ORIGINAL ARTICLE

International Journal of
EATING DISORDERS WILEY

Development of transdiagnostic clinical risk prediction models for 12-month onset and course of eating disorders among adolescents in the community

Deborah Mitchison PhD [1,2] | Shirley B. Wang AM [3] | Tracey Wade PhD [4] |
Ann F. Haynos PhD [5] | Kay Bussey PhD [2] | Nora Trompeter PhD [2] |
Alexandra Lonergan PhD [2] | Jack Tame MNeuroPsych [1,2] | Phillipa Hay MD [1,6]

[1]Eating Disorder and Body Image Network, Translational Health Research Institute, School of Medicine, Western Sydney University, Penrith, New South Wales, Australia

[2]Centre for Emotional Health, School of Psychological Sciences, Macquarie University, North Ryde, New South Wales, Australia

[3]Department of Psychology, Harvard University, Cambridge, Massachusetts, USA

[4]Flinders Institute for Mental Health and Wellbeing, Flinders University, Adelaide, South Australia, Australia

[5]Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA

[6]Mental Health Services, South West Sydney Local Health District, Campbelltown, Australia

Correspondence
Deborah Mitchison, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia.
Email: deborah.mitchison@westernsydney.edu.au

Abstract

Objective: To develop and internally validate risk prediction models for adolescent onset and persistence of eating disorders.

Methods: $N = 963$ Australian adolescents (11–19 years) in the EveryBODY Study cohort completed online surveys in 2018 and 2019. Models were built to predict 12-month risk of (1) onset, and (2) persistence of a DSM-5 eating disorder.

Results: *Onset Model.* Of the $n = 687$ adolescents without an eating disorder at baseline, 16.9% were identified with an eating disorder after 12 months. The prediction model was based on evidence-based risk factors for eating disorder onset available within the dataset (sex, body mass index percentile, strict weight loss dieting, history of bullying, psychological distress, weight/shape concerns). This model showed fair discriminative performance (mean AUC = .75). The most important factors were psychological distress, weight and shape concerns, and female sex. *Diagnostic Persistence Model.* Of the $n = 276$ adolescents with an eating disorder at baseline, 74.6% were identified as continuing to meet criteria for an eating disorder after 12 months. The prediction model for diagnostic persistence was based on available evidence-based risk factors for eating disorder persistence (purging, distress, social impairment). This model showed poor discriminative performance (mean AUC = .65). The most important factors were psychological distress and self-induced vomiting for weight control.

Discussion: We found preliminary evidence for the utility of a parsimonious model for 12-month onset of an eating disorder among adolescents in the community. Future research should include additional evidence-based risk factors and validate models beyond the original sample.

Public Significance: This study demonstrated the feasibility of developing parsimonious and accurate models for the prediction of future onset of an eating disorder among adolescents. The most important predictors in this model included psychological distress and weight and shape concerns. This study has laid the ground work for future research to build and test more accurate prediction models in diverse samples,

prior to translation into a clinical tool for use in real world settings to aid decisions about referral to early intervention.

1 | INTRODUCTION

Despite eating disorders (EDs) affecting 21.0%–36.9% of adolescents, access to specialized treatment is extremely low (Hammerle et al., 2016; Micali et al., 2015; Mitchison et al., 2019), primarily due to a lack of detection and referral. This is true even for the better recognized conditions such as anorexia nervosa and bulimia nervosa, in which only 10%–27% of those affected are estimated to have accessed ED specific healthcare (Fatt et al., 2019; Swanson et al., 2011). Established screening instruments and methods have good sensitivity and specificity for detecting a *current* ED (Hill et al., 2010; Maguen et al., 2018; Solmi et al., 2015); however, their ability to predict *future* outcomes, such as onset and persistence of an ED, is unclear. Key risk factors, such as weight and shape concerns and other sociodemographic (e.g., female gender) and clinical features (e.g., strict dieting), have well-known prognostic value in predicting variance in ED onset (Bakalar et al., 2015; Culbert et al., 2015; Day, Bussey, Trompeter, & Mitchison, 2021; de Portela Santana et al., 2012; Ghaderi, 2001; Glashouwer et al., 2019; Jacobi et al., 2004; Keel & Forney, 2013; Lie et al., 2019; Mazzeo & Bulik, 2009; N. Micali, 2005; Stice, 2002; Vall & Wade, 2015). However, to date these have not been integrated into algorithmically-driven screening procedures that would optimize precise prediction of these key outcomes, and facilitate early intervention.

In other areas of psychiatry, prognostic model research (also known as "clinical risk prediction modeling") has been applied to predict outcomes such as likelihood to develop a new onset of psychosis over a 5-year period in the UK National Health Service with good accuracy (Fusar-Poli et al., 2018, 2019; for a summary of model performance indicators and their interpretation, see Supporting Information File 1). Prognostic modeling is a specific methodology (Steyerberg et al., 2013) with recommended reporting guidelines (the TRIPOD Statement [Collins et al., 2015]) that makes use of the evidence base from risk factor and early modeling (Fairburn et al., 2005) research to build models that provide an overall risk score for each individual. The translational outcome of such research, online risk calculators that can be used by clinicians and consumers to guide help-seeking, allocation to early intervention, and treatment decisions (as part of evidence-based practice alongside clinician expertise and client preferences), would be a considerable advancement to the ED field. The value to these models is capturing people before they develop an ED, so that we can make better use of early interventions, which have been shown to result in better outcomes than treatment as usual (Richards et al., 2022). Another ED outcome for which accurate prediction would be of high value is the likelihood of persistence of an ED among those already affected in the community who are yet to receive treatment. For instance, predicting likelihood of ED persistence could assist in the primary care setting regarding decisions about the level of intervention to recommend following screening. This could matched to level of risk for persistence from a watch-and-wait approach, to guided-self-help, to referral to ED specific treatment—ultimately improving the efficiency of EDs healthcare.

The development of prognostic models using methodologies that align with the TRIPOD Statement (Collins et al., 2015) has only recently commenced for outcomes related to EDs. These models include the prediction of future persistence of ED diagnosis based on a wide range of clinical and sociodemographic factors (Haynos, Wang et al., 2021), current ED diagnosis based on internet activity (Sadeh-Sharvit et al., 2020), and response to ED treatment based on pretreatment clinical variables (Espel-Huynh et al., 2021). In a study by Haynos, Wang et al. (2021), risk prediction models were developed for ED persistence and presence of specific symptoms after 1 and 2 years using a pre-existing dataset of 320 females with an established ED. The models predicting the future presence of specific ED symptoms showed "fair" to "good" accuracy (mean "area under the receiver operating characteristic curve"—AUC range: 0.71–0.89) and better than the models predicting persistence of EDs (mean AUC range: 0.61–0.62, "poor" accuracy). Models built utilizing a machine learning method (ML; elastic net) also outperformed models built using traditional logistic regression. It is important to note here that the extent to which these models performed well in part depended on the large number of risk factors (33 in total) included, as more information in models improves precision. However, it is also true that more complex models pose a challenge for eventual translation into real world settings as they involve greater administrative burden (Steyerberg et al., 2013). Nevertheless, these first studies demonstrate promise for the application of ML methods to the development of prediction models for ED outcomes, including using self-report data.

The aims of the current study were to develop the first prognostic models for ED onset (Aim 1) and persistence (Aim 2) in the community. Data were from adolescents (the peak age of ED onset) in the Every-BODY cohort who completed self-report surveys over 12 months. A previous study with this cohort observed 18.2% new onset cases (inclusive of subthreshold disorders) over a 12 month period (Prnjak et al., 2021). With the goal to develop models that may eventually be translated into clinical practice, only the best available predictors were selected for the models, based on evidence from systematic reviews and meta-analyses. While a body of studies suggest ML is comparable to logistic regression for building risk models (e.g., Espel-Huynh et al., 2021), emerging evidence suggests ML is superior when models are more complex (e.g., more predictors; Haynos, Wang et al., 2021) or when accuracy of prediction is a priority (Wang, 2021). Hence, we tested both approaches in order to further inform methodological decision-making in this area. The study was conducted in alignment

with the reporting guidelines outlined in the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) Statement (Collins et al., 2015).

2 | MATERIALS AND METHODS

2.1 | Data source

The data were from Wave 1 and Wave 2 (i.e., the first and second follow-up surveys) of the EveryBODY Cohort, a representative population-based self-report study of EDs among Australian adolescents. These waves were chosen as they were the most inclusive of the predictors and outcomes of interest.

2.2 | Study population

Participants were those who attended one of eight secondary schools in New South Wales, Australia, and participated in the online surveys in both Wave 1 (2018) and Wave 2 (2019) of the EveryBODY study. Four public and four private schools were included, and the sample were demographically representative of adolescents in the State of New South Wales. Surveys were administered at schools between May and December based on the school's preference for roll-out. Details of the recruitment procedures and the EveryBODY cohort have been published previously (Prnjak et al., 2021; Trompeter et al., 2018). The retention rate from Wave 1 to Wave 2 was 61.6%. When comparing those who did versus did not participate in Wave 2, there was no difference in gender distribution, however, Wave 2 completers were slightly younger (14.62 vs. 15.08 years on average; $p < .001$, $\eta_p^2 = 0.021$), had a slightly lower BMI percentile (50.92 vs. 53.77 on average; $p = 0.014$, $\eta_p^2 = 0.002$), and were slightly more likely to have been born in Australia (84.8% vs. 80.4%; $p = 0.45$, V = 0.056). These effect sizes were all small, and largely explained by the fact that students who were in their final year of school at Wave 1 were very difficult to recruit for the Wave 2 survey when they were no longer in school. For Aim 1, we developed models with data from participants who did not meet criteria for an ED at Wave 1 and who had complete data for diagnosis at Wave 2 (N = 687). For Aim 2, we developed models with data from participants who did meet criteria for an ED at Wave 1 and had complete data for diagnosis at Wave 2 available (N = 276). Approval for the study was granted by the Macquarie University Human Research Ethics Committee and the New South Wales Department of Education. All participants assented to the study and their parents/guardians provided passive informed consent.

2.3 | Study measures

2.3.1 | Outcomes

The two outcomes were (1) ED onset (meeting criteria for a probable ED at Wave 2, but not Wave 1), and (2) ED persistence (meeting criteria for a probable ED at Wave 1 and Wave 2). As described previously, diagnoses were determined based on self-report responses to a range of standardized measures (including the Eating Disorder Examination Questionnaire, Fairburn & Beglin, 2008; K10 Psychological Distress Scale, Kessler et al., 2002; and Pediatric Quality of Life Scale, Varni et al., 2003) and specific questions designed by the investigators to assess specific diagnostic criteria (see Supporting Information File 2 for full information; Mitchison et al., 2019).

2.3.2 | Predictors

We selected predictors based on replicated evidence of their prognostic value for ED onset and persistence, respectively, as summarized in meta-analyses, systematic and expert narrative reviews (Bakalar et al., 2015; Culbert et al., 2015; Day, Bussey, Trompeter, & Mitchison, 2021; de Portela Santana et al., 2012; Ghaderi, 2001; Glashouwer et al., 2019; Jacobi et al., 2004; Keel & Forney, 2013; Lie et al., 2019; Mazzeo & Bulik, 2009; N. Micali, 2005; Stice, 2002; Vall & Wade, 2015). As recommended, continuous variables were not dichotomized (Fusar-Poli et al., 2018; Fusar-Poli et al., 2019).

Predictors for the models in Aim 1 (ED onset) included binary sex (Weissman, 2019), body mass index (kg/m^2; BMI) percentile adjusted for child age and sex (Bakalar et al., 2015; Stice, 2002), days of weight loss dieting over the preceding one month (Bakalar et al., 2015; Stice, 2002), history of being a victim of bullying at school (Bakalar et al., 2015; Day, Bussey, Trompeter, & Mitchison, 2021; Lie et al., 2019), frequency of psychological distress over the preceding one month (Bakalar et al., 2015; Culbert et al., 2015; Keel & Forney, 2013; Stice, 2002), and severity of weight/shape concerns over the preceding one month (Bakalar et al., 2015; Glashouwer et al., 2019; Keel & Forney, 2013). Predictors for the models in Aim 2 (ED persistence) came primarily from two reviews based on anorexia nervosa and bulimia nervosa (H.-C. Steinhausen, 2002; H. Steinhausen & Weber, 2009) and included frequency of purging, psychological distress, and perceived social functioning over the past month.

BMI percentile was based on self-reported height and weight and adjusted for the adolescent's sex and age. Self-report of height and weight data in adolescents has been recommended when direct measurement is impractical, because even though there is a slight tendency for underestimation of weight, self-report and directly measured data remain highly correlated (Kee et al., 2017; Sherry et al., 2007).

Weight loss dieting was measured using an author-derived question, "*Over the past 28 days (4 weeks) how many days have you been on a very strict weight loss diet?*" and participants indicated their response by free text. Evidence of the convergent validity of this item include the significant association between scores on the item and scores on measures of other disordered eating behaviors (Aouad et al., 2019); (Pursey et al., 2020), and weight-related bullying (Day, Bussey, Trompeter, Hay, et al., 2021).

History of bullying was assessed with the author-derived question "*Have you ever been bullied at school?*" and participants indicated their response by selecting "no" or "yes." This is a similar question to that

employed in other adolescent population-based surveys. However, in the present study the question was prefaced with a standard definition of bullying, as recommended by bullying research experts to improve construct validity (Day, Bussey, Trompeter, Hay, et al., 2021; Day, Bussey, Trompeter, & Mitchison, 2021).

Psychological distress was measured using the *Kessler Psychological Distress Scale* (K-10) (Kessler et al., 2002), a well-validated questionnaire frequently used in population studies, which assesses symptoms of anxiety and depression over the last 4 weeks on 10 Likert-type questions. Scores range from 10 to 50, with higher scores indicative of greater distress. Internal consistency in the Aim 1 sample was $\omega = 0.89$ and in the Aim 2 sample was $\omega = 0.93$.

Weight and shape concerns were assessed using the combined items of the *Weight Concern* and *Shape Concern* subscales of the *ED Examination Questionnaire (EDE-Q)* (Christopher G Fairburn & Beglin, 2008). This combined scale has been well-validated in Australian adolescents (J. Mond et al., 2014) and consists of 12 Likert-type items assessing dissatisfaction, preoccupation and overvaluation with weight/shape over the past 4 weeks. Scores are averaged across the items, ranging from 0 to 6, with higher scores indicative of greater weight/shape concerns. Internal consistency in the Aim 1 sample was $\omega = 0.92$.

Purging was assessed using two separate items in the *EDE-Q* (Christopher G Fairburn & Beglin, 2008) that asked participants to indicate the frequency of self-induced vomiting and laxative use for weight control purposes over the previous 28 days. These items have been used previously in Australian adult (Mitchison et al., 2012, 2014) and adolescent (Fatt et al., 2019; Trompeter et al., 2020) population studies.

Social functioning was assessed using the *Social Functioning* subscale of the *Pediatric Quality of Life Scale (PedsQL) short form* (Varni et al., 2003). The well-validated subscale consists of 3 Likert-type items that assess frequency of social functioning impairment over the previous 28 days. Scores range from 0 to 100, with higher scores indicating lower levels of social functioning impairment. Internal consistency in the Aim 2 sample was $\omega = 0.85$.

2.4 | Statistical analyses

Preliminary analyses included computation of descriptive statistics for each sample. Univariate logistic regression analyses were conducted to examine the unadjusted relationships between each candidate predictor and the outcomes and are presented in Supporting Information File 3.

2.4.1 | Modeling approach

Analyses were performed in R version 3.6.1 via caret (Kuhn, 2021) and glmnet (Friedman et al., 2010) packages. To predict EDs onset and persistence, we conducted elastic net regularized logistic regressions. We chose the elastic net algorithm given its well-established accuracy and robustness, its ability to maintain clinical interpretability compared to less transparent ML algorithms (e.g., random forests,

neural networks), and its validity in prediction ED risk models in a prior paper (Haynos, Wang et al., 2021). For each outcome (i.e., in Aim 1 and Aim 2), we also compared predictive accuracy of elastic net models with nonregularized logistic regression models, as has been done previously (Haynos, Wang et al., 2021). As per the additional aim of this study, the ML versus logistic regression models were compared based on the model performance metrics listed below (see Supporting Information File for further information on interpretation of these metrics).

All models included participants with complete data available for outcomes (i.e., ED diagnostic information at Wave 2). Regarding missing data for predictors, there was 5.28% missingness for Aim 1 (predicting probable ED onset) and 0.87% missingness for Aim 2 (predicting persistence of probable ED). We used K-nearest neighbor imputation during data preprocessing for all models. Given that, our data were imbalanced (with fewer individuals meeting criteria for a probable ED at Wave 2 than those not meeting criteria for a probable ED), we used upsampling to improve the balance across classes in all models (Kuhn & Johnson, 2013).

To obtain metrics of predictive accuracy, we followed recommendations (Kuhn & Johnson, 2013), and used 10-fold cross-validation with three repetitions to select the optimal λ (shrinkage) and α (mixing) parameters for each elastic net model; we also used the same cross-validation procedure for logistic regression models. Repeating the training and testing process in this way can provide more reasonable estimates of model performance for future datasets than splitting a sample into a single training and testing set, particularly for smaller samples. Finally, we evaluated variable importance for all models with the varImp() function in caret.

2.4.2 | Model performance

A standard metric for examining a model's performance is the area under the receiver operating characteristic curve (AUC, which measures area under a curve with 1—specificity on the x-axis and sensitivity on the y-axis). An AUC of 0.5 indicates chance-level predictive accuracy and an AUC of 1.0 indicates perfect classification (50–0.59 = extremely poor; 0.60–0.69 = poor; 0.70–0.79 = fair; 0.80–0.89 = good; 0.90–1.00 = excellent). We also evaluated several other classification metrics, including the average cross-validation estimates of: area under the precision-recall curve (AUPRC), accuracy, positive predictive value (PPV), sensitivity, specificity, and Brier score For more information on each of these metrics and their interpretation, see Supporting Information File 1.

3 | RESULTS

3.1 | Sample characteristics

Demographic and clinical characteristics of the subsamples included in this study are presented in Table 1. A description of the demographic

TABLE 1 Demographic characteristics and scores on predictors at Wave 1 for participants used for Aim 1 (predicting eating disorder onset) and Aim 2 (predicting eating disorder persistence).

	Eating disorder onset sample (Aim 1), N = 687	Eating disorder persistence sample (Aim 2), N = 276
	n (%)	
Onset of eating disorder after 1 year	16.9	n/a
Persistence of eating disorder at 1 year	n/a	74.6
Sex[a]		
Male	401 (58.4)	54 (24.5)
Female	286 (41.6)	166 (75.5)
Country of birth		
Australia	577 (84.0)	232 (84.1)
Other	110 (16.0)	44 (15.9)
Bullying history	257 (41.9)	149 (64.5)
	Mean (SD)	
Age	14.4 (3.6)	15.0 (1.5)
BMI percentile	47.2 (30.3)	62.0 (31.0)
Strict weight loss dieting days past month	0.5 (2.4)	4.3 (7.8)
Psychological distress	14.5 (5.8)	30.5 (10.3)
Weight and shape concerns	0.8 (1.0)	3.7 (1.7)
Purging episodes past month		
Self-induced vomiting	0.1 (0.7)	1.3 (4.4)
Laxatives	0.1 (1.0)	0.8 (4.3)
Social functioning	91.0 (15.5)	71.3 (26.6)

[a]Includes responses to "what is your gender: male, female" for one school who did not want to include separate questions on biological sex and gender.

characteristics of the full EveryBODY cohort, which closely approximate the general adolescent population in New South Wales, have been published elsewhere (Mitchison et al., 2022).

3.1.1 | Aim 1: ED onset

There were n = 687 participants without an ED at Wave 1, comprising the sample for Aim 1. Just over half of this sample was male, mostly born in Australia (84%), with an average age of 14.4 years, and with an average BMI percentile within the Center for Disease Control (CDC) definition of "healthy." Wave 1 disordered eating was low in the sample with only 1.2% scoring above the cut-off for extreme weight and shape concerns (scoring ≥4 on the EDE-Q combined Weight and Shape Concerns scale; Mond et al., 2014),

and at-least-weekly fasting, objective binge eating, self-induced vomiting, and laxative use reported by only 4.5%, 5.4%, 0.6% and 0.9% of participants, respectively. At Wave 2, 116 (16.9%, 95% CI: 14.3–19.9%) of these participants had developed a probable ED. This rate was expected based on a previous study (Prnjak et al., 2021) using earlier waves of this cohort, and when considering the high rates of prevalence (19%–37%) for the full spectrum of EDs globally (Hammerle et al., 2016; Micali et al., 2015).

3.1.2 | Aim 2: ED persistence

There were n = 276 participants with a probable ED at Wave 1, comprising the sample for Aim 2. Three quarters of this sample was female, mostly born in Australia (84%), with an average age of 15 years, and an average BMI percentile that was within the CDC definition of "healthy." The probable EDs that were observed in this sample included n = 11 with anorexia nervosa, n = 66 with bulimia nervosa, n = 21 with binge ED, n = 39 with atypical anorexia nervosa, n = 35 with subthreshold bulimia nervosa, n = 11 with subthreshold binge ED, n = 37 with purging disorder, n = 81 with night eating syndrome, and n = 12 with unspecified feeding/ED. Only 13.8% reported having ever had seen a health professional for a problem with their body image. At Wave 2, when assessed 1 year later, 206 (74.6%, 95% CI: 69.2–79.4%) participants in this sample continued to meet criteria for a probable ED. Access to treatment for a body image problem remained low, increased only marginally to 14.9% by Wave 2.
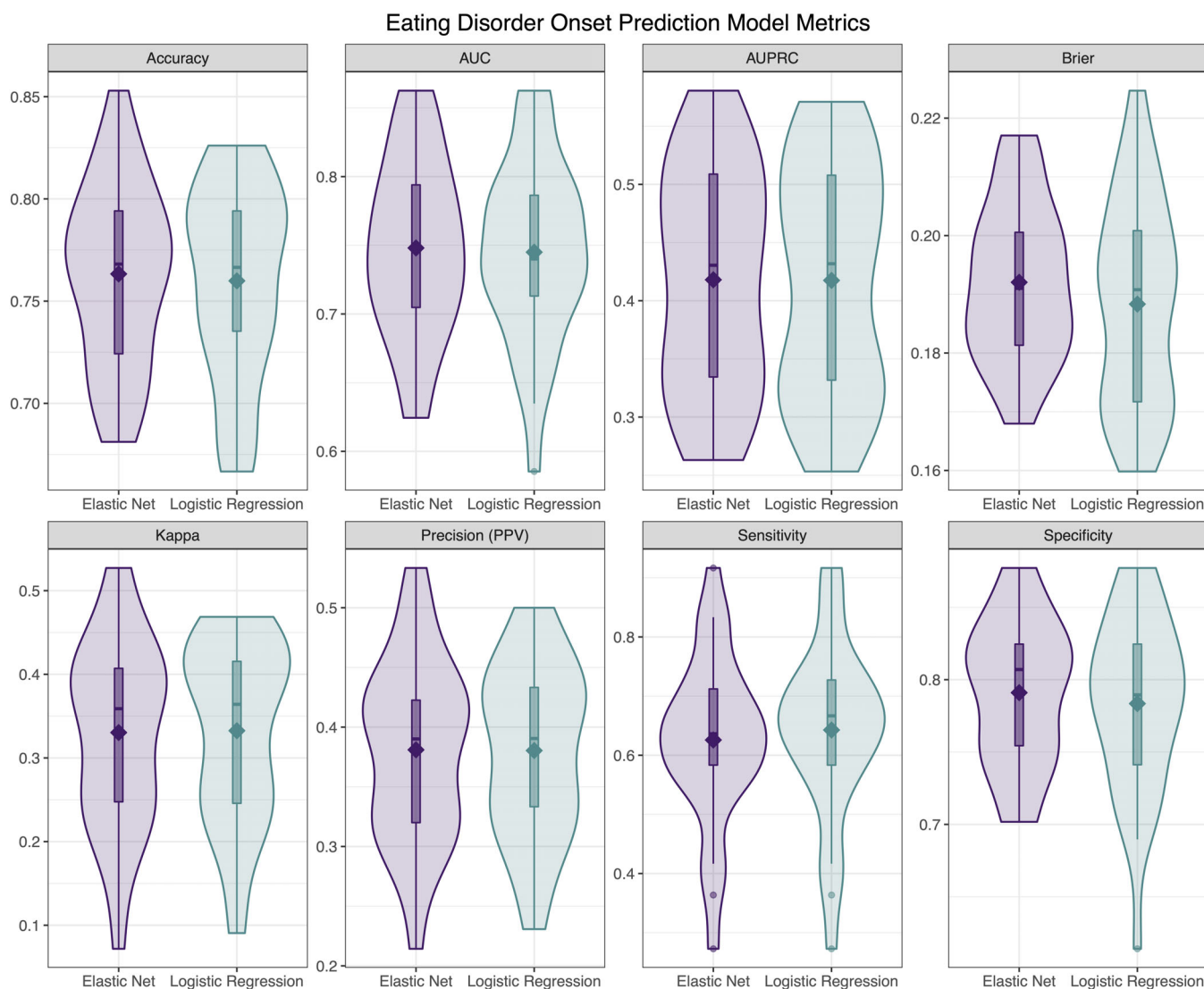
3.2 | Model performance

3.2.1 | Aim 1: Predicting EDs onset

Model performance metrics for the logistic regression and elastic net models predicting probable ED onset are presented in Figure 1. The logistic regression model provided fair AUC (mean cross-validated AUC = 0.75), with poor AUPRC (M = 0.42), fair accuracy (M = 0.76), fair PPV (M = 0.38), poor sensitivity (M = 0.64), fair specificity (M = 0.78), and fair Brier score (M = 0.19). The most important variable in the logistic regression model was psychological distress (β = 8.01), followed by weight/shape concerns (β = 5.81), female sex (β = 3.50), being a bully victim (β = 2.98), BMI percentile (β = 1.82), and strict weight loss dieting (β = 0.16).

The elastic net model showed nearly identical performance to the logistic regression model, with fair AUC (M = 0.75), poor AUPRC (M = 0.42), fair accuracy (M = 0.76), fair PPV (M = 0.38), poor sensitivity (M = 0.63), fair specificity (M = 0.79), and fair Brier score (M = 0.19). The most important variable was also psychological distress (β = 0.54), followed by weight/shape concerns (β = 0.37), female sex (β = 0.22), being a bully victim (β = 0.14), BMI percentile (β = 0.02), and strict weight loss dieting (β < 0.001).

FIGURE 1 Descriptive visualizations of the distribution of cross-validated model performance metrics of elastic net and logistic regression models predicting eating disorders onset at Wave 2 based on Wave 1 data ($N = 687$). AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve.

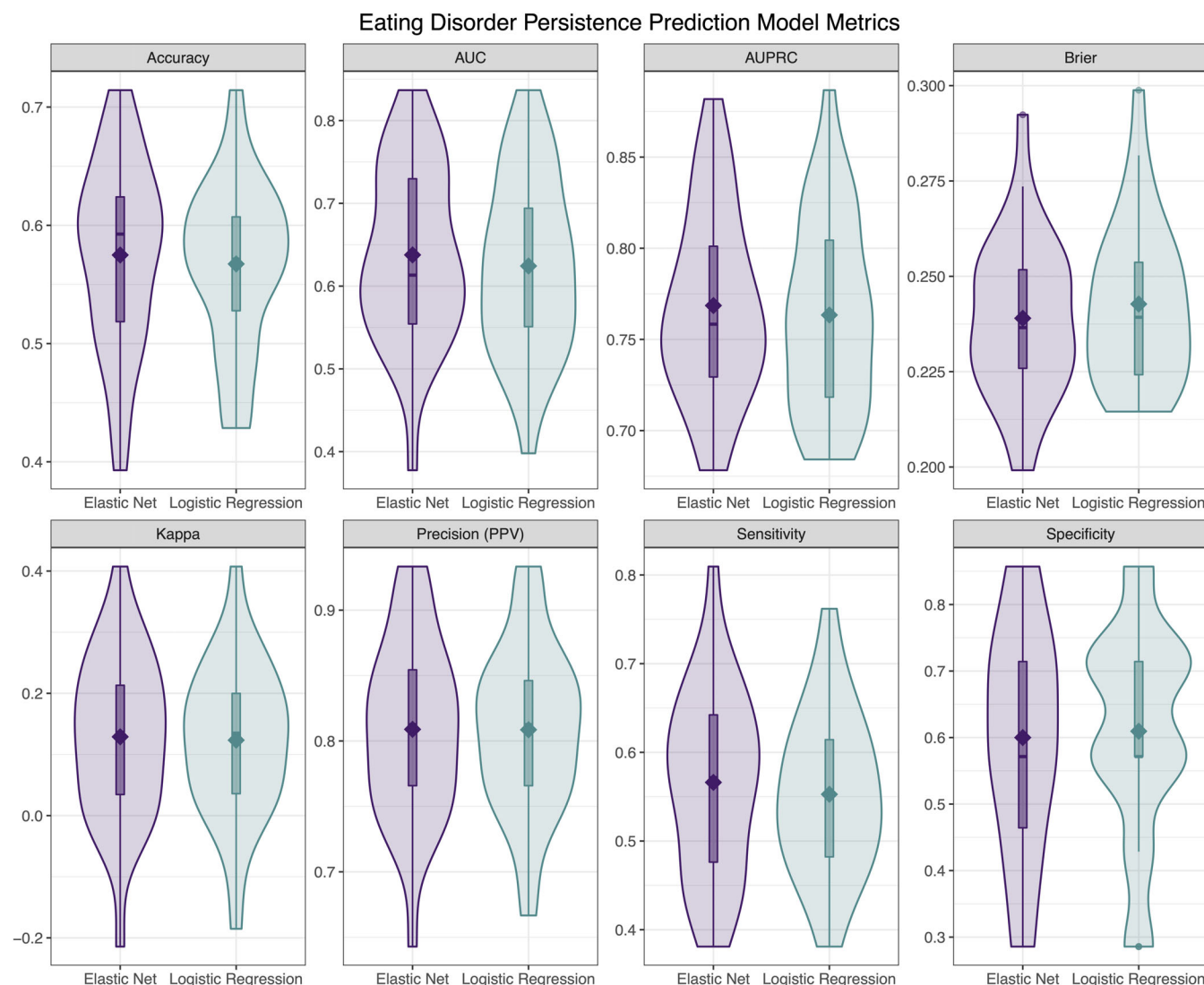### 3.2.2 | Aim 2: Predicting EDs persistence

Model performance metrics for the logistic regression and elastic net models predicting persistence of a probable ED are presented in Figure 2. The logistic regression model provided poor AUC ($M = 0.62$), with fair AUPRC ($M = 0.76$), poor accuracy ($M = 0.57$), good PPV ($M = 0.81$), poor sensitivity ($M = 0.55$), fair specificity ($M = 0.61$), and poor Brier score ($M = 0.24$). The most important variable in the logistic regression model was psychological distress ($\beta = 3.07$), followed by self-induced vomiting ($\beta = 1.45$), social functioning ($\beta = 1.27$), and laxative use ($\beta = 0.31$).

The elastic net model showed nearly identical performance to the logistic regression model, with poor AUC ($M = 0.64$), fair AUPRC ($M = 0.77$), poor accuracy ($M = 0.57$), good PPV ($M = 0.81$), poor sensitivity ($M = 0.57$), fair specificity ($M = 0.60$), and poor Brier score ($M = 0.24$). The most important variable was also psychological

distress ($\beta = 0.34$), followed by self-induced vomiting ($\beta = 0.19$), social functioning ($\beta = 0.14$), and laxative use ($\beta = 0.03$).

4 | DISCUSSION

This study presents, to our knowledge, the first general population mixed gender prognostic models for probable ED onset and persistence among adolescents. A focus was on model parsimony to enhance translatability for eventual use in the real world. Our model predicting probable ED onset performed well, accurately discriminating between future cases versus noncases 75% of the time, and positively identifying 64% of future cases (sensitivity) and 79% of future noncases (specificity). However, the model predicting persistence of a probable ED was poor, with accurate discrimination between future cases versus noncases closer to chance levels (64%). This is likely due

Eating Disorder Persistence Prediction Model Metrics

FIGURE 2 Descriptive visualizations of the distribution of cross-validated model performance metrics of elastic net and logistic regression models predicting eating disorders persistence at Wave 2 based on Wave 1 data ($N = 276$). AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve.

to there being a lack of relevant predictors included in this model, an important focus for future research.

The performance metrics for the model predicting probable ED onset were in the range observed for established and implemented prognostic models within other areas of health and medicine (e.g., cancer and heart disease [Fusar-Poli et al., 2015]). Although the values for discrimination, sensitivity and specificity were nominally labeled as fair, poor, and fair, respectively, the acceptability of these values in practice depends on a number of factors, including (i) improvement in prediction provided by the model over current practice and (ii) the cost–benefit analysis of correct versus false positive identification. In regards to EDs, we know that mental health literacy (Mond, 2014), detection and treatment access (Hart et al., 2011), and screening practice is extremely low. Furthermore, although screening tools can detect *current* EDs (Hill et al., 2010; Maguen et al., 2018), their utility in predicting *future* outcomes is unclear. If, by way of example, a screening tool based on the model of onset in the

present study was to be implemented in a school setting (e.g., an online questionnaire with just the predictors of importance, linked to the model algorithm to determine risk score), it would have the potential to correctly discriminate between 75% of adolescents who will go on to develop an ED in the next 12 months—a significant improvement upon current school screening practices.

Psychological distress was found to be the most influential predictor within the ED onset model. Distress may be indicative of general lack of psychological wellbeing or the presence of psychiatric illness, including disorders other than EDs. Univariate analyses demonstrated that for every 5 points scored higher on a measure of psychological distress (the K-10, scoring range 0–50), the risk for an individual developing a probable ED within the next year increased by 75%. The role of distress in predicting ED onset is in line with the well-established evidence of distress (American Psychiatric Association, 2013; Hay & Williams, 2013; Mitchison et al., 2015; Stice, 2002) and psychiatric comorbidity

(Ahn et al., 2019; Singleton et al., 2019) as transdiagnostic risk factors and correlates across many psychiatric disorders, include EDs.

On the other hand, ED-specific risk factors had variable importance in the models developed. Whereas weight and shape concerns were found to be important in the model for ED onset, weight loss dieting and weight status were not. These findings are at odds with the body of literature focusing on dieting and higher weight as risk factors for development of an ED, and targets for ED prevention programs. The current findings, however, do align with major EDetiological theory which posits that weight and shape concerns emerge prior to the development of ED behaviors, which subsequently further entrench weight and shape concerns, resulting in a vicious cycle (e.g., cognitive behavioral theory; Fairburn et al., 2003). Thus, focusing on cognitive variables (e.g., distress and weight and shape concerns) in screening, may enable the casting of an adequately "broad net" to pick up the majority of those who will go on to develop an ED before that event actually occurs. On the other hand, it should be acknowledged that due to our goal for model parsimony we included only one item assessing dieting. This single item may not have tapped into the construct of "risky dieting" as well as a full scale that has greater scope to capture the multi-dimensionality of this behavior. In regards to weight status not contributing significantly to onset prediction, this may reflect that EDs in the current sample were associated with the full spectrum of weight status from very low to very high weights (Mitchison et al., 2019), and underscores the need to focus less on absolute weight, (as opposed to weight *change*, which may be associated with disordered eating) and cognitive preoccupation with weight, when screening for ED risk.

This study found little discernible difference in the models developed using a traditional logistic regression approach as opposed to ML. This is similar to previous prediction research, including within the field of EDs predicting response to treatment (Espel-Huynh et al., 2021) and ED caseness (Krug et al., 2021). These findings suggest that, for many purposes, simpler analytic methods may be acceptable for interrogating questions pertaining to outcome prediction. On the other hand, studies with larger numbers of predictor variables have found an advantage to ML approaches in terms of accuracy (Haynos, Wang et al., 2021; Sadeh-Sharvit et al., 2020) and reducing model factors to achieve parsimony (Krug et al., 2021). It should be noted that this study applied only one ML approach (elastic net). Researchers in the ED field have started to address the question of which ML approach works best under which conditions (Krug et al., 2021), which should lead to greater clarity for statistical planning in future.

4.1 | Strengths and limitations

A strength of this study is the application of prognostic modeling methodology (Steyerberg et al., 2013), including the use of ML as an analytical tool. We used a prospective design, which counters problems with retrospective recall (e.g., Krug et al., 2021), and selected only predictors with known prognostic value, to balance model accuracy with parsimony. This pilot research also benefited from a large demographically diverse and phenotypically-rich community sample of adolescents with outcome measures inclusive of the full spectrum of EDs.

A primary limitation of this study was the absence of several predictors which could have improved model performance. The present study relied on secondary analyses of existing data, which precluded preselection of the full suite of evidence-based risk factors (Steyerberg et al., 2013) for ED onset (e.g., duration of illness, perfectionism, and psychiatric comorbidity) or ED persistence (e.g., perfectionism, perceived pressure to be thin, and age at puberty/menarche for ED onset) (Bakalar et al., 2015; Culbert et al., 2015; Day, Bussey, Trompeter, & Mitchison, 2021; de Portela Santana et al., 2012; Ghaderi, 2001; Glashouwer et al., 2019; Jacobi et al., 2004; Keel & Forney, 2013; Lie et al., 2019; Mazzeo & Bulik, 2009; Micali, 2005; Stice, 2002; Vall & Wade, 2015). Other limitations of the current study include lack of preregistration, the relatively small sample size compared to other risk model development studies, which precluded investigation into diagnostic and gender spectrum differences; the use of EDE-Q data for both predictor and outcome information, which may have artificially inflated the strength of observed relationships; and other limitations noted in previous studies using these data, including self-report and single-item measurement of some variables, and measurement of some of the diagnostic criteria over a 1 month as opposed to the DSM-5 3 month time period. Of note, the models developed in this study used population-based data and are only generalizable to community-based settings. Further, while the model as described in this study for ED persistence, once improved, will be useful in guiding decisions about *whether* to intervene, it cannot guide clinicians on the type of treatments that may ultimately be beneficial. A model that can guide treatment decisions in this way would also be of value, and should be a goal of future research, making use of evidence-based predictors of treatment moderators.

4.2 | Clinical and public health implications and future research

The finding that prediction of outcomes worked best when considering both transdiagnostic and ED-specific predictors emphasizes the need to move away from a focus on single risk factors in ED prediction, screening and interventions, and rather to consider multivariable approaches capitalizing on best known risk factors, whether they be disorder-specific or transdiagnostic. Of note, current screening instruments have tended to be ED specific (Hill et al., 2010; Maguen et al., 2018). The aim at the heart of clinical risk prediction is to translate evidence into practice by producing prediction tools. These for instance may be developed in the form of online risk calculators, such as the well-known Framingham Risk Scores for heart disease, that can be easily used by consumers and clinicians to guide shared decision-making about if, when and how to commence intervention. Such instruments are readily available for a variety of medical and psychiatric (Fusar-Poli et al., 2019) outcomes but are not yet developed for EDs. We recommend that researchers seeking to develop such

prediction tools follow established prognostic modeling guidelines (Hemingway et al., 2013; Hingorani et al., 2013; Riley et al., 2013; Steyerberg et al., 2013). This involves producing highly accurate models using large cohort data that is inclusive of all relevant evidence-based risk factors. The next step prior to implementation of the risk model is external validation (and re-calibration if needed) in new cohorts which will define the parameters of the generalizability of the prediction model, including along dimensions of gender, age, and ethnicity/race. According to systematic reviews, external validation is an often overlooked step (Schmidt et al., 2017), limiting the utility of many models. The final step is impact testing which involves the transformation of statistical models into online calculators and testing their uptake and effectiveness in real world settings.

5 | CONCLUSIONS

This study has provided further evidence of the feasibility of developing risk prediction models for ED outcomes and developed the first models in a general population-based mixed gender sample of adolescents. The models performed relatively well considering the limitations of pre-existing data, giving confidence in future modeling work. Researchers are encouraged to continue this work, with eventual translation of evidence-based models to improve prevention and treatment for youth with EDs in mind.

AUTHOR CONTRIBUTIONS

Deborah Mitchison: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; writing – original draft. **Shirley B Wang:** Formal analysis; investigation; visualization; writing – review and editing. **Tracey Wade:** Conceptualization; investigation; writing – review and editing. **Ann Frances Haynos:** Investigation; writing – review and editing. **Kay Bussey:** Data curation; investigation; methodology; project administration; writing – review and editing. **Nora Trompeter:** Data curation; methodology; project administration; writing – review and editing. **Alexandra Lonergan:** Data curation; methodology; project administration; writing – review and editing. **Jack Tame:** Data curation; project administration; writing – review and editing. **Phillipa Hay:** Conceptualization; investigation; writing – review and editing.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Deborah Mitchison https://orcid.org/0000-0002-6736-7937
Shirley B. Wang https://orcid.org/0000-0002-8583-3014
Tracey Wade https://orcid.org/0000-0003-4402-770X
Ann F. Haynos https://orcid.org/0000-0002-2548-1731
Nora Trompeter https://orcid.org/0000-0001-5800-8679
Alexandra Lonergan https://orcid.org/0000-0002-7203-503X
Phillipa Hay https://orcid.org/0000-0003-0296-6856

REFERENCES

Ahn, J., Lee, J.-H., & Jung, Y.-C. (2019). Predictors of suicide attempts in individuals with eating disorders. *Suicide and Life-Threatening Behavior*, *49*(3), 789–797. https://doi.org/10.1111/sltb.12477

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Publishing Inc.

Aouad, P., Hay, P., Soh, N., Touyz, S., Mannan, H., & Mitchison, D. (2019). Chew and spit (CHSP) in a large adolescent sample: Prevalence, impact on health-related quality of life, and relation to other disordered eating features. *Eating Disorders*, *1-14*, 509–522. https://doi.org/10.1080/10640266.2019.1695449

Bakalar, J. L., Shank, L. M., Vannucci, A., Radin, R. M., & Tanofsky-Kraff, M. (2015). Recent advances in developmental and risk factor research on eating disorders. *Current Psychiatry Reports*, *17*(6), 42. https://doi.org/10.1007/s11920-015-0585-x

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ*, *350*, g7594. https://doi.org/10.1136/bmj.g7594

Culbert, K. M., Racine, S. E., & Klump, K. L. (2015). What we have learned about the causes of eating disorders—A synthesis of sociocultural, psychological, and biological research. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(11), 1141–1164. https://doi.org/10.1111/jcpp.12441

Day, S., Bussey, K., Trompeter, N., Hay, P., Lonergan, A., & Mitchison, D. (2021). Associations of weight- or shape-related bullying with diverse disordered eating behaviors in adolescents. *International Journal of Eating Disorders*, *54*, 1641–1651.

Day, S., Bussey, K., Trompeter, N., & Mitchison, D. (2021). The impact of teasing and bullying victimization on disordered eating and body image disturbance among adolescents: A systematic review. *Trauma, Violence, & Abuse*, 23(3), 985–1006. https://doi.org/10.1177/1524838020985534

de Portela Santana, M. L., da Costa Ribeiro, H., Jr., Mora Giral, M., & Raich, R. M. (2012). Epidemiology and risk factors of eating disorder in adolescence—A review. *Nutricion Hospitalaria*, 27(2), 391–401. https://doi.org/10.3305/nh.2012.27.2.5573

Espel-Huynh, H., Zhang, F., Thomas, J. G., Boswell, J. F., Thompson-Brenner, H., Juarascio, A. S., & Lowe, M. R. (2021). Prediction of eating disorder treatment response trajectories via machine learning does not improve performance versus a simpler regression approach. *International Journal of Eating Disorders*, 54(7), 1250–1259. https://doi.org/10.1002/eat.23510

Fairburn, C. G., & Beglin, S. J. (2008). Eating disorder examination questionnaire (EDE-Q 6.0). In C. G. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 309–314). Guilford Press.

Fairburn, C. G., Cooper, Z., Doll, H. A., & Davies, B. (2005). Identifying dieters who will develop an eating disorder: A prospective, population-based study. *American Journal of Psychiatry*, 162(12), 2249–2255. https://doi.org/10.1176/appi.ajp.162.12.2249

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, 41(5), 509–528.

Fatt, S. J., Mond, J., Bussey, K., Griffiths, S., Murray, S. B., Lonergan, A., Hay, P., Trompeter, N., & Mitchison, D. (2019). Help-seeking for body image problems among adolescents with eating disorders: Findings from the EveryBODY study. *Eating and Weight Disorders*, 25, 1267–1275. https://doi.org/10.1007/s40519-019-00759-9

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22.

Fusar-Poli, P., Cappucciati, M., Rutigliano, G., Schultze-Lutter, F., Bonoldi, I., Borgwardt, S., Riecher-Rössler, A., Addington, J., Perkins, D., Woods, S. W., McGlashan, T. H., Lee, J., Klosterkötter, J., Yung, A. R., & McGuire, P. (2015). At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. *World Psychiatry*, 14(3), 322–332. https://doi.org/10.1002/wps.20250

Fusar-Poli, P., Hijazi, Z., Stahl, D., & Steyerberg, E. W. (2018). The science of prognosis in psychiatry: A review. *JAMA Psychiatry*, 75(12), 1289–1297.

Fusar-Poli, P., Werbeloff, N., Rutigliano, G., Oliver, D., Davies, C., Stahl, D., McGuire, P., & Osborn, D. (2019). Transdiagnostic risk calculator for the automatic detection of individuals at risk and the prediction of psychosis: Second replication in an independent national health service trust. *Schizophrenia Bulletin*, 45(3), 562–570.

Ghaderi, A. (2001). Review of risk factors for eating disorders: Implications for primary prevention and cognitive behavioural therapy. *Scandinavian Journal of Behaviour Therapy*, 30(2), 57–74. https://doi.org/10.1080/02845710117890

Glashouwer, K. A., van der Veer, R. M., Adipatria, F., de Jong, P. J., & Vocks, S. (2019). The role of body image disturbance in the onset, maintenance, and relapse of anorexia nervosa: A systematic review. *Clinical Psychology Review*, 74, 101771.

Hammerle, F., Huss, M., Ernst, V., & Bürger, A. (2016). Thinking dimensional: Prevalence of DSM-5 early adolescent full syndrome, partial and subthreshold eating disorders in a cross-sectional survey in German schools. *BMJ Open*, 6(5), e010843. https://doi.org/10.1136/bmjopen-2015-010843

Hart, L. M., Granillo, M. T., Jorm, A. F., & Paxton, S. J. (2011). Unmet need for treatment in the eating disorders: A systematic review of eating disorder specific treatment seeking among community cases. *Clinical Psychology Review*, 31(5), 727–735.

Hay, P., & Williams, S. E. (2013). Exploring relationships over time between psychological distress, perceived stress, life events and immature defense style on disordered eating pathology. *BMC Psychology*, 1(1), 27. https://doi.org/10.1186/2050-7283-1-27

Haynos, A. F., Wang, S. B., Lipson, S., Peterson, C. B., Mitchell, J. E., Halmi, K. A., & Crow, S. (2021). Machine learning enhances prediction of illness course: A longitudinal study in eating disorders. *Psychological Medicine*, 51(8), 1392–1402. https://doi.org/10.1017/S0033291720000227

Hemingway, H., Croft, P., Perel, P., Hayden, J. A., Abrams, K., Timmis, A., Briggs, A., Udumyan, R., Moons, K. G. M., Steyerberg, E. W., Roberts, I., Schroter, S., Altman, D. G., Riley, R. D., & for the PROGRESS Group. (2013). Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ*, 346.

Hill, L. S., Reid, F., Morgan, J. F., & Lacey, J. H. (2010). SCOFF, the development of an eating disorder screening questionnaire. *International Journal of Eating Disorders*, 43(4), 344–351. https://doi.org/10.1002/eat.20679

Hingorani, A. D., van der Windt, D. A., Riley, R. D., Abrams, K., Moons, K. G., Steyerberg, E. W., Schroter, S., Sauerbrei, W., Altman, D. G., Hemingway, H., & for the PROGRESS Group. (2013). Prognosis research strategy (PROGRESS) 4: Stratified medicine research. *BMJ*, 346, e5793.

Jacobi, C., Hayward, C., De Zwaan, M., Kraemer, H. C., & Agras, W. S. (2004). Coming to terms with risk factors for eating disorders: Application of risk terminology and suggestions for a general taxonomy. *Psychological Bulletin*, 130(1), 19–65. https://doi.org/10.1037/0033-2909.130.1.19

Kee, C. C., Lim, K. H., Sumarni, M. G., Teh, C. H., Chan, Y. Y., Nuur Hafizah, M. I., Cheah, Y. K., Tee, E. O., Ahmad Faudzi, Y., & Amal Nasir, M. (2017). Validity of self-reported weight and height: A cross-sectional study among Malaysian adolescents. *BMC Medical Research Methodology*, 17(1), 85. https://doi.org/10.1186/s12874-017-0362-0

Keel, P. K., & Forney, K. J. (2013). Psychosocial risk factors for eating disorders. *International Journal of Eating Disorders*, 46(5), 433–439. https://doi.org/10.1002/eat.22094

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L. T., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, 32(6), 959–976.

Krug, I., Linardon, J., Greenwood, C., Youssef, G., Treasure, J., Fernandez-Aranda, F., Karwautz, A., Wagner, G., Collier, D., Anderluh, M., Tchanturia, K., Ricca, V., Sorbi, S., Nacmias, B., Bellodi, L., & Fuller-Tyszkiewicz, M. (2021). A proof-of-concept study applying machine learning methods to putative risk factors for eating disorders: Results from the multi-centre European project on healthy eating. *Psychological Medicine*, 1-10, 1–10. https://doi.org/10.1017/S003329172100489X

Kuhn, M. (2021). caret: Classification and Regression Training. R package version 6.0-88. Retrieved from https://CRAN.R-project.org/package=caret

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26, p. 13). Springer.

Lie, S. Ø., Rø, Ø., & Bang, L. (2019). Is bullying and teasing associated with eating disorders? A systematic review and meta-analysis. *International Journal of Eating Disorders*, 52(5), 497–514. https://doi.org/10.1002/eat.23035

Maguen, S., Hebenstreit, C., Li, Y., Dinh, J. V., Donalson, R., Dalton, S., Rubin, E., & Masheb, R. (2018). Screen for Disordered Eating: Improving the accuracy of eating disorder screening in primary care. *General Hospital Psychiatry*, 50, 20–25. https://doi.org/10.1016/j.genhosppsych.2017.09.004

Mazzeo, S. E., & Bulik, C. M. (2009). Environmental and genetic risk factors for eating disorders: What the clinician needs to know. *Child and Adolescent Psychiatric Clinics of North America*, 18(1), 67–82. https://doi.org/10.1016/j.chc.2008.07.003

Micali, N. (2005). Childhood risk factors: Longitudinal continuities and eating disorders. *Journal of Mental Health*, 14(6), 567–574. https://doi.org/10.1080/09638230500347566

Micali, N., Solmi, F., Horton, N. J., Crosby, R. D., Eddy, K. T., Calzo, J. P., Sonneville, K. R., Swanson, S. A., & Field, A. E. (2015). Adolescent eating disorders predict psychiatric, high-risk behaviors and weight outcomes in young adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(8), 652–659.e651. https://doi.org/10.1016/j.jaac.2015.05.009

Mitchison, D., Hay, P., Slewa-Younan, S., & Mond, J. (2012). Time trends in population prevalence of eating disorder behaviors and their relationship to quality of life. *PLoS One*, *7*(11), e48450.

Mitchison, D., Hay, P., Slewa-Younan, S., & Mond, J. (2014). The changing demographic profile of eating disorder behaviors in the community. *BMC Public Health*, *14*. https://doi.org/10.1186/1471-2458-14-943

Mitchison, D., Mond, J., Bussey, K., Trompeter, N., Lonergan, A., Griffiths, S., Trompeter, N., Lonergan, A., Pike, K. M., Murray, S. B., & Hay, P. (2019). DSM-5 full syndrome, other specified, and unspecified eating disorders in 5000 adolescents: Prevalence and clinical significance. *Psychological Medicine*, *50*(6), 981–990. https://doi.org/10.1017/S0033291719000898

Mitchison, D., Mond, J., Griffiths, S., Hay, P., Nagata, J. M., Bussey, K., Trompeter, N., Lonergan, A., & Murray, S. B. (2022). Prevalence of muscle dysmorphia in adolescents: Findings from the EveryBODY study. *Psychological Medicine*, *52*(14), 3142–3149. https://doi.org/10.1017/s0033291720005206

Mitchison, D., Morin, A., Mond, J., Slewa-Younan, S., & Hay, P. (2015). The bidirectional relationship between quality of life and eating disorder symptoms: A 9-year community-based study of Australian women. *PLoS One*, *10*(3), e0120591. https://doi.org/10.1371/journal.pone.0120591

Mond, J., Hall, A., Bentley, C., Harrison, C., Gratwick-Sarll, K., & Lewis, V. (2014). Eating-disordered behavior in adolescent boys: Eating disorder examination questionnaire norms. *International Journal of Eating Disorders*, *47*(4), 335–341. https://doi.org/10.1002/eat.22237

Mond, J. M. (2014). Eating disorders "mental health literacy": An introduction. *Journal of Mental Health*, *23*(2), 51–54. https://doi.org/10.3109/09638237.2014.889286

Prnjak, K., Hay, P., Mond, J., Bussey, K., Trompeter, N., Lonergan, A., & Mitchison, D. (2021). The distinct role of body image aspects in predicting eating disorder onset in adolescents after one year. *Journal of Abnormal Psychology*, *130*(3), 236–247. https://doi.org/10.1037/abn0000537

Pursey, K. M., Hay, P., Bussey, K., Trompeter, N., Lonergan, A., Pike, K. M., Mond, J., & Mitchison, D. (2020). Diabetes and disordered eating behaviours in a community-based sample of Australian adolescents. *Journal of Eating Disorders*, *8*(1), 5. https://doi.org/10.1186/s40337-020-0282-y

Richards, K. L., Hyam, L., Allen, K. L., Glennon, D., Di Clemente, G., Semple, A., Jackson, A., Belli, S. R., Dodge, E., Kilonzo, C., Holland, L., & Schmidt, U. (2022). National roll-out of early intervention for eating disorders: Process and clinical outcomes from first episode rapid early intervention for eating disorders. *Early Intervention in Psychiatry*, *17*(2), 202–211.

Riley, R. D., Hayden, J. A., Steyerberg, E. W., Moons, K. G., Abrams, K., Kyzas, P. A., Malats, N., Briggs, A., Schroter, S., Altman, D. G., Hemingway, H., & for the PROGRESS Group. (2013). Prognosis research strategy (PROGRESS) 2: Prognostic factor research. *PLoS Medicine*, *10*(2), e1001380.

Sadeh-Sharvit, S., Fitzsimmons-Craft, E. E., Taylor, C. B., & Yom-Tov, E. (2020). Predicting eating disorders from Internet activity. *International Journal of Eating Disorders*, *53*(9), 1526–1533. https://doi.org/10.1002/eat.23338

Schmidt, A., Cappucciati, M., Radua, J., Rutigliano, G., Rocchetti, M., Dell'Osso, L., Politi, P., Borgwardt, S., Reilly, T., Valmaggia, L., McGuire, P., & Fusar-Poli, P. (2017). Improving prognostic accuracy in subjects at clinical high risk for psychosis: Systematic review of predictive models and meta-analytical sequential testing simulation. *Schizophrenia Bulletin*, *43*(2), 375–388. https://doi.org/10.1093/schbul/sbw098

Sherry, B., Jefferds, M. E., & Grummer-Strawn, L. M. (2007). Accuracy of adolescent self-report of height and weight in assessing overweight status: A literature review. *Archives of Pediatrics & Adolescent Medicine*, *161*(12), 1154–1161. https://doi.org/10.1001/archpedi.161.12.1154

Singleton, C., Kenny, T. E., Hallett, D., & Carter, J. C. (2019). Depression partially mediates the association between binge eating disorder and health-related quality of life. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00209

Solmi, F., Hatch, S. L., Hotopf, M., Treasure, J., & Micali, N. (2015). Validation of the SCOFF questionnaire for eating disorders in a multiethnic general population sample. *International Journal of Eating Disorders*, *48*(3), 312–316. https://doi.org/10.1002/eat.22373

Steinhausen, H., & Weber, S. (2009). The outcome of bulimia nervosa: Findings from one-quarter century of research. *The American Journal of Psychiatry*, *166*, 1331–1341.

Steinhausen, H.-C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry*, *159*(8), 1284–1293.

Steyerberg, E. W., Moons, K. G., van der Windt, D. A., Hayden, J. A., Perel, P., Schroter, S., for the PROGRESS Group, Riley, R. D., Hemingway, H., & Altman, D. G. (2013). Prognosis Research Strategy (PROGRESS) 3: Prognostic model research. *PLoS Medicine*, *10*(2), e1001381.

Stice, E. (2002). Risk and maintenance factors for eating pathology: A meta-analytic review. *Psychological Bulletin*, *128*(5), 825–848.

Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and correlates of eating disorders in adolescents: Results from the national comorbidity survey replication adolescent supplement. *Archives of General Psychiatry*, *68*(7), 714–723.

Trompeter, N., Bussey, K., Forbes, M. K., Mond, J., Hay, P., Basten, C., Goldstein, M., Thornton, C., Heruc, G., Byrne, S., Griffiths, S., Lonergan, A., Touyz, S., & Mitchison, D. (2020). Differences between Australian adolescents with eating disorder symptoms who are in treatment or not in treatment for an eating disorder. *Early Intervention in Psychiatry*, *1-7*, 882–888. https://doi.org/10.1111/eip.13027

Trompeter, N., Bussey, K., Hay, P., Mond, J., Murray, S. B., Lonergan, A., Griffiths, S., Pike, K., & Mitchison, D. (2018). Fear of negative evaluation and weight/shape concerns among adolescents: The moderating effects of gender and weight status. *Journal of Youth and Adolescence*, *47*(7), 1398–1408. https://doi.org/10.1007/s10964-018-0872-z

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders*, *48*(7), 946–971.

Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. *Ambulatory Pediatrics*, *3*(6), 329–341.

Wang, S. B. (2021). Machine learning to advance the prediction, prevention, and treatment of eating disorders. *European Eating Disorders Review*, *29*, 683–691.

Weissman, R. S. (2019). The role of sociocultural factors in the etiology of eating disorders. *Psychiatric Clinics of North America*, *42*(1), 121–144. https://doi.org/10.1016/j.psc.2018.10.009

SUPPORTING INFORMATION

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