Exercise addiction in athletes: Comparing two assessment instruments and willingness to stop exercise after medical advice

(Article begins on next page)

Exercise addiction in athletes: comparing two assessment instruments and willingness to stop exercise after medical advice

Abstract

Exercise is overwhelmingly beneficial for physical and mental health, but for some people exercise addiction (EA) can develop and negatively impact an individual. This study sought to 1) compare the latent structure of two instruments assessing EA and 2) examine differences in attitudes towards stopping exercise, if required to on medical grounds, among exercise-addicted and non-addicted athletes.

In a cross-sectional study, 1,011 athletes competing at different levels completed an anonymous on-line survey. The survey contained Exercise Dependence Scale-Revised (EDS-R), Exercise Addiction Inventory (EAI), and questions on adherence to medical prescriptions to stop exercise. We tested the latent structure of EDS-R and EAI with multigroup confirmatory factor analyses (CFA), across gender and competition level. Finally, we measured the difference of athletes' attitudes towards stopping exercise, if prescribed by a physician. Both instruments showed good fit indexes, even across gender. CFAs on EAI scores showed some violations of measurement invariance across competition level (ΔCFI= .03; ΔRMSEA= .02). On the contrary, CFAs on EDS-R scores did not show invariance violations across competition level (ΔCFI= < .01; ΔRMSEA= < .01). Finally, athletes who reached thresholds for exercise addiction, by means of EDS-R, were more prone to not follow medical prescriptions to cease exercise, independently of the competition level. These results suggest that athletes' answers on the EDS-R seems to be less affected by competition level, compared to EAI. Moreover, EDS-R outcomes could be used to identify individuals who may be unlikely to cease exercise for medical reasons, independently of their competition level.

Public Significance Statement. The present study suggests that the answers to Exercise Dependence Scale-Revised are not influenced by either the gender or the competition level of an athlete. On the contrary, the answers to Exercise Addiction Inventory can be influenced by athletes' competition level. Furthermore, Exercise Dependence Scale-Revised appears to be more effective than the Exercise Addiction Inventory in detecting athletes' attitude toward medical prescription to stop exercise.

Keywords: Exercise Dependence Scale; Exercise Addiction Inventory; Measurement invariance; Multigroup CFA; Psychometric Comparison.

Exercise addiction in athletes: comparing two assessment instruments and willingness to stop exercise after medical advice

There is overwhelming evidence that exercise is beneficial for physical and mental health (Ashdown-Franks et al., 2020; Vina et al., 2012). Moderate levels of exercise have been shown **to have a protective effect on** myocardial infarction and several types of cancer (Kim et al., 2019). Exercise is also effective in improving a wide range of symptomatic outcomes ranging from primary symptoms in mental disorders to cognition (Ashdown-Franks et al., 2020; Stubbs et al., 2018). However, exercise can also become obsessive and compulsive to an extent in which it can negatively impact on a person's physical and mental health. Such an obsessive-compulsive relationship with exercise has been termed exercise addiction (Szabo et al., 2015).

Exercise addiction was first described by Frederick Baekeland who found that people exercising for more than three days per week suffered from sleep and psychological symptoms akin to substance withdrawal symptoms upon cessation of exercise (Baekeland, 1970). Exercise addiction can cause an impairment in physical (e.g., spine injuries or anemia), and social functioning even **in the absence of** injuries (Hausenblas et al., 2017; Wouthuyzen-Bakker & van Assen, 2015). In the general population, estimates of prevalence of exercise addiction are extremely heterogeneous, with reports ranging from 4.0% in adolescent school athletes (Lichtenstein et al., 2018) to 42% in selected at-risk populations, such as those attending gyms (Lejoyeux et al., 2008).

A common component of every type of addiction, either to substances or to specific behaviors, is the difficulty to cease the activity despite negative consequences. For athletes, such difficulty **is a delicate topic** (Vina et al., 2012), for which the tradeoff **between the amount of** exercise to maintain appropriate preparation and excessive exercise is not clearly defined. Among athletes, several arrhythmogenic conditions including cardiomyopathies and coronary artery disease can contraindicate exercise due to the risk of sudden cardiac death (Maron, 2007; Zorzi et al., 2020). In several sports for example, all competitive athletes **must undergo an annual** medical examination to screen for such conditions by law (Vessella et al., 2020), and if a potentially fatal cardiac anomaly is detected, the athlete is disqualified from competitive events and she/he is recommended to stop or reduce the sporting activity. Athletes with exercise addiction, however, would struggle to follow this prescription. In a previous study, people assessed as exercise addicted reported to be more likely to continue practicing despite a potentially worsening heart disease, compared to non-exercise addicted athletes (Zorzi et al., 2020). Hence, it is clinically relevant to use valid tools detecting, or at least screening, exercise addiction among competitive athletes, to have a broader picture of their sport attitude and their resistances to stop or reduce it. To date, the most widely used questionnaires to measure exercise addiction are the Exercise Addiction Inventory (EAI; Terry et al., 2004) and the Exercise Dependence Scale- Revised (EDS-R: Costa et al., 2012; Hausenblas & Symons Downs, 2002a, 2002c). These tools have been extensively validated and the psychometric properties tested, especially during the last decade. For instance, Mónok et al., (2012) examined the two tools' properties on a nationally representative population and concluded that both instruments are valid to screen potential exercise addiction (Mónok et al., 2012). On the other hand, in a recent systematic review (Di Lodovico et al., 2019), the EAI was found to be preferable to the EDS-R in screening the risk of exercise addiction in sport exercisers, because it identifies higher proportions of at-risk individuals. Despite the steps forward, both examples suggest how the literature base is still developing, since several issues still need to be addressed to provide a definitive consensus on which instrument should be used to screen for exercise addiction (Di Lodovico et al., 2019; Mónok et al., 2012). For instance, the above examples refer to the general population or specific types of exercisers (e.g., endurance exercisers, bodybuilders or fitness attendees), however the responses to both instruments should be analyzed comparing amatorial and professional athletes or, more in general, on differing types of competition levels. In the same way, it should be tested if the response trends of male and female athletes could be different (Di Lodovico et al., 2019). The structural and psychometric validity of EA questionnaires is a relevant issue in specific athletic populations and has not been extensively explored to date. Furthermore, exercise addiction can be a serious problem for athletes, since it could push athletes beyond their physiological limits, being unable to modulate training or to stop it when indicated on medical grounds. To date, no study has assessed how available questionnaires are valid to predict adherence to medical prescriptions to stop exercise, and importantly, how available questionnaires perform across competition level, and gender.

Therefore, the present study has two main aims:

1. To test the latent structure of EAI and EDS-R across competition level and gender, and to understand if one of both tools are influenced by such features.
2. To test the differences between addicted versus non-addicted exercisers (measured by means EAI and EDS-R) on adherence to medical prescription to stop exercise, even across competition level.

Methods

The sample

A sample of athletes were invited (via social media groups dedicated to athletes) to fill an anonymous questionnaire asking their opinion both on the importance of sports activity for themselves and about preparticipation evaluation. Athletes were defined as those individuals engaged in an organized sports program requiring regular training and competition. According to the Italian law (Decree of Italian Ministry of Health: Rules for the Health Care of Competitive Sport Activities, 1982), these individuals are required to undergo mandatory preparticipation evaluation by a specialized sports medicine physician. At the beginning of the questionnaire, participants were asked if they were currently practicing competitive sports requiring mandatory medical evaluation: those who replied that they did not need the certificate of fitness were excluded from the analysis. Moreover, another exclusion criterion was applied to define participants as athletes: those exercising less than 4 hours per week were excluded. This classification was made accordingly the recent guidelines applied in sport cardiology (Pelliccia et al., 2020). The final sample consisted of 1011 athletes (mean age 33.9 $\pm$ 13.75 years; 23.83% females). The link to the survey was posted in social networks for one month. Participants gave their informed consent to participate to the study, that was carried out according to the Declaration of Helsinki. The Institutional Review Board of the Padua Hospital, after a preliminary consultation, specified that ethical approval was not necessary for this study[1], since it consisted in an observational study being carried out using anonymous questionnaires on the general population. Participants did not receive any form of payment to participate.

Measures

Gender, age, the level of competition (amateur, local, regional, national, international), years of experience, hours of weekly practice and kind of sport practiced were collected. Table 1 displays the descriptive statistics of the present study. Categories of practised sport are in Supplmenetary material (Table S1).

[PLEASE, INSERT TABLE 1 HERE]

Exercise Dependence Scale- Revised (EDS-R) and Exercise Addiction Inventory (EAI) were used to measure exercise addiction. The EDS-R is a questionnaire composed of 21 questions rated on a six-point Likert response scale (1= Never; 6 =Always). The scale is divided into seven sub-categories (three questions per scale), based on the DSM IV (American Psychiatric Association [APA], 2000) criteria for substance addiction. EDS-R categories are: withdrawal, continuance, tolerance, control loss, decrease of other activities, time, and effect of intention. For each category, a total score is obtained by summing the ratings to each item. In general, higher score suggest higher chance to be at risk of exercise addiction. A score greater than 14 suggests an at-risk individual, between 7 and 14 a nondependent-symptomatic individual, and less than 7 denotes a nondependent-asymptomatic individual. According to the EDS-R manual (Hausenblas & Symons Downs, 2002b), an exercise dependent profile is suggested whenever at least three subscales present an 'at-risk' score; a nondependent-asymptomatic profile whenever a nondependent-asymptomatic score occurs in at least four subscales; a nondependent-symptomatic profile otherwise.

---

[1] Consequently, a protocol number was not provided, since the *Institutional Review Board (IRB)* judged it as unnecessary.

The EAI is a short questionnaire composed of six items rated on a five-point Likert scale (1 = strongly disagree, 5= strongly agree), and investigates six main addiction domains: salience, mood modification, tolerance, withdrawal symptoms, conflict and relapse (Griffiths et al., 2015; Mónok et al., 2012; Sicilia et al., 2013; Terry et al., 2004). Table 2 contains the items of both instruments. As with the EDS-R, higher scores indicate a higher risk of exercise addiction. A score greater than 23 suggests at risk individuals; a score between 13 and suggests a potentially symptomatic person while a score less than 13 suggests an asymptomatic person (Griffiths et al., 2015). Table S2 of supplementary material contains the frequencies of each level of response, for each item of both instruments.

The following questions were used to assess participants adherence to hypothetical medical prescription**s** made by a *physician*[2]: *"How would you react to a medical prescription to temporarily interrupt exercise* (Q1 in the sequel)*; "After a medical investigation, it emerged that you have a critical cardiac anomaly that could cause a deadly heart attack. Unfortunately, it is necessary to permanently stop exercise. How would you react*[3]*?"* (Q2 in the sequel). Both questions were rated on a dichotomous scale indicating the intention to stop (0) or to continue sport activities (1) despite medical prescription.

Analytic plan

We performed confirmatory factor analysis (CFA) on each instrument. Considering that data were skewed (see Table S2), we considered them as ordinal and adopted a Diagonally Weighted Least Squares (Li, 2016) robust estimator in all the CFAs. As suggested by several authors (Forero et al., 2009; Kyriazos, 2018), the sample size for studies including confirmatory factor analysis

[2] The fact that the medical prescriptions were made by a physician was an information that participants derived from (a) the instruction of the section containing the questions within the online form and (b) from the word used. In particular, the adjective "medical" in Italian means that the recommendation or the exam is made by a physician, that is the only one that can make a diagnosis or prescribe to stop exercise.
[3] This question was the last of a series of question investigating the reaction to medical prescription. The question "how would you react?" was a short for of the question "*How would you react to this medical prescription?*"

should range from 200 to 500 participants. Considering that we performed two multigroup CFAs, we found it reasonable to collect at least 1000 participants (Comrey & Lee, 2013).

We applied a listwise deletion of missing data (i.e., only participants who both reported the gender/competition level and responded to all the items of the scale were included). In this way, we used the same set of complete cases to run all the CFAs. **Furthermore, we checked if missing data were equally and randomly distributed across items or if systematic patterns of missing data occurred for specific items. After the data-elimination, we also checked if there were differences in age between included and excluded participants. This gave us the chance to consider also potential influences of age on the CFAs results.** As suggested by several authors (Mónok et al., 2012; Schermelleh-Engel et al., 2003; Weston & Gore, 2006), we used multiple goodness-of-fit statistics in order to interpret the models' fit. For the Comparative Fit Index (CFI: Bentler, 1990) values ≥ .95 represent a good fit, values ≥ .90 for an adequate fit, while values < .90 a not acceptable fit. **For Root Mean Square Error of Approximation's (RMSEA: Steiger, 1990) values ≤ .05 indicate a good fit, values between .05 and .10 an adequate fit while values > .10 a not acceptable fit** . For the Standardized Root Mean Square Residual (SRMR: Hu & Bentler, 1999), values ≤ .05 represent a good fit, values between 0.5 and 0.8 an adequate fit while values > .8 a not acceptable fit. We used the R statistical environment (R Core Team, 2020) to run all analyses and the *lavaan* package (Rosseel, 2012) for the CFAs. Furthermore, we used Cronbach's alpha to test the internal consistency of both questionnaires, by means of the *ltm* package (Rizopoulos, 2006).

We tested measurement invariance of both questionnaires' items score to understand if some external characteristics could affect the participants answers. Considering that our main goal was to test the EDS-R and the EAI latent structures across competition level and gender, we tested a series of multigroup CFAs on these variables, separately. As recommended by several authors (Putnick & Bornstein, 2016; Vandenberg & Lance, 2000), we focused on the configural, metric, scalar and residual invariances. **As a first step, we performed an omnibus test to both fit a baseline model and to test the general structure of both instruments**. After testing separate model for each group, we tested the configural invariance, to understand if the factor structure was equal across groups (van de Schoot et al., 2012). Configural invariance was supported if the model presented (at least) adequate goodness-of-fit indices and the pattern of loadings were significant in all groups. Then, we constrained the factor loadings to be equal across groups to test the metric invariance. Such test allows to understand if the contribution (i.e., the loading) of each item to its factor is equivalent among groups. Metric invariance was supported if the model fit was not worse than the fit of configural model. As suggested by several authors (Cheung & Rensvold, 2002; Gilson et al., 2013) we focused on two criteria to compare both models fit and, therefore, to discuss about the presence of measurement invariance: (1), the difference (Δ) between fit indices of the models, where we considered a $\Delta CFI > .01$ and a $\Delta RMSEA > .015$ as indicative of invariance violations; (2), the overall goodness of fit of the singular models: not acceptable fit suggested us potential non-invariance (Beaujean et al., 2011). As a fourth step, we constrained both the factor loadings and the **item intercepts/thresholds**[4] to be equal across groups, to test scalar invariance. In this way, it was possible to understand if both the starting level toward the construct and the threshold of response were equivalent among groups. We compared the overall fit of this model model with the metric invariance one. Finally, we tested the residual invariance (i.e., by constraining factor loadings, item intercepts/thresholds and residual variances to be equal across groups) to understand if the specific and error variances of the items were equivalent among groups (Putnick & Bornstein, 2016). In this case, model fit was compared with the scalar invariance one. Among the comparisons, whenever measurement non-invariance was found, we decided to test also partial measurement invariance. In particular, we constrained the parameter of the invariant items, relaxing the other parameters and comparing again the models. In this way, it was possible to identify which groups could interpret the construct assessed by a specific item in different ways. We used the *lavaan* package (Rosseel, 2012)

[4] In case of ordinal scale as the ones in the present study, an item threshold can be defined as the level of latent trait required to endorse a point of the response scale more likely than the previous response point. For ordinal and Likert-like scale, each item will have n-1 thresholds. Since item belonging to EAI are rated on a 5-point Likert scale, they present four thresholds. Items belonging to EDS-R present five thresholds.

to test measurement invariance. The procedure we used to define and compare models followed the recommendations of Hirschfeld and von Brachel (2014).

As a further comparison between instruments, we also tested the specificity and sensitivity of both tool by means of receiver operating characteristic (ROC) curve. This analysis also yielded potential cutoffs for both tools, comparing them with the ones emerged from the original scoring algorithms through the kappa statistics (Cohen, 1960). We used two external criteria to run ROC analysis, namely the answers to the question Q1 and the percentiles on the hours of weekly practice. Participants who both (a) expressed the will to continue the sport activities despite medical prescription and (b) were over the 95% percentile along the distribution of the training hours were considered as at risk of addiction. We determined the cut off on the total scores of both instruments, in line with previous works testing the structure of the instruments (Müller et al., 2013). We used the pROC package (Robin et al., 2011)**.**

To understand whether the EDS-R and EAI scores were influenced by gender, years of practice and average time devoted to exercise every week, we adopted multiple linear regressions. We also considered both gender × years of practice and gender × average time devoted to exercise every week interactions. We used partial eta squared values (ηp2) to quantify the effect size, by means of the *heplots* package (Friendly, 2010).

Finally, based on EDS-R and EAI outcomes, we tested the difference between addicted versus non addicted exercisers on the adherence to medical prescription. We performed a multiple logistic regression, setting Q2 responses as dependent variable and EDS-R/EAI outcomes and the competition level as categorical predictors. Regarding the EDS-R/EAI outcomes, we were interested in comparing participants considered at risk of exercise addiction to participants who were either symptomatic-not addicted or asymptomatic. We also considered this comparison across competition levels. Moreover, odds ratios were calculated for each coefficient. **About the sample size used for this last group of analyses (i.e., logistic, multiple regressions and ROC), we used only those participants who provided data on the target categorical variables. For instance, in the multiple regressions, we**

used all those individuals that provided complete data on gender, years of practice and average practice time. To calculate scores in case of missing data on specific items of both EAI and EDS-R, we applied a mean imputation method, but only for those subjects who had at least the ~90% of items responses. As can be seen in Table S2, we imputed a very small set of cases, considering that the amount of missing data (per item) was negligible. This strategy allowed us to reduce the loss of useful data.

Results

Prevalence and descriptive statistics

The calculated prevalence of exercise addiction was 6% and 11% when measured with the EDS-R and EAI respectively. **Considering mean differences on scales and subscales' scores, we observed that, in general, male participants obtained lower scores than female ones in all the overall scales and subscales, except for tolerance (p= .54), lack of control (p= .07), reduction in other activities (p=.21) and intention effects EDS-R subscales (p=.59). Furthermore, we observed that participants who reported an intention to continue exercising (i.e.,data taken from answer to both questions Q1 and Q2), despite medical prescriptions or suggestions, obtained higher scores in all the overall scales and subscales. We did not observe mean differences on the withdrawal subscale between respondents to the Q1 question (p= .31). Finally, we observed that participants competing at international level obtained the highest scores (especially compared to amatorial participants) in all the scales and subscale and, such result, was statistically significant for both EAI and EDS-R total score and almost all the EDS-R subscales (see Table 1).**

Latent structure and measurement invariance

The sample used to run all the CFAs consisted of 943 participants (i.e., composed by participants who reported their gender, level of competition and all the responses to EAI and EDS-R). **Descriptive**

statistics of this subsample can be found within supplementary materials (see Table S3). We did not find systematic patterns of missingness among items (all the items presented very low and similar rates of missingness). We also did not find age differences among participants who were included or excluded. Regarding the EAI, the CFA suggested a good fit of the entire model to the data ($\chi 2(9)= 29.33$; CFI= .98; SRMR= .03; RMSEA= .05[.01-.05]). Standardized factor loadings ranged from .33 to .69 and are reported in Table 2, and internal consistency was α= .68. Independent CFAs on each competition type obtained good to acceptable fit indexes, (range CFI: .95-.99; range SRMR: .04-.06; range RMSEA: .05-.09) except for the RMSEA of amatorial athletes (.09; see Table 3). As for stratified models on competition levels, CFAs on gender revealed good fit indices (female: $\chi 2(9)= 6.79$; CFI= 1; SRMR= .03; RMSEA= .00[.00-.06]; male: $\chi 2(9)= 28$; CFI= .98; SRMR= .03; RMSEA= .05[.03-.07]). Multigroup CFA to test measurement invariance showed that the model on competition level was good in terms of configural and metric invariance, but some invariance violations emerged comparing the metric with the scalar model (ΔCFI=.03; ΔRMSEA=.02). **Once we found the violation, we applied the Lagrange Multiplier test (by jointly using the *lavTestScore()* and the *parTable()* functions of the lavaan package), for releasing single constraints. This test allowed us to test if item thresholds statistically differed among groups.** Examining such items thresholds of the scalar model, we found that the threshold four of the item three of EAI (i.e., "I use exercise as a way of changing my mood") of athletes competing at international level was the lowest (0.14) compared to the thresholds of athletes competing at different levels (ranging from 0.22 to 0.25). Likewise, the threshold three for of item four of EAI (i.e., "Over time I have increased the amount of exercise I do in a day") of athletes competing at regional level was the lowest (|-0.57|) compared to the thresholds of athletes competing at different levels (ranging from |-0.61| to |-0.69|). Therefore, we relaxed the parameters of those thresholds and tested the partial scalar invariance. As shown in Table 3, the goodness-of-fit indices and the deltas improved (ΔCFI < .01; ΔRMSEA < .01).

On the contrary, the model across gender suggested configural invariance, with good indices of fit ($\chi2(18)$= 32.44; CFI= .99; SRMR= .03; RMSEA= .04[.02-.06]), suggesting the same factor structure across female and male participants. Similar results emerged in terms of metric, scalar and residual invariance. Overall, the results indicate that the EAI showed the same factor loadings, the same item intercepts and the same residual variances across female and male athletes.

[PLEASE, INSERT TABLE 2 HERE]

Regarding the EDS-R, as reported in Table 3, the fit of the entire model to the data resulted as good ($\chi2(168)$ = 805.06; CFI= .97; SRMR= .05; RMSEA= .06[.06-.07]). Standardized factor loadings ranged from .54 to .97, as shown in Table 2. EDS-R showed very good internal consistency both as an entire scale (α= .83), and within each subscale (α withdrawal= .92; α continuity= .80; α tolerance= .82; α control loss= .85; α reduction= .57; α time= .79; α intention= .82) except for the reduction in other activities subscale. In the multigroup CFAs, all the models on competition levels ranged from good to acceptable indices (range CFI: .97-.98; range SRMR: .06-.08; range RMSEA: .06-.07). Similarly, we found acceptable goodness-of-fit indices on gender models (female: $\chi2(168)$= 348.47; CFI= .97; SRMR= .06; RMSEA= .07[.06-.08] male: $\chi2(168)$= 634.08; CFI= .98; SRMR= .05; RMSEA= .06[.06-.07]). We observed configural invariance both on competition level and gender, with also metric, scalar and residual invariances supported by small changes in models fit (see Table 3).

[PLEASE, INSERT TABLE 3 HERE]

On the ROC analysis **(n= 1006)**, we observed that a cutoff of 19.5 for the EAI was sufficient to determine people at risk to develop exercise addiction. The specificity was .56, the sensitivity was .67 and the area under the curve (AUC) was equal to .67 **(Figure 1, upper panel)**. Based on this new cut off, 439 participants were screened as at risk (43.22%). The kappa coefficient between the new cut off and the one suggested by Hausenblas and Symons Downs (2002b) was .28, suggesting minimal agreement (Cohen, 1960). In regard to the EDS-R, we observed that participants with a score higher than 76.5 were screened as at risk, with a specificity of .82, a sensitivity of .67 and an AUC of

.73 **(Figure 1, lower panel)**. Based on this new cutoff, 181 participants were screened as at risk (i.e., 17.90%). In this case, the agreement between the new and the original cutoffs was weak (κ= .41).

[PLEASE, INSERT FIGURE 1 HERE]

Effects of gender, years of practice and average practice time on exercise addiction

The multiple regression **(n= 999)** on EAI scores yielded no significant effects on gender, years of practice and the interaction effects. The effect of the average time per week spent in training emerged as small ($\eta p2$= .03) but statistically significant: a one-unit increase in hours per week spent in training led to an increase of .11 in EAI score.

[PLEASE, INSERT TABLE 4 HERE]

The multiple regression on the EDS-R total scores **(n= 999)** also yielded no significant effects across years of practice or interaction effects. There was, however, a small but significant effect of gender ($\eta p2$< .01). Moreover, we found that a one-unit increase in hours per week spent in training led to an increase of .39 in EDS-R score ($\eta p2$= .047).

Finally, logistic regression **(n= 976)** suggested that only the EDS-R showed a statistically significant difference between participants at risk of exercise addiction and non-addicted. We observed that the odds of continuing exercising (even in case of a critical cardiac anomaly) in participants who were screened at risk of exercise addiction were 2.91 times greater than the odds of continuing exercising in participants who were screened as exercise non-addicted (OR=2.91 CI=1.01-8.94 p= .04; see Table 5), a results that was not found in the EAI. In this case, the odds of continuing exercising (even in case of a critical cardiac anomaly) in participants who competed at local level were 0.62 times lower than the odds of continuing exercising in participants who competed at amatorial level (OR=0.62 CI=0.4-0.96 p= .03).

[PLEASE, INSERT TABLE 5 HERE]

Discussion

During the past ten years, exercise addiction has received increasing attention (Di Lodovico et al., 2019). There is the possibility that sport could be compulsive to the point that it may modify the adherence to medical prescription to cease exercise, if required. Athletes may find this particularly difficult, considering that most of the time athletes are guided by a strong passion that could moderate the relevance of any exercise addiction (Vega et al., 2016). Both main and differential diagnoses can be difficult (Freimuth et al., 2011), as exercise addition has been reported to exist in a continuum, from beneficial exercise to exercise addiction. Furthermore, a number of comorbid disorders, such as eating disorders in particular, and other addictive or impulsive behaviors can co-exist (Berczik et al., 2012; Freimuth et al., 2011; Lichtenstein et al., 2018; Pinna et al., 2015; Trott, Jackson, Firth, Jacob, et al., 2020). Therefore, it is fundamental to build and test assessment instruments, whose structure must be stable not only in general populations, but also in different athletic populations across several different competition levels. Moreover, a tradeoff between brevity and exhaustivity must be reached for such assessment tools. Currently there are two well-validated questionnaires, the EDS-R and the EAI, with the EDS-R having the advantage of investigating more aspects related to exercise addiction, and the EAI being faster to administer. Although some studies have compared these tools in terms of validity, usability and factorial structure, these validation studies were not stratified by gender or athletic level of competition. The present study tried to add new insight onto such issues debated in literature (Di Lodovico et al., 2019; Mónok et al., 2012), retesting the latent structures of EAI and EDS-R across competition level and gender. Furthermore, the present study tried to understand if, based on those tools' outcome, those identified as at risk of exercise addiction would be less adherent the medical prescriptions, compared to non-addicted individuals.

The difference in prevalence rates was large, and in general agreement with previous studies that have found the EAI to yield higher prevalence rates than the EDS-R (Trott, Jackson, Firth, Fisher, et al., 2020; Zorzi et al., 2020), possibly because of the differing number of items and underlying theoretical structures the tools are based on. The ratio between prevalence rates was confirmed also by the ROC analyses, which suggested that EAI screened more participants as "at risk" than the EDS-R, even using new external criteria. These results suggest that the EDS-R should be preferred to the EAI in cases where more conservative screening for exercise addiction is necessary.

We observed that both instruments showed a good overall factorial structure, in line with previous works (Mónok et al., 2012; Terry et al., 2004). Moreover, we found that both the EAI and EDS-R can be independently used by female and male athletes **(even if male participant reported, in descriptive terms, lower scores in almost all the tested scales and subscales)**, as suggested by their measurement invariance. These results suggest, therefore, that both instruments can be used in scenarios where screening of exercise addiction is required, controlling for gender. Some potential issues in terms of measurement invariance emerged for the EAI where violations of scalar invariance emerged**. Recalling that an item threshold can be defined as the level of latent trait necessary to endorse a point of the response scale more likely than the previous response point, we found that the fourth threshold of item three (i.e., "I use exercise as a way of changing my mood") and the third threshold of item four (i.e., "Over time I have increased the amount of exercise I do in a day") were the two sources of such scalar non-invariance**. **In the former, we observed that athletes competing at international level had the lowest fourth threshold, suggesting for these athletes it is necessary for a lower level of the latent trait to answer a "5" (the top score) to this item over a "4". In other words, such athletes seem to be more prone to score the maximum on this item than other athletes.** In line with previous works with similar results (e.g. Rhudy et al., 2020), this result could reflect either a real difference toward the tendency of using exercise as a way to change mood or a response bias caused by different endorsement of the response among groups. In the latter, we observed that the lowest third threshold belonged to the athletes competing at regional level, suggesting the same dichotomy found for item three. On the contrary, the structure of the EDS-R emerged as invariant across competition levels. Although two thresholds of 24 (four thresholds × 6 items of the EAI) could be interpreted as trivial from a measurement point of view (Rhudy et al., 2020), it could be preferable to use an instrument that has fewer critical points. Therefore, it appears that the EDS-R is preferable to the EAI in this population.

A slight difference between the instruments emerged in the ROC analysis. Based on the two external criteria, we observed that the EAI yielded more participants at risk of exercise addiction than the EDS-R did, but with less sensitivity. The ratio between our new prevalence rates were broadly in line with the ratio found by calculating the prevalence rates using the methods proposed by the instrument's authors. It appears that the EAI consistently yields almost double the prevalence of participants at risk of EA than the EDS-R. It is important to stress, however, that the results of the ROC analyses should be taken with caution. Both minimal (for EAI) and weak (for EDS-R) agreements with the original scoring algorithm suggested that the measure used as external criteria are not free from bias, due to the self-reporting nature of the tools. This constitutes a limitation of the study but also paves also the way for future studies aimed at finding more objective criteria.

Another aspect highlighting a difference between the EAI and the EDS-R came after testing the difference on adherence to medical prescriptions. **In general, we observed that participants who reported an intention to continue exercising obtained higher scores in both EAI and EDS-R scales and almost all EDS-R subscales. Beyond such descriptive results**, the analysis the logistic regression suggested that, when the medical prescription was to permanently stop exercising due to a potential (and hypothetical) cardiac anomaly, athletes at risk of addiction were less prone to adhere to the prescription, but only when exercise addiction was measured by the EDS-R. This result was independent of competition level. Further study should aim to replicate these results that, currently, are merely exploratory.

Such results have several potential clinical implications, beyond the psychometric aspects. First, the risk of exercise addiction appears to be not so infrequent, and should be routinely assessed in competitive athletes, given that it can lead to excessive exercising with higher risk of injuries, impairment in occupational, and social functioning (Hausenblas et al., 2017). Considering the method of exercise addiction measurement, the present results suggest that, in case of competitive athletes, the EDS-R is preferable to the EAI, mainly because the responses to the questions in this tool appear to not be influenced by the competition level of an athlete. Levels of exercise among athletes are of course higher than those in the general population, and it is reasonable that anaerobic and cardio-respiratory training of the athlete be judged as excessive from the general population point of view. Nonetheless, when occupational and social functioning becomes impaired due to excessive time and mental focus dedicated to exercise, which is often the criteria for other forms of addiction, exercising can assume a pathological entity in this population (Hausenblas et al., 2017). Reducing exercise and replacing this with other stimulating activities should be attempted once a person is judged as addicted to exercise. Interventions to prevent occupational, social functioning impairment and recurrent injuries occurring from exercise addiction should be ideally implemented systemically, from the lowest level of competitions, since more engrained addictive behaviours could emerge when the athlete career improves.

Secondly, we observed differences between addicted and non-addicted athletes in terms of adherence to medical prescriptions. When the medical prescription was to permanently stop exercise due to a cardiac anomaly, this difference was detected only by assessing the risk of exercise addiction with the EDS-R, in line with a previous research (Zorzi et al., 2020). Such result suggests that the EDS-R could detect athletes' attitude more effectively than the EAI. This result is particularly relevant among subjects who may receive a contraindication to continue exercise. If they continue exercising despite the presence of a cardiac disease at risk of sudden cardiac death, the risk goes beyond social and occupational impairment or injuries. Hence, one further clinical implication of the present findings is that exercise addiction should be assessed in people with cardiac anomalies.

Thirdly, our results regarding the psychometric properties of the EDS-R across gender and competition level, provide new data in the literature comparing it with EAI (Di Lodovico et al., 2019; Mónok et al., 2012). It appears that the EDS-R is a valuable candidate to assess exercise addiction, even if administered in non-psychological contexts. This aspect is not trivial: in general, the medical assessment to allow or not allow a person practicing competitive exercise is performed by a medical doctor specialized in cardiology or sport medicine. Both may not be very familiar with questionnaires or scales and need to assess several biological variables during an often-small amount of time. Hence, they need an exhaustive tool to routinely administer it as a part of the whole evaluation. Our results showed that EDS-R can provide useful results, and information less influenced by external features, although it contains a greater number of items and dimensions than EAI.

The present work has some limitations. The operational definition of athletes we used in this work follows Italian law (Decree of Italian Ministry of Health: Rules for the Health Care of Competitive Sport Activities, 1982), and guidelines derived from sport cardiology (Pelliccia et al., 2020). Therefore, such a definition could be context-specific, namely it does not completely overlap with the ones that can be found in literature (Araújo & Scharhag, 2016). It would be interesting to understand if our results could be replicated adopting more international definitions. Furthermore, this study is cross-sectional, hence no causal inference can be made. The question on adherence to medical prescription assessed an intention not to comply, asked in a self-report modality. Given all the limitations of such a self-report tool, other and more objective measures could be used in future studies. Moreover, multigroup CFAs of competition type either across or separately for gender would have been beneficial, but it would require a sample fitting 10 groups (5 competition levels × 2). Our sample size, although consisting of 943 participants with complete data, would have needed to be higher to provide reliable results. The present sample was also unbalanced across both gender and competition types, increasing the risk of mis specified and underfitted models. To address this, future studies should collect more responses from female athletes at different competition levels. Finally, it would be interesting to screen athletes for sub-threshold OCD or eating disorders considering that such disorders are often comorbid or primary to exercise addiction (Berczik et al., 2012; Freimuth et al., 2011; Lichtenstein et al., 2018; Pinna et al., 2015; Trott, Jackson, Firth, Jacob, et al., 2020).

In summary, this study presents new useful findings. Our results indicate that the usefulness of the EAI may be limited to screening of exercise addiction requiring responses not influenced by gender. Furthermore, the structure of the EDS-R appears to be less influenced by gender or competition level of the athlete. The EAI identifies more individuals reaching thresholds of exercise addiction (Di Lodovico et al., 2019), possibly overestimating exercise addiction. The EDS-R can provide a comprehensive and multidimensional picture of addiction related symptoms, and appears to be less affected by athletes' characteristics. Finally, the outcome provided by EDS-R can be used to detect differences on athletes' attitude to comply to medical prescription to stop exercise, independently of their competition level.

We therefore conclude that the EDS-R is a good candidate to be administered in several contexts, including offices of general practitioners (Di Lodovico et al., 2019; Griffiths et al., 2015; Mónok et al., 2012).

Acknowledgments:

Brendon Stubbs is supported by a Clinical Lectureship (ICA-CL-2017-03-001) jointly funded by Health Education England (HEE) and the National Institute for Health Research (NIHR). Brendon Stubbs is part funded by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust. Brendon Stubbs also holds active grants with the Medical Research Council (GCRF and multimorbidity calls) and Guys and St Thomas Charity (GSTT). Brendon Stubbs has received consultancy fees from ASICS Europe BV. The views expressed are those of the author(s) and not necessarily those of the (partner organisation), the NHS, the NIHR, the Department of Health and Social Care, the MRC or GSTT.

Joseph Firth is supported by a University of Manchester Presidential Fellowship (P123958) and a UK Research and Innovation Future Leaders Fellowship (MR/T021780/1).

All other authors have nothing to declare.

References

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders-IV Text Revision, APA. In *Washington, DC*.

Araújo, C. G. S., & Scharhag, J. (2016). Athlete: a working definition for medical and health sciences research. *Scandinavian Journal of Medicine & Science in Sports*, *26*(1), 4–7. https://doi.org/10.1111/sms.12632

Ashdown-Franks, G., Firth, J., Carney, R., Carvalho, A. F., Hallgren, M., Koyanagi, A., Rosenbaum, S., Schuch, F. B., Smith, L., Solmi, M., Vancampfort, D., & Stubbs, B. (2020). Exercise as Medicine for Mental and Substance Use Disorders: A Meta-review of the Benefits for Neuropsychiatric and Cognitive Outcomes. *Sports Medicine*, *50*(1), 151–170. https://doi.org/10.1007/s40279-019-01187-6

Baekeland, F. (1970). Exercise Deprivation: Sleep and Psychological Reactions. *Archives of General Psychiatry*, *22*(4), 365–369. https://doi.org/10.1001/archpsyc.1970.01740280077014

Beaujean, A. A., Freeman, M. J., Youngstrom, E., & Carlson, G. (2011). The Structure of Cognitive Abilities in Youths With Manic Symptoms: A Factorial Invariance Study. *Assessment*, *19*(4), 462–471. https://doi.org/10.1177/1073191111399037

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Berczik, K., Szabó, A., Griffiths, M. D., Kurimay, T., Kun, B., Urbán, R., & Demetrovics, Z. (2012). Exercise Addiction: Symptoms, Diagnosis, Epidemiology, and Etiology. *Substance Use & Misuse*, *47*(4), 403–417. https://doi.org/10.3109/10826084.2011.639120

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology press.

Costa, S., Cuzzocrea, F., Hausenblas, H. A., Larcan, R., & Oliva, P. (2012). Psychometric examination and factorial validity of the Exercise Dependence Scale-Revised in Italian exercisers. *Journal of Behavioral Addictions JBA*, *1*(4), 186–190. https://doi.org/10.1556/jba.1.2012.009

Decree of Italian Ministry of Health: Rules for the health care of competitive sport activities, (1982).

Di Lodovico, L., Poulnais, S., & Gorwood, P. (2019). Which sports are more at risk of physical exercise addiction: A systematic review. *Addictive Behaviors*, *93*, 257–262. https://doi.org/https://doi.org/10.1016/j.addbeh.2018.12.030

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573

Freimuth, M., Moniz, S., & Kim, S. R. (2011). Clarifying Exercise Addiction: Differential Diagnosis, Co-occurring Disorders, and Phases of Addiction. *International Journal of Environmental Research and Public Health*, *8*(10), 4069–4081. https://doi.org/10.3390/ijerph8104069

Friendly, M. (2010). HE Plots for Repeated Measures Designs. *Journal of Statistical Software; Vol 1, Issue 4 (2010)* . https://doi.org/10.18637/jss.v037.i04

Gilson, K.-M., Bryant, C., Bei, B., Komiti, A., Jackson, H., & Judd, F. (2013). Validation of the Drinking Motives Questionnaire (DMQ) in older adults. *Addictive Behaviors*, *38*(5), 2196–

2202. https://doi.org/https://doi.org/10.1016/j.addbeh.2013.01.021

Griffiths, M. D., Urbán, R., Demetrovics, Z., Lichtenstein, M. B., de la Vega, R., Kun, B., Ruiz-Barquín, R., Youngman, J., & Szabo, A. (2015). A cross-cultural re-evaluation of the Exercise Addiction Inventory (EAI) in five countries. *Sports Medicine - Open*, *1*(1), 5. https://doi.org/10.1186/s40798-014-0005-5

Hausenblas, H. A., Schreiber, K., & Smoliga, J. M. (2017). Addiction to exercise. *BMJ*, *357*, j1745. https://doi.org/10.1136/bmj.j1745

Hausenblas, H. A., & Symons Downs, D. (2002a). Exercise dependence: a systematic review. *Psychology of Sport and Exercise*, *3*(2), 89–123. https://doi.org/https://doi.org/10.1016/S1469-0292(00)00015-7

Hausenblas, H. A., & Symons Downs, D. (2002b). Exercise Dependence Scale-21 Manual. *Recuperado de Http://Www. Personal. Psu. Edu/Dsd11/EDS/EDS21Manual. Pdf*.

Hausenblas, H. A., & Symons Downs, D. (2002c). How Much is Too Much? The Development and Validation of the Exercise Dependence Scale. *Psychology & Health*, *17*(4), 387–404. https://doi.org/10.1080/0887044022000004894

Hirschfeld, G., & Von Brachel, R. (2014). Improving Multiple-Group confirmatory factor analysis in R–A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation*, *19*(1), 7.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kim, Y., Sharp, S., Hwang, S., & Jee, S. H. (2019). Exercise and incidence of myocardial infarction, stroke, hypertension, type 2 diabetes and site-specific cancers: prospective cohort study of 257 854 adults in South Korea. *BMJ Open*, *9*(3), e025590.

https://doi.org/10.1136/bmjopen-2018-025590

Kyriazos, T. A. (2018). Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*, *09*(08), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lejoyeux, M., Avril, M., Richoux, C., Embouazza, H., & Nivoli, F. (2008). Prevalence of exercise dependence and other behavioral addictions among clients of a Parisian fitness room. *Comprehensive Psychiatry*, *49*(4), 353–358. https://doi.org/https://doi.org/10.1016/j.comppsych.2007.12.005

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lichtenstein, M. B., Griffiths, M. D., Hemmingsen, S. D., & Støving, R. K. (2018). Exercise addiction in adolescents and emerging adults – Validation of a youth version of the Exercise Addiction Inventory. *Journal of Behavioral Addictions J Behav Addict*, *7*(1), 117–125. https://doi.org/10.1556/2006.7.2018.01

Maron, B. J. (2007). Hypertrophic Cardiomyopathy and Other Causes of Sudden Cardiac Death in Young Competitive Athletes, with Considerations for Preparticipation Screening and Criteria for Disqualification. *Cardiology Clinics*, *25*(3), 399–414. https://doi.org/https://doi.org/10.1016/j.ccl.2007.07.006

Mónok, K., Berczik, K., Urbán, R., Szabo, A., Griffiths, M. D., Farkas, J., Magi, A., Eisinger, A., Kurimay, T., Kökönyei, G., Kun, B., Paksi, B., & Demetrovics, Z. (2012). Psychometric properties and concurrent validity of two exercise addiction measures: A population wide study. *Psychology of Sport and Exercise*, *13*(6), 739–746. https://doi.org/https://doi.org/10.1016/j.psychsport.2012.06.003

Müller, A., Claes, L., Smits, D., Gefeller, O., Hilbert, A., Herberg, A., Müller, V., Hofmeister, D., & de Zwaan, M. (2013). Validation of the German Version of the Exercise Dependence Scale. *European Journal of Psychological Assessment*, *29*(3), 213–219. https://doi.org/10.1027/1015-5759/a000144

Pelliccia, A., Sharma, S., Gati, S., Bäck, M., Börjesson, M., Caselli, S., Collet, J.-P., Corrado, D., Drezner, J. A., Halle, M., Hansen, D., Heidbuchel, H., Myers, J., Niebauer, J., Papadakis, M., Piepoli, M. F., Prescott, E., Roos-Hesselink, J. W., Graham Stuart, A., … Wilhelm, M. (2020). 2020 ESC Guidelines on sports cardiology and exercise in patients with cardiovascular disease. *European Heart Journal*. https://doi.org/10.1093/eurheartj/ehaa605

Pinna, F., Dell'Osso, B., Di Nicola, M., Janiri, L., Altamura, A. C., Carpiniello, B., & Hollander, E. (2015). Behavioural addictions and the transition from DSM-IV-TR to DSM-5. *Journal of Psychopathology*, *21*(4), 380–389.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rhudy, J. L., Arnau, R. C., Huber, F. A., Lannon, E. W., Kuhn, B. L., Palit, S., Payne, M. F., Sturycz, C. A., Hellman, N., Guereca, Y. M., Toledo, T. A., & Shadlow, J. O. (2020). Examining Configural, Metric, and Scalar Invariance of the Pain Catastrophizing Scale in Native American and Non-Hispanic White Adults in the Oklahoma Study of Native American Pain Risk (OK-SNAP). *Journal of Pain Research*, *13*, 961–969. https://doi.org/10.2147/JPR.S242126

Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response

Analysis. *Journal of Statistical Software; Vol 1, Issue 5 (2007)* . https://doi.org/10.18637/jss.v017.i05

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. https://doi.org/10.1186/1471-2105-12-77

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software; Vol 1, Issue 2 (2012)* . https://doi.org/10.18637/jss.v048.i02

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Sicilia, A., Alías-Garcia, A., Ferriz, R., & Moreno-Murcia, J. A. (2013). Adaptación y validación al español del Inventario de Adicción al Ejercicio (EAI). *Psicothema*, *25*(3), 377–384.

Steiger, J. H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

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 . *European Psychiatry*, *54*, 124–144. https://doi.org/DOI: 10.1016/j.eurpsy.2018.07.004

Szabo, A., Griffiths, M. D., de La Vega Marcos, R., Mervó, B., & Demetrovics, Z. (2015). Methodological and Conceptual Limitations in Exercise Addiction Research. *The Yale Journal* *of Biology and Medicine*, *88*(3), 303–308. https://pubmed.ncbi.nlm.nih.gov/26339214

Terry, A., Szabo, A., & Griffiths, M. (2004). The exercise addiction inventory: a new brief screening tool. *Addiction Research & Theory*, *12*(5), 489–499. http://10.0.4.56/16066350310001637363

Trott, M., Jackson, S. E., Firth, J., Fisher, A., Johnstone, J., Mistry, A., Stubbs, B., & Smith, L. (2020). Exercise Addiction Prevalence and Correlates in the Absence of Eating Disorder Symptomology: A Systematic Review and Meta-analysis. *Journal of Addiction Medicine*, *Publish Ah*. https://journals.lww.com/journaladdictionmedicine/Fulltext/9000/Exercise_Addiction_Prevalence_and_Correlates_in.99216.aspx

Trott, M., Jackson, S. E., Firth, J., Jacob, L., Grabovac, I., Mistry, A., Stubbs, B., & Smith, L. (2020). A comparative meta-analysis of the prevalence of exercise addiction in adults with and without indicated eating disorders. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*. https://doi.org/10.1007/s40519-019-00842-1

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Vega, R. de la, Parastatidou, I. S., Ruíz-Barquín, R., & Szabo, A. (2016). Exercise Addiction in Athletes and Leisure Exercisers: The Moderating Role of Passion. *Journal of Behavioral Addictions J Behav Addict*, *5*(2), 325–331. https://doi.org/10.1556/2006.5.2016.043

Vessella, T., Zorzi, A., Merlo, L., Pegoraro, C., Giorgiano, F., Trevisanato, M., Viel, M., Formentini, P., Corrado, D., & Sarto, P. (2020). The Italian preparticipation evaluation programme: diagnostic yield, rate of disqualification and cost analysis. *British Journal of Sports Medicine*, *54*(4), 231 LP – 237. https://doi.org/10.1136/bjsports-2018-100293

Vina, J., Sanchis-Gomar, F., Martinez-Bello, V., & Gomez-Cabrera, M. C. (2012). Exercise acts as a drug; the pharmacological benefits of exercise. *British Journal of Pharmacology*, *167*(1), 1–12. https://doi.org/10.1111/j.1476-5381.2012.01970.x

Weston, R., & Gore, P. A. (2006). A Brief Guide to Structural Equation Modeling. *The Counseling Psychologist*, *34*(5), 719–751. https://doi.org/10.1177/0011000006286345

Wouthuyzen-Bakker, M., & van Assen, S. (2015). Exercise-induced anaemia: a forgotten cause of iron deficiency anaemia in young adults. *British Journal of General Practice*, *65*(634), 268 LP – 269. https://doi.org/10.3399/bjgp15X685069

Zorzi, A., Cardaioli, F., Cipriani, A., D’Ascenzi, F., Corrado, D., & Solmi, M. (2020). Correlation between exercise addiction and attitude of competitive athletes towards cardiac preparticipation evaluation. *Journal of Cardiovascular Medicine (Hagerstown, Md.)*. https://doi.org/10.2459/jcm.0000000000000997

Tables

Table 1
Descriptives of the sample

Variable		EDS-R (M; SD)								EAI (M; SD)
	%	Total	Withdrawal	Continuity	Tolerance	Control	Reduction	Time	Intention	Total
Gender										
Missing	0									
Females[a]	23.83	65.5	11.93	8.77	9.76	7.53	7.57	12.83	7.1	19.58
		(15.28)	(4.3)	(4.1)	(3.72)	(3.72)	(2.65)	(3.01)	(3.28)	(4.41)
Males	76.17	62.15	11.16	7.85	9.58	7.03	7.3	12	7.22	18.39
		(14.16)**	(4.39)*	(3.99)**	(3.48)	(3.39)	(2.66)	(3.01)**	(3.11)	(4.63)**
Level of practiced sport										
Missing	3.59	63.74	11.83	8.56	9.77	7.31	7.44	11.8	7.03	17.23
		(17.62)	(5.02)	(3.95)	(3.41)	(3.13)	(2.45)	(3.8)	(3.49)	(5.06)
Amateur[a]	26.23	60.55	11.6	7.5	9.11	6.71	7.14	11.26	7.22	17.78
		(14.8)	(4.34)	(3.86)	(3.6)	(3.28)	(2.7)	(3.04)	(3.22)	(4.52)
Local	18.65	61.68	11.85	8.53	8.87	6.76	7.22	11.57	6.89	18.47
		(13.66)	(4.17)	(4.02)*	(3.22)	(3.22)	(2.52)	(3.1)	(3.18)	(4.7)
Regional	19.98	62.45	11.31	796	9.32	6.94	7.66	12.39	6.86	19.17
		(12.16)	(4.49)	(3.93)	(3.4)	(3.18)	(2.66)	(2.65)****	(2.83)	(4.46)**
National	24.49	64	10.96	8.13	9.92	7.56	7.39	12.73	7.3	18.65
		(14.96)*	(4.34)	(4.1)	(3.39)*	(3.78)*	(2.69)	(2.81)****	(3.2)	(4.46)
International	10.66	69.51	10.69	8.63	12.05	8.3	7.54	14.19	8.11	20.87
		(15.37)****	(4.44)	(4.45)*	(3.49)****	(4)****	(2.77)	(2.62)****	(3.19)*	(4.36)****
Answer to Q1										
Missing	0									
"I would stop exercising" (0)[a]	64.39	62.31	11.32	7.95	9.53	7.02	7.27	12.14	7.07	18.55
		(14.1)	(4.36)	(4)	(3.49)	(3.42)	(2.56)	(3.02)	(3.08)	(4.57)
"I would not stop exercising" (1)	35.61	75.26	11.96	10.47	11.27	9.4	9.27	13.43	9.46	21.19
		(16.6)****	(4.74)	(3.8)****	(4)***	(3.9)****	(3.62)****	(3.15)***	(3.63)****	(4.7)****

Answer to Q2										
Missing	0									
"I would stop exercising" (0)[a]	65.88	60.83 (13.59)	11.1 (4.37)	7.64 (3.87)	9.42 (3.5)	6.84 (3.36)	6.99 (2.52)	11.9 (2.99)	6.94 (3.04)	18.21 (4.62)
"I would not stop exercising" (1)	34.12	67.14 (15.34)****	11.83 (4.35)*	8.93 (4.22)****	10 (3.57)*	7.75 (3.63)****	8.1 (2.78)****	12.81 (3.03)****	7.69 (3.32)****	19.61 (4.45)****

Note. EAI, exercise addiction inventory; EDS-R, exercise dependence scale, 21 items. [a]**= this level is assumed as the reference one for analyses testing mean differences among groups. *= p< .05; **= p< .01; ***= p< .001; ****= p< .0001. P values are adjusted by applying a False Discovery Rate correction.**

Table 2

Standardized factor loadings of EDS21 and EAI items

	Item Description	Withdrawal	Continuity	Tolerance	Control	Reduction	Time	Intention	EAI
EDS-R									
1	I exercise to avoid feeling irritable	.97							
8	I exercise to avoid feeling anxious	.88							
15	I exercise to avoid feeling tense	.9							
2	I exercise despite recurring physical problems		.9						
9	I exercise when injured		.66						
16	I exercise despite persistent physical problems		.86						
3	I continually increase my exercise intensity to achieve the desired effects/benefits			.81					
10	I continually increase my exercise frequency to achieve the desired effects/benefits			.79					
17	I continually increase my exercise duration to achieve the desired effects/benefits			.81					
4	I am unable to reduce how long I exercise				.88				
11	I am unable to reduce how often I exercise				.83				
18	I am unable to reduce how intense I exercise				.86				
5	I would rather exercise than spend time with family/friends					.68			
12	I think about exercise when I should be concentrating on school/work					.61			
19	I choose to exercise so that I can get out of spending time with family/friends					.54			
6	I spend a lot of time exercising						.84		
13	I spend most of my free time exercising						.66		
20	A great deal of my time is spent exercising						.85		
7	I exercise longer than I intend							.91	
14	I exercise longer than I expect							.8	
21	I exercise longer than I plan							.75	
EAI									
1	Exercise is the most important thing in my life								.58

2	Conflicts have arisen between me and my family and/or my partner about the amount of exercise	.48
3	I use exercise as a way of changing my mood	.63
4	Over time I have increased the amount of exercise I do in a day	.65
5	If I have to miss an exercise session, I feel moody and irritable	.69
6	If I cut down the amount of exercise I do, and then start again, I always end up exercising as often as I did before	.33

Note. EAI: Exercise Addiction Inventory, 6 items; EDS-R: Exercise Dependence Scale- Revised, 21 items.

Table 3

Multigroup confirmatory factor analyses for EDS21 and EAI across gender and competition type

#	Model	χ^2	df	SRMR	CFI	ΔCFI	RMSEA	ΔRMSEA
	EAI: Competition							
1	Omnibus	29.33	9	.03	.98		.05	
2	Amateur	29.58	9	.05	.95		.09	
3	Local	11.57	9	.04	.99		.04	
4	Regional	18.80	9	.05	.96		.08	
5	National	17.67	9	.04	.97		.06	
6	International	12.09	9	.06	.97		.06	
7	Conf. inv.	55.71	45	.05	.99		.04	
8	Metric i. (7 vs 8)	79.55	65	.06	.99	< .01	.03	< .01
9	Scalar i. (8 vs 9)	202.75	133	.06	.96	.03	.05	.02
10	Par.s.i. (8 vs 10)	155.66	125	.06	.98	< .01	.04	< .01
11	Resid.i. (9 vs 11)	202.75	133	.06	.96	0	.05	0
12	Par.r.i. (10 vs 12)	155.66	125	.06	.98	0	.04	0
	EAI: Gender							
13	Female	6.79	9	.03	1		0	
14	Male	28	9	.03	.98		.05	
15	Conf. inv.	33.99	18	.03	.99		.04	
16	Metric i. (15 vs 16)	31.90	23	.03	.99	< .01	.04	< .01
17	Scalar i. (16 vs 17)	55.53	40	.03	.99	< .01	.03	< .01
18	Resid.i. (17 vs 18)	55.53	40	.03	.99	0	.03	0
	EDS-R: Competition							
19	Omnibus	805.06	168	.05	.97		.06	
20	Amateur	365.88	168	.06	.97		.07	
21	Local	284.42	168	.07	.97		.06	
22	Regional	287.32	168	.07	.97		.06	
23	National	311.01	168	.06	.98		.06	
24	International	264.64	168	.08	.97		.07	
25	Conf. inv.	1000.12	840	.06	.99		.03	
26	Metric i. (25 vs 26)	1203.28	896	.07	.99	< .01	.04	.01
27	Scalar i. (26 vs 27)	1346.17	1156	.07	.99	< .01	.03	.01
28	Resid.i. (27 vs 28)	1354	1172	.07	.99	0	.03	< .01
	EDS-R: Gender							
29	Female	348.47	168	.06	.97		.07	
30	Male	634.08	168	.05	.98		.06	
31	Conf. inv.	957.72	336	.05	.98		.06	
32	Metric i. (31 vs 32)	908.68	350	.05	.98	< .01	.06	< .01
33	Scalar i. (32 vs 33)	1010.32	424	.05	.98	< .01	.05	< .01
34	Resid.i. (33 vs 34)	1010.70	424	.05	.98	0	.05	0

Note. EAI: Exercise Addiction Inventory, 6 items; EDS-R: Exercise Dependence Scale- Revised, 21 items. Conf. Inv.: Configural invariance. Metric i. : Metric invariance. Scalar i. : Scalar invariance. Resid i. : Residual invariance. Par.s.i: partial scalar invariance Par.r.i.: partial residual invariance.

Table 4

Multiple regression results.

	Response variable: EAI			Response variable: EDS-R		
Predictors	**β**	**p**	**ηp2**	**β**	**p**	**ηp2**
Gender (Male)	-1.06	.12	.01	-4.98	.02	.01
Years of practice	-0.02	.3	.01	-0.03	.73	< .01
Average practice time	0.11	< .01	.03	0.34	< .01	.05
Gender X Years of practice	-0.02	.58	< .01	0.01	.98	< .01
Gender X Average practice time	0.02	.58	< .01	0.14	.2	< .01

Note: p: p value; ηp2: partial eta squared

Table 5

Results of multiple logistic regressions of EAI and EDS-R on question evaluating the adherence to medical prescriptions

		Response variable: Q2	
Predictors	**β**	**p value**	**OR(CI)**
EAI			
(Reference: Addicted)	0.64	.19	1.9(.72-5.05)
Competition level			
(Reference: Amateur)			
Local	-0.48	**.03**	**.62(.4-.96)**
Regional	-0.2	.34	.82(.53-1.24)
National	-0.06	.77	.94(.64-1.39)
International	0.28	.29	1.32(.79-2.2)
EDS-R			
(Reference: Addicted)	1.07	**.04**	**2.9(1.01-8.9)**
Competition level			
(Reference: Amateur)			
Local	-0.35	.1	.7(.45-.107)
Regional	-0.12	.53	.88(.58-1.32)
National	-0.01	.98	.99(.67-1.46)
International	0.13	.6	1.14(.68-1.9)

Notes. Q2: *"A medical investigation revealed that you have a critical cardiac anomaly that could cause a deadly heart attack. Unfortunately, it is necessary to permanently stop exercise. How would you react?"*. Statistically significant effects are reported in bold.

Figures

Figure 1

ROC curves for EAI (upper panel) and EDS-R (lower panel).