The association between muscle dysmorphia and eating disorder symptomatology: A systematic review and meta-analysis

LAURA BADENES-RIBERA[1], MARIA RUBIO-APARICIO[2], JULIO SÁNCHEZ-MECA[3],
MATTEO ANGELO FABRIS[4]* and CLAUDIO LONGOBARDI[4]

[1]Departamento de Methodologia and Behavioral Science, University of Valencia, Valencia, Spain
[2]Departamento de Psicología de la Salud, University of Valencia, Alicante, Spain
[3]Departamento de Psicología Básica y Metodología, University of Murcia, Murcia, Spain
[4]Dipartimento di Psicologia, Università degli Studi di Torino, Turin, Italy

Background and aims: Research shows inconsistent findings about the link between muscle dysmorphia (MD) and eating disorder (ED) symptomatology. The aim of this study is to synthesize the scientific evidence available on this topic, the researchers conducted a systematic review and meta-analysis. *Methods:* The literature search enabled us to identify 39 published articles, which provided 36 independent estimations of the correlation between the two variables. *Results:* Our analysis found a positive association between MD and ED symptoms ($r_+ = .36$; 95% CI = 0.30, 0.41). Moderator analyses showed that the type of sample and the tools for assessing MD and ED were statistically associated with the MD–ED effect sizes. The methodological quality of the studies exhibited a positive, statistically significant association with the MD–ED effect sizes. *Conclusions:* Higher levels of MD were related to greater ED symptomatology, but several study characteristics may moderate the association between the two variables. In this study, we discuss limitations and implications for clinical practice and future research.

Keywords: Adonis complex, reverse anorexia, eating behaviors, meta-analysis, muscle dysmorphia

INTRODUCTION

In a study of the psychiatric effects of steroid consumption, Pope, Katz, and Hudson (1993) described a new syndrome called "reverse anorexia," due to its similarities with anorexia nervosa (AN) and subsequently renamed "muscle dysmorphia" (MD; Pope, Gruber, Choi, Olivardia, & Phillips, 1997). These authors suggested that MD is a form of body dysmorphic disorder (BDD) characterized by an obsessive preoccupation with the size and shape of one's muscles, causing significant distress or impairment in daily functioning. Individuals with MD have a preoccupation with not being sufficiently lean and muscular. They believe that their muscles are smaller than they truly are (Grieve, Truba, & Bowersox, 2009; Olivardia, 2001), and they perceive themselves as puny and unattractive (Olivardia, Pope, & Hudson, 2000). Consequently, they seek even greater musculature and greater leanness to enhance the visibility of their muscularity (Choi, Pope, & Olivardia, 2002).

Additional symptoms associated with MD include the following: poor quality of life (Pope et al., 2005; Tod & Edwards, 2015a; Tod, Edwards, & Cranswick, 2016); higher rates of mood and anxiety disorders (Cafri, Olivardia, & Thompson, 2008); risk of obsessive–compulsive symptomatology and interpersonal sensitivity (Longobardi, Prino, Fabris, & Settani, 2017); lower self-esteem and self-perception (Chaney, 2008; Mitchell, Murray, Cobley, et al., 2017);

increased feelings of loneliness (Chaney, 2008); impairments in social and occupational functioning (Olivardia et al., 2000; Pope, Khalsa, & Bhasin, 2017; Tod et al., 2016); and depression, neuroticism, and perfectionism (Mitchell, Murray, Cobley, et al., 2017).

MD is more prevalent in males, particularly those who engage in sports that emphasize increased muscle mass or power gain, such as weightlifting or bodybuilding (American Psychiatric Association [APA], 2013; Cella, Iannaccone, & Cotrufo, 2012; Fabris, Longobardi, Prino, & Settani, 2017; Pope, Phillips, & Olivardia, 2000). However, studies have also observed MD symptoms in females (Hale, Diehl, Weaver, & Briggs, 2013; Readdy, Cardinal, & Watkins, 2011; Robert, Monroe-Chandler, & Gammage, 2009; Tod et al., 2016). Thus, prior research has shown that both female and male bodybuilders are at extremely high risk of having MD symptoms (Hale et al., 2013). The average age of onset of MD is approximately 19–20 years (Cafri et al., 2008; Olivardia, 2001), and its etiology is not yet well known (Grieve, 2007).

At present, the American Psychiatric Association (APA, 2013) recognizes MD as a specifier for BDD in the fifth

* Corresponding author: Matteo Angelo Fabris; Dipartimento di Psicologia, Università degli Studi di Torino, Via Verdi 10, Turin 10124, Italy; Phone: +39 011 6703056; Fax: +39 011 8126231; E-mail: matteoangelo.fabris@unito.it

edition of *Diagnostic and Statistical Manual of Mental Health Disorders* (DSM-5); therefore, it belongs to the obsessive–compulsive disorder spectrum. According to the DSM-5, individuals with MD are preoccupied with the idea that their body build is too small or insufficiently lean or muscular, even though they have a normal-looking body or are quite muscular. Consequently, they perform repetitive behaviors (e.g., mirror checking, excessive grooming, skin picking, and reassurance seeking) or mental acts (e.g., comparing their appearance with that of others) in response to their appearance concerns. Therefore, the body, level of muscularity, and leanness are the obsession, and the compulsion is the desire or drive to gain more muscularity or leanness (Pope et al., 2000; Sandgren & Lavalle, 2018). Other DSM-5 diagnostic criteria specify the following necessary traits of the disorder: (a) appearance preoccupation should cause clinically significant distress or impairment in social, occupational, or other important areas of functioning; and (b) it should not be better explained by concerns with body fat or weight in an individual whose symptoms meet diagnostic criteria for an eating disorder.

Several studies have found an association between MD symptoms and higher levels of ED symptomatology (Giardino & Procidano, 2012; Goodale, Watkins, & Cardinal, 2001; Hildebrandt, Schlundt, Langenbucher, & Chung, 2006; Klimek & Hildebrandt, 2018; Lopez, Pollack, Gonzales, Pona, & Lundgren, 2015; Mitchell, Murray, Hoon, et al., 2017; Murray et al., 2012; Olivardia et al., 2000), little or no control over one's dietary regime, compulsive exercise (Olivardia et al., 2000), diet pill use (McFarland & Kaminski, 2009), vomiting (Hildebrandt et al., 2006; McFarland & Kaminski, 2009), laxative use (Hildebrandt et al., 2006), and a prior history of EDs (Olivardia et al., 2000). Nevertheless, other studies have failed to find an association between MD and ED symptomatology (e.g., Cafri et al., 2008; Maida & Armstrong, 2005). For instance, Cafri et al. (2008) examined the link between MD and EDs in a sample of weightlifting males who met the current criteria for MD, past MD, or no history of MD. They administered the Structured Clinical Interview for DSM Disorders (SCID) to assess MD and EDs. Their findings revealed that the MD group did not show higher rates of EDs, any steroid use, or steroid abuse/dependence than controls. It is possible that the SCID was less sensitive than other assessment tools in detecting the association between MD and ED. Along the same lines, Maida and Armstrong (2005) carried out a study with a sample of males who lift weights and manifest a broad range of attitudes about their bodies, from those falling within the mainstream to those whose worries may be classified as pathological. These authors did not find a relationship between MD symptomatology and the bulimia subscale from the Eating Disorders Inventory (EDI), but they found that MD was positively related to variables measuring symptoms of body dissatisfaction, obsessive–compulsive disorder, depression, and anxiety.

This variability in the findings might be partly due to the different sampling and methodological characteristics of the studies (Nieuwoudt, Zhou, Coutts, & Booker, 2012). Across studies, differences exist in the definition of DM and the measurement tools used to assess MD and ED

symptomatology (Lavender, Brown, & Murray, 2017; Lopez-Cuautle, Vazquez-Arevalo, & Mancilla-Diaz, 2016; Mitchell, Murray, Cobley, et al., 2017; Sandgreen & Lavalle, 2018). Thus, to assess MD, studies have used different self-report questionnaires and interview schedules, such as the Muscle Dysmorphic Disorder Inventory (MDDI; Hildebrandt, Langenbucher, & Schlundt, 2004), three different versions of the Muscle Dysmorphia Inventory (MDI; Lantz, Rhea, & Cornelius, 2002; Rhea, Lantz, & Cornelius, 2004; Short, 2005), the Muscle Appearance Satisfaction Scale (MASS; Mayville, Williamson, White, Netemeyer, & Drab, 2002), the Muscle Dysmorphia Symptom Questionnaire (MDSQ; Olivardia et al., 2000), Hale's Scale (Hale, 2008), and the Muscle Dysmorphia Questionnaire (Cubberley, 2009). To assess ED symptomatology, studies have used the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994), the EDI (Garner, Olmstead, & Polivy, 1983), or the Eating Attitudes Test-26 (EAT-26; Garner, Olmstead, Bohr, & Garfinkel, 1982).

The studies also differ in the sampling procedure (e.g., convenience sample vs. probabilistic sample), the target population investigated [e.g., non-clinical samples: university students, athletes, and bodybuilders; clinical samples: MD-diagnosed participants, AN-diagnosed participants, and people who use anabolic androgenic steroid (AAS)], the gender of the participants (e.g., only males, only females, or mixed samples), the geographical location (USA, Australia, Mexico, Italy, Spain, etc.), and the ethnicity of the participants (Sandgren & Lavalle, 2018; Suffolk, Dovey, Goodwin, & Meyer, 2013).

To the best of our knowledge, no one has yet systematically studied the association between MD and ED symptomatology using a meta-analysis. A prior meta-analysis synthesized the evidence about the relationship between the drive for muscularity (DFM) and ED symptomatology, obtaining an average correlation of $r_+ = .27$ and indicating that higher levels of DFM were associated with higher levels of ED psychopathology (Tod & Edwards, 2015b). Based on Cohen (1988), correlation coefficients of about .1, .3, and .5 can be interpreted as reflecting small, moderate, and large associations, respectively. Therefore, Todd and Edwards's (2015b) findings showed a small to moderate association between DFM and ED symptomatology. Nevertheless, DFM and MD are different but related constructs. The former describes individual motivation to become more muscular, whereas the latter is a psychological disorder that represents the pathological pursuit of muscularity and leanness (Cafri et al., 2005). Hence, the desire to increase muscularity (i.e., DFM) does not imply distress (i.e., MD; Kimmel & Mahalik, 2004; Morrison, Morrison, & McCann, 2006). Therefore, people can present a high DFM without necessarily being distressed by perceived inadequacy or developing MD. Thus, the research has shown that individuals with MD are different from normal weightlifters in their symptom pathology and psychiatric comorbidity (Cafri et al., 2008; Olivardia et al., 2000). Recent evidence suggests that the former has a greater presence of eating disordered attitudes and beliefs than the latter (Mitchell, Murray, Hoon, et al., 2017; Murray, Nagata, et al., 2017). In addition, research also suggests that these constructs may have different neuropsychological correlates (Griffiths, Murray, & Touyz, 2013).

Therefore, we primarily aimed to conduct a systematic review and meta-analysis to synthesize the scientific evidence about the relationships between MD and ED symptomatology, given the inconsistencies in the results from empirical studies that have previously investigated this association. A second aim was to identify the study characteristics that might moderate the heterogeneous results that emerged, such as the geographic location, type of sample, measurement instruments for MD and ED, age, gender, and ethnicity, among others.

METHODS

We conducted a systematic review and meta-analysis on the relationship between MD and ED symptomatology following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). Because PRISMA guidelines were developed to be applied to meta-analyses studying intervention efficacy, some of their items were not applicable to our meta-analysis. Consequently, we also adhered to the guidelines recently proposed by the American Psychological Association Publications and Communications Board Task Force (Appelbaum et al., 2018, Table 9, p. 2123).

Study selection criteria

The studies had to meet the following inclusion criteria: (a) the study had to be published in a peer-reviewed journal, (b) it had to be an original and quantitative investigation, (c) it had to assess MD using measurement instruments that specifically evaluated the symptoms and diagnostic criteria of MD (e.g., MASS, MDDI, MDI, etc.), (d) it had to assess ED using measurement instruments that specifically evaluated the symptoms and/or diagnostic criteria of ED (e.g., EAT, EDI, and EDE), and (e) it had to empirically examine the relationship between MD and ED symptoms. We established no date, language, or participant's age limits. We excluded qualitative studies, literature reviews, systematic reviews, meta-analyses, commentaries, editorials, and studies that did not assess MD and its relationship with ED symptomatology.

Search strategy

The researchers carried out electronic searches in June 2018 in the Medline (via PubMed), PsycInfo, Science Direct, and Web of Science databases, using the following terms in ALL FIELDS: muscle dysmorph*, MDI, reverse anorexia, bigorexia, vigorexia, and Adonis complex. Furthermore, we conducted manual searches of lists of references from the retrieved studies to identify additional studies that met the selection criteria. This search yielded two additional eligible studies. In addition, the researchers screened lists of references from previous reviews and meta-analyses (e.g., Dos Santos, Tirico, Stefano, Touyz, & Claudino, 2015; Mitchell, Murray, Cobley, et al., 2017; Sandegren & Lavalle, 2018; Suffolk et al., 2013; Tod & Edwards, 2015b) to find studies that met the inclusion criteria for the present

meta-analysis. Finally, in order to locate unpublished papers, the researchers sent e-mails to 12 of the most prolific authors in the field. Of the six responses we received, none led to the discovery of an unpublished study.

Two researchers carried out the eligibility process for this meta-analysis independently, resolving disagreements between raters by consensus. Figure 1 presents a flowchart of the study screening and selection process.

Coding of the studies

The researchers produced a protocol for extracting the characteristics of the studies and applied it to each study. The following characteristics were coded: year of the study, geographical location, sampling method, type of sample (e.g., clinical, community, gymnasiums, university, and AAS users), sample size, gender, mean age, sexual orientation, ethnicity, education level, tools used for assessing MD and ED, and statistics reported to calculate the effect sizes. In addition, the researchers assessed the methodological quality of the studies by applying an ad hoc 10-item checklist, given that in the literature there is no clear methodological checklist to use in studies with a cross-sectional design where the purpose is to analyze the relationships between two variables, as in this study. Most quality checklists are designed for randomized trials and other types of research, particularly evaluations of interventions (e.g., Cochrane checklist). Therefore, their criteria focus on interventions and follow-up assessments. Other checklists are focused on cohort studies (e.g., Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies from National Institutes of Health); however, most of their items are also related to interventions and follow-up assessments. Therefore, we developed our own criteria to assess the methodological quality of primary studies.

Appendix A shows each of the 10 items on the quality checklist. Each item scored 1 when the study met the criteria and 0 otherwise. We calculated a total quality score (TQS) for each study by adding up all the corresponding quality item scores (range: 0–10), with a higher score indicating higher overall quality.

The researchers carried out the coding process in a standardized and systematic manner and two reviewers independently extracted the data. Interrater reliability was satisfactory, with a mean intraclass correlation of .98 ($SD = 0.036$), ranging from 0.90 to 1 for continuous variables, and with a mean κ coefficient of 0.92 ($SD = 0.078$), ranging from 0.81 to 1 for qualitative variables. Inconsistencies between coders were resolved by consensus.

Computing effect sizes

The effect size index was a correlation coefficient (e.g., the Pearson's correlation coefficient, its Fisher's Z transformation, the φ coefficient, the point-biserial correlation coefficient) calculated between an MD scale and an ED scale (Borenstein, Hedges, Higgins, & Rothstein, 2009). For each study, the researchers translated the Pearson's correlations to Fisher's Z in order to normalize their distributions and stabilize their variances. Then, the researchers back-translated the Fisher's Z values for the individual effect sizes, as well as
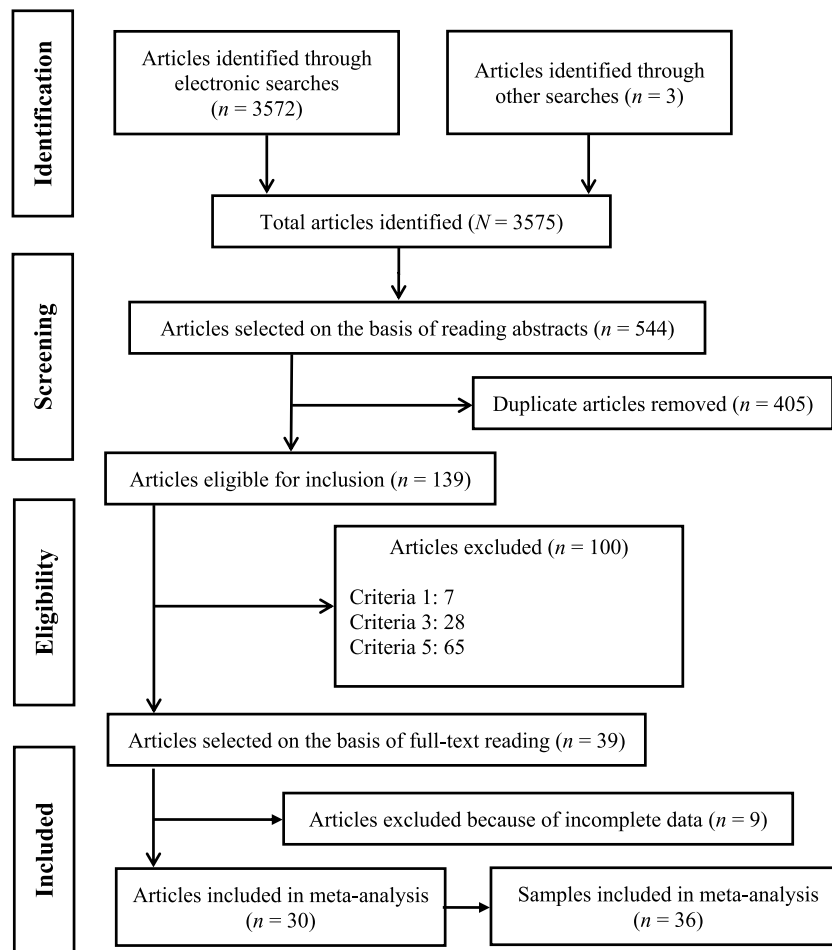
Figure 1. Flowchart of the study selection process in the systematic review and meta-analysis of MD and ED symptomatology

those for the mean effect sizes and their confidence limits, into the Pearson's correlation metric in order to make their interpretation easier (Borenstein et al., 2009).

In studies that reported several correlations for the MD–ED relationship, because researchers used multiple measures of MD and/or ED, we calculated their average to avoid statistical dependence.

In studies that did not directly report correlation coefficients, we applied appropriate translations between effect sizes. Thus, several studied reported odds ratios (ORs; Bo et al., 2014; Cafri et al., 2008; Olivardia et al., 2000; Pope et al., 2005), and we applied a formula to transform them into correlation coefficients. For each study, the ORs were translated into d, with $d = \frac{\mathrm{Log}_e(\mathrm{OR})}{1.65}$, Log_e being the natural logarithm. Then, the ds were translated into r, with $r = \frac{d}{\sqrt{d^2 + a}}$, where a was estimated with $a = \frac{(n_1 + n_2)^2}{n_1 \cdot n_2}$, and n_1, n_2 were the sample sizes in the two conditions (Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003).

Statistical analyses

The researchers carried out a meta-analysis to assess the relationship between MD and ED symptomatology. In order to accommodate the variability shown by the effect sizes, we assumed a random-effects model (Borenstein et al., 2009;

Sánchez-Meca & Marín-Martínez, 2008). We calculated a pooled correlation coefficient and its corresponding 95% confidence interval (CI). In addition, we assessed the statistical significance of the pooled correlation using the Z test.

The researchers constructed a forest plot to represent the individual and pooled effect size estimates, with their 95% CIs, and to allow visual inspection of effect size heterogeneity. We calculated both Cochran's Q statistic and the I^2 index to assess the consistency of the effect sizes (Borenstein et al., 2009; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). A Q statistic with $p < .05$ was indicative of heterogeneity among the effect sizes. We estimated the degree of this heterogeneity using the I^2 index. I^2 values of around 25%, 50%, and 75% denoted low, moderate, and large heterogeneity, respectively.

To examine the influence of moderator variables on effect size variability, the researchers calculated analyses of variance (ANOVAs) and meta-regressions for categorical and continuous variables, respectively, by assuming a mixed-effects model (Borenstein et al., 2009; López-López, Marín-Martínez, Sánchez-Meca, van den Noortgate, & Viechtbauer, 2014). We assessed the statistical significance of qualitative and continuous moderators with the Q_B and Q_R statistics, respectively. In addition, we calculated an estimate of the proportion of variance accounted for by the moderator variable following Raudenbush's (2009) proposal (López-López et al., 2014). We assessed the model

misspecification with the Q_W and Q_E statistics for qualitative and continuous moderators, respectively.

Finally, to assess publication bias, the researchers used both a funnel plot with Duval and Tweedie's trim-and-fill method for imputing missing data and the Egger's test (Duval & Tweedie, 2000; Sterne & Egger, 2005). A funnel plot with Duval and Tweedie's trim-and-fill method uses available data to impute missing (unreported) studies and recalculates the overall effect that would be observed with their inclusion (Duval & Tweedie, 2000). The Egger's test is an unweighted simple regression, which takes the precision of each study as the independent variable (precision being defined as the inverse of the standard error of each effect size) and the effect size divided by its standard error as the dependent variable. A non-statistically significant result of the *t*-test for the hypothesis of an intercept equal to zero means that publication bias can be eliminated as a threat to the validity of the pooled effect (Sterne & Egger, 2005).

The researchers interpreted all statistical tests assuming a significance level of 5% ($p < .05$). The statistical analyses were performed with the Comprehensive Meta-analysis software program, version 3.0 (Biostat, Englewood, NJ, USA; Borenstein, Hedges, Higgins, & Rothstein, 2014).

RESULTS

Study selection

The search strategy produced a total of 3,575 manuscripts. After scanning the titles and abstracts of the 3,575 identified manuscripts, the researchers preselected 544 relevant studies based on the inclusion and exclusion criteria. Then, eliminating duplicates ($n = 405$) left a total of 139 studies to review. After a review of the full text of the remaining articles, 39 fulfilled the selection criteria. Nevertheless, nine articles did not report a correlation between MD and ED or statistical data to calculate it. In these cases, the researchers sent e-mails to the authors of the studies to obtain these data, but none of them replied to our request. Therefore, we excluded these studies from the meta-analysis (Babusa &Túry, 2012; Babusa, Urban, Czeglédi, & Túry, 2012; Guidi, Clementi, & Grandi, 2013; Hale et al., 2013; Hildebrandt, Langenbucher, Karmin, Loeg, & Hollander, 2011; Hildebrandt et al., 2006; Hildebrandt, Walker, Alfano, Delinsky, & Bannon, 2010; Kanayama, Barry, Hudson, & Pope, 2006; Magallares, 2016). Thus, the meta-analysis included 30 published articles; all of these studies were published in a peer-reviewed journal between 2000 (Olivardia et al., 2000) and 2018 (Klimek, Murray, Brown, Gonzales, & Blashill, 2018). Three of the 30 articles provided data on four, two, and three independent samples, respectively (Giardino & Procidano, 2012; Lamanna, Grieve, Derryberry, Hakman, & McClure, 2010; Segura García et al., 2010). Thus, the data set for our meta-analysis was composed of a total of 36 independent samples.

Study characteristics

Appendix B presents descriptive characteristics of the articles included ($n = 39$) in the systematic review, yielding a total of 45 studies or independent samples. Overall, the majority of the studies used a convenience sample (e.g., gymnasiums, vitamin stores, bodybuilding discussion forums, Facebook, and university) and a cross-sectional design. Only one study used a probabilistic sample (Castro-Lopez, Cachón-Zagalaz, Molero, & Zagalaz-Sánchez, 2013). A large number of studies were conducted in the USA ($n = 23$), with 14 carried out in Europe, 4 in Australia, and 2 in Mexico; one was conducted in China, and a multinational was study carried out in the UK, the USA, Australia, and Singapore.

The 45 independent samples included 8,516 participants (mean = 189, range = 11, and 648 participants). The weighted mean age of all participants was 25.89 years. Of the 45 independent samples, 31 were composed exclusively of men and 7 of women, whereas both males and females participated in the remaining 7 samples. Overall, the majority of the participants was Caucasian, with the presence of varying percentages of other racial and ethnic groups. Finally, the majority of the participants had received at least some college education.

The studies used a variety of instruments to measure MD and ED. The tools most commonly used to assess MD were the MDDI (Hildebrandt et al., 2004; $n = 16$), the MDI (Rhea et al., 2004; $n = 7$), and the MASS (Mayville et al., 2002; $n = 11$). The assessment instruments most frequently used in the ED studies were the EAT-26 (Garner et al., 1982; $n = 15$), the EDI (Garner et al., 1983; $n = 12$), and the EDE-Q (Fairburn & Beglin, 1994; $n = 10$).

Assessment of methodological quality

Appendix C presents the results obtained from the assessment of the methodological quality of the studies. Specifically, the appendix shows the scores of individual studies on each quality item. The researchers also obtained a TQS by adding up the 1s and 0s for the checklist items (range = 0–10). On average, the studies had a methodological quality of 5.28 ($SD = 1.22$), with a range of 3–7 points.

Most of the studies met the following criteria: (a) they used participant's samples that allowed an appropriate representation of a clinical or at-risk MD or ED population ($n = 31$); (b) they employed measurement instruments with good psychometric properties (validity and reliability) to assess MD in the study sample ($n = 30$); (c) they employed measurement instruments with good psychometric properties (validity and reliability) to assess ED in the study sample ($n = 28$); (d) they did not dichotomize the assessment of MD and ED ($n = 39$); (e) they reported on results related to all the MD and ED instruments described in "Methods" section ($n = 44$); (f) they applied appropriate statistical tests (non-parametric vs. parametric methods) to assess the relationships between MD and ED ($n = 24$); and (g) they had no private financial support ($n = 39$). Nevertheless, most of the studies did not provide data on power analysis or report CIs around the effect sizes measured.

Finally, the following criteria were least likely to be met by the studies: (a) use of representative sampling procedures ($n = 1$) and (b) a priori determination of the sample size to identify an effect ($n = 2$). In addition, none of the studies specified whether there were dropouts or if these dropouts

had similar sociodemographic characteristics to those of the final sample.

Synthesis of results

This meta-analysis estimated the relationship between MD and ED symptomatology. Figure 2 presents a forest plot of the MD–ED associations found in each individual study and their 95% CIs, as well as the mean correlation coefficient resulting from pooling all the studies and its corresponding 95% CI.

As Figure 2 shows, the pooled effect size for the relationship between MD and ED was $r_+ = .356$ (95% CI = 0.298, 0.411, $\kappa = 36$). In addition, the analysis found considerable heterogeneity among individual effect sizes, [$Q(35) = 200.925$, $p < .0001$, $I^2 = 82.7\%$].

Analysis of publication bias

Because all the studies included were published papers, we carried out analyses to determine whether publication bias might be a threat to the validity of the results of the meta-analysis. The Duval and Tweedie (2000) trim-and-fill method did not impute any effect size (see Appendix D). In addition, the Egger's test applied to the intercept of a simple regression model of the effect sizes did not reach statistical significance [intercept = 1.09, $t(34) = 1.176$, $p = .248$]. Therefore, we can reasonably rule out publication bias as a serious threat to our meta-analytic findings.

Moderator analyses of the relationships between MD and ED symptomatology

The effect sizes showed a large amount of heterogeneity ($I^2 = 82.7\%$). Consequently, the researchers conducted statistical analyses to identify the study characteristics that might explain this heterogeneity. Specifically, we used weighted ANOVAs and simple meta-regressions for categorical and continuous moderator variables, respectively, taking the correlation coefficients between MD and ED symptomatology as the dependent variable.

Table 1 presents the results of the ANOVAs conducted on the categorical variables, such as country of the study, type of sample, MD and ED measurement instrument, and type of statistics reported. Three moderator variables showed a statistically significant association with effect sizes: (a) the type of sample used in the study, (b) the measurement instrument used to assess MD, and (c) the measurement instrument used to assess ED. The type of sample used in the study, classified as people who reported current AAS use versus other samples, accounted for 26% of the variance ($p < .001$). Specifically, a study found stronger MD–ED associations when it was conducted among people who used AAS ($r_+ = .710$), compared to other samples ($r_+ = .342$). In addition, the tools used to assess ED explained 26% of the variance ($p < .001$), with stronger MD–ED associations found when the EDE-Q Modified was the instrument used to assess ED ($r_+ = .642$), in comparison with other instruments ($r_+ = .336$). Finally, studies that used
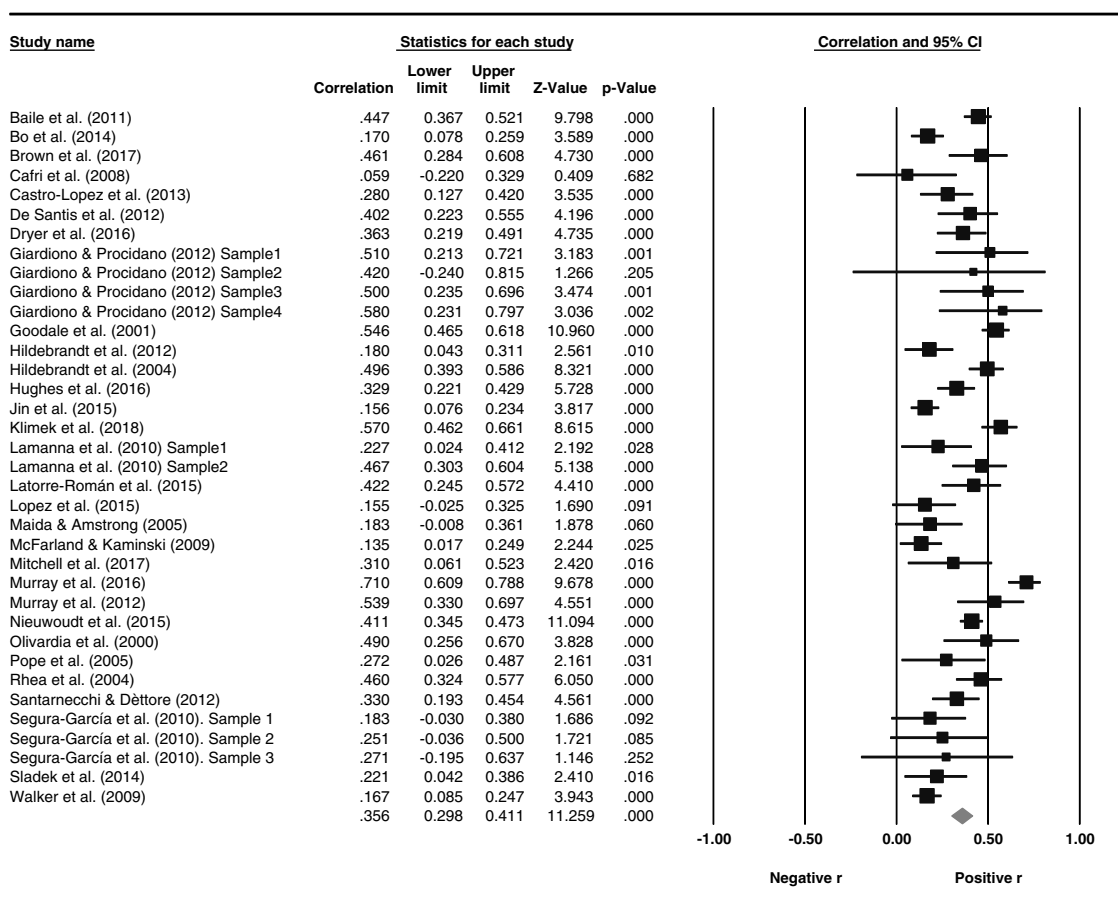

Study name	Statistics for each study					Correlation and 95% CI
	Correlation	Lower limit	Upper limit	Z-Value	p-Value	
Baile et al. (2011)	.447	0.367	0.521	9.798	.000	
Bo et al. (2014)	.170	0.078	0.259	3.589	.000	
Brown et al. (2017)	.461	0.284	0.608	4.730	.000	
Cafri et al. (2008)	.059	-0.220	0.329	0.409	.682	
Castro-Lopez et al. (2013)	.280	0.127	0.420	3.535	.000	
De Santis et al. (2012)	.402	0.223	0.555	4.196	.000	
Dryer et al. (2016)	.363	0.219	0.491	4.735	.000	
Giardiono & Procidano (2012) Sample1	.510	0.213	0.721	3.183	.001	
Giardiono & Procidano (2012) Sample2	.420	-0.240	0.815	1.266	.205	
Giardiono & Procidano (2012) Sample3	.500	0.235	0.696	3.474	.001	
Giardiono & Procidano (2012) Sample4	.580	0.231	0.797	3.036	.002	
Goodale et al. (2001)	.546	0.465	0.618	10.960	.000	
Hildebrandt et al. (2012)	.180	0.043	0.311	2.561	.010	
Hildebrandt et al. (2004)	.496	0.393	0.586	8.321	.000	
Hughes et al. (2016)	.329	0.221	0.429	5.728	.000	
Jin et al. (2015)	.156	0.076	0.234	3.817	.000	
Klimek et al. (2018)	.570	0.462	0.661	8.615	.000	
Lamanna et al. (2010) Sample1	.227	0.024	0.412	2.192	.028	
Lamanna et al. (2010) Sample2	.467	0.303	0.604	5.138	.000	
Latorre-Román et al. (2015)	.422	0.245	0.572	4.410	.000	
Lopez et al. (2015)	.155	-0.025	0.325	1.690	.091	
Maida & Amstrong (2005)	.183	-0.008	0.361	1.878	.060	
McFarland & Kaminski (2009)	.135	0.017	0.249	2.244	.025	
Mitchell et al. (2017)	.310	0.061	0.523	2.420	.016	
Murray et al. (2016)	.710	0.609	0.788	9.678	.000	
Murray et al. (2012)	.539	0.330	0.697	4.551	.000	
Nieuwoudt et al. (2015)	.411	0.345	0.473	11.094	.000	
Olivardia et al. (2000)	.490	0.256	0.670	3.828	.000	
Pope et al. (2005)	.272	0.026	0.487	2.161	.031	
Rhea et al. (2004)	.460	0.324	0.577	6.050	.000	
Santarnecchi & Dèttore (2012)	.330	0.193	0.454	4.561	.000	
Segura-García et al. (2010). Sample 1	.183	-0.030	0.380	1.686	.092	
Segura-García et al. (2010). Sample 2	.251	-0.036	0.500	1.721	.085	
Segura-García et al. (2010). Sample 3	.271	-0.195	0.637	1.146	.252	
Sladek et al. (2014)	.221	0.042	0.386	2.410	.016	
Walker et al. (2009)	.167	0.085	0.247	3.943	.000	
	.356	0.298	0.411	11.259	.000	

Figure 2. Forest plot of the association between MD and ED symptomatology

Table 1. Results of the weighted ANOVAs for the influence of categorical variables on the effect sizes

Moderator variable	k	r_+	95% CI		ANOVA results
			rl	ru	
Geographic location					$Q_B(1) = 0.051, p = .821$
USA	18	.363	.274	.445	$R^2 = .00$
Other countries	18	.349	.270	.424	$Q_W(34) = 195.879, p < .001$
Sample setting					$Q_B(5) = 31.793, p < .001$
Gymnasiums	18	.359	.284	.430	$R^2 = .18$
University	9	.301	.169	.423	$Q_W(30) = 153.370, p < .001$
Area metropolitan	3	.352	.276	.425	
AAS users	1	.710	.609	.788	
Clinical (BDD or EDS)	2	.272	.055	.464	
University and other	3	.387	.240	.516	
Sample setting					$Q_B(1) = 30.137, p < .001$
AAS users	1	.710	.609	.788	$R^2 = .26$
Other	35	.342	.287	.394	$Q_W(34) = 167.137, p < .001$
MD measurement					
MDDI	12	.396	.278	.503	$Q_B(7) = 17.087, p = .017$
MASS	7	.367	.220	.497	$R^2 = .00$
SCID	3	.285	.031	.504	$Q_W(28) = 171.339, p < .001$
ACQ	4	.398	.320	.471	
MDS	1	.135	.017	.249	
MDI	6	.293	.185	.394	
MDQ	1	.363	.219	.491	
MDSQ	2	.386	−.012	.678	
Type of MD measurement					$Q_B(1) = 0.385, p = .535$
Diagnostic	3	.285	.031	.504	$R^2 = .00$
Symptomatology	33	+.361	.301	.419	$Q_W(34) = 202.228, p < .001$
ED measurement[a]					
EAT	14	.345	.263	.423	$Q_B(6) = 30.585, p < .001$
EDI	10	.348	.268	.423	$R^2 = .23$
EDE-Q	7	.348	.187	.491	$Q_W(30) = 141.207, p < .001$
EDE-Q Modified	2	.642	.449	.777	
SCID	2	.177	−.036	.374	
MEBBIE	1	.135	.017	.249	
CHAA	1	.447	.367	.521	
EDQ Modified vs. others[a]					$Q_B(1) = 8.087, p = .007$
EDE-Q Modified	2	.642	.449	.777	$R^2 = .26$
Others	35	.336	.281	.389	$Q_W(35) = 167.335, p < .001$
Type of ED measurement[a]					$Q_B(1) = 3.129, p = .077$
Diagnostic	2	.177	−.036	.374	$R^2 = .01$
Symptomatology	35	.364	.305	.419	$Q_W(35) = 200.840, p < .001$
Statistics reported					$Q_B(1) = 2.274, p = .132$
Correlation coefficient	32	.368	.308	.426	$R^2 = .03$
Odds ratio	4	.242	.078	.393	$Q_W(34) = 189.439, p < .001$

Note. k: number of studies. r_+: mean effect size. rl and ru: 95% lower and upper confidence limits around r_+; Q_B: between-categories Q statistic; Q_W: within-categories Q statistic; R^2: proportion of variance accounted for by the moderator variable; ED: eating disorder; EDE-Q: Eating Disorder Examination Questionnaire; MEBBIE: Male Eating Behavior and Body Image Evaluation; CHAA: Cuestionario de Hábitos de Alimentación Alterados; SCID: structured clinical Interview for DSM-IV; MDDI: Muscle Dysmorphic Disorder Inventory; MASS: Muscle Appearance Satisfaction Scale; MDI: Muscle Dysmorphia Inventory; AAS: anabolic androgenic steroid; ACQ: Adonis Complex Questionnaire; MDSQ: Muscle Dysmorphia Symptom Questionnaire; MDQ: Muscle Dysmorphia Questionnaire; CI: confidence interval; ANOVA: analysis of variance; BDD: body dysmorphic disorder.
[a]Hughes et al.'s (2016) study used two tools for assessing EDs (EAT-26 and EDEQ).

quantitative measurement instruments to assess ED symptomatology, such as EDE-Q or EAT, that had a larger mean effect size ($r_+ = .364$) than the studies that used other instruments ($r_+ = .177$), although this difference was only marginally statistically significant ($p = .077, R^2 = .01$).

Table 2 presents the simple meta-regressions conducted on continuous variables, such as mean age, gender (% male), and ethnicity (% Caucasian) of the sample, as well as the methodological TQS. The only moderator variable that exhibited a statistically significant association with effect sizes was the methodological TQS for the individual studies, accounting for 25% of the variance ($p < .05$).

In addition, in order to identify the individual items on the quality checklist that were statistically associated with

Table 2. Results of the simple weighted meta-regressions of continuous moderators on the effect sizes

Moderator variable	k	b_j	Q_R	Q_E	R^2
Mean age (years)	36	0.0015	0.05	196.51**	.0
Gender (% males)	36	0.0000	0.45	199.26**	.0
Ethnicity (% Caucasian)	19	−0.0003	0.03	101.58**	.0
Methodology quality checklist (0–10)	36	0.0776	10.03*	176.31**	.25

Note. k: number of studies; b_j: unstandardized regression coefficient; Q_R: Q statistic for testing the statistical significance of the moderator variable; Q_E: Q statistic for testing the model misspecification; R^2: proportion of variance accounted for by the moderator variable. *$p < .05$. **$p < .001$.

the effect sizes, we applied ANOVAs for each individual item (see Table 3). As Table 3 shows, items 4, 5, and 6 were statistically associated with the effect sizes. Thus, studies that reported any estimate of the reliability/validity of the MD measurement instrument (item 4) or any estimate of the reliability/validity of the ED measurement used in the study sample (item 5) showed a mean effect size ($r_+ = .403$ and $r_+ = .416$, respectively) larger than those that did not ($r_+ = .255$ and $r_+ = .196$, respectively), accounting for a large proportion of variance ($R^2 = .13$ and $R^2 = .31$, respectively). In addition, studies that did not dichotomize MD and/or ED variables showed a mean effect size ($r_+ = .376$) larger than those that did so ($r_+ = .219$), accounting for a moderate proportion of variance ($R^2 = .06$).

DISCUSSION

This is the first meta-analytic study on the relationship between MD and ED symptomatology. Our results showed that there is a positive and statistically significant relationship between MD and ED symptomatology, indicating that higher levels of MD were associated with higher levels of ED symptomatology. Following Cohen's (1988) criteria, a correlation coefficient of $r_+ = .356$ can be interpreted as reflecting a relationship of moderate magnitude. Our findings are consistent with the findings from Tod and Edwards's (2015b) review, which found an association between DFM and ED symptomatology. However, the strength of the correlation was greater between MD and

Table 3. Results of the weighted ANOVAs for the influence of methodological quality items on the effect sizes

Quality item	k	r_+	95% CI		ANOVA results
			rl	ru	
1. Probabilistic sampling					$Q_B(1) = 0.988, p = .320$
Yes	1	.280	.127	.420	$R^2 = .00$
No	35	.359	.299	.415	$Q_W(34) = 202.124, p < .001$
2. Adequate target population					$Q_B(1) = 1.118, p = .290$
Yes	24	.380	.310	.446	$R^2 = .01$
No	12	.316	.215	.411	$Q_W(34) = 191.814, p < .001$
4. Reliability/validity MD measurement tool					$Q_B(1) = 6.928, p = .008$
Yes	24	.403	.331	.470	$R^2 = .13$
No	12	.255	.167	.339	$Q_W(34) = 192.883, p < .001$
5. Reliability/validity ED measurement tool					$Q_B(1) = 26.172, p < .001$
Yes	24	.416	.349	.479	$R^2 = .31$
No	12	.196	.147	.244	$Q_W(34) = 159.323, p < .001$
6. Absence of dichotomous DM and ED					$Q_B(1) = 6.412, p = .011$
Yes	30	.376	.314	.434	$R^2 = .06$
No	6	.219	.109	.324	$Q_W(34) = 186.515, p < .001$
8. Statistical test appropriate					$Q_B(1) = 0.058, p = .810$
Yes	21	.360	.280	.435	$R^2 = .00$
No	15	.346	.265	.423	$Q_W(34) = 202.325, p < .001$
9. Private financial support					$Q_B(1) = 0.138, p = .711$
Yes	4	.373	.288	.452	$R^2 = .00$
No	32	.354	.290	.414	$Q_W(34) = 201.692, p < .001$
10. Statistical power					$Q_B(1) = 0.426, p = .514$
Yes	2	.251	−.101	.548	$R^2 = .00$
No	34	.361	.302	.418	$Q_W(34) = 202.397, p < .001$

Note. Results for item 3 (dropouts were similar in socio-demographic characteristics to those of the final sample) and item 7 (absence of reporting bias) were not analyzed because all studies scored 0 and 1, respectively. k: number of studies; r_+: mean effect size; rl and ru: 95% lower and upper confidence limits around r_+; Q_B: between-categories Q statistic; Q_W: within-categories Q statistic; R^2: proportion of variance accounted for by the moderator variable; ANOVA: analysis of variance; ED: eating disorder; MD: muscle dysmorphia.

ED symptomatology ($r = .36$) than between DFM and ED ($r = .27$). These findings suggest that both MD and DFM share an association with EDs. People with DFM and MD want to gain lean muscle mass, which is achieved not only through physical training but also by maintaining a diet high in lean protein and using food supplements. Therefore, dietary practices are an important factor in both DFM and MD. Furthermore, research suggests that the behavior and diet subscales of the DFM predict MD characteristics (Robert et al., 2009), which might explain that both MD and DFM share an association with EDs. In other research, DFM has been identified as a precursor to the development of MD (Olivardia et al., 2000). Therefore, those who scored high on a measure of MD displayed a higher degree of DFM. From this point of view, MD might be considered an extreme variant of DFM.

We ruled out publication bias as a threat to our meta-analytic results, but we found great variability among the individual effect sizes. In order to identify the study features that could explain at least part of the effect size variability, we carried out ANOVAs and meta-regressions. Several study characteristics were statistically related to effect size, such as the type of sample, the methodological quality, and the MD and ED measurement instruments used. First, our findings showed stronger associations between MD and ED symptomatology in males reporting current AAS use than in other people (undergraduates, athletes, etc.). However, this result must be interpreted very cautiously because AAS samples were represented in only one study. Therefore, it seems plausible that males who use AAS exhibit elevated levels of MD symptomatology. Indeed, some of the most common motivations for using AAS are to increase muscle mass and strength (Cohen, Collins, Darkes, & Gwartney, 2007), to improve physical appearance (Kimergård, 2015), and to decrease body dissatisfaction and MD symptoms (Grogan, Shepherd, Evans, Wright, & Hunter, 2006). In this way, research has identified MD symptoms as a predictor of the use of substances such as anabolic steroids, prohormones, and ephedrine (Cafri, van den Berg, & Thompson, 2006). Thus, people who use AAS might be at greater risk of developing MD and ED symptomatology. These findings suggest the need to implement health promotion and prevention programs to reduce AAS use.

Furthermore, our results showed that of all the ED measurement instruments, the EDE-Q Modified (Murray et al., 2012) was the most sensitive tool in identifying relationships with MD symptomatology. Some conceptual overlapping between the EDE-Q Modified and the measures of MD is likely. It is important to note that the EDE-Q Modified is a version of the EDE-Q with the gender-relevant items identified and reversed in polarity to enhance its sensitivity in indexing male concerns (Murray et al., 2012). For instance, the item on the EDE-Q "Have you had a definite fear that you might gain weight or become fat?" was reversed on the EDE-Q Modified to "Have you had a definite fear that you might lose weight or become less muscular?" Other examples of reversed items are "Have you been deliberately trying to *increase* the amount of food you eat to influence your shape or weight?" and "Has thinking about food *or its protein content* made it more difficult to concentrate on things you are interested in; for example,

reading, watching TV, or following a conversation?" (modifications italicized). Such overlaps might influence the strength of the relationships between MD and ED symptomatology. In addition, it is plausible that the strength of the association between MD and the EDE-Q Modified might be affected by the type of sample used to examine this relationship. In this regard, the only two studies that applied the EDE-Q Modified (Murray, Griffiths, Mond, Kean, & Blashill, 2016; Murray et al., 2012) used samples of participants who were not from a community population. In one of them; the participants were AAS users (Murray et al., 2016), and in the other study, they were clinical participants (Murray et al., 2012). Therefore, clinicians should be aware of this fact and take it into account when identifying co-occurrence.

Moreover, our findings revealed that of all the ED assessment instruments, the SCID was the least sensitive in identifying relationships between MD and ED symptomatology, which might explain why the study by Cafri et al. (2008) did not reveal a relationship between MD and ED symptomatology. One possible explanation is that the SCID provides a more thorough and conservative assessment of an ED, having less overlap with MD than other ED measures. This might suggest that MD is associated with certain ED symptoms but not strongly or consistently with a separate ED diagnosis.

Finally, our findings suggest that studies with higher methodological quality are more sensitive in identifying relationships between MD and ED symptomatology. In addition, we found stronger MD–ED associations in the studies that assessed the psychometric properties of the MD and ED instruments with the data found than in those that did not assess these properties. These results agree with psychometric theory, which proposes that using measurement instruments with good psychometric properties will provide greater sensitivity in revealing statistical associations between variables than using non-reliable and/or non-valid instruments. Finally, we found lower MD–ED associations in the studies that dichotomized the MD and ED measures (i.e., reporting ORs instead of Pearson's correlation coefficients). These results agree with statistical theory, which proposes that dichotomizing variables produce a loss of sensitivity in detecting statistical associations among variables (Schmidt & Hunter, 2015). Therefore, one would expect higher correlations in studies with higher methodological quality that used quantitative measurement instruments for MD assessment, such as MASS and MDDI, and for ED assessment, such as EAT or EDE-Q.

Limitations

This meta-analysis has several limitations. Five studies could not be included in the final analysis because the statistical data reported did not allow us to calculate the effect sizes, and their authors did not reply to the request for data. Furthermore, most of the studies included were conducted in the USA, which might limit the generalization of the results. In addition, given the existence of great heterogeneity among the effect sizes, it is possible that other moderator variables not considered in our meta-analysis might be relevant in explaining this heterogeneity

(Menees, Grieve, Mienaltowski, & Pope, 2013). These moderators are likely to be psychosocial constructs such as socially prescribed perfectionism (Dryer, Farr, Hiramatsu, & Quinton, 2016) or masculinity-related constructs (Tod et al., 2016). Increased understanding and awareness of MD and ED symptomatology might contribute to identifying and referring people at risk, so that they can receive help from mental health services. Finally, it is important to remember that all the studies included in this meta-analysis were cross-sectional, which means we cannot draw causal inferences about the relationship between MD and ED symptomatology.

Recommendations for future research

Most research on MD–ED has used non-clinical samples, often university students, who, in general, have low psychopathology. Thus, future research should investigate the MD–ED association in clinical samples, that is, in people diagnosed with MD and/or ED. These studies might improve our knowledge to develop more optimal treatment approaches. On other hand, to find out whether the EDE-Q Modified (Murray et al., 2012) is really the most sensitive tool for assessing the association between MD and ED symptomatology, additional studies are needed to analyze this relationship in community samples.

Moreover, given that most of the studies used cross-sectional designs, which do not allow us to draw causal inferences about the relationship between MD and ED symptomatology, future research should implement longitudinal studies. These studies might help to clarify the relationship between MD and ED symptomatology in chronological terms, for instance, whether ED precedes MD or vice versa, and they might clarify the stability of MD and/or ED symptomatology.

Finally, given that most of the studies were conducted in the USA and with male participants, future research should study this topic in other social contexts and in the female population in order to investigate potential differences based on cultural factors and/or sex differences.

CONCLUSIONS

This is the first meta-analytic study on the relationship between MD and ED symptomatology, contributing to a more accurate view of this phenomenon, despite the limitations described above. Furthermore, our findings may have some implications for clinical practice. They suggest that the symptoms and behaviors characteristic of MD tend to co-occur with ED symptomatology in males and females. Thus, when evaluating individuals presenting with MD, clinicians should investigate the possibility of the presence of ED, and vice versa, particularly in males. In this regard, a recent case study showed the transition from thinness-oriented to muscularity-oriented disordered eating during the course of treatment for AN (Murray, Griffiths, Mitchison, & Mond, 2017). As the authors of this study pointed out, the emergence of muscularity-oriented disordered eating during the treatment of AN might be misinterpreted as healthy, due to the move away from

dangerously low body weight, which is the objective of AN treatment.

People with MD symptomatology may also have ED symptomatology, which appears to be associated with greater psychopathology (Chandler, Grieve, Derryberry, & Pegg, 2009; Ebbeck, Watkins, Concepcion, Cardinal, & Hammermeister, 2009; Grieve & Shacklette, 2012; Pope et al., 2005). This situation, in addition to the serious limitations in social and work performance caused by MD and the associated unhealthy behaviors, makes it necessary to use strategies for identifying at-risk individuals and for prevention, such as inviting them to participate in psychotherapy. Because MD and ED symptomatology, and the co-occurrence of these symptoms, can occur more often among athletes, the development of such intervention strategies may be of special interest for mental health professionals working in the sports area. We do not believe that participation in sports is inherently pathological, but it is likely that some individuals take up sports for reasons (which may be pathological) that correlate with MD and ED symptomatology or increase the risk of developing symptoms of these two disorders. A psychological assessment conducted by an experienced sports psychologist could identify the athletes in need of psychological treatment. Based on our data, we recommend that clinicians use the MDDI or MASS to assess MD and the EDE-Q Modified to assess ED when examining the co-occurrence of MD and ED symptomatology. These are the most sensitive measurement instruments in identifying relationships between MD and ED (considering the previous discussion about the EDEQ Modified).

Conflict of interest: The authors declare no potential conflict of interest with regard to the research, authorship, and/or publication of this article.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

American Psychiatric Association [APA]. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington, DC: American Psychiatric Association.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board Task Force report. *American Psychologist, 73*(1), 3–25. doi:10.1037/amp0000191

*Babusa, B., & Túry, F. (2012). Muscle dysmorphia in Hungarian non-competitive male bodybuilders. *Eating Weight Disorders, 17,* e49–e53. doi:10.1007/BF03325327

*Babusa, B., Urbán, R., Czeglédi, E., & Túry, F. (2012). Psychometric properties and construct validity of the Muscle Appearance Satisfaction Scale among Hungarian men. *Body Image, 9*(1), 155–162. doi:10.1016/j.bodyim.2011.08.005

*Baile, J. I., González, A., Ramírez, C., & Suárez, P. (2011). Imagen corporal, hábitos alimentarios y hábitos de ejercicio físico en hombres usuarios de gimnasio y hombres universitarios no usuarios [Body image, eating behaviours and exercise training among male gym users and male non-users at the university]. *Revista de Psicología del Deporte, 20,* 353–366. Retrieved from http://www.rpd-online.com/article/view/783

*Bo, S., Zoccali, R., Ponzo, V., Soldati, L., De Carli, L., Benso, A., Fea, E., Rainoldi, A., Durazzo, M., Fassino, S., & Abbate-Daga, G. (2014). University courses, eating problems and muscle dysmorphia: Are there any associations? *Journal of Translational Medicine, 12,* 221. doi:10.1186/s12967-014-0221-2

Borenstein, M, J, Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Borenstein, M. J., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2014). *Comprehensive meta-analysis* (version 3). Englewood, NJ: Biostat.

*Brown, T. A., Forney, K. J., Pinner, D., & Keel, P. K. (2017). A randomized controlled trial of The Body Project: More Than Muscles for men with body dissatisfaction. *International Journal of Eating Disorders, 50,* 873–883. doi:10.1002/eat.22724

*Cafri, G., Olivardia, R., & Thompson, J. K. (2008). Symptom characteristics and psychiatric comorbidity among males with muscle dysmorphia. *Comprehensive Psychiatry, 49,* 374–379. doi:10.1016/j.comppsych.2008.01.003

Cafri, G., Thompson, J. K., Ricciardelli, L., McCabe, M., Smolak, L., & Yesalis, C. (2005). Pursuit if the muscular ideal: Physical and psychological consequences and putative risk factors. *Clinical Psychology Review, 25*(2), 215–239. doi:10.1016/j.cpr.2004.09.003

Cafri, G., van den Berg, P., & Thompson, J. K. (2006). Pursuit of muscularity in adolescent boys: Relations among biopsycho-social variables and clinical outcomes. *Journal of Clinical Child and Adolescent Psychology, 35*(2), 283–291. doi:10.1207/s15374424jccp3502_12

*Castro-Lopez, R., Cachón-Zagalaz, J., Molero, D., & Zagalaz-Sánchez, M. L. (2013) Dismorfia muscular y su relación con síntomas de trastornos de la conducta alimentaria [Muscle dysmorphia and its relationship with the symptoms of eating disorders]. *Revista Mexicana de Trastornos Alimentarios [Mexican Journal of Eating Disorders], 4,* 31–36. Retrieved from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-15232013000100004&lng=es

Cella, S., Iannaccone, M., & Cotrufo, P. (2012). Muscle dysmorphia: A comparison between competitive bodybuilders and fitness practitioners. *Journal of Nutritional Therapeutics, 1,* 12–18. Retrieved from http://www.lifescienceglobal.com/pms/index.php/jnt/article/view/451/pdf

Chandler, C. G., Grieve, F. G., Derryberry, W. P., & Pegg, P. O. (2009). Are anxiety and obsessive-compulsive symptoms related to muscle dysmorphia? *International Journal of Men's Health, 8,* 143–154. Retrieved from http://www.mensstudies.info/OJS/index.php/IJMH/article/view/568

Chaney, M. P. (2008). Muscle dysmorphia, self-esteem, and loneliness among gay and bisexual men. *International Journal of Men's Health, 7*(2), 157–170. doi:10.3149/jmh.0702.157

Choi, P. Y. L., Pope, H., & Olivardia, R. (2002). Muscle dysmorphia: A new syndrome in weightlifters. *British Journal of Sports Medicine, 36*(5), 375–376. doi:10.1136/bjsm.36.5.375

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Collins, R., Darkes, J., & Gwartney, D. (2007). A league of their own: Demographics, motivations and patterns of use of 1,955 male adult non-medical anabolic steroid users in the United States. *Journal of the International Society of Sports Nutrition, 4*(1), 12. doi:10.1186/1550-2783-4-12

Cubberley, R. (2009). *Evaluating the reliability and validity of the Muscle Dysmorphia Inventory* (Master's thesis). Western Kentucky University, Bowling Green, Kentucky. Retrieved from http://digitalcommons.wku.edu/theses/121/

*De Santis, J. P., Layerla, D. M., Barroso, S., Gattamorta, K. A., Sanchez, M., & Prado, G. J. (2011). Predictors of eating attitudes and behaviors among gay Hispanic men. *Archives of Psychiatric Nursing, 26,* 11–126. doi:10.1016/j.apnu.2011.06.003

Dos Santos, C. A., Tirico, P. P., Stefano, S. C., Touyz, S. W., & Claudino, A. M. (2015). Systematic review of the diagnostic category muscle dysmorphia. *Australian and New Zealand Journal of Psychiatry, 50*(4), 322–333. doi:10.1177/0004867415614106

*Dryer, R., Farr, M., Hiramatsu, I., & Quinton, S. (2016). The role of sociocultural influences on symptoms of muscle dysmorphia and eating disorders in men, and the mediating effects of perfectionism. *Behavioral Medicine, 42*(3), 174–182. doi:10.1080/08964289.2015.1122570

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot- based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. doi:10.1111/j.0006-341X.2000.00455.x

Ebbeck, V., Watkins, P. L., Concepcion, R. Y., Cardinal, B. J., & Hammermeister, J. (2009). Muscle dysmorphia symptoms and their relationships to self-concept and negative affect among college recreational exercisers. *Journal of Applied Sport Psychology, 21*(3), 262–275. doi:10.1080/10413200903019376

Fabris, M. A., Longobardi, C., Prino, L. E., & Settani, M. (2017). Attachment style and risk of muscle dysmorphia in a sample of male bodybuilders. *Psychology of Men & Masculinity, 19*(2), 273–281. doi:10.1037/men0000096

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370. doi:10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23

Garner, D. M., Olmstead, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine, 12*(4), 871–878. doi:10.1017/S0033291700049163

Garner, D., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a Multidimensional Eating Disorder Inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2*(2), 15–34. doi:10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

*Giardino, J. C., & Procidano, M. E. (2012). Muscle dysmorphia symptomatology: A cross-cultural study in Mexico and the

United States. *International Journal of Men's Health, 11*(1), 83–103. doi:10.3149/jmh.1101.83

*Goodale, K. R., Watkins, P. L., & Cardinal, B. J. (2001). Muscle dysmorphia: A new form of eating disorder? *American Journal of Health Education, 32*(5), 260–266. doi:10.1080/19325037.2001.10603480

Grieve, F. G. (2007). A conceptual model of factors contributing to the development of muscle dysmorphia. *Eating Disorders, 15*(1), 63–80. doi:10.1080/10640260601044535

Grieve, F. G., & Shacklette, M. D. (2012). Brief report on men's bodies and mood: Correlates between depressive symptoms and muscle dysmorphia symptoms. *North American Journal of Psychology, 14,* 563. Retrieved from http://www.freepatentsonline.com/article/North-American-Journal-Psychology/312401777.html

Grieve, F. G., Truba, N., & Bowersox, S. (2009). Etiology, assessment, and treatment of muscle dysmorphia. *Journal of Cognitive Psychotherapy, 23*(4), 306–314. doi:10.1891/0889-8391.23.4.306

Griffiths, S., Murray, S., & Touyz, S. (2013). Drive for muscularity and muscularity-oriented disordered eating in men: The role of set shifting difficulties and weak central coherence. *Body Image, 10*(4), 636–639. doi:10.1016/j.bodyim.2013.04.002

Grogan, S., Shepherd, S., Evans, R., Wright, S., & Hunter, G. (2006). Experiences of anabolic steroid use: In-depth interviews with men and women body builders. *Journal of Health Psychology, 11*(6), 845–856. doi:10.1177/1359105306069080

*Guidi, J., Clementi, C., & Grandi, C. (2013). Valutazione del disagio psicologico e delle caratteristiche di personalità nella dipendenza da esercizio fisico primaria [Psychological distress and personality characteristics among individuals with primary exercise dependence]. *Rivista di Psichiatria, 48,* 121–129. doi:10.1708/1272.14036

Hale, W. D. (2008). *Scale development in muscle dysmorphia* (Unpublished doctoral dissertation). Oklahoma State University, Stillwater, OK.

*Hale, B., Diehl, D., Weaver, K., & Briggs, M. (2013). Exercise dependence and muscle dysmorphia in novice and experienced female bodybuilders. *Journal of Behavioral Addictions, 2*(4), 244–248. doi:10.1556/JBA.2.2013.4.8

*Hildebrandt, T., Harty, S., & Langenbucher, J. W. (2012). Fitness supplements as a gateway substance for anabolic-androgenic steroid use. *Psychology of Addictive Behaviors, 26*(4), 955–962. doi:10.1037/a0027877

*Hildebrandt, T., Langenbucher, J. W., Karmin, J. K., Loeg, K. L., & Hollander, E. (2011). Development and validation of Appearance and Performance Enhancing Drug Use Schedule. *Addictive Behaviors, 36*(10), 949–958. doi:10.1016/j.addbeh.2011.05.002

*Hildebrandt, T., Langenbucher, J., & Schlundt, D. G. (2004). Muscularity concerns among men: Development of attitudinal and perceptual measures. *Body Image, 1*(2), 169–181. doi:10.1016/j.bodyim.2004.01.001

*Hildebrandt, T., Schlundt, D., Langenbucher, J., & Chung, T. (2006). Presence of muscle dysmorphia symptomatology among male weightlifters. *Comprehensive Psychiatry, 47*(2), 127–135. doi:10.1016/j.comppsych.2005.06.001

*Hildebrandt, T., Walker, C. D., Alfano, L., Delinsky, S., & Bannon, K. (2010). Development and validation of a Male Specific Body Checking Questionnaire. *International Journal of Eating Disorders, 43,* 77–87. doi:10.1002/eat.20669

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods, 11*(2), 193–206. doi:10.1037/1082-989X.11.2.193

*Hughes, E. K., Dean, C., & Allen, J. S. (2016). Measures of eating disorder symptoms, drive for muscularity, and muscle dysmorphia: Norms and typologies of Australian men. *Australian Journal of Psychology, 68*(4), 270–280. doi:10.1111/ajpy.12105

*Jin, X., Jin, Y., Zhou, S., Li, X., Yang, S., Yang, D., Nieuwoudt, J. E., & Yao, J. (2015). The Muscle Appearance Satisfaction Scale: A factorial analysis of validity and reliability for its use on adult Chinese male weightlifters. *Body Image, 14,* 94–101. doi:10.1016/j.bodyim.2015.04.004

Kaminski, P. L., McFarland, M. B., & Chapman, B. P. (2008). *The Muscle Dysmorphia Scale (MDS)*. Available from the author: Department of Psychology, University of North Texas, Denton, TX.

Kaminski, P. L., Slaton, S. R., Caster, J., Own, L., Baker, K., & Chapman, B. P. (2002, November). *The Male Eating Behavior and Body Image Evaluation (MEBBIE): A scale to measure eating, exercise, and body image concerns in men*. Poster presented at the annual convention of the Texas Psychological Association, San Antonio, TX.

*Kanayama, G., Barry, S., Hudson, J. I., & Pope, H. G. (2006). Body image and attitudes toward male roles in anabolic-androgenic steroid users. *American Journal of Psychiatry, 163*(4), 697–703. doi:10.1176/ajp.2006.163.4.697

Kimergård, A. (2015). A qualitative study of anabolic steroid use amongst gym users in the United Kingdom: Motives, beliefs and experiences. *Journal of Substance Use, 20*(4), 288–294. doi:10.3109/14659891.2014.911977

Kimmel, S. B., & Mahalik, J. R. (2004). Measuring masculine body ideal distress: Development of a measure. *International Journal of Men's Health, 3*(1), 1–10. doi:10.3149/jmh.0301.1

Klimek, P., & Hildebrandt, T. (2018). Psychosocial correlates of gap time to anabolic-androgenic steroid use. *International Journal of Eating Disorders, 51*(6), 535–541. doi:10.1002/eat.22859

*Klimek, P., Murray, S. B., Brown, T., Gonzales, M., & Blashill, A. J. (2018). Thinnes and muscularity internalization: Associations with disordered eating and muscle dysmorphia in men. *International Journal of Eating Disorders, 51*(4), 352–357. doi:10.1002/eat.22844

*Lamanna, J., Grieve, F. G., Derryberry, W. P., Hakman, M., & McClure, A. (2010). Antecedents of eating disorders and muscle dysmorphia in a non-clinical sample. *Eating Weight Disorders, 15*(1-2), e23–e33. doi:10.1007/BF03325277

Lantz, C. D., Rhea, D. J., & Cornelius, A. E. (2002). Muscle dysmorphia in elite-level power lifters and bodybuilders: A test of differences within a conceptual model. *Journal of Strength and Conditioning Research, 16*(4), 649e655. doi:10.1519/00124278-200211000-00026

*Latorre-Román, P. A., Garrido-Ruiz, A., & García-Pinillos, F. (2015). Versión española del cuestionario del complejo de Adonis: un cuestionario para el análisis del dimorfismo muscular o vigorexia [Spanish version of the Adonis Complex Questionnaire: A questionnaire for the analysis of muscle dysmorphia or bigorexia]. *Nutrición Hospitalaria, 31,* 1246–1253. doi:10.3305/nh.2015.31.3.8292

Lavender, J. M., Brown, T. A., & Murray, S. B. (2017). Men, muscles, and eating disorders: An overview of traditional and muscularity-oriented disordered eating. *Current Psychiatry Reports, 19*(6), 32. doi:10.1007/s11920-017-0787-5

Longobardi, C., Prino, L. E., Fabris, M. A., & Settani, M. (2017). Muscle dysmorphia and psychopathology: Findings from an Italian sample of male bodybuilders. *Psychiatric Research, 256,* 231–236. doi:10.1016/j.psychres.2017.06.065

*Lopez, A., Pollack, L., Gonzales, S., Pona, A. A., & Lundgren, J. D. (2015). Psychosocial correlates of muscle dysmorphia among collegiate males. *Journal of Psychological Inquiry, 20,* 58–66.

Lopez-Cuautle, C., Vazquez-Arevalo, R., & Mancilla-Diaz, J. M. (2016). Evaluación diagnóstica de la dismorfia muscular [Diagnostic evaluation of muscular dysmorphia]. *Anales de Psicología, 32*(2), 405–4016. doi:10.6018/analesps.32.2.203871

López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J., van den Noortgate, W., & Viechtbauer, W. (2014). Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *British Journal of Mathematical and Statistical Psychology, 67*(1), 30–48. doi:10.1111/bmsp.12002

*Magallares, A. (2016). Drive for thinness and pursuit of muscularity: The role of gender ideologies. *Universitas Psychologica, 15*(2), 353–360. doi:10.11144/Javeriana.upsy15-2.dtpm

*Maida, D. M., & Armstrong, S. L. (2005). The classification of muscle dysmorphia. *International Journal of Men's Health, 4*(1), 73–91. doi:10.3149/jmh.0401.73

Mayville, S. B., Williamson, D. A., White, M. A., Netemeyer, R. G., & Drab, D. L. (2002). Development of the Muscle Appearance Satisfaction Scale: A self-report measure for the assessment of muscle dysmorphia symptoms. *Assessment, 9*(4), 351–360. doi:10.1177/1073191102238156

*McFarland, M. B., & Kaminski, P. L. (2009). Men, muscles and mood: The relation between self-concept, dysphoria, and body image disturbances. *Eating Behaviors, 10*(1), 68–70. doi:10.1016/j.eatbeh.2008.10.007

Menees, L., Grieve, F. G., Mienaltowski, A., & Pope, J. (2013). Critical comments about the body and muscle dysmorphia symptoms in collegiate men. *International Journal of Men's Health, 12*(1), 17–28. doi:10.3149/jmh.1201.17

Mitchell, L., Murray, S. B., Cobley, S., Hackett, D., Gifford, J., Capling, L., & O'Connor, H. (2017). Muscle dysmorphia symptomatology and associated psychological features in bodybuilders and non-bodybuilder resistance trainers: A systematic review and meta-analysis. *Sports Medicine, 47*(2), 233–259. doi:10.1007/s40279-016-0564-3

*Mitchell, L., Murray, S. B., Hoon, M., Hackett, D., Prvan, T., & O'Connor, H. (2017). Correlates of muscle dysmorphia symptomatology in natural bodybuilders: Distinguishing factors in the pursuit of hyper-muscularity. *Body Image, 22,* 1–5. doi:10.1016/j.bodyim.2017.04.003

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269. doi:10.7326/0003-4819-151-4-200908180-00135

Morrison, T. G., Morrison, M. A., & McCann, L. (2006). Striving for bodily perfection? An overview of the drive for muscularity. In M. V. Kindes (Ed.), *Body image: New research* (pp. 1–34). Hauppauge, NY: Nova Science.

Murray, S. B., Griffiths, S., Mitchison, D., & Mond, J. M. (2017). The transition from thinness-oriented to muscularity-oriented disordered eating in adolescent males: A clinical observation. *Journal of Adolescent Health, 60*(3), 353–355. doi:10.1016/j.jadohealth.2016.10.014

*Murray, S. B., Griffiths, S., Mond, J. M., Kean, J., & Blashill, A. J. (2016). Anabolic steroid use and body image psychopathology in men: Delineating between appearance-versus performance-driven motivations. *Drug and Alcohol Dependence, 165,* 198–202. doi:10.1016/j.drugalcdep.2016.06.008

Murray, S. B., Nagata, J. M., Griffiths, S., Calzo, J. P., Brown, T. A., Mitchison, D., Blashill, A. J., & Mond, J. M. (2017). The enigma of male eating disorders: A critical review and synthesis. *Clinical Psychology Review, 57,* 1–11. doi:10.1016/j.cpr.2017.08.001

*Murray, S. B., Rieger, E., Hildebrand, T., Karlov, L., Russell, J., Boon, E., Dawson, R. T., & Touyz, S. W. (2012). A comparison of eating, exercise, shape, and weight related symptomatology in males with muscle dysmorphia and anorexia nervosa. *Body Image, 9*(2), 193–200. doi:10.1016/j.bodyim.2012.01.008

Nieuwoudt, J. E., Zhou, S., Coutts, R. A., & Booker, R. (2012). Muscle dysmorphia: Current research and potential classification as a disorder. *Psychology of Sport and Exercise, 13*(5), 569–577. doi:10.1016/j.psychsport.2012.03.006

*Nieuwoudt, J. E., Zhou, S., Coutts, R. A., & Booker, R. (2015). Symptoms of muscle dysmorphia, body dysmorphic disorder, and eating disorder in a nonclinical population of adult male weightlifters of Australia. *Journal of Strength and Conditioning Research, 29*(5), 1406–1414. doi:10.1519/JSC.0000000000000763

Olivardia, R. (2001). Mirror, mirror on the wall, who's the largest of them all? The features and phenomenology of muscle dysmorphia. *Harvard Review of Psychiatry, 9*(5), 254–259. doi:10.1080/10673220127900

*Olivardia, R., Pope, H. G., & Hudson, J. I. (2000). Muscle dysmorphia in male weightlifters: A case-control study. *American Journal of Psychiatry, 157*(8), 1291–1296. doi:10.1176/appi.ajp.157.8.1291

Pope, H. G., Gruber, A. J., Choi, P., Olivardia, R., & Phillips, K. A. (1997). Muscle dysmorphia: An underrecognized form of body dysmorphic disorder. *Psychosomatics, 38*(6), 548–557. doi:10.1016/S0033-3182(97)71400-2

Pope, H. G., Katz, D. L., & Hudson, J. I. (1993). Anorexia nervosa and "reverse anorexia" among 108 male bodybuilders. *Comprehensive Psychiatry, 34*(6), 406–409. doi:10.1016/0010-440X(93)90066-D

Pope, H. G., Khalsa, J. H., & Bhasin, S. (2017). Body image disorders and abuse of anabolic-androgenic steroids among men. *JAMA-Journal of the American Medical Association, 317*(1), 23–24. doi:10.1001/jama.2016.17441

Pope, H. G., Phillips, K. A., & Olivardia, R. (2000). *The Adonis complex: The secret crisis of male body obsession*. New York, NY: The Free Press.

*Pope, C. G., Pope, H. G., Menard, W., Fay, C., Olivardia, R., & Phillips, K. A. (2005). Clinical features of muscle dysmorphia among males with body dysmorphic disorder. *Body Image, 2,* 395–400. doi:10.1016/j.bodyim.2005.09.001

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 295–315). New York: Russell Sage Foundation.

Readdy, T., Cardinal, B. J., & Watkins, P. L. (2011). Muscle dysmorphia, gender role stress, and sociocultural influences: An exploratory study. *Research Quarterly for Exercise and Sport, 82,* 310–319. doi:10.1080/02701367.2011.10599759

*Rhea, D. J., Lantz, C. D., & Cornelius, A. E. (2004). Development of the Muscle Dysmorphia Inventory (MDI). *Journal of Sports Medicine and Physical Fitness, 44,* 428–435. Retrieved from http://www.drdebbierhea.com/wp-content/uploads/2013/05/Development-of-the-MDI-2005.pdf

Robert, C. A., Munroe-Chandler, K. J., & Gammage, K. L. (2009). The relationship between the drive for muscularity and muscle dysmorphia in male and female weight trainers. *The Journal of Strength and Conditioning Research, 23*(6), 1656–1662. doi:10.1519/JSC.0b013e3181b3dc2f

Sánchez-Meca, J., & Marín-Martínez, F. (2008). Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychological Methods, 13*(1), 31–48. doi:10.1037/1082-989X.13.1.31

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448–467. doi:10.1037/1082-989X.8.4.448

Sandgren, S. S., & Lavallee, D. (2018). Muscle dysmorphia research neglects DSM-5 diagnostic criteria. *Journal of Loss and Trauma, 23*(3), 211–243. doi:10.1080/15325024.2018.1428484

*Santarnecchi, E., & Dèttore, D. (2012). Muscle dysmorphia in different degrees of bodybuilding activities: Validation of the Italian version of Muscle Dysmorphia Disorder Inventory and bodybuilder image grid. *Body Image, 9*(3), 396–403. doi:10.1016/j.bodyim.2012.03.006

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research synthesis* (3rd ed.). Thousand Oaks, CA: Sage.

*Segura-García, C., Ammendolia, A., Procopio, L., Papaianni, M. C., Sinopoli, F., Bianco, C., De Fazio, P., & Capranica, L. (2010). Body uneasiness, eating disorders and muscle dysmorphia in individuals who overexercise. *Journal of Strength and Conditioning Research, 24*(11), 3098–3104. doi:10.1519/JSC.0b013e3181d0a575

Short, J. (2005). *Creating a tool to measure muscle dysmorphia* (Unpublished master's thesis). Western Kentucky University, Bowling Green, KY.

*Sladek, M. R., Engeln, R., & Miller, S. A. (2014). Development and validation of the Male Body Talk Scale: A psychometric investigation. *Body Image, 11*(3), 233–244. doi:10.1016/j.bodyim.2014.02.005

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 99–100). Chichester, UK: Wiley.

Suffolk, M. T., Dovey, T. M., Goodwin, H., & Meyer, C. (2013). Muscle dysmorphia: Methodological issues, implication for research. *Eating Disorders, 21*(5), 437–457. doi:10.1080/10640266.2013.828520

Tod, D., & Edwards, C. (2015a). Relationships among muscle dysmorphia characteristics, body image, quality of life, and coping in males. *Journal of Science and Medicine in Sport, 18*(5), 585–589. doi:10.1016/j.jsams.2014.07.015

Tod, D., & Edwards, C. (2015b). A meta-analysis of the drive for muscularity's relationships with exercise behaviour, disordered eating, supplement consumption, and exercise dependence. *International Review of Sport and Exercise Psychology, 8*(1), 185–203. doi:10.1080/1750984X.2015.1052089

Tod, D., Edwards, C., & Cranswick, I. (2016). Muscle dysmorphia: Current insights. *Journal of Psychology Research and Behavior Management, 9,* 179–188. doi:10.2147/PRBM.S97404

*Walker, D. C., Anderson, D. A., & Hildebrandt, T. (2009). Body checking behaviors in men. *Body Image, 63,* 164–170. doi:10.1016/j.bodyim.2009.05.001

APPENDIX A: CHECKLIST FOR ASSESSING THE METHODOLOGICAL QUALITY OF THE STUDIES

Item	Description
1	Representative sampling procedures (Yes: 1; No: 0)
2	The inclusion and exclusion criteria of participants were suitable to represent clinical or at risk population for MD and ED (Yes: 1; No: 0)
3	Dropouts were similar in sociodemographic characteristics to those of the final sample (completers) (Yes: 1; No: 0; Unable to determine: 0)
4	The measurement instrument used to assess MD show good psychometric properties (validity and reliability) in the sample of study (Yes: 1; No: 0; Unable to determine: 0)
5	The measurement instrument used to assess ED shows good psychometric properties (validity and reliability) in the sample of study (Yes: 1; No: 0; Unable to determine: 0)
6	The assessment of DM and ED was not dichotomized (Yes: 1; No: 0)
7	Absence of reporting bias: results for all MD and ED instruments described in "Methods" section were reported (Yes: 1; No: 0)
8	The statistical test used to assess the relationships between MD and ED was appropriate to the data (non-parametric methods vs. parametric methods) (Yes: 1; No: 0; Unable to determine: 0)
9	Absence of private financial support (Yes: 1; No: 0)
10	Statistical power: sample sizes have been calculated to detect an effect (Yes: 1; No: 0; Unable to determine: 0)

Note. The methodological quality checklist consisted of 10 criteria with a dichotomous response scale. The presence of the criterion is given 1 point and its absence 0 points. The total score was 10 points. MD: muscle dysmorphia; ED: eating disorders.

APPENDIX B: CHARACTERISTICS OF THE STUDIES OR INDEPENDENT SAMPLES INCLUDED IN THE SYSTEMATIC REVIEW AND/OR META-ANALYSIS ($N = 45$)

Study/country	N	Sampling method	Age (mean)	Males (%)	Sexual orientation (%)	Ethnicity (%)	Education level (%)	MD	ED	ES (r)
Babusa and Tury (2012)/ Hungary	120	Convenience sample of bodybuilders recruited in gymnasiums and undergraduates non bodybuilders	27.75	100	NR	NR	NR	MASS	EDI	NR
Babusa et al. (2012)/Hungary	529	Convenience sample of weight lifters and undergraduates	24.37	100	NR	NR	>Undergraduate	MASS	EDI	NR
Baile, González, Ramírez, and Suárez (2011)/Spain	418	Convenience sample from gymnasiums and university	20.86	100	NR	NR	>Undergraduate (81.3)	ACQ	CHAA	.447
Bo et al. (2014)/Italy	440	Convenience sample from university	19.80	45.9	NR	NR	>Undergraduate	MDDI	EAT-26	.170
Brown, Forney, Pinner, and Keel (2017)/USA	93	Convenience sample from university and local community	20.37	100	>H (90.3)	NR	>Undergraduate (86)	MDDI	EDE-Q	.461
Cafri et al. (2008)/USA	51	Convenience sample of weightlifters from gymnasiums and vitamin stores	25.71, Range: 18–40	100	>H (96.1)	>C (80.4)	NR	SCID-NP	SCID-NP	.059
Castro-López et al. (2013)/Spain	154	Probabilistic sample from bodybuilding gymnasiums	24.97	92.2	NR	NR	NR	ACQ	EDI-2	.280
De Santis et al. (2011)/USA	100	Convenience sample from bars, clubs, and beaches	32.47, Range: 18–51	100	G = 100	LH = 100	NR	ACQ	EAT-26	.402
Dryer et al. (2016)/Australia	158	Convenience sample from metropolitan area	26.94, Range: 18–36	100	H = 76 G = 15 B = 9	NR	University degree (45)	MDQ	EDI-3	.363
Sample 1: Giardino and Procidano (2012)/Mexico	35	Convenience sample of weightlifters from university gymnasiums	23.34, Range: 19–35	100	NR	LH = 71.7 MR = 8.7 C = 4.3	>Undergraduate (89.2)	MASS	EAT-26	.510
Sample 2: Giardino and Procidano (2012)/Mexico	11	Convenience sample of weightlifters from university gymnasiums	22.18, Range: 18–27	0	NR	LH = 71.7 MR = 8.7 C = 4.3	>Undergraduate (89.2)	MASS	EAT-26	.420
Sample 3: Giardino and Procidano (2012)/USA	43	Convenience sample of weightlifters from university gymnasiums	20.47, Range: 18–30	100	NR	C = 77.5 LH = 6 AA = 6 O = 10.5	>Undergraduate (91.2)	MASS	EAT-26	.500
Sample 4: Giardino and Procidano (2012)/USA	24	Convenience sample of weightlifters from university gymnasiums	20.17, Range: 18–22	0	NR	C = 77.5 LH = 6 AA = 6 O = 10.5	>Undergraduate (91.2)	MASS	EAT-26	.580
Goodale et al. (2001)/USA	323	Convenience sample from university	20.0, Range: 18–41	34.9	NR	>C (80.4)	Undergraduate	MDSQ	EDE-Q	.546
Guidi et al. (2013)/Italy	79	Convenience sample from gymnasiums	30	43	NR	NR	Undergraduate (30.4)	MDQ	EDI-2	NR

Study/Country	N	Sample	Age				Education	MDI	ED measure	r
Hale et al. (2013)/USA	74	Convenience sample from gymnasiums	Range: 18–48	0	NR	NR	NR	MDI	EDI-2 (drive for thinness scale)	NR
Hildebrandt, Harty, and Langenbucher (2012)/USA	201	Convenience sample from University	19.17	49.8	>H (97.5)	C = 55.7, AA = 19.9, AsA = 15.9, O = 8.5	Undergraduate	MDDI	EDE–Q	.180
Hildebrandt et al. (2011)/USA	85	Convenience sample from gymnasiums, nutritional supplements stores, discussion boards, newspaper advertisements, and general volunteer websites to complete a series of assessments	36.63	83.7	H = 82.3, G = 11.9, B = 5.8	C = 54.5, LH = 21.7, AA = 15.6, AsA = 6.9, O = 1.3	Undergraduate (28.1)	MDDI	EDE-Q	NR
Hildebrandt et al. (2004)/USA	237	Convenience sample of weightlifters from gymnasiums and vitamin stores	32.64, Range: 18–72	100	NR	NR	NR	MDDI	EDI	.496
Hildebrandt et al. (2006)/USA	237	Convenience sample of weightlifters from gymnasiums and vitamin stores	32.64, Range: 18–72	100	NR	NR	NR	MDDI	EDI	NR
Hildebrandt et al. (2010)/USA	342	Convenience sample from university	19.76	57.3	NR	C = 53.3, LH = 17, AA = 7.6, O = 22.2	Undergradate	MDDI	EDE-Q / EAT-26	NR
Hughes, Dean, and Allen (2016)/Australia	284	Convenience sample from metropolitan area	32.6, Range: 19–84	100	NR	C = 84.9, A = 9.9, I = 1.7, O = 3.5	NR	MDI	EAT-26 / EDE-Q	.387 / .271
Jin et al. (2015)/China	592	Convenience sample from fitness centers and gymnasiums	27, Range: 18–58	100	NR	NR	NR	MASS	EAT-26	.156
Kanayama et al. (2006)/USA	89	Convenience sample from gymnasiums	29.67	100	NR	>C (80.9)	College graduate (42.7)	Own	EDI	NR
Klimek et al. (2018)/USA	180	Convenience sample from university	19.6, Range: 18–33	100	87.2	C = 55	Undergraduate	MDDI	EDE-Q	.570
Sample 1: Lamanna et al. (2010)/USA	93	Convenience sample from university	19.13, Range: 17–44 (data from all sample)	100	NR	C = 87.4, AA = 6, LH = 1.1, NA = 1.1, AsA = 1.1, O = 3.3 (data from all sample)	Undergraduate	MDDI	EAT-26	.227

(Continued)

(Continued)

Study/country	N	Sampling method	Age (mean)	Males (%)	Sexual orientation (%)	Ethnicity (%)	Education level (%)	MD	ED	ES (r)
Sample 2: Lamanna et al. (2010)/USA	106	Convenience sample from university	19.13, Range: 17–44 (data from all sample)	0	NR	C = 87.4 AA = 6 LH = 1.1 NA = 1.1 AsA = 1.1 O = 3.3 (data from all sample)	Undergraduate	MDDI	EAT-26	.467
Latorre-Román et al. (2015)/ Spain	99	Convenience sample from bodybuilding gymnasiums	25.45	100	NR	NR	NR	ACQ	EAT-26	.422
Lopez et al. (2015)/USA	120	Convenience sample from University	23.24	100	NR	C = 60 AA = 20 LH = 8.3 O = 11.7	Undergraduate	MDI	EDE-Q	.155
Magallares (2016)/Spain	615	Convenience sample from university	25.56, Range: 18–30	0	NR	NR	Undergraduate	MASS	EDI-2	NR
Maida and Armstrong (2005)/ USA	106	Convenience sample of weightlifters from gymnasiums	Range: 18–45	100	NR	NR	>had attended college for 1–3 years (56)	MDSQ	EDI	.183
McFarland and Kaminski (2009)/USA	276	Convenience sample from university	Median: 20 (data from all sample)	100	H = 89.8 G/B = 7.5	EA = 66.7 LH = 12.2 AA = 10.5	Undergraduate	MDS	MEBBIE	.135
Mitchell, Murray, Hoon, et al. (2017)/Australia	60	Convenience sample of competitive natural bodybuilders	29.6	100	NR	NR	NR	MDDI	EAT	.31
Murray et al. (2016)/England	122	Convenience sample of people reporting current anabolic androgenic steroid use from bodybuilding gymnasiums and specialist needle exchange for self-administered anabolic–androgenic steroids injections	29.38 Range: 17–48	100	H = 83.6 G/B = 13.1	NR	NR	MDDI	EDE-Q Modified	.710
Murray et al. (2012)/UK/USA/ Australia/Singapore	60	Convenience sample from gymnasiums and clinical settings	26.58	100	NR	NR	NR	MDDI	EDE-Q Modified	.539
Nieuwoudt, Zhou, Coutts, and Booker (2015)/Australia	648	Convenience sample from weightlifting/ bodybuilding discussion forums, facebook	29.5 Range: 18–65	100	>H (93.4)	>C (85.3)	NR	MASS	EAT-26	.411
Olivardia et al. (2000)/USA	54	Convenience sample from gymnasiums	25.4	100	NR	>C (90.7)	Graduate: 48.14	SCID-P	EDI	.490
Pope et al. (2005)/USA	63	Convenience sample from clinical settings and advertisement	34.91	100	NR	>C (90.5)	>at least some college (over 69.8)	SCID-NP	SCID-NP	.272
Rhea et al. (2004)/USA	151	Convenience sample from bodybuilders (n = 58) and weight lifters (n = 93)	28.48	91.4	NR	C = 90 O = 10	NR	MDI	EDI	.460

Study/Country	N	Sample	Age	% male	Sexual orientation	Ethnicity	Education	MD measure	ED measure	ES
Santarnecchi and Dettore (2012)/Italy	180	Convenience sample from gymnasiums	32.67 Range: 23–41	100	NR	NR	NR	MDDI	EDI	.330
Sample 1: Segura-García et al. (2010)/Italy	86	Convenience sample from fitness center	27.67	100	NR	NR	NR	MDI	EDI-2	.183
Sample 2: Segura-García et al. (2010)/Italy	48	Convenience sample from fitness center	28.6	0	NR	NR	NR	MDI	EDI-2	.251
Sample 3: Segura-García et al. (2010)/Italy	20	Convenience sample from clinical setting (women with EDs)	22.1	0	NR	NR	NR	MDI	EDI-2	.271
Sladek, Engeln, and Miller (2014)/USA	160	Convenience sample from university and via Amazon's Mechanical Turk website	26.62 Range: 18–64	100	>H (97.1)	>C (68.8)	>Undergraduate (57.2)	MASS	EAT-26	NR
Walker, Anderson, and Hildebrandt (2009)/Albania	550	Convenience sample from university	18.98 Range: 16–30	100	NR	C = 68.9 AA = 8.7 A = 7.7 H = 5.8 MR = 4.6 O = 4.3	Undergraduate	MDDI	EDE-Q	.167

Note. MD: muscle dysmorphia; ED: eating disorder; ES: effect size statistics; NR: not reported. *Ethnicity:* A: Asian; AA: African American; AsA: Asian American; C: Caucasian; EA: European American; H: Hispanic; I: Indian; LH: Latin Hispanic; MR: multiracial; NA: Native American; O: others. *Sexual Orientation:* H: heterosexual; G: gay, B: bisexual. *Muscle dysmorphia:* ACQ: Adonis Complex Questionnaire (Pope et al., 2000); MDS-Q: Muscle Dysmorphia Symptom Questionnaire (Olivardia et al., 2000); MASS: Muscle Appearance Satisfaction Scale (Mayville et al., 2002); MDDI: Muscle Dysmorphic Disorder Inventory (Hildebrandt et al., 2004); MDI: Muscle Dysmorphia Inventory (Rhea et al., 2004); MD-Q: Muscle Dysmorphia Symptom Questionnaire (Grieve et al., 2012); MDS: Muscle Dysmorphia Scale (Kaminski, McFarland, & Chapman, 2008); SCID-NP: structured clinical interview for DSM-IV – Non patient version; SCID-P: structured clinical interview for DSM-IV – Patient version. *Eating disorders:* CHAA: Cuestionario de Hábitos de Alimentación Alterados; EAT-26: Eating Attitudes Test (Garner, Olmstead, Bohr, & Garfinkel, 1982); EDI: Eating Disorders Inventory (Garner et al., 1983); EDE-Q: Eating Disorder Examination Questionnaire (Fairburn & Beglin, 1994); EDE-Q Modified: Eating Disorder Examination Questionnaire Modified (Murray et al., 2012); EDI: Eating Disorders Inventory (Garner et al., 1983); MEBBIE: Male Eating Behavior and Body Image Evaluation (Kaminski et al., 2002).

APPENDIX C: METHODOLOGICAL QUALITY OF THE STUDIES

Study	Methodological quality criteria										
	1	2	3	4	5	6	7	8	9	10	Total (0–10)
Babusa and Túry (2012)	0	1	0	1	0	1	1	1	0	0	5
Babusa et al. (2012)	0	1	0	1	0	1	1	1	1	0	6
Baile et al. (2011)	0	1	0	0	1	1	1	0	1	0	5
Bo et al. (2014)	0	0	0	0	0	0	1	1	1	0	3
Brown et al. (2017)	0	1	0	1	1	1	1	1	0	0	6
Cafri et al. (2008)	0	1	0	0	0	0	1	1	1	1	5
Castro López et al. (2013)	1	1	0	0	0	1	1	0	1	0	5
De Santis et al. (2011)	0	0	0	1	1	1	1	1	0	1	6
Dryer et al. (2016)	0	0	0	1	1	1	1	1	1	0	6
Giardino and Procidano (2012): Study 1	0	1	0	1	1	1	1	0	1	0	6
Giardino and Procidano (2012): Study 2	0	1	0	1	1	1	1	0	1	0	6
Giardino and Procidano (2012): Study 3	0	1	0	1	1	1	1	0	1	0	6
Giardino and Procidano (2012): Study 4	0	1	0	1	1	1	1	0	1	0	6
Goodale et al. (2001)	0	0	0	1	1	1	1	1	1	0	6
Guidi et al. (2013)	0	1	0	0	0	1	1	0	1	0	4
Hale et al. (2013)	0	1	0	1	1	1	1	0	1	0	6
Hildebrandt et al. (2004)	0	1	0	1	1	1	1	0	1	0	6
Hildebrandt et al. (2006)	0	1	0	1	1	1	0	1	1	0	6
Hildebrandt et al (2010)	0	0	0	0	0	1	1	0	1	0	3
Hildelbrandt et al. (2011)	0	1	0	1	1	1	1	0	1	0	6
Hildebrandt et al. (2012)	0	0	0	1	1	1	1	1	1	0	6
Hughes et al. (2016)	0	0	0	1	1	1	1	1	1	0	6
Jin et al. (2015)	0	1	0	1	1	1	1	1	1	0	7
Kanayama et al. (2006)	0	1	0	0	0	1	1	0	0	0	3
Klimek et al. (2018)	0	0	0	1	1	1	1	1	1	0	6
Lamanna et al. (2010): Study 1	0	0	0	1	1	1	1	1	1	0	6
Lamanna et al. (2010): Study 2	0	0	0	1	1	1	1	1	1	0	6
Latorre-Román et al. (2014)	0	1	0	1	1	1	1	0	1	0	6
Lopez et al. (2015)	0	0	0	0	0	1	1	0	1	0	3
Magallares (2016)	0	0	0	1	1	1	1	0	1	0	5
Maida et al. (2005)	0	1	0	0	0	1	1	0	1	0	4
McFarland and Kaminski (2009)	0	0	0	1	0	1	1	0	1	0	4
Mitchell, Murray, Cobley, et al. (2017)	0	1	0	1	1	1	1	1	0	0	6
Murray et al. (2012)	0	1	0	1	1	1	1	1	1	0	7
Murray et al. (2016)	0	1	0	1	1	1	1	1	1	0	7
Nieuwoudt et al. (2015)	0	1	0	1	1	1	1	1	1	0	7
Olivardia et al. (2000)	0	1	0	0	0	0	1	0	1	0	3
Pope et al. (2005)	0	1	0	0	0	1	1	0	1	0	4
Rhea et al. (2004)	0	1	0	1	1	1	1	1	1	0	7
Santarnecchi and Dettore (2012)	0	1	0	1	1	1	1	0	0	0	5
Segura-García et al. (2010): Sample 1	0	1	0	0	0	0	1	1	1	0	4
Segura-García et al. (2010): Sample 2	0	1	0	0	0	0	1	1	1	0	4
Segura-García et al. (2010): Sample 3	0	1	0	0	0	0	1	1	1	0	4
Sladek et al. (2014)	0	1	0	0	0	1	1	0	1	0	4
Walker et al. (2009)	0	0	0	1	1	1	1	1	1	0	6

Note. The methodological quality checklist consists of 10 criteria with a dichotomous response scale. The presence of the criterion is given 1 point and its absence 0 points. The total quality score for each study was obtained by summing the 1s and 0s through the 10 quality items (see Appendix A for a description of the quality items).

APPENDIX D

Funnel plot of the MD–ED effect sizes to assess publication bias. White circles represent each of the studies included. The absence of black circles indicates that the trim-and-fill method had not to impute missing data to symmetrize the funnel plot.

Funnel Plot of Standard Error by Fisher's Z