

ORIGINAL ARTICLE

Identifying research priorities for the study of atypical anorexia nervosa: A Delphi study

Mattias Strand MD¹  | Johan Zvrskovec MSc^{2,3}  |
 Christopher Hübel MD PhD^{2,3,4,5}  | Christine M Peat PhD⁶ |
 Cynthia M Bulik PhD^{2,6,7} | Andreas Birgegård PhD^{1,2} 

¹Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Stockholm, Sweden

²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

³Social Genetic and Developmental Psychiatric Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

⁴UK National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley Hospital, London, UK

⁵National Centre for Register-Based Research, Aarhus University, Aarhus, Denmark

⁶Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁷Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Correspondence

Mattias Strand, Transkulturellt Centrum, Solnavägen 4, 113 65 Stockholm, Sweden.
 Email: mattias.strand@ki.se

Action Editor: Ruth Weissman

Funding information

National Institute of Mental Health, Grant/Award Numbers: R01MH119084, R01MH120170; Vetenskapsrådet, Grant/Award Number: 538-2013-8864

Abstract

Objective: Individuals meeting all criteria for anorexia nervosa (AN) except that weight falls within or above the normal range despite significant weight loss are categorized as having atypical AN (AAN). Existing research has provided mixed evidence concerning the diagnostic demarcation of AN and AAN. The aim of the present study was to identify research priorities for furthering the understanding of AN and AAN as diagnostic entities.

Method: Employing the Delphi methodology, experts in the field were invited to suggest research questions that need to be explored in the demarcation of AN from AAN. This yielded 24 research areas, that were presented in subsequent rounds where panelists were asked to prioritize areas of primary interest.

Results: Fifty-three panelists completed all three Delphi rounds. Consensus was only reached on three items considered to be of primary interest: medical, neurobiological, and neurological factors; epidemiology and natural course; and treatment response in AAN compared to AN. In contrast, questions of premorbid weight and determining the need for and nature of a body mass index cutoff differentiating between AAN and AN were seen as being of low priority.

Discussion: These findings reveal a relatively low degree of consensus on the demarcation of AN from AAN in the field of eating disorders. A reason could be that the definition and use of the AAN category vary in research and clinical practice. In order to achieve further diagnostic clarity, research on the demarcation of AAN and AN should focus on the identified prioritized research areas.

KEYWORDS

body mass index, feeding and eating disorders, interdisciplinary research, obesity, overweight, stereotyping

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *International Journal of Eating Disorders* published by Wiley Periodicals LLC.

1 | BACKGROUND

In the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), meeting all criteria for anorexia nervosa (AN) except that weight falls within the normal range despite significant weight loss was subsumed under the Eating Disorder Not Otherwise Specified category (American Psychiatric Association, 1994). With the introduction of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) in 2013, this clinical condition—i.e., AN symptomatology in the absence of underweight—was named atypical anorexia nervosa (AAN) and recognized as one of the examples within the Other Specified Feeding and Eating Disorders category (American Psychiatric Association, 2013). Notably, in contrast to the previous edition, the DSM-5 criteria also explicitly specify that the weight of an individual with AAN may be above the normal range.

Clinicians and researchers have noted an anecdotal increase in the number of patients presenting for treatment with a restricting-type eating disorder who meet all diagnostic criteria for AN except that their body mass index (BMI) is within the normal, overweight, or obese weight ranges. In a study on adolescents admitted to hospital because of a restrictive eating disorder, 8% fell in the AAN category in 2005 compared to 47% in 2009 (Whitelaw, Gilbertson, Lee, & Sawyer, 2014). AN symptomatology in the absence of underweight has increasingly been described as an overlooked concern, both in the scientific literature (Neumark-Sztainer, 2015) and in popular media (Dennett, 2019; Tait, 2015).

A retrospective cohort study of adolescent patients with a restrictive eating disorder presenting for intake evaluation showed that the group of patients that fell in the overweight or obese weight range before the onset of eating disorder symptoms (36.7% of the sample) had a mean duration of illness of almost 20 months before intake compared with just over 11 months for the group of patients with a pre-onset weight within the normal range—i.e., detection of eating pathology was substantially longer for those with BMIs in the overweight/obese range (Lebow, Sim, & Kransdorf, 2015). Similarly, a case series described how two adolescent patients with severe restrictive eating disorders and marked weight loss went largely unrecognized by their primary care physicians, primarily because their BMIs were within the obese range at the onset of illness (Sim, Lebow, & Billings, 2013). There appears to be an obvious risk that weight loss in an individual with overweight or obesity is seen by default as solely positive against a backdrop of “the obesity epidemic” and the potential medical complications associated with obesity. Moreover, both the affected individual as well as family members may be less inclined to consult a physician when dramatic weight loss occurs in a context of previous or concurrent overweight or obesity.

Within this context, several authors and individuals with lived experience call for clinical recognition of individuals who experience extreme weight loss, evidence of under- or malnutrition, and cognitive and behavioral signs and symptoms of an eating disorder regardless of weight (Forney, Brown, Holland-Carter, Kennedy, & Keel, 2017; Neumark-Sztainer, 2015; Sawyer, Whitelaw, Le Grange, Yeo, &

Hughes, 2016; Sim et al., 2013; Whitelaw et al., 2014; Whitelaw, Lee, Gilbertson, & Sawyer, 2018).

1.1 | How different are AN and AAN?

Some studies have pointed to overall similarities between AN and AAN patients in terms of illness severity. A study on adolescents with restrictive eating in a specialized eating disorders program found that compared to patients with AN, the group of adolescents with AAN were more often overweight or obese premorbidly and had lost more weight over a longer period of time at presentation (Sawyer et al., 2016). No group differences were observed between AN and AAN patients in terms of resting pulse rate, frequency of bradycardia, marked orthostatic changes, hypothermia, or need for hospital admission at presentation. Systolic blood pressure was higher and amenorrhea was less frequent in the AAN group. The AAN group reported significantly more severe eating disorder symptoms on all subscales of the Eating Disorders Examination Questionnaire (EDE-Q)—the authors argue that this is an expected finding, given that a higher BMI is generally associated with greater body image concerns. There were no significant group differences regarding the occurrence of binge eating, vomiting, laxative misuse, or compulsive exercise. Also, the prevalence of psychiatric comorbidity, self-harm or suicidal ideation, depressive symptomatology, obsessive-compulsive traits, or psychotropic medication did not differ between the AN and AAN groups.

Another study by the same group of researchers investigated medical complications such as bradycardia or hypophosphatemia in adolescents with either AN or AAN hospitalized because of a restrictive eating disorder (Whitelaw et al., 2018). In sum, the authors conclude that total weight loss and recent weight loss were more accurate predictors of somatic complications of restrictive eating than the admission weight per se.

In contrast, in a recent study on bone health in men with AN, AAN, and avoidant/restrictive food intake disorder (ARFID), the AN group was at higher risk of lower bone mineral density and hip strength than the other groups although bone mineral density was also reduced in AAN and ARFID (Schorr et al., 2019). Here, the authors conclude that from a bone health perspective, the absence of a history of underweight may be relatively protective.

Small-scale studies of brain structure and function have pointed to potential pathogenetic differences between AN and AAN. For example, in contrast to AN, alterations in gray and white matter microstructure do not appear to be common features of AAN; however, this may reflect less severe undernutrition in AAN (Olivo et al., 2018; Olivo et al., 2019). Moreover, food-related neural responses in AAN may differ from what is typically seen in AN (Olivo et al., 2019).

The flexibility in defining significant weight loss in AAN inherent to the DSM-5 criteria may be reasonable in clinical settings, but is difficult to operationalize in research. In a study of an adult nonclinical cohort, three definitions of AAN with different levels of lifetime weight loss (>5%, >10%, and >15%) were tested (Forney et al., 2017).

Comparisons were also made with participants who had lost weight but did not have cognitive eating disorder concerns (i.e., criteria B and C of DSM-5 AN), as well as those with cognitive concerns but no weight loss. A vast majority of participants categorized as AAN reported previous weights in the overweight and obese ranges. At all three definitions of significant weight loss, the AAN group displayed elevated eating pathology and distress compared with healthy controls. In comparison with other DSM-5 eating disorders, women in the AAN group displayed somewhat less severe eating pathology and distress, whereas these levels were similar for men in the AAN and other DSM-5 eating disorders groups. The AAN group was also clearly distinguishable from the groups with only weight loss or cognitive concern regarding eating pathology and distress. It should be noted that this sample was nonclinical and that lifetime weight loss was studied, not recent weight loss. Moreover, a clinical observation is that the impaired insight into the severity of their symptoms that characterize some individuals with AN (Gorwood, Duriez, Lengvenyte, Guillaume, & Criquillion, 2019; Konstantakopoulos, Tchanturia, Surguladze, & David, 2011) could lead them to underreport on cognitive measures of AN. Whether this also holds true for individuals in the AAN category is uncertain.

Furthermore, diagnostic overlap between AAN and purging disorder has also been suggested. In the above study, 57% of participants with purging disorder and 52% with bulimia nervosa also met AAN criteria at the 5% weight loss definition (Forney et al., 2017).

Understanding the differences and similarities between AN and AAN is important as clinicians caring for patients with AAN may have to weigh several factors when developing a treatment plan. The approach to renourishment may differ when weight gain is not a medical necessity. Simply transferring recommended renourishment protocols for AN to individuals with AAN may not be the optimal approach to achieve remission and health (Nagata, Garber, & Buckelew, 2018). In a small treatment study of family-based treatment (FBT) for families of adolescents with AAN, eating disorder cognitions and behaviors improved but weight increased only slightly (Hughes, Le Grange, Court, & Sawyer, 2017). The mean BMI was 21.0 kg/m² at baseline and 21.6 kg/m² post-treatment. Remission rates were similar or even superior to those reported in clinical trials of FBT for adolescent AN; however, the authors note that assessing treatment success in AAN is complex since the criterion of rapid weight restoration used in evaluation of FBT in AN does not apply. It is possible that weight restoration to premorbid levels in AAN could further improve remission rates, but the authors' experience is that healthcare providers, patients, and families are reluctant to consider this option because of the medical risks associated with overweight and obesity. Another recent study on weight-gain trajectories in FBT for adolescents with restricting eating disorders found that for AAN patients, the most common trajectory was an initial shallow slope of weight gain with a subsequent return to their pre-intervention BMI and an overall weight maintenance at the 60th percentile throughout treatment (Lebow et al., 2019).

Also of note is that in a large retrospective review of DSM-5 eating disorder diagnoses, 27.3% of patients with an AAN diagnosis had

a < 90% median BMI at intake and 13.7% of patients with an AN diagnosis had a ≥ 90% median BMI at intake (Forman et al., 2014). Clearly, the AAN and AN diagnostic criteria are not uniformly applied in clinical settings.

Importantly, AAN is defined differently in the tenth edition of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-10) (World Health Organization, 2004) than in the DSM-5, which affects the study populations in investigations of AAN utilizing these diagnostic systems. Studies that define AAN according to the ICD-10 criteria will thus include individuals with AN-type restrictive eating without underweight, but also individuals without amenorrhea (categorized as AN in the DSM-5) or those without marked fear of gaining weight (possibly categorized as ARFID in the DSM-5; for an example, see Silén et al., 2015).

In sum, some studies indicate that AN and AAN are very similar in terms of clinical characteristics, such as level of cognitive restraint or medical consequences of the restrictive energy intake (Sawyer et al., 2016; Whitelaw et al., 2018). Based on findings such as these, some argue that the most important body weight measure in any restrictive eating disorder is the total amount of body weight that has been lost, rather than the actual current body weight. Others, however, maintain that there are important differences between AN and AAN and that patients do not benefit from these diagnoses being conceptualized as a single clinical entity (Olivo, Zhukovsky, et al., 2019; Schorr et al., 2019). Importantly, in some healthcare systems the current diagnostic division affects insurance coverage and may result in reduced access to treatment for patient groups relegated to the "atypical" sphere, underscoring the need of examining the validity of this demarcation.

1.2 | Objective

The aim of the present study was to identify research priorities for furthering the understanding of AN and AAN as diagnostic entities. By employing the Delphi methodology (described in detail in the Methods section), we surveyed which research questions that experts within the field—researchers, clinicians, and patient advocates—think need to be asked and answered in order to determine whether AN and AAN are the same or different conditions. We expected that the Delphi panel would generate a substantial number of potentially important research questions regarding AAN and that there would be a certain level of disagreement among the panel members regarding which of these questions are most relevant and should be prioritized, but also that they would be able to reach consensus (i.e., ≥85% agreement as described below) on several of these issues.

2 | METHODS

2.1 | The Delphi method

The Delphi methodology was developed in the 1960s as a method for analyzing and achieving consensus in questions concerning future

needs and developments (Helmer-Hirschberg, 1967). Over the years, the Delphi method has increasingly been used in medical research for charting levels of consensus regarding clinical practice and research priorities (Hasson, Keeney, & McKenna, 2000; Hsu & Sanford, 2007; Keeney, Hasson, & McKenna, 2006; Powell, 2003; Sinha, Smyth, & Williamson, 2011). In a Delphi study, a panel of experts is invited (through mail or, more commonly, email or an online survey platform) to answer one or more predefined questions at issue concerning topical developments in their field of expertise. Recurring themes within their answers are then identified through qualitative content analysis or similar approaches. In a second Delphi round, these synthesized themes are fed back to the panel members, who are then asked to answer whether they agree or disagree on a number of statements derived from the panel's initial answers. This process yields quantitative data on level of agreement within the group. These data are once more fed back to the panel in a third Delphi round, whereby the respondents are also informed about how the group as a whole has answered on the various questions. Respondents are then asked to revisit their position based on this information; if they maintain an opinion that differs markedly from the rest of the panel, they are asked to provide a rationale. Under certain circumstances, a fourth Delphi round may be used in order to achieve a maximum level of consensus, but previous experience shows that this is rarely necessary.

The purpose of this procedure is to explore which questions the panel members are able to reach consensus on (usually defined as at least 85% agreement within the group). Of course, the final outcome may also be a stabilized non-consensual, polarized distribution rather than consensus for some (or all) items. The anonymity within the panel makes it easier for those who hold an opinion that differs from the rest of the group to be heard and to argue for their position. In the field of eating disorders research, the Delphi methodology has previously been employed in order to, for example, explore consensus regarding recommendations on physical exercise for individuals with eating disorders (Noetel, Dawson, Hay, & Touyz, 2017), dieticians' opinions regarding nutritional recommendations (Mittnacht & Bulik, 2015), or steps of care and staging of illness in AN (Buchman, Attia, Dawson, & Steinglass, 2019; Steinglass, Glasofer, Dalack, & Attia, 2020).

Existing guidelines for Delphi studies do not provide a definitive recommended sample size (i.e., how many experts that should be invited to participate in the Delphi expert panel). Based on the topic in question and the scope of the research field, a Delphi panel may include anywhere from a couple of experts to several hundred respondents. A common approach is to invite between 40 and 80 experts, with the aim of securing an expert panel of 20 to 40 persons after dropout.

2.2 | Participants and procedures

In the present study, the Delphi methodology was employed with an aim of three Delphi rounds in total. Panel members were invited

based on having displayed an established interest and expertise in AN, AAN, and/or conceptual diagnostic questions, as demonstrated by either:

1. current Fellow of the Academy for Eating Disorders (AED) status;
2. membership of the Eating Disorders Research Society (EDRS);
3. appointment as a Professor or Associate Professor in the field of eating disorders;
4. spent 10 years or more working in the field;
5. published peer-reviewed journal article(s) and/or book(s) that are focused on AN, AAN, and/or conceptual diagnostic questions;
6. awarded the AED Meehan/Hartley Leadership Award for Public Service and/or Advocacy; or
7. participation as committee members in working groups for the development of DSM-5 or ICD-11 diagnostic criteria.

The above criteria and key publications in the field were used to develop a list of potential panelists, who were then invited by email to participate in the Delphi survey. All panelists were anonymous to each other. In the invitation email, the Delphi procedures were thoroughly explained; for example, the importance of committing to responding to all three rounds was stressed, but it was also explicitly stated that participation was fully voluntary and that the respondents were free to opt out of further participation at any time during the study.

In Delphi round 1, 127 invited panelists were sent a link to an online questionnaire where they were asked to respond in free text to the following question (after having been provided a description of current AN and AAN diagnostic criteria): "In your opinion, which research questions need to be asked and answered in order to determine whether AN and AAN are the same or different conditions?" The answers from the 78 responding panelists (61% of those invited) were processed by employing a qualitative content analysis approach (Hsieh & Shannon, 2005), conducted in parallel by two of the authors (AB and MS). The data were coded and labelled inductively according to a bottom-up principle, and then re-coded in a top-down approach to make sure that the coding scheme was accurate and reliable. The overarching perspective in responses (as per instructions) was comparisons between AN and AAN. Responses were grouped into broad themes and narrow categories which then became the basis for formulating 24 items for Delphi round 2.

In Delphi round 2, the instruction text for the questionnaire was: "In your opinion, which of the following research areas are the most important and timely to focus on in order to determine whether AN and AAN are the same or different conditions? Although many research questions may be important, try to in some sense prioritize the areas that you believe are of primary interest at this point." Respondents were then provided Likert scales ranging from 1 to 6 (i.e., from "Strongly disagree" to "Strongly agree") for the 24 items identified in round 1.

The round 2 questionnaire was completed by 62 respondents (79% of round 1 respondents). In the subsequent round 3, the same items were repeated again with the following instruction: "In this third

and final round, we present you with the 24 items from round 2 together with the distribution of responses from all participants in round 2. In this round, we ask you to revisit your answers using the same scale." 53 respondents (68% of round 1 respondents; 85% of round 2 respondents) completed the round 3 questionnaire. For item 10, >85% consensus was already reached in round 2 and respondents scoring outside the consensus range in round 3 were asked to comment on their choice.

For the present study, consensus was defined as achieving $\geq 85\%$ of panelists choosing the rankings "Strongly agree" or "Agree" (i.e., positive consensus) or "Strongly disagree" or "Disagree" (i.e., negative consensus). Near consensus was declared when between 75% and 85% of panelists showed either agreement or disagreement, and no consensus was declared when less than 75% of panelists showed either agreement or disagreement. No data were

excluded based on attrition; i.e., every contribution to every round was included in the synthesis. However, the final consensus ratios were calculated based on the answers from those panelists who actually completed the Delphi round 3.

Data collection was administered through the online survey platform Confront, which is a standard web platform employed by Karolinska Institutet. Study data were handled in accordance with Swedish Data Security Regulations (GDPR) and the integrity and safety policies at Karolinska Institutet. The study was approved by the Swedish Ethical Review Authority (No. 2019-02518). The study protocol was preregistered at the Open Science Framework (<https://osf.io/efp63/>). Two analyses were subsequently added to this protocol. First, because of attrition between rounds, a sensitivity analysis was performed whereby answers from round 2 were carried forward to an artificial round 3 in which we simulated that the responses from round

TABLE 1 Proportion of respondents in Delphi round 3 who scored 1–2 and 5–6 respectively (where $\geq 85\%$ in either is consensus by the a priori criterion)

Item ("Research should focus on...")	Round 3 (n = 53)		
	% ≤ 2	% ≥ 5	Median
...treatment response in AAN compared to AN	0.0	94.7	6
...medical, neurobiological, and neurological factors in AAN compared to AN	5.7	90.6	6
...the epidemiology and natural course of AAN compared to AN	3.8	86.8	5
...the degree of functional impairment in AAN compared to AN	5.7	83.0	6
...psychiatric comorbidity in AAN compared to AN	0.0	83.0	5
...cognitive and behavioral eating disorder psychopathology in AAN compared to AN	5.7	81.1	6
...psychological features and risk factors such as personality, compulsivity, emotion regulation, impulsivity etc. in AAN compared to AN	5.7	81.1	5
...the longitudinal association between AAN and AN	5.7	77.4	5
...neurocognitive impairment in AAN compared to AN	1.9	77.4	5
...the metabolic factors in AAN compared to AN	3.8	71.7	6
...qualitative research into patient experiences of AAN	11.3	71.7	5
...proportion of weight loss in AAN compared to AN	7.5	69.8	5
...weight suppression in AAN compared to AN	13.2	67.9	5
...whether weight loss is best determined by the individual's own weight loss or in comparison to population norms	7.5	67.9	5
...comparisons with other eating disorders (e.g., BN, EDNOS, BED)	3.8	66.0	5
...speed of weight loss in AAN compared to AN	5.7	66.0	5
...the genetic overlap between AAN and AN	11.3	64.2	5
...familial risk factors in AAN compared to AN	13.2	56.6	5
...whether there are subgroups of AAN corresponding to restrictive vs. binge/purge AN	24.5	54.7	5
...premorbid weight in AAN compared to AN	11.3	47.2	4
...the degree, nature, and consequences of weight stigma in AAN compared to AN	30.2	45.3	4
...the demographic, developmental, and sociocultural risk factors in AAN compared to AN	1.9	43.4	4
...determining the need for, and nature of, a BMI cutoff between AAN and AN	39.6	37.7	3
...what societal benefits or harms may result from determining if AAN is different from AN, and who the stakeholders are (patients, advocacy, health care, research etc.)	5.7	30.2	4

Note: Items are ranked in order of positive responses. Also displaying median value. Items on which consensus and near consensus were reached are marked in grey.

Abbreviations: AAN: atypical anorexia nervosa; AN: anorexia nervosa; BED: binge-eating disorder; BMI: body mass index; BN: bulimia nervosa; EDNOS: eating disorder not otherwise specified.

2 for the non-participants of round 3 remained unchanged. We also investigated how using the median of the responses as cutoff value would affect results.

3 | RESULTS

In Table 1, the 24 items are ranked in order of positive round 3 responses. The distributions of panelist responses are shown in Figure 1. Examples of panelist comments that the items are based on are provided in Table S1. Positive consensus (i.e., $\geq 85\%$ of panelists agree or strongly agree) was reached for three items: (a) medical, neurobiological, and neurological factors in AAN compared to AN; (b) the epidemiology and natural course of AAN compared to AN; and (c) treatment response in AAN compared to AN.

Near positive consensus (i.e., between 75% and 85% positive agreement) was observed for six additional items: (a) the degree of functional impairment in AAN compared to AN; (b) cognitive and behavioral eating disorder psychopathology in AAN compared to AN; (c) psychiatric comorbidity in AAN compared to AN; (d) psychological features and risk factors such as personality, compulsivity, emotion regulation, impulsivity etc. in AAN compared to AN; (e) the longitudinal association between AAN and AN; and (f) neurocognitive impairment in AAN compared to AN.

Negative consensus (i.e., $\geq 85\%$ of the panelists disagree or strongly disagree) or near negative consensus (i.e., between 75% and 85% disagree or strongly disagree) was not seen for any item. The five items that were furthest from positive consensus—indicating that panelists had low consensus (i.e., high variability of responses even in round 3) in determining the differences between AAN and AN—included: (a) demographic, developmental, and sociocultural risk factors in AAN compared to AN; (b) premorbid weight in AAN compared to AN; (c) determining the need for, and nature of, a BMI cutoff between AAN and AN; (d) the degree, nature, and consequences of weight stigma in AAN compared to AN; and (e) what societal benefits or harms may result from determining if AAN is different from AN, and who the stakeholders are (patients, advocacy, health care, research etc.). The item with the largest proportion of negative responses, although still far from negative consensus (39.6%), concerned determining the need for, and nature of, a BMI cutoff between AAN and AN. As seen in Figure 1, the panelist responses tended to be centered around the middle of the scale for some of these items, whereas a more polarized distribution of responses was seen for others—e.g., concerning weight stigma and BMI cutoff—where slight peaks were seen at both ends of the scale.

Since consensus was reached on relatively few items, we also explored the use of median = 6 (meaning that $\geq 50\%$ of responses had a maximum rating of strongly agree) as a cutoff value. This criterion

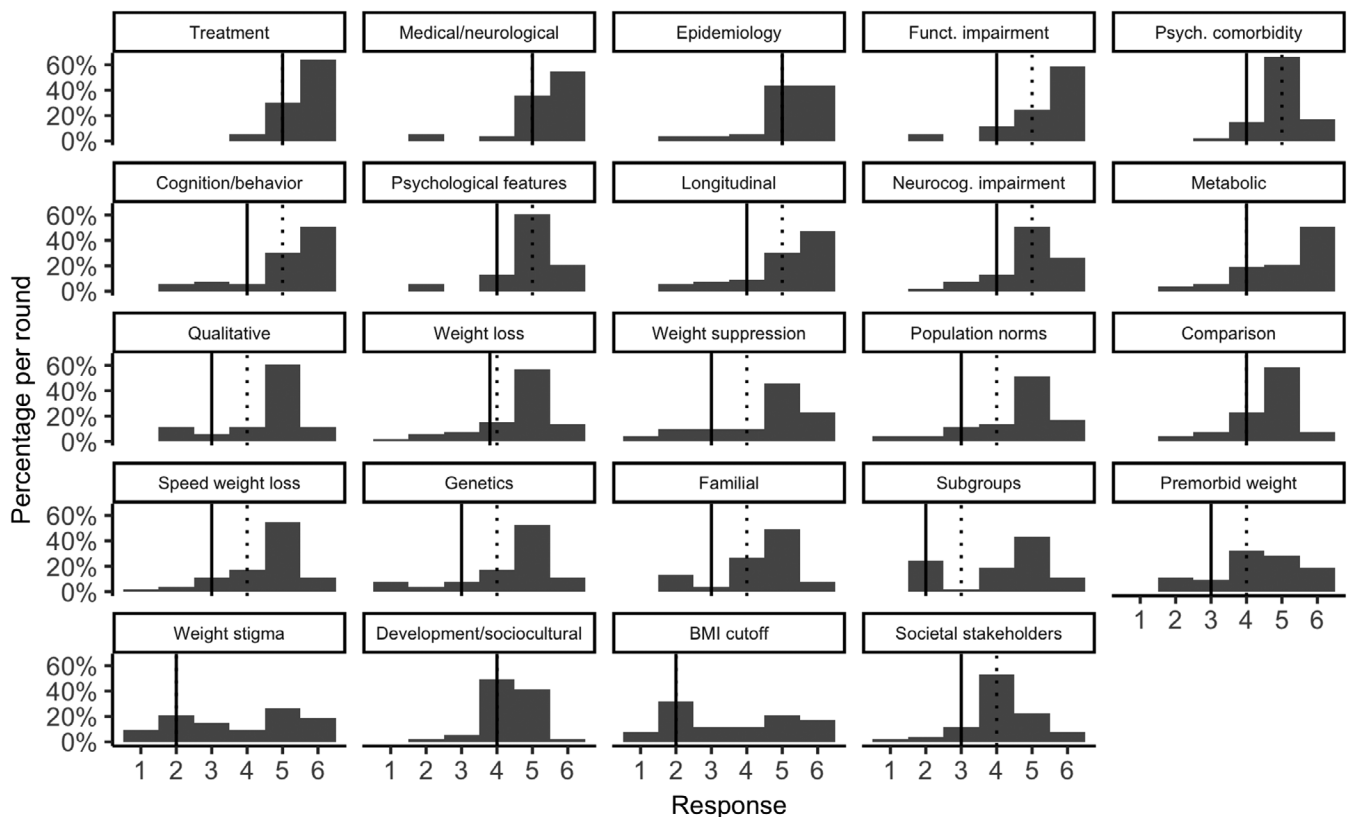


FIGURE 1 Distribution of panelist responses ($n = 53$) in Delphi round 3 (on a Likert scale of 1–6, where 1 is “Strongly disagree” and 6 is “Strongly agree”). Solid vertical lines indicate levels of 85% agreement. Dotted vertical lines indicate levels of 75% agreement. Only when these lines fall on responses 5 or 6 is positive consensus or near consensus achieved

would have resulted in positive consensus for three additional items (some of which already reached near consensus): (a) the degree of functional impairment in AAN compared to AN; (b) cognitive and behavioral eating disorder psychopathology in AAN compared to AN; and (c) the metabolic factors in AAN compared to AN.

The sensitivity analysis performed to assess the impact of attrition did not alter the pattern of the findings.

4 | DISCUSSION

Existing research has provided mixed evidence concerning the diagnostic demarcation of AN and AAN. The present study reveals that there is indeed relatively little consensus on this topic in the field of eating disorders research and advocacy. Out of 24 identified items of interest, 53 Delphi panelists only reached consensus, defined as $\geq 85\%$ agreement, on three items that they considered to be of primary interest for follow-up: medical, neurobiological, and neurological factors in AAN compared to AN; the epidemiology and natural course of AAN compared to AN; and treatment response in AAN compared to AN. These items can be described as fairly traditional in the context of medical diagnostics—they are, for example, all included in the Feighner criteria for determining the validity of diagnoses (Feighner et al., 1972)—and it is therefore not surprising that a high level of agreement was seen.

Negative consensus, that is, $\geq 85\%$ agreement about the relative unimportance of particular items, was not found. However, on five items, less than half of the panelists agreed or strongly agreed, suggesting that these areas had low consensus in determining whether they should be prioritized in determining the differences between AAN and AN. Notably, these included the question of premorbid weight in AAN compared to AN and determining the need for and nature of a BMI cutoff differentiating between AAN and AN—in fact, the latter was the item with the largest proportion of negative responses (39.6%). This is a surprising finding given the central nature of body weight in the demarcation of AN and AAN in DSM definitions as well as in popular debate (Dennett, 2019; Tait, 2015). Likewise, neither positive nor near positive consensus was achieved for the items concerning proportion and speed of weight loss in AAN compared to AN or whether weight loss is best determined by the individual's own weight loss or in comparison to population norms. Here, it should be noted that the question posed to the Delphi panelists was not whether they consider a certain item, such as body weight, as relevant to the diagnosis of AN and/or AAN, but whether they think it would be of primary interest to determine how the potential research priority impacts the demarcation of AN and AAN. Thus, the question of body weight and a weight cutoff was not seen as very important in elucidating the nature of AAN in relation to AN in the current study.

Moreover, neither positive nor near positive consensus was achieved for the item concerning the degree, nature, and consequences of weight stigma in AAN compared to AN—with 30.2% negative responses, this was the second most negatively rated item. This is also somewhat surprising, given the emphasis on potential weight

stigma in popular media reporting as well as in academic papers focusing on how clinicians' preconceptions may delay adequate assessment and treatment for individuals with overweight (Lebow et al., 2015; Sim et al., 2013). Likewise, the item concerning the demographic, developmental, and sociocultural risk factors in AAN compared to AN was also among the lowest ranking research priorities. Since the problem of weight stigma seems to be exacerbated by a low socioeconomic status (Ciciurkaite & Perry, 2018), it is noteworthy that both items relating to these issues received comparably low rankings.

In recent years, considerable advances have been made in the field of the genetics of eating disorders in general (Breithaupt, Hübel, & Bulik, 2018; Bulik, Kleiman, & Yilmaz, 2016) and AN in particular (Hübel et al., 2019; Watson et al., 2019). Even so, in the present study, positive or near positive consensus was not achieved for the item concerning a genetic overlap between AAN and AN, meaning panelists did not emphasize degree of genetic similarity to be a high priority in defining the boundaries between AN and AAN. Considering that the item concerning the demographic, developmental, and socio-cultural risk factors was also among the lowest ranking, research questions concerning the etiology of AAN and AN were clearly not prioritized, although near consensus was achieved for the item concerning the longitudinal association between AAN and AN.

The findings presented here should be considered in light of a number of limitations. Even if a list of predefined eligibility criteria is adhered to, a Delphi study inevitably involves a selective sample of panelists, which may affect the outcome. Naturally, although the rationale behind inviting panelists based on organizational fellowship, professorial status etc. is to ensure scientific merit, this may also introduce demographic and socioeconomic bias. Another potential limitation is the attrition rate with 53 respondents (i.e., 42% of the 127 invited respondents and 68% of the 78 round 1 panelists) remaining in the panel until round 3. Between rounds 2 and 3, however, a good response rate was achieved with 85% of round 2 respondents completing round 3. Bearing in mind the general recommendation of securing an expert panel of at least 20 to 40 persons after dropout, our study adhered to this guideline across all rounds. A sensitivity analysis simulating an artificial round 3 with all round 2 participants did not alter the results.

Furthermore, respondents were explicitly asked to prioritize the areas that they consider to be of primary interest for research at this point. Therefore, a lack of positive consensus regarding a certain item does not readily implicate that it is unimportant, merely that it is not seen as being of primary importance relative to the other suggested items. Indeed, the overall lack of negative consensus means that there were no research questions that were seen as unimportant. Moreover, there are a number of hypothetical response biases inherent to Likert-type survey research, such as order effects bias (i.e., the impact of the order in which questions appear), end aversion (i.e., the tendency to avoid end-of-scale answers; this would make consensus less likely), or an opposite tendency for extreme end-of-scale answers (which would inflate consensus).

Despite efforts to include respondents from eating disorder advocacy organizations in the Delphi panel, only a small number ($n < 5$)

identified themselves as primarily advocates. The majority of participating panelists were included based on their membership in professional organizations for researchers and clinicians in the field. It is important to note that the lines between these categories are in no way absolute. For example, many clinicians and researchers are involved in advocacy work or have lived experience and many patients and family members go on to receive clinical degrees and participate in or conduct research. The question of the nature of AAN in relation to AN has at times been seen as controversial (e.g., the challenges of individuals with AAN in receiving treatment viewed as secondary to weight stigma). Although a survey of research priorities may appear as a benign effort in this regard, it is possible that the act of participation in the study could be experienced as legitimizing a contested mainstream biomedical view for some invited panelists. This is a shortcoming of the study since achieving consensus in the field would ideally also involve participation of non-researcher experts and experts by lived experience. In a broader sense, it has been underscored that the polarization within the field is lamentable, “creating tension that has important ramifications for the ways in which we structure research, intervention, and treatment efforts” (Trainer, Brewis, Hruschka, & Williams, 2015, p. 62) and ultimately hindering progress towards a better understanding of all individuals with restrictive eating disorders. Increased collaboration between the eating disorder and obesity research fields could aid in exploring the intersection between eating disorders and overweight/obesity (Neumark-Sztainer, 2009).

5 | CONCLUSION

In sum, the findings from the present Delphi study reveal relatively little consensus within the field of eating disorders research and advocacy on the topic of which research questions need to be answered in order to further elucidate the relationship between AAN and AN. Out of a list of 24 research topics identified by the Delphi panel, consensus (defined as $\geq 85\%$ agreement) were reached on three items—medical, neurobiological, and neurological factors in AAN compared to AN; the epidemiology and natural course of AAN compared to AN; and treatment response in AAN compared to AN—which can be described as fairly traditional in a context of medical diagnostics.

Naturally, prioritizing research in these high-consensus areas does not imply that low-consensus areas, such as a weight cutoff, should not be explored. For example, in the high-priority treatment response area, studies could focus on non-weight-based outcomes while still acknowledging that changes in body weight do represent a central component of recovery in some eating disorders.

A reason behind the overall lack of consensus could be that the definition and use of the AAN category vary in research and clinical practice (Forman et al., 2014; Forney et al., 2017; Silén et al., 2015). It has been noted that certain internal inconsistencies in the AAN and AN diagnostic criteria in the DSM-5 exist (Birgegård, Groß, de Man Lapidoth, & Norring, 2013), which may give rise to different

interpretations and misunderstandings. Therefore, a clearer diagnostic operationalization of AAN may be needed. At this point, there appears to be little consensual support for the concept of AAN being adequately researched in order for it to be lifted out of the Other Specified Feeding and Eating Disorders category in the DSM-5 to become a separate diagnosis in its own right. Based on the present study, in order to achieve further diagnostic clarity, research on the demarcation between AAN and AN should focus on medical and neurobiological factors, on their respective epidemiological characteristics and natural course, and on how treatment response differs in AAN compared to AN.

ACKNOWLEDGEMENTS

Dr. Bulik acknowledges funding from the Swedish Research Council (Vetenskapsrådet) grant no. 538-2013-8864 and National Institute of Mental Health (R01MH120170 and R01MH119084).

CONFLICTS OF INTEREST

Dr. Bulik reports: Shire (grant recipient, Scientific Advisory Board member); Idorsia (consultant); Pearson (author, royalty recipient).

DATA AVAILABILITY STATEMENT

Anonymized study data are available from the corresponding author on reasonable request.

ORCID

Mattias Strand  <https://orcid.org/0000-0003-4799-0737>

Johan Zvrskovec  <https://orcid.org/0000-0002-8862-0874>

Christopher Hübel  <https://orcid.org/0000-0002-1267-8287>

Andreas Birgegård  <https://orcid.org/0000-0003-1220-9680>

REFERENCES

- American Psychiatric Association. (1994). *DSM-IV: Diagnostic and Statistical Manual of Mental Illness* (4th ed.). Washington, D.C: American Psychiatry Publishing.
- American Psychiatric Association. (2013). *DSM-5: Diagnostic and Statistical Manual of Mental Illness* (5th ed.). Washington, DC: American Psychiatry Publishing.
- Birgegård, A., Groß, G., de Man Lapidoth, J., & Norring, C. (2013). DSM-5: Problems and suggestions. *Journal of Eating Disorders*, 1(1), 8. <https://doi.org/10.1186/2050-2974-1-8>
- Breithaupt, L., Hübel, C., & Bulik, C. M. (2018). Updates on genome-wide association findings in eating disorders and future application to precision medicine. *Current Neuropharmacology*, 16, 1102–1110. <https://doi.org/10.2174/1570159X16666180222163450>
- Buchman, S., Attia, E., Dawson, L., & Steinglass, J. E. (2019). Steps of care for adolescents with anorexia nervosa—A Delphi study. *International Journal of Eating Disorders*, 52(7), 777–785. <https://doi.org/10.1002/eat.23088>
- Bulik, C. M., Kleiman, S. C., & Yilmaz, Z. (2016). Genetic epidemiology of eating disorders. *Current Opinion in Psychiatry*, 29(6), 383–388. <https://doi.org/10.1097/YCO.0000000000000275>
- Ciciurkaite, G., & Perry, B. L. (2018). Body weight, perceived weight stigma and mental health among women at the intersection of race/ethnicity and socioeconomic status: Insights from the modified labelling approach. *Sociology of Health & Illness*, 40(1), 18–37. <https://doi.org/10.1111/1467-9566.12619>

- Dennett, C. (2019, February 20). Anorexia knows no body type - and thinking otherwise can be a barrier to treatment. *Seattle Times*.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A. J., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26(1), 57–63. Retrieved from: <https://doi.org/10.1001/archpsyc.1972.01750190059011>
- Forman, S. F., McKenzie, N., Hehn, R., Monge, M. C., Kapphahn, C. J., Mammel, K. A., ... Woods, E. R. (2014). Predictors of outcome at 1 year in adolescents with DSM-5 restrictive eating disorders: Report of the National Eating Disorders Quality Improvement Collaborative. *Journal of Adolescent Health*, 55(6), 750–756. <https://doi.org/10.1016/j.jadohealth.2014.06.014>
- Forney, K. J., Brown, T. A., Holland-Carter, L. A., Kennedy, G. A., & Keel, P. K. (2017). Defining “significant weight loss” in atypical anorexia nervosa. *International Journal of Eating Disorders*, 50(8), 952–962. <https://doi.org/10.1002/eat.22717>
- Gorwood, P., Duriez, P., Lengvenyte, A., Guillaume, S., & Criquillon, S. (2019). Clinical insight in anorexia nervosa: Associated and predictive factors. *Psychiatry Research*, 281, 112561. <https://doi.org/10.1016/j.psychres.2019.112561>
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>
- Helmer-Hirschberg, O. (1967). Analysis of the Future: The Delphi Method. <https://www.rand.org/pubs/papers/P3558.html>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Hsu, C.-C., & Sanford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1–8.
- Hübel, C., Gaspar, H. A., Coleman, J. R. I., Finucane, H., Purves, K. L., Hanscombe, K. B., ... Breen, G. (2019). Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 180(6), 428–438. <https://doi.org/10.1002/ajmg.b.32709>
- Hughes, E. K., Le Grange, D., Court, A., & Sawyer, S. M. (2017). A case series of family-based treatment for adolescents with atypical anorexia nervosa. *International Journal of Eating Disorders*, 50(4), 424–432. <https://doi.org/10.1002/eat.22662>
- Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, 53(2), 205–212. <https://doi.org/10.1111/j.1365-2648.2006.03716.x>
- Konstantakopoulos, G., Tchaturia, K., Surguladze, S. A., & David, A. S. (2011). Insight in eating disorders: Clinical and cognitive correlates. *Psychological Medicine*, 41(9), 1951–1961. <https://doi.org/10.1017/S0033291710002539>
- Lebow, J., Sim, L. A., & Kransdorf, L. N. (2015). Prevalence of a history of overweight and obesity in adolescents with restrictive eating disorders. *Journal of Adolescent Health*, 56(1), 19–24. <https://doi.org/10.1016/j.jadohealth.2014.06.005>
- Lebow, J., Sim, L., Crosby, R. D., Goldschmidt, A. B., Le Grange, D., & Accurso, E. C. (2019). Weight gain trajectories during outpatient family-based treatment for adolescents with anorexia nervosa. *International Journal of Eating Disorders*, 52(1), 88–94. <https://doi.org/10.1002/eat.23000>
- Mittnacht, A. M., & Bulik, C. M. (2015). Best nutrition counseling practices for the treatment of anorexia nervosa: A Delphi study. *International Journal of Eating Disorders*, 48(1), 111–122. <https://doi.org/10.1002/eat.22319>
- Nagata, J. M., Garber, A. K., & Buckelew, S. M. (2018). Weight restoration in atypical anorexia nervosa: A clinical conundrum. *International Journal of Eating Disorders*, 51(11), 1290–1293. <https://doi.org/10.1002/eat.22953>
- Neumark-Sztainer, D. (2009). The interface between the eating disorders and obesity fields: Moving toward a model of shared knowledge and collaboration. *Eating and Weight Disorders*, 14(1), 51–58. <https://doi.org/10.1007/BF03327795>
- Neumark-Sztainer, D. (2015). Higher weight status and restrictive eating disorders: An overlooked concern. *Journal of Adolescent Health*, 56(1), 1–2. <https://doi.org/10.1016/j.jadohealth.2014.10.261>
- Noetel, M., Dawson, L., Hay, P., & Touyz, S. (2017). The assessment and treatment of unhealthy exercise in adolescents with anorexia nervosa: A Delphi study to synthesize clinical knowledge. *International Journal of Eating Disorders*, 50(4), 378–388. <https://doi.org/10.1002/eat.22657>
- Olivo, G., Solstrand Dahlberg, L., Wiemerslage, L., Swenne, I., Zhukovsky, C., Salonen-Ros, H., ... Schiöth, H. B. (2018). Atypical anorexia nervosa is not related to brain structural changes in newly diagnosed adolescent patients. *International Journal of Eating Disorders*, 51(1), 39–45. <https://doi.org/10.1002/eat.22805>
- Olivo, G., Swenne, I., Zhukovsky, C., Tuunainen, A.-K., Saaid, A., Salonen-Ros, H., ... Schiöth, H. B. (2019). Preserved white matter microstructure in adolescent patients with atypical anorexia nervosa. *International Journal of Eating Disorders*, 52(2), 166–174. <https://doi.org/10.1002/eat.23012>
- Olivo, G., Zhukovsky, C., Salonen-Ros, H., Larsson, E.-M., Brooks, S., & Schiöth, H. B. (2019). Functional connectivity underlying hedonic response to food in female adolescents with atypical AN: The role of somatosensory and salience networks. *Translational Psychiatry*, 9(1), 276. <https://doi.org/10.1038/s41398-019-0617-0>
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4), 376–382. <https://doi.org/10.1046/j.1365-2648.2003.02537.x>
- Sawyer, S. M., Whitelaw, M., Le Grange, D., Yeo, M., & Hughes, E. K. (2016). Physical and psychological morbidity in adolescents with atypical anorexia nervosa. *Pediatrics*, 137(4), e20154080. <https://doi.org/10.1542/peds.2015-4080>
- Schorr, M., Drabkin, A., Rothman, M. S., Meenaghan, E., Lashen, G. T., Mascolo, M., ... Miller, K. K. (2019). Bone mineral density and estimated hip strength in men with anorexia nervosa, atypical anorexia nervosa and avoidant/restrictive food intake disorder. *Clinical Endocrinology*, 90(6), 789–797. <https://doi.org/10.1111/cen.13960>
- Silén, Y., Raevuori, A., Jüriloo, E., Tainio, V.-M., Marttunen, M., & Keski-Rahkonen, A. (2015). Typical versus atypical anorexia nervosa among adolescents: Clinical characteristics and implications for ICD-11. *European Eating Disorders Review*, 23(5), 345–351. <https://doi.org/10.1002/erv.2370>
- Sim, L. A., Lebow, J., & Billings, M. (2013). Eating disorders in adolescents with a history of obesity. *Pediatrics*, 132(4), e1026–e1030. <https://doi.org/10.1542/peds.2012-3940>
- Sinha, I. P., Smyth, R. L., & Williamson, P. R. (2011). Using the Delphi technique to determine which outcomes to measure in clinical trials: Recommendations for the future based on a systematic review of existing studies. *PLoS Medicine*, 8(1), 1–5. <https://doi.org/10.1371/journal.pmed.1000393>
- Steinglass, J. E., Glasofer, D. R., Dalack, M., & Attia, E. (2020). Between wellness, relapse, and remission: Stages of illness in anorexia nervosa. *International Journal of Eating Disorders*, 53, 1088–1096. <https://doi.org/10.1002/eat.23237>
- Tait, A. (2015). When you're both overweight and anorexic. Retrieved March 1, 2019, from Broadly website: https://broadly.vice.com/en_us/article/gvz7m7/when-youre-both-overweight-and-anorexic
- Trainer, S., Brewis, A., Hruschka, D., & Williams, D. (2015). Translating obesity: Navigating the front lines of the “war on fat.”. *American Journal of Human Biology*, 27(1), 61–68. <https://doi.org/10.1002/ajhb.22623>
- Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R. I., Gaspar, H. A., ... Consortium, E. D. W. G. of the P. G. (2019). Genome-wide association study identifies eight risk loci and implicates

- metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, 51(8), 1207–1214. <https://doi.org/10.1038/s41588-019-0439-2>
- Whitelaw, M., Gilbertson, H., Lee, K. J., & Sawyer, S. M. (2014). Restrictive eating disorders among adolescent inpatients. *Pediatrics*, 134(3), e758–e764. <https://doi.org/10.1542/peds.2014-0070>
- Whitelaw, M., Lee, K. J., Gilbertson, H., & Sawyer, S. M. (2018). Predictors of complications in anorexia nervosa and atypical anorexia nervosa: Degree of underweight or extent and recency of weight loss? *Journal of Adolescent Health*, 63(6), 717–723. <https://doi.org/10.1016/j.jadohealth.2018.08.019>
- World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems, 10th revision (2nd ed.)*. Geneva, Switzerland: World Health Organization.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Strand M, Zvrskovec J, Hübel C, Peat CM, Bulik CM, Birgegård A. Identifying research priorities for the study of atypical anorexia nervosa: A Delphi study. *Int J Eat Disord*. 2020;53:1729–1738. <https://doi.org/10.1002/eat.23358>