: Recommendations and rationale. Am. J. Psychiatry 170, 834–851. https://doi.org/10.1176/appi.ajp.2013.12060782

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 27, 159–184. https://doi.org/10.1136/bmj.327.7414.557

Ho, P.A., Dahle, D.N., Noordsy, D.L., 2018. Why Do People With Schizophrenia Exercise? A Mixed Methods Analysis Among Community Dwelling Regular Exercisers. Front. Psychiatry 9, 1–8. https://doi.org/10.3389/fpsyt.2018.00596

Knapen, J., Vliet, P.V.A.N.D.E., Coppenolle, H.V.A.N., 2003. Evaluation of cardio-respiratory fitness and perceived exertion for patients with depressive and anxiety disorders : a study on reliability. https://doi.org/10.1080/09638280310001616277

Kodama, S., 2009. CLINICIAN ' S CORNER Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events. J. Am. Med. Assoc. 301, 2024–2035.

Lucas, N.P., Macaskill, P., Irwig, L., Bogduk, N., 2010. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). J. Clin. Epidemiol. 63, 854–861. https://doi.org/10.1016/j.jclinepi.2009.10.002

Meyer, T., Broocks, A., 2000. Therapeutic impact of exercise on psychiatric diseases. Guidelines for exercise testing and prescription. Sport. Med. 30, 269–279. https://doi.org/10.2165/00007256-200030040-00003

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ 339, 332–336. https://doi.org/10.1136/bmj.b2535

Muñoz-Bermejo, L., Pérez-Gómez, J., Manzano, F., Collado-Mateo, D., Villafaina, S., Adsuar, J.C., 2019. Reliability of isokinetic knee strength measurements in children: A systematic review and meta-analysis. PLoS One 14, 1–15. https://doi.org/10.1371/journal.pone.0226274

Perez-Cruzado, D., Cuesta-Vargas, A.I., Vera-Garcia, E., Mayoral-Cleries, F., 2017. Physical fitness and levels of physical activity in people with severe mental illness: A cross-sectional study. BMC Sports Sci. Med. Rehabil. 9, 1–6. https://doi.org/10.1186/s13102-017-0082-0

Reznik, I., Volchek, L., Mester, R., Kotler, M., Sarova-Pinhas, I., Spivak, B., Weizman, A., 2000. Myotoxicity and neurotoxicity during clozapine treatment. Clin. Neuropharmacol. 23, 276–280. https://doi.org/10.1097/00002826-200009000-00007

Scalco, J.C., Martins, R., Keil, P.M.R., Mayer, A.F., Schivinski, C.I.S., 2018. Psychometric properties of functional capacity tests in children and adolescents: Systematic review. Rev. Paul. Pediatr. 36, 500–510. https://doi.org/10.1590/1984-0462/;2018;36;4;00002

Solway, S., Brooks, D., Lacasse, Y., Thomas, S., 2001. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. Chest 119, 256–270. https://doi.org/10.1378/chest.119.1.256

Stubbs, B., Vancampfort, D., Hallgren, M., Firth, J., Veronese, N., Solmi, M., Brand, S., Cordes, J., Malchow, B., Gerber, M., Schmitt, A., Correll, C.U., De Hert, M., Gaughran, F., Schneider, F., Kinnafick, F., Falkai, P., Möller, H.J., Kahl, K.G., 2018. EPA guidance on physical activity as a treatment for severe mental illness: a meta-review of the evidence and Position Statement from the European Psychiatric

Association (EPA), supported by the International Organization of Physical Therapists in Mental . Eur. Psychiatry 54, 124–144. https://doi.org/10.1016/j.eurpsy.2018.07.004

Taylor, R.R., Jason, L.A., Jahn, S.C., 2003. Chronic fatigue and sociodemographic characteristics as predictors of psychiatric disorders in a community-based sample. Psychosom. Med. 65, 896–901. https://doi.org/10.1097/01.PSY.0000088580.28749.7F

Vancampfort, D., Basangwa, D., Kimbowa, S., Firth, J., Schuch, F., Van Damme, T., Mugisha, J., 2019a. Test–retest reliability, validity, and correlates of the 2-min walk test in outpatients with depression. Physiother. Res. Int. 1–7. https://doi.org/10.1002/pri.1821

Vancampfort, D., Buys, R., Sienaert, P., Wyckaert, S., Herdt, A. De, Hert, M. De, Probst, M., 2015a. Validity of the 6 min walk test in outpatients with bipolar disorder. Psychiatry Res. 230, 664–667. https://doi.org/10.1016/j.psychres.2015.09.039

Vancampfort, D., Correll, C.U., Galling, B., Probst, M., De Hert, M., Ward, P.B., Rosenbaum, S., Gaughran, F., Lally, J., Stubbs, B., 2016a. Diabetes mellitus in people with schizophrenia, bipolar disorder and major depressive disorder: A systematic review and large scale meta-analysis. World Psychiatry 15, 166–174. https://doi.org/10.1002/wps.20309

Vancampfort, D., Guelinckx, H., Hert, M. De, Stubbs, B., Soundy, A., Rosenbaum, S., Schepper, E. De, Probst, M., 2014. Reliability and clinical correlates of the Astrand – Rhyming sub-maximal exercise test in patients with schizophrenia or schizoaffective disorder. Psychiatry Res. 220, 778–783. https://doi.org/10.1016/j.psychres.2014.08.049

Vancampfort, D., Guelinckx, H., Probst, M., Stubbs, B., Rosenbaum, S., Ward, P.B., De Hert, M., 2015b. Health-related quality of life and aerobic fitness in people with schizophrenia. Int. J. Ment. Health Nurs. 24, 394–402. https://doi.org/10.1111/inm.12145

Vancampfort, D., Hallgren, M., Vandael, H., Probst, M., Stubbs, B., Raymaekers, S., Van Damme, T., 2019b. Test-retest reliability and clinical correlates of the Eurofit test battery in people with alcohol use disorders. Psychiatry Res. 271, 208–213. https://doi.org/10.1016/j.psychres.2018.11.052

Vancampfort, D., Kimbowa, S., Basangwa, D., Smith, L., Stubbs, B., Damme, T. Van, Hert, M. De, Mugisha, J., 2019c. Test-retest reliability, concurrent validity and correlates of the two-minute walk test in outpatients with psychosis. Psychiatry Res. 112619. https://doi.org/10.1016/j.psychres.2019.112619

Vancampfort, D., Probst, M., Daenen, A., Damme, T. Van, De Hert, M., Rosenbaum, S., Bruyninckx, D., 2016b. Impact of antipsychotic medication on physical activity and physical fitness in adolescents: An exploratory study. Psychiatry Res. 242, 192–197. https://doi.org/10.1016/j.psychres.2016.05.042

Vancampfort, D., Probst, M., De Herdt, A., Corredeira, R.M.N., Carraro, A., De Wachter, D., De Hert, M., 2013a. An impaired health related muscular fitness contributes to a reduced walking capacity in patients with schizophrenia: A cross-sectional study. BMC Psychiatry 13, 1–8. https://doi.org/10.1186/1471-244X-13-5

Vancampfort, D., Probst, M., Scheewe, T., De Herdt, A., Sweers, K., Knapen, J., van Winkel, R., De Hert, M., 2013b. Relationships between physical fitness, physical activity, smoking and metabolic and mental health parameters in people with schizophrenia. Psychiatry Res. 207, 25–32. https://doi.org/10.1016/j.psychres.2012.09.026

Vancampfort, D., Probst, M., Sweers, K., Maurissen, K., Knapen, J., 2011. Eurofit test battery in patients with schizophrenia or schizoaffective disorder : Reliability and clinical correlates. Eur. Psychiatry 27, 416–421. https://doi.org/10.1016/j.eurpsy.2011.01.009

Vancampfort, D., Probst, M., Sweers, K., Maurissen, K., Knapen, J., Hert, M. De, 2010. Reliability, minimal detectable changes, practice effects and correlates of the 6-min walk test in patients with schizophrenia. Psychiatry Res. 187, 62–67. https://doi.org/10.1016/j.psychres.2010.11.027

Vancampfort, D., Rosenbaum, S., Schuch, F., Ward, P.B., Richards, J., Mugisha, J., Probst, M., Stubbs, B., 2017. Cardiorespiratory Fitness in Severe Mental Illness: A Systematic Review and Meta-analysis. Sport. Med. 47, 343–352. https://doi.org/10.1007/s40279-016-0574-1

Vancampfort, D., Sienaert, P., Wyckaert, S., Hert, M. De, Stubbs, B., Rosenbaum, S., Buys, R., Probst, M., 2015c. Test-retest reliability, feasibility and clinical correlates of the Eurofit test battery in people with bipolar disorder. Psychiatry Res. 1–6. https://doi.org/10.1016/j.psychres.2015.05.042

Vancampfort, D., Stubbs, B., Mitchell, A.J., De Hert, M., Wampers, M., Ward, P.B., Rosenbaum, S., Correll, C.U., 2015d. Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: A systematic review and meta-analysis. World Psychiatry 14, 339–347. https://doi.org/10.1002/wps.20252

Vancampfort, D., Wyckaert, S., Sienaert, P., Hert, M. De, Soundy, A., Rosenbaum, S., Richards, J., Probst, M., 2016c. Test-retest study of the six-minute walk test in people with bipolar disorder 28, 39–44.

Vanhees, L., Lefevre, J., Philippaerts, R., Martens, M., Huygens, W., Troosters, T., Beunen, G., 2005. How to assess physical activity? How to assess physical fitness? Eur. J. Prev. Cardiol. 12, 102–114. https://doi.org/10.1097/01.hjr.0000161551.73095.9c

Vliet, P. Van de, Coppenolle, H. Van, Knapen, J., 1999. Physical measures, perceived physical ability, and body acceptance of adult psychiatric patients. Adapt. Phys. Act. Q. 16, 113–125. https://doi.org/10.1123/apaq.16.2.113

Walker, J., & Almond, P., 2010. Interpreting statistical findings: a guide for health professionals and students. McGraw-Hill Educ. (UK).

Walker, E.R., McGee, R.E., Druss, B.G., 2015. Mortality in Mental Disorders and Global Disease Burden Implications: A Systematic Review and Meta-analysis. Physiol. Behav. 176, 139–148. https://doi.org/10.1001/jamapsychiatry.2014.2502

World Health Organization, 2016. International Statistical Classification of Diseases and Related Health Problems. World Heal. Organ. 10, xii, 362 p. https://doi.org/10.1002/1520-6505(2000)9:5<201::AID-EVAN2>3.3.CO;2-P

Table 1. General characteristics of studies with validity and reliability of fitness variables measures in patients' with mental disorders.

Note: PWCf: Physical work capacity female; PCWm: Physical work capacity men; 6MWT: six minute walk test; 2MWT: two minutes walk test; VO_2max: maximal oxygen uptake.

Author	Diagnoses	Participants (n)	Women %	Age (yrs)	Submaximal test 1	Submaximal test 2	Maximal test	Retest Period
Knapen et al, 2003	Major Depression	29	-	35.6 ± 11.5	PWCm 130; PWCm 150; PWCf 130; PWCf 150	-	-	7 days
Vancampfort et al., 2010	Schizophrenia	71	44	37.4 ± 10.1	6MWT	-	-	3 days
Vancampfort et al., 2011	Schizophrenia	50	40	32.9 ± 9.3	EUROFIT TEST	-	-	3 days
Vancampfort et al., 2014	Schizophrenia	47	28	34.2 ± 11.2	Astrand–Rhyming test	-	-	3 days
Vancampfort et al., 2015a	Bipolar Disorder	19	-	47.1 ± 8.3	6MWT	-	Cycle ergometer maximal test	3 days
Vancampfort et al., 2015c	Bipolar Disorder	46	48	43.4 ± 11.6	EUROFIT TEST	-	-	3 days
Vancampfort et al., 2016b	Bipolar Disorder	46	50	42.9 ± 11.5	6MWT	-	-	3 days
Gomes et al, 2016	Schizophrenia	51	24	40.7 ± 7.1	6MWT	-	Treadmill maximal protocol	3 days
Vancampfort et al., 2019c	Psychosis	50	45	33.5 ± 14.3	2MWT	6MWT	-	3 days
Vancampfort et al., 2019a	Major Depression	50	30	29 ± 14.2	2MWT	6MWT	-	3 days
Vancampfort et al., 2019b	Alcohol and drugs	45	28.9	41.3 ± 12.9	EUROFIT TEST	-	-	3 days

Table 2. Subgroup analysis of tests used in Test-Retest Reliability of Fitness Tests.

Analysis	**Number of RCTs**	**Meta-analysis**			**Heterogeneity**
		SMD	**95% CI**	**P value**	**I²**
Test protocol					
2MWT	2	0.95	0.907 0.974	<0.01	1.51
6MWT	3	0.96	0.938 0.978	<0.01	72.31
Astrand-Rhyming	1	0.92	0.805 0.968	<0.01	0
PWC 130	2	0.81	0.609 0.850	<0.01	0
PWC 150	2	0.84	0.934 0.977	<0.01	86.13
Flamingo Balance - EFT	3	0.75	0.812 0.933	<0.01	34.56
Handgrip Strength - EFT	3	0.96	0.931 0.976	<0.01	0
Plate Tapping - EFT	3	0.88	0.938 0.978	<0.01	83.02
Sit-And-Reach - EFT	3	0.96	0.887 0.960	<0.01	0
Standing Broad Jump - EFT	3	0.96	0.903 0.966	<0.01	52.85
Shuttle Run - EFT	3	0.93	0.654 0.906	<0.01	35.47
Sit-Ups - EFT	3	0.94	0.679 0.924	<0.01	66.32

Note：RCT: Randomized Clinical Trials. SMD: Standardized mean difference. 2MWT: Two minutes walk test; 6MWT: Two minutes walk test; PWC: Physical Work Capacity; EFT: Eurofit Test