

**Table S1. Scoring criteria for the quality assessment of individual studies**

Criteria	Scale name	Detailed description of questions relevant to this report	Scoring
<b>C1</b>	Clear description of research question	Was the specific independent variable(s) identified?	+ Φ –
		Was the dependent variable(s) clearly indicated?	+ Φ –
		Were hypotheses clearly stated?	+ Φ –
		Were the target population and setting specified?	+ Φ --
<b>C2</b>	Sampling selection free from bias	Grading according to: Randomly sampled from defined population (+), Stratified sampling from a defined population (+), Cluster sampling (Φ), Convenience sampling (Φ), Unclear (--)	+ Φ –
<b>C3</b>	Sampling adequately described (Originally: Sample selection free from bias, see C2)	Were inclusion/exclusion criteria specified (e.g., risk, diagnostic criteria), and with sufficient detail and without omitting criteria critical to the study?	+ Φ –
		Were criteria applied equally to all study groups?	+ Φ –
		Were health, demographics, and other characteristics of subjects described?	+ Φ –
<b>C4</b>	Comparable study groups	Were concurrent controls used? (Concurrent preferred over historical controls.) Given if sampled concurrently.	+ Φ –
		If cohort study or cross-sectional study, were groups comparable on important confounding factors and/or were preexisting differences (e.g., age, BMI, SES)	+ Φ –
<b>C5</b>	Participants withdrawals or response rate described	Was the number, characteristics of withdrawals (i.e., dropouts, lost to follow up, attrition rate) and/or response rate (cross-sectional studies) described for each group?	+ Φ –
		Were all enrolled subjects/patients (in the original sample) accounted for? If not, was their exclusion comprehensible?	+ Φ –
		If group comparison: Were reasons for withdrawals similar across groups?	+ Φ –
<b>C6</b>	Appropriateness and clear description of data collection procedures	In observational study, were study settings, and data collection procedures clearly described?	+ Φ –
		<i>(Added) Was the approach appropriate for the research question?</i>	+ Φ –
		<i>(Added) Was the sampling strategy appropriate for the research question?</i>	+ Φ –
<b>C7</b>	Clearly defined, valid outcomes of orthorexic eating	Were orthorexic eating measures appropriate to question and outcomes of concern?	+ Φ –
		Were the observations and measurements based on standard, valid, and reliable data collection/instruments/tests/procedures? (Grading according to: TOS/DOS/EHQ (+), ORTO/BOT (Φ), other (--))	+ Φ –
		Were other factors accounted for (measured) that could affect outcomes?	+ Φ –
		Were the measurements conducted consistently across groups?	+ Φ –
		Is the validity of the ON measure mentioned?	+ Φ --

<b>C8</b>	Score reliability estimate given	Are score reliability estimates given? (Grading according to: any estimate from current sample (+), any estimate from another study (Φ), None given (--))	+ Φ –
<b>C9</b>	Clearly defined, valid, reliable outcomes of exercise (addiction)	Were exercise (addiction) measures appropriate to question and outcomes of concern?	+ Φ –
		Were the observations and measurements based on standard, valid, and reliable data collection instruments/tests/procedures?	+ Φ –
		Are validity and reliability of the exercise (addiction) measure mentioned?	+ Φ –
		Were other factors accounted for (measured) that could affect outcomes?	+ Φ –
		Were the measurements conducted consistently across groups?	+ Φ --
<b>C10</b>	Appropriate statistical analysis	Were statistical analyses adequately described and the results reported appropriately?	+ Φ –
		Were correct statistical tests used and assumptions of test not violated?	+ Φ –
		Were statistics reported with levels of significance and/or confidence intervals (added) and/or effect size?	+ Φ –
		Were adequate adjustments made for effects of confounding factors that might have affected the outcomes (e.g., multivariate analyses)?	+ Φ –
		If negative findings, was a power calculation reported to address type 2 error?	+ Φ --
<b>C11</b>	Conclusions supported by results	Is there a discussion of findings? ( <i>Added: Are the claims made supported by sufficient evidence?</i> )	+ Φ –
		Are biases and study limitations identified and discussed?	+ Φ --
<b>C12</b>	Unlikely funding bias	Were sources of funding and investigators' affiliations described?	+ Φ –
		Was there no apparent conflict of interest?	+ Φ --

This is a modified set of the quality criteria for primary research as proposed in the Evidence Analysis Manual of the Academy of Nutrition and Dietetics (2016, Retrieved from [https://www.andean.org/vault/2440/web/files/2016\\_April\\_EA\\_Manual.pdf](https://www.andean.org/vault/2440/web/files/2016_April_EA_Manual.pdf)).

As the association between orthorexia/obsessive healthy eating and exercise (addiction) was not an objective of all included primary studies, the quality score was computed as the mean of responses across criteria that could be evaluated.

+ Positive (= 2 points): Indicates that the report has clearly addressed these issues.

Φ Neutral (= 1 point): Indicates that the presence of this criterion is ambiguous or that the report is neither exceptionally strong nor exceptionally weak.

-- Negative (= 0 points): Indicates that these issues have not been adequately addressed.

**Table S2. Quality assessment of individual studies included in the meta-analysis on the association between orthorexic eating and exercise (addiction)**

<b>First author (year)</b>	<b>C1</b> Clear research question	<b>C2</b> Sampling bias	<b>C3</b> Sample description	<b>C4</b> Comparable study groups	<b>C5</b> Withdrawals/ response rates	<b>C6</b> Data collection procedures	<b>C7</b> Outcome of ON	<b>C8</b> ON measure reliability	<b>C9</b> Outcome of Ex(Add)	<b>C10</b> Statistics	<b>C11</b> Conclusions	<b>C12</b> Funding bias	<b>Final score<sup>1</sup></b>
Aksoydan (2009)	1	1	1	0	1	2	1	0	N/A*	1	1	0	<b>0.82</b>
Almeida (2018)	1	1	1	N/A	0	2	1	2	N/A*	1	2	1	<b>1.20</b>
Bert (2019)	2	1	2	1	1	2	1	1	1	1	2	2	<b>1.42<sup>2</sup></b>
	2	1	2	1	1	2	2	1	1	1	2	2	<b>1.50<sup>3</sup></b>
Bo (2014)	2	1	2	1	1	2	1	0	N/A*	2	2	1	<b>1.36</b>
Bóna (2019)	1	1	2	N/A	2	2	1	2	1	1	2	2	<b>1.55</b>
Çiçekoğlu (2018)	2	1	1	1	0	2	1	1	N/A*	1	0	0	<b>0.91</b>
Clifford (2019)	1	1	2	2	0	2	1	1	1	1	2	2	<b>1.33</b>
Dunn (2017)	2	1	2	N/A	2	2	2	0	N/A*	1	2	1	<b>1.50</b>
Duran (2020)	1	1	1	1	1	2	1	2	0	1	1	1	<b>1.08</b>
Freire (2020)	2	1	2	N/A	0	2	1	2	2	1	2	2	<b>1.55</b>
Gorrasi (2019)	2	1	2	2	1	2	2	1	N/A*	2	2	2	<b>1.73</b>
Henriksson (2014)	2	2	2	N/A	2	2	1	0	1	0	1	0	<b>1.18</b>
Kiss-Leizer (2019)	2	1	2	N/A	0	2	1	2	1	2	2	2	<b>1.55</b>
Lewis (2012)	2	1	2	1	0	2	1	2	2	2	2	0	<b>1.42</b>

Supplementary material for Strahler et al. Obsessive healthy eating and orthorexic eating tendencies in sport contexts: A systematic review and meta-analysis

Malmborg (2017)	2	1	2	2	1	2	1	0	2	1	1	1	<b>1.33</b>
Özdengül (2020)	1	1	2	2	1	2	2	2	1	2	2	2	<b>1.67</b>
Oberle (2018)	2	1	2	N/A	2	2	2	1	2	1	2	1	<b>1.64</b>
Roncero (2017)	2	1	2	N/A	2	2	2	2	N/A*	2	2	0	<b>1.70</b>
Rudolph (2017)	2	1	1	N/A	2	2	2	2	1	1	2	1	<b>1.55</b>
Rudolph (2018)	1	1	2	N/A	1	2	2	2	2	1	2	1	<b>1.55</b>
Segura-Garcia (2012)	2	1	2	1	2	2	1	2	N/A	2	1	0	<b>1.45</b>
Strahler (2018)	1	1	2	N/A	2	2	2	2	1*	2	2	2	<b>1.73</b>
Surala 2020	2	1	1	N/A	0	1	1	1	0	1	2	2	<b>1.09</b>
White 2020	1	1	1	N/A	1	2	2	2	2	1	2	1	<b>1.45</b>
Worsfold 2020	2	1	2	1	1	2	1	2	N/A*	2	2	0	<b>1.45</b>

ON; Orthorexic eating, Orthorexia nervosa, or obsession with healthy eating; Ex(Add), exercise or exercise addiction measure.

<sup>1</sup> As the association between orthorexia/obsessive healthy eating and exercise (addiction) was not an objective of all included primary studies, the quality score was computed as the mean of responses across criteria that could be evaluated.

<sup>2</sup> ORTO-15 as outcome measure

<sup>3</sup> EHQ as outcome measure

\* The association between orthorexia and exercise was none of the paper's objectives.

# Exercise (yes/no) one variable among many other health behaviors.