



OPEN ACCESS

EDITED BY

Ayhan Bilgiç,
Izmir University of Economics, Turkey

REVIEWED BY

Stefano Palomba,
Magna Græcia University, Italy
Hurşit Ferahkaya,
Dr. Ali Kemal Belviranlı Gynecology
and Pediatric Hospital, Turkey

*CORRESPONDENCE

Lianhong Wang
2726912415@qq.com

SPECIALTY SECTION

This article was submitted to
Public Mental Health,
a section of the journal
Frontiers in Psychiatry

RECEIVED 02 July 2022

ACCEPTED 15 August 2022

PUBLISHED 06 September 2022

CITATION

Ding R, Zhou H, Yan X, Liu Y, Guo Y,
Tan H, Wang X, Wang Y and Wang L
(2022) Development and validation of
a prediction model for depression in
adolescents with polycystic ovary
syndrome: A study protocol.
Front. Psychiatry 13:984653.
doi: 10.3389/fpsy.2022.984653

COPYRIGHT

© 2022 Ding, Zhou, Yan, Liu, Guo, Tan,
Wang, Wang and Wang. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Development and validation of a prediction model for depression in adolescents with polycystic ovary syndrome: A study protocol

Rui Ding^{1,2}, Heng Zhou³, Xin Yan^{1,2}, Ying Liu^{1,2},
Yunmei Guo^{1,2}, Huiwen Tan^{1,2}, Xueting Wang^{1,2},
Yousha Wang^{1,2} and Lianhong Wang^{1,2*}

¹Nursing Department, Affiliated Hospital of Zunyi Medical University, Zunyi, China, ²Nursing College, Zunyi Medical University, Zunyi, China, ³Reproductive Medicine Department, Affiliated Hospital of Zunyi Medical University, Zunyi, China

Introduction: The high prevalence and severity of depression in adolescents with polycystic ovary syndrome (PCOS) is a critical health threat that must be taken seriously. The identification of high-risk groups for depression in adolescents with PCOS is essential to preventing its development and improving its prognosis. At present, the routine screening of depression in adolescents with PCOS is mainly performed using scales, and there is no early identification method for high-risk groups of PCOS depression in adolescents. It is necessary to use a warning model to identify high-risk groups for depression with PCOS in adolescents.

Methods and analysis: Model development and validation will be conducted using a retrospective study. The study will involve normal adolescent girls as the control group and adolescent PCOS patients as the experimental group. We will collect not only general factors such as individual susceptibility factors, biological factors, and psychosocial environmental factors of depression in adolescence, but will also examine the pathological factors, illness perception factors, diagnosis and treatment factors, and symptom-related factors of PCOS, as well as the outcome of depression. LASSO will be used to fit a multivariate warning model of depression risk. Data collected between January 2022 and August 2022 will be used to develop and validate the model internally, and data collected between September 2022 and December 2022 will be used for external validation. We will use the C-statistic to measure the model's discrimination, the calibration plot to measure the model's risk prediction ability for depression, and the nomogram to visualize the model.

Discussion: The ability to calculate the absolute risk of depression outcomes in adolescents with PCOS would enable early and accurate predictions of depression risk among adolescents with PCOS, and provide the basis for the formulation of depression prevention and control strategies, which have important theoretical and practical implications.

Trial registration number: [ChiCTR2100050123]; Pre-results.

KEYWORDS

polycystic ovary syndrome, adolescents, depression, prediction model, LASSO regression

Introduction

Polycystic Ovary Syndrome (PCOS) is a female endocrine metabolic disease of unknown etiology, with a prevalence of 5.6 to 11.04% in adolescence (1) and 10.26% in the Chinese adolescent population (2). Symptoms of PCOS include abnormal menstruation, infertility, hyperandrogenism, and polycystic ovarian changes (3). It has been controversial and challenging to diagnose PCOS during adolescence due to the overlap between regular pubertal physiological changes (irregular menstrual cycles, acne, and polycystic ovarian morphology on pelvic ultrasound) and adult PCOS diagnostic criteria. Consensus statements on adult and pediatric health have acknowledged these challenges (4–6). To be more specific, challenges include the risk of under-diagnosis, delayed and/or poor diagnosis experiences (7), and over-diagnosis, as well as the additional risk of the use of inconsistent non-evidence-based approaches to PCOS diagnosis and management (8). PCOS generally begins after the first menarche in adolescence and lasts through the reproductive years and perimenopause. Besides affecting reproductive health, PCOS increases the risk of diseases such as type 2 diabetes, cardiovascular disease and endometrial cancer (9), as well as the development of psychiatric disorders. Adolescent girls with PCOS are 2.4 times more likely to suffer from depression than normal girls of the same age, about 50–60% (10, 11). A depressive disorder can result in great psychological and physical distress in adolescents, leading to suicide in over half of them, and it can affect the health and psychosocial functioning of adults as well (12). Studies have shown that PCOS can also lead to infertility in the long run (13). Nevertheless, recent studies suggest that PCOS on its own is not necessarily a cause of infertility (14). It is possible that infertility can be exacerbated by emotional and psychiatric problems in PCOS patients due to illness and fear of future infertility (14). Thus, it is recommended that psychological, emotional, and other non-medical measures be optimized in order to treat infertility, as well as increasing the attention given to depression in individuals with PCOS. PCOS, however, has no identified etiology and no specific treatment. Life management is recommended by international evidence-based guidelines as the ideal treatment for PCOS (15), but patients with the disorder have poor life management skills (16). The guideline declares that patients with PCOS can participate in better life management by improving their psychological status (4). There is no doubt that depression is a significant health issue

that must be addressed in the management of adolescents with PCOS.

Depression high-risk groups are individuals or subgroups who suffer from certain depressive symptoms but do not meet the diagnostic criteria for depression (17). They constitute a mental sub-health state between health and depression. Given the high prevalence of depression in adolescents with PCOS, warning of its onset risk and identification of high-risk groups is significant for preventing its development and improving its prognosis (18). There is currently no early detection approach for high-risk groups of depression in adolescents with PCOS, and most routine screening uses questionnaires. Hence, it's necessary to develop a warning model to identify adolescents with PCOS at high risk for depression.

The development of warning models often involves the incorporation of influencing factors. However, a variety of factors affect depression in adolescents, including: (i) biological factors, such as hormones and individual development during adolescence, which modifies neuroendocrine regulation and brain structure in diverse ways, contribute to adolescent depression. For instance, altering the activity of the hypothalamic-pituitary-adrenocortical (HPA) axis, affecting the release of neurotransmitters, altering the REM sleep latency, and so on, can all have an impact on the neurological structure and psychological state of the adolescent population, eventually leading to depression (19); (ii) individual susceptibility factors, such as genetic factors, with heredity influence approximately 40–50% of depressed people. Depressed people's first-degree relatives have a 10–13% chance of developing depression (20); (iii) psychosocial environmental factors, such as family environments, social development and social role changes, and negative behavior patterns, contribute to a sense of helplessness, which leads to negative attributions and ultimately to depression in the long run (21). PCOS may also be associated with several factors, such as insulin resistance and obesity, which may further aggravate the endocrine disorder during adolescence and increase the dysregulation of the HPA axis. Likewise, PCOS also leads to higher levels of androgen in the body, which is closely related to the risk factors of mental diseases (22), so there is speculation that pathological factors such as insulin resistance, obesity and high levels of androgen play a role in the occurrence of depression in PCOS (23). Hormonal drugs, like Ethinylestradiol cyproterone acetate tablets, are often used to treat PCOS. These drugs affect the cortical and subcortical areas of the brain that control emotions. These areas are immature

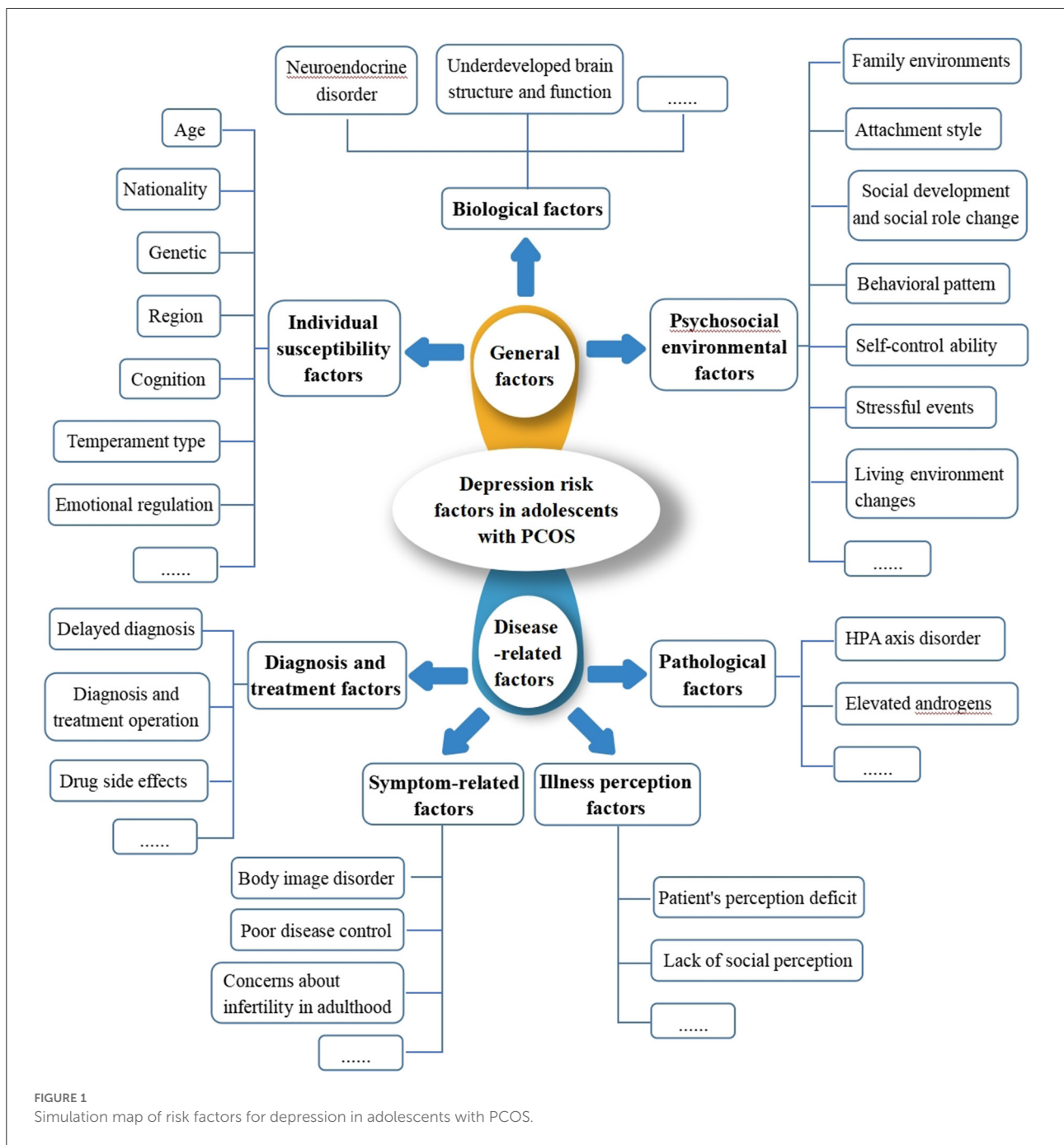


FIGURE 1 Simulation map of risk factors for depression in adolescents with PCOS.

and oversensitive in adolescents, which greatly increases the risk of depression (24). A lack of control over menstrual irregularities (11), as well as a lack of knowledge and information about the disease, which increases the psychological burden of the patients and may worsen depression (25). It can be seen that depression in PCOS is linked to the above-mentioned factors directly or indirectly. It's difficult to predict depression risk from one or more of these factors with any certainty because the number and magnitude of influencing factors are still inconclusive

(Figure 1). Thus, developing a warning depression model for adolescents with PCOS through the interaction of complex influencing factors is a critical problem that requires immediate research, and there is no such report currently available.

As previously noted, the occurrence of depression in adolescents with PCOS involves multiple dimensions and contains numerous independent variables. To minimize the bias caused by the absence of important independent variables, it's important to select as many variables as possible

at the beginning of the warning model construction process. Traditional regression analysis methods using statistical software such as SPSS, Jamovi, etc., can quantitatively analyze the strength of a predictor based on the regression coefficient. However, conventional regression analysis methods cannot handle collinearity problems and produce high variance when there are numerous independent variables. Least Absolute Shrinkage and Selection Operator (LASSO) is a type of machine learning method introduced by Tibshirani in 1996 (26). On the one hand, LASSO can select features automatically. When there are numerous independent variables, LASSO can selectively integrate powerful explanatory factors into the model. On the other hand, LASSO can compress the coefficients of some minor or insignificant independent variables to zero and pick more relevant independent variables to prevent overfitting. As compared to traditional regression analysis methods, LASSO can process all independent variables simultaneously, which greatly increases the model's stability. LASSO has been gradually applied to the development of warning models in the medical area in recent years (27, 28), as it offers high stability, fast calculation speed, and ease of interpretation in building warning models. To date, LASSO has been applied to the construction of diagnostic prediction models, early risk identification of ischemic stroke, for example (27). LASSO has also been applied to the construction of prognostic prediction models, for instance, survival after ampullary adenocarcinoma resection (28). Additionally, a study based on LASSO used electronic primary care data to develop a predictive model of common mental disorders among adolescents and early adults (29). But the study noted that primary care data underestimated common mental disorders. Due to race and PCOS, the model cannot predict depression with PCOS in adolescents. Thus, we will explore the causes of depression in teenagers with PCOS and the influence of a variety of risk variables, as well as develop and test a depression risk prediction model for adolescents with PCOS. Not only will this model be able to identify high-risk groups for depression with PCOS in adolescence, but also for high-risk groups of depression in general adolescent girls.

Objectives

The aims of the study are to: (i) Develop a depression prediction model for adolescents with PCOS to identify at-risk populations early; (ii) Validate the accuracy and validity of the model and to assess the clinical utility of the model.

Methods and analysis

Preparation

A multidisciplinary research team was established, consisting of obstetricians and gynecologists ($n = 2$),

psychiatrists ($n = 2$), psychologists ($n = 1$), nursing specialists ($n = 1$), statisticians ($n = 1$) and research assistants ($n = 2$). Members of the research team have received uniform training in PCOS clinical practice and adolescents as well as extensive experience in their respective fields. Firstly, the team derived an understanding of depression among adolescent PCOS patients by retrieving evidence-based information and integrating clinical practices. Secondly, the research team reviewed the factors associated with depression in adolescents in the general population, and then the pathological factors of PCOS, the perception factors of PCOS, the diagnosis and treatment factors of PCOS, and the symptoms-related factors of PCOS, to summarize the risk factors for depression in adolescents with PCOS.

Study design

We will use a retrospective cohort design for predictive model development and validation studies. It will follow expert guidance on model development and validation (30, 31), and reported per the Transparent Reporting of multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement (32).

Participants

The study will be conducted at the Affiliated Hospital of Zunyi Medical University in Zunyi, Guizhou Province. It is a 2,800-bed general hospital with over two million outpatient attendances per year. An average of 80–100 adolescent PCOS patients are seen at the gynecological clinic per month, providing the sample size required for this study.

Patient recruitment

Adolescents with PCOS who attend the gynecology clinic of the Affiliated Hospital of Zunyi Medical University from January 2022 to August 2022 will be selected as the potential participants using a consecutive enrollment method. The final participants will be identified based on inclusion and exclusion criteria, and data are collected as both a training dataset and a validation dataset for the model construction. Data collected from September 2022 to December 2022 as an external validation cohort. Participants listed above will comprise the experimental group. Additionally, a control group will be recruited from the gynecology clinic among healthy, age-matched adolescent girls.

Eligibility criteria

Inclusion criteria

1. Aged 10–19 years old (33).
2. According to “Chinese guidelines for diagnosis and treatment of polycystic ovary syndrome,” which was developed by the Endocrinology Group of the Chinese Medical Association Obstetrics and Gynecology Branch in 2018 (34), PCOS participants in the study must meet all three indicators in the Rotterdam diagnostic criteria at the same time (35), including: (i) oligomenorrhea persisting for at least 2 years after menarche, or amenorrhea; (ii) clinical manifestations of hyperandrogenism or hyperandrogenemia; (iii) the diagnosis of PCO on ultra sound includes increased ovarian size ($>10\text{ cm}^3$).
3. Able to self-report (verbally understandable and articulate).
4. Participants and their families volunteer to participate in the study.

Exclusion criteria

1. Participants with cognitive impairment and major mental illness (36)(schizophrenia, schizoaffective disorder, bipolar disorder, etc.) (37).
2. Participants with serious comorbidity(Other diseases that seriously affect the patient’s normal life, such as acute myocardial infarction, systemic lupus erythematosus complicated by severe renal dysfunction, etc.).
3. Other conditions that lead to elevated androgen levels (congenital adrenal cortical hyperplasia, Cushing syndrome, androgen-secreting tumors, etc.) and cause ovulation disorders (hyperprolactinemia, premature ovarian failure, functional hypothalamic amenorrhea, thyroid dysfunction, etc.).

Determinations and variables

After discussion and conclusion, the research team obtained indicators of independent and dependent variables affecting depression in PCOS adolescents, as follows.

Dependent variables

Depression will be measured by the Chinese version of the Children’s Depression Scale(CDI) (38). The scale is comprised of 27 items, each of which offers three possible replies based on the respondent’s actual mood during the previous 2 weeks. The project consists of five dimensions: anhedonia, poor efficacy, low self-esteem, negative emotions, and interpersonal problems. The intensity of each item determined whether it was rated zero, one, or two, and the negative word statement was scored in reverse. A score of 19 indicated the presence of depressive symptoms. It has

been reported that the Chinese version of the CDI scale has been applied to children and adolescents ranging from 7 to 18 years of age in China (38–40). CDI has the advantage of having the lowest reading level of all depression measurement tools (only the first grade level). The Chinese version of the CDI scale has a good level of reliability and validity, as well as a high clinical diagnostic value for depression, and can be used to measure depression in children and adolescents in China. Utilize the total score to reflect the severity of depressive symptoms or diagnostic threshold due to its operability and clinical applicability.

We will use the Kiddie-Schedule for Affective Disorders and Schizophrenia-Epidemiological version (K-SADS-E) to confirm the diagnosis of depression and exclude possible comorbidities. The validated Mandarin version will be used in this study (41). It is a semi-structured diagnostic interview questionnaire designed to assess the current and past mental health status of children and adolescents aged 6-18 in accordance with the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5). In this study, two psychiatrists trained in K-SADS-E interview harmonization will conduct interviews to diagnose depression and exclude possible comorbidities. It will take approximately 1–1.5 h for each participant and their parents to complete the interview. It has been demonstrated that K-SADS-E has good psychometric properties (41). The diagnoses of K-SADS-E demonstrated good convergent and divergent validity with most corresponding clinical questionnaires.

Independent variables

The predictors for building the model are shown in [Table 1](#).

Data collection

Two uniformly trained researchers will identified participants based on inclusion and exclusion criteria. Firstly, a researcher will guide participants and their parents to a quiet outpatient room that is conducive to dialogue and data collection. Another researcher will explain the purpose, significance, and implementation of the study to participants and their parents in a one-on-one format, assuring them that their privacy would be protected. With the consent of participants and their parents, researchers will sign informed consent forms with them. Afterwards, psychiatrists within the study team will conduct semi-structured interviews with participants and their parents, focusing on the diagnosis of depression and excluding possible comorbidities. The researchers will distribute the QR code of the scale that was included in the questionnaire platform in advance to participants who are eligible, and the participants will scan the QR code on WeChat to fill out the form online. Finally, the researchers will give the participants health education and

TABLE 1 Model-building predictors: definition/measuring tools, variable type and units/categories.

Candidate predictor	Definition/Measuring tools	Variable type	Units/categories
Individual susceptibility factors			
Age	Adolescent's age	Continuous & Categorical	Years (age classified into approximately three to four categories)
Nationality	Adolescent's nationality	Binary	0 "the Han nationality" 1 "minority nationality"
Body mass index(BMI)	Body mass divided by the square of the body height	Continuous	kg/m ²
Waist-to-hip ratio (WHR)	WHR is found by dividing circumference of the waist by the circumference of the hips	Continuous	
Living residence	Current residence of adolescents	Binary	0 "City" 1 "Countryside"
Education level	Current educational level of adolescents	Categorical	Education level classified into approximately five categories
Single parent family	Adolescents live with only one parent	Binary	0 "No" 1 "Yes"
Family history of mental illness	Mental illness within three generations in the paternal, maternal, direct and collateral lines in adolescents	Binary	0 "No" 1 "Yes"
Duration of illness	Time since first diagnosis of PCOS	Categorical	Duration of illness classified into approximately five to six categories
Personality	Eysenck Personality Questionnaire-Junior (42)	Continuous	
Emotional regulation skills	Emotion Regulation Scale (43)	Continuous	
Diagnosis and treatment factors			
Medication use	Adolescents with PCOS taking some medications, such as metformin, oral contraceptives and so on. If yes, please specify the medicine	Binary & Categorical	0 "No" 1 "Yes" (medications classified into approximately four to five categories)
Duration of medicine use	Duration of medicine use since first diagnosis of PCOS	Categorical	Duration of medicine use classified into approximately five to six categories
Biological factors			
Sleep	Athens Insomnia Scale (44)	Continuous	
Psychosocial environmental factors			
Interpersonal sensitivity	Symptom Checklist-90 (45)	Continuous	
Social support	Medical Outcomes Study Social Support Survey (46)	Continuous	
Clinical symptoms			
Hirsutism	Ferriman-Gallwey Rating Scale (47)	Continuous	
Acne	Rosenfield Scale (48)	Continuous	
Body image	The Negative Physical Self Scale (49)	Continuous	
Illness perception factors			
Illness perception	The Brief Illness Perception Questionnaire (50)	Continuous	
Laboratory investigations			
Sex hormone-binding globulin	Electro chemiluminescent immunoassay(ECLIA)	Continuous	nmol/L
Free testosterone index		Continuous	
Estradiol		Continuous	pmol/L
Follicle stimulating hormone		Continuous	mIU/mL
Luteinizing hormone		Continuous	mIU/mL
Progesterone		Continuous	nmol/L
Prolactin		Continuous	mIU/L
Total testosterone		Continuous	nmol/L
Dehydroepiandrosterone sulphate		Continuous	μmol/L
Third-generation thyrotropin		Continuous	μIU/mL
Insulin		Continuous	μIU/mL

prescription advice, as well as recording laboratory indicators including sex hormones and insulin. The total planned time for this process will be around 2 h. The data will be extracted *via* a questionnaire platform and collated into Excel format, with the two researchers collecting the information independently and negotiating any disagreements.

Quality control

1. The screening process will strictly adhere to the inclusion and exclusion criteria.
2. Prior to conducting formal research, researchers will receive standardized training.
3. Before the survey, researchers will explain the contents and precautions of the questionnaire to participants and their families. To ensure participants understand items, researchers will use uniform language to explain them, while avoiding leading words.
4. The questionnaire must be completed without omissions to ensure completeness of content and efficient recall.
5. Two researchers check and correct any discrepancies with the original data to ensure accuracy.

Sample size and calculation

In the case of a binary outcome, an approximate 95% confidence interval for the total outcome proportion (ϕ) is

$$\hat{\phi} \pm 1.96 \sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{n}},$$

Therefore, the absolute margin of error (δ) is $1.96 \sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{n}}$.

For this reason, in order to get an accurate estimate of the overall outcome probability in the target population, based on the expected outcome proportion ($\hat{\phi}$) and the desired margin of error, the required sample size is calculated as: $n = (\frac{1.96}{\delta})^2 \hat{\phi}(1-\hat{\phi})$.

In general, we advise aiming for a margin of error of <0.05 . Then assuming an anticipated outcome proportion in the study population of 0.4, then a minimum of 369 participants is needed to aim for an estimation error of no more than 0.05 around the actual value of 0.4 (51).

The study will include all adolescents with PCOS who attend the gynecology clinic of the Affiliated Hospital of Zunyi Medical University between January 2022 and December 2022, with an estimated sample size of 800. This study will recruit 400 healthy adolescent girls as the control group.

Statistical analysis methods

The data obtained will be analyzed using SPSS Statistics version 18 (IBM Corp) software and evaluated using comparative and descriptive statistics. Student *t*-test will be used for normally distributed parametric data, while Mann-Whitney U-test will be used for non-normally distributed data. The χ^2 test and Fisher exact test will be used to evaluate categorical data. A correlation between data will be evaluated using Pearson correlation coefficients for parametric data and Spearman correlation coefficients for non-parametric data. Descriptive statistics will be presented as median (range), mean \pm SD, or number/frequency (percentage). Statistics will be considered significant when the *P*-value <0.05 .

Handling of predictors

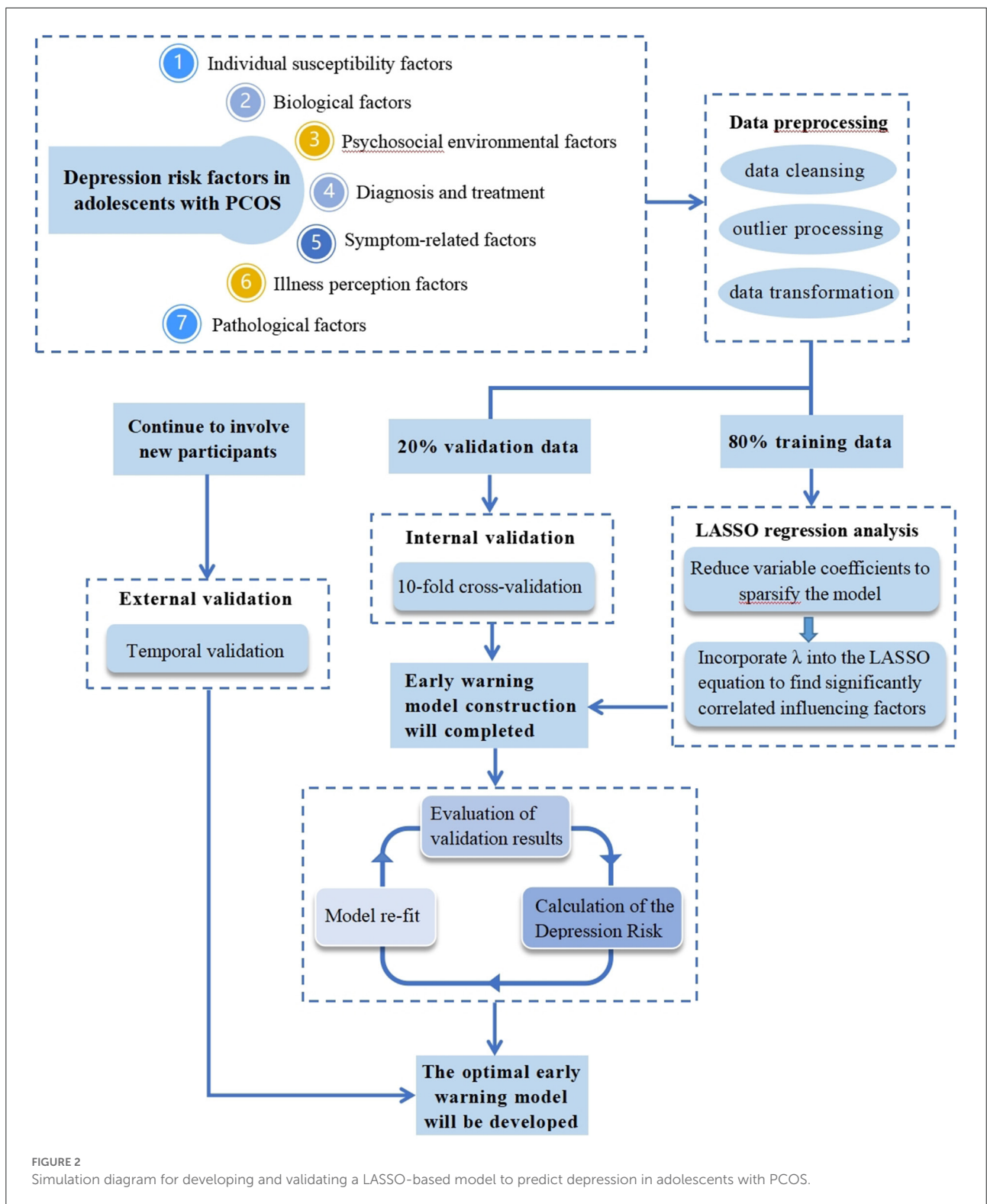
The collected data will first be preprocessed, including data cleansing, outlier processing, and data transformation. Data cleaning is the process of deleting redundant information and correcting information with obvious errors. Based on the data situation, we will choose between deletion, imputation, or substitution for outlier processing. The data will be transformed in the following ways: continuous variables will be kept in the model as continuous variables if they are linear or will be combined with similar segments if they are non-linear (treated as binary categories or ordered categories); ordered multicategorical variables will be treated as rank variables or dummy variables as appropriate, and when not linearly related to the outcome, optimal scale regression will be used to explore effect inflection points; unordered multicategorical variables need to be split into multiple binary variables (Figure 2).

Preprocessed data will be assigned variable values according to various variable types and a data set in CSV format will then be generated.

Model-building procedures (including predictor selection)

After the data set is imported into the R software, it will be divided into 80% training data and 20% validation data. The training data will be used to build the warning model, and the validation data will be used for internal validation.

To begin with, we will use R software to read the features of the dataset. After that, we will load the two R packages “corrplot” and “glmnet.” By using the `cor()` function, we will create a correlation object “corrlate” and visualize the correlation coefficient. The predictors and outcome variables in the training and validation sets will be encoded into matrices, the predictors will be stored in X vectors, and the outcome variables in Y vectors. A LASSO regression will be performed when alpha



is set to one. The program will then be run. When the parameter lambda is greater, the estimated parameters will also be compressed to a smaller degree. As lambda reaches a certain value, some variables that are not important to the model will

be compressed to zero, indicating that these variables will be eliminated. Thus, the lambda corresponding to the optimal result, the influencing factors of preliminary screening, and the regression coefficient will be obtained.

Internal validation and assessment of model performance

Internal validation will be conducted using the 20% validation set samples and will utilize a cross-validation method based on 10 folds. Then, the result of LASSO cross-validation will be obtained, and the lambda when it is one standard error away from the mean square error will be considered optimal (at this time, the mean square error is small, and the number of independent variables is also small). For that lambda case, we will examine the regression coefficients and exclude variables with a regression coefficient of zero. Those variables with regression coefficients other than zero are selected, thereby identifying the influencing factors that are significantly associated with depression in PCOS adolescents. After screening the variables, a multivariate linear model will be constructed.

The prediction accuracy of the model will be reflected by discrimination and calibration. In this study, the C-statistic is used to measure discrimination, followed by a calibration plot to measure the accuracy of the absolute risk predictions of the warning model. An evaluation of the fit will be performed by drawing a scatter plot.

External validation

External validation will be conducted using temporal validation with data collected from September 2022 to December 2022, to ensure model portability and generalizability. This means that model risk scores will be calculated rigorously and the model will be refitted in the validation cohort.

After the model is defined, we will consider the nomogram as the final presentation form in order to facilitate clinical application of the model and assist medical personnel in identifying high-risk groups of depression in adolescents with PCOS.

All statistical analysis will be performed using R software (Windows version 4.1.2).

Data storage and management

Data will be entered and stored through a third-party electronic platform, and all documents and data will be password-protected on a network within the platform, accessible only to researchers. Data will be exported in Excel format for further analysis when processed.

Discussion

This research protocol aims to construct and validate a warning model for depression in adolescents with PCOS. It will answer three questions: what factors influence depression risk in adolescents with PCOS; what role do various types of

depression risk factors play in the development of depression in adolescents with PCOS; and how interpretable and accurate is the LASSO-based warning system.

There are many factors associated with depression in adolescents with PCOS. For the first time, a depression warning model will be constructed that takes into account not only general factors like individual susceptibility factors, biological factors, and psychosocial environmental factors of depression in adolescence, but also the pathological factors, illness perception factors, diagnosis and treatment factors, and symptom-related factors of PCOS. This study will precisely and early identify depression at-risk groups for PCOS in adolescence by including the influencing factors of several dimensions of depression into the model and leveraging LASSO's ability to automatically screen variables and boost model interpretability.

So far, it has been difficult to identify adolescents at high-risk of depression, which has a severe impact on adolescent mental health. In this study, the research population consists of adolescents with PCOS, who have a significantly greater prevalence of depression and more complicated impacts than general adolescent girls, an important representative sample. Through this study, we hope to remove the bottleneck of identifying adolescents at risk of depression. The success of this study can not only identify the high-risk groups of depression in adolescents with PCOS, but also demonstrate the risk factors and key links associated with depression in adolescents with PCOS. At the same time, it can be extended and applied to the management of high-risk groups of depression throughout adolescent girls, thereby contributing to reducing the prevalence of depression in adolescent girls and implementing comprehensive and effective prevention measures.

The study is a retrospective cohort study, and the questionnaire will be partially completed by patients' recollection of recent events, which may be subject to a degree of recall bias. In prospective cohort studies, depression outcomes and factors contributing to depression can be longitudinally tracked over a period of time of high-risk groups with PCOS in adolescence, which can help to make models more robust and representative. Data for this study will be collected at one hospital, so it will be a single-center study. It is proposed that future studies should collect data from multiple hospitals simultaneously to increase sample size and persuasiveness. In this study, participants will be selected in accordance with the diagnostic criteria prescribed in the "Chinese guidelines for diagnosis and treatment of polycystic ovary syndrome." Due to the inconsistency in diagnostic criteria among different ethnic groups, the generalizability of this study is limited to some degree. It is recommended that metformin and oral contraceptives be used to treat PCOS in adolescence, in accordance with international evidence-based guidelines (4). It has been reported, that oral contraceptives may increase participants' risk of depression (24), whereas metformin can help to alleviate depression (52). Hence, confounding

effects may occur if participants take metformin and oral contraceptives concurrently. As part of our investigation, we will assess the medication status of the participants. However, we will not be able to eliminate the possibility that there may be a confounding effect. The model was externally validated using temporal validation, as the data source is the same as the model development cohort, which may be less transferable and generalizable than geographical validation (which validates the model against data from other centers or even other countries).

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Zunyi Medical University: Zunhe Lun Review [2021] 1-093. Written informed consent to participate in this study will be provided by the participant's legal guardian/next of kin. Written informed consent will be obtained from the individuals and minor's legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

RD was involved in writing the protocol, editing the manuscript, setting up the trial, and acquisition of data. HZ, XY, YG, YL, HT, XW, and YW participated in the design of the trial, analysis, and interpretation of the data. LW carried out the conception, design, and final approval of the version to be published and agreed to be accountable for all aspects of the work by ensuring that any questions regarding the accuracy or integrity of any part of the work are appropriately investigated

References

1. Naz M, Tehrani FR, Majd HA, Ahmadi F, Ozgoli G, Fakari FR, et al. The prevalence of polycystic ovary syndrome in adolescents: a systematic review and meta-analysis. *Int J Reprod Biomed.* (2019) 17:533–42. doi: 10.18502/ijrm.v17i8.4818
2. Wu Q, Gao J, Bai D, Yang Z, Liao Q. The prevalence of polycystic ovarian syndrome in Chinese women: a meta-analysis. *Ann Palliat Med.* (2021) 10:74–87. doi: 10.21037/apm-20-1893
3. Ding T, Hardiman PJ, Petersen I, Wang FF, Qu F, Baio G. The prevalence of polycystic ovary syndrome in reproductive-aged women of different ethnicity: a systematic review and meta-analysis. *Oncotarget.* (2017) 8:96351–8. doi: 10.18632/oncotarget.19180
4. Pena AS, Witchel SF, Hoeger KM, Oberfield SE, Vogiatzi MG, Misso M, et al. Adolescent polycystic ovary syndrome according to the international evidence-based guideline. *BMC Med.* (2020) 18:72. doi: 10.1186/s12916-020-01516-x
5. Hoeger KM, Dokras A, Piltonen T. Update on PCOS: Consequences, challenges, and guiding treatment. *J Clin Endocrinol Metab.* (2021) 106:e1071–83. doi: 10.1210/clinem/dgaa839
6. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations from the international evidence-based guideline for the

and resolved. All authors contributed to the article and approved the submitted version.

Funding

The work is supported by Zunyi Science and Technology Planning Project [Zun Shi Ke He HZ Zi (2021) No. 21]. The Science and Technology Department of Guizhou Province, China [Grant No. Qian Ke He (2017) 5733-077]. The Health Commission of Guizhou Province, China (Grant No. Gzwjkj2019-1-018).

Acknowledgments

The authors acknowledge the adolescents with PCOS who participate in the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

assessment and management of polycystic ovary syndrome. *Fertil Steril.* (2018) 110:364–79. doi: 10.1016/j.fertnstert.2018.05.004

7. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed diagnosis and a lack of information associated with dissatisfaction in women with polycystic ovary syndrome. *J Clin Endocrinol Metab.* (2017) 102:604–12. doi: 10.1210/jc.2016-2963

8. Sebastian MR, Wiemann CM, Bacha F, Alston TS. Diagnostic evaluation, comorbidity screening, and treatment of polycystic ovary syndrome in adolescents in three specialty clinics. *J Pediatr Adolesc Gynecol.* (2018) 31:367–71. doi: 10.1016/j.jpaga.2018.01.007

9. Wang J, Wang Q, Cui C, Zhu S, Li H, Zhang C. Research progress on the effect of polycystic ovarian syndrome on the development of the offspring. *Chin J Reprod Contracep.* (2020) 40:423–6. doi: 10.3760/cma.j.cn101441-20190516-00197

10. Coban OG, Tulaci OD, Adanir AS, Onder A. Psychiatric disorders, self-esteem, and quality of life in adolescents with polycystic ovary syndrome. *J Pediatr Adolesc Gynecol.* (2019) 32:600–4. doi: 10.1016/j.jpaga.2019.07.008

11. Sari SA, Celik N, Uzun CA. Body perception, self-esteem, and comorbid psychiatric disorders in adolescents diagnosed with polycystic ovary syndrome. *J Pediatr Adolesc Gynecol.* (2020) 33:691–6. doi: 10.1016/j.jpaga.2020.08.018

12. Kiviruusu O, Strandholm T, Karlsson L, Marttunen M. Outcome of depressive mood disorder among adolescent outpatients in an eight-year follow-up. *J Affect Disord.* (2020) 266:520–7. doi: 10.1016/j.jad.2020.01.174
13. Rafique M, Nuzhat A, Al-Jaroudi D. Risk of infertility index in women with polycystic ovarian syndrome. *J Coll Physicians Surg Pak.* (2020) 30:1188–92. doi: 10.29271/jcpsp.2020.11.1188
14. Palomba S. Is fertility reduced in ovulatory women with polycystic ovary syndrome? An opinion paper. *Hum Reprod.* (2021) 36:2421–8. doi: 10.1093/humrep/deab181
15. Moran LJ, Tassone EC, Boyle J, Brennan L, Harrison CL, Hirschberg AL, et al. Evidence summaries and recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome: lifestyle management. *Obes Rev.* (2020) 21:e13046. doi: 10.1111/obr.13046
16. Lie FS, Douma A, Verhaeghe J. Implementing the international evidence-based guideline of assessment and management of polycystic ovary syndrome (PCOS): how to achieve weight loss in overweight and obese women with PCOS? *J Gynecol Obstet Hum Reprod.* (2021) 50:101894. doi: 10.1016/j.jogoh.2020.101894
17. Noyes BK, Munoz DP, Khalid-Khan S, Brietzke E, Booj L. Is subthreshold depression in adolescence clinically relevant? *J Affect Disord.* (2022) 309:123–30. doi: 10.1016/j.jad.2022.04.067
18. Benjet C, Albor YC, Bocanegra ES, Borges G, Mendez E, Casanova L, et al. Incidence and recurrence of depression from adolescence to early adulthood: a longitudinal follow-up of the Mexican adolescent mental health survey. *J Affect Disord.* (2020) 263:540–6. doi: 10.1016/j.jad.2019.11.010
19. Colasanto M, Madigan S, Korczak DJ. Depression and inflammation among children and adolescents: A meta-analysis. *J Affect Disord.* (2020) 277:940–8. doi: 10.1016/j.jad.2020.09.025
20. Elpida B. Prodromal symptoms and initial episode of depression-genetic and emotional correlations. *Procedia - Soc Behav Sci.* (2013) 78:451–5. doi: 10.1016/j.sbspro.2013.04.329
21. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet.* (2012) 379:1056–67. doi: 10.1016/S0140-6736(11)60871-4
22. de Wit AE, Giltay EJ, de Boer MK, Bosker FJ, Cohn AY, Nolen WA, et al. Plasma androgens and the presence and course of depression in a large cohort of women. *Transl Psychiatry.* (2021) 11:124. doi: 10.1038/s41398-021-01249-2
23. Besenek M, Gurlek B. Hyperandrogenism in polycystic ovary syndrome affects psychological wellbeing of adolescents. *J Obstet Gynaecol Res.* (2021) 47:137–46. doi: 10.1111/jog.14444
24. de Wit A E, Booj S H, Giltay E J, Hadine J, Robert AS, Albertine JO. Association of use of oral contraceptives with depressive symptoms among adolescents and young women. *JAMA Psychiat.* (2020) 77:52–9. doi: 10.1001/jamapsychiatry.2019.2838
25. Hopkins CS, Kimble LP, Hodges HF, Koci AF, Mills BB. A mixed-methods study of coping and depression in adolescent girls with polycystic ovary syndrome. *J Am Assoc Nurse Pract.* (2019) 31:189–97. doi: 10.1097/JXX.0000000000000125
26. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B.* (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x
27. Meng Z, Wang M, Guo S, Zhou Y, Zheng M, Liu M, et al. Development and validation of a LASSO prediction model for better identification of ischemic stroke: a case-control study in china. *Front Aging Neurosci.* (2021) 13:630437. doi: 10.3389/fnagi.2021.630437
28. Moekotte AL, van Roessel S, Malleo G, Rajak R, Ecker BL, Fontana M, et al. Development and external validation of a prediction model for survival in patients with resected ampullary adenocarcinoma. *Eur J Surg Oncol.* (2020) 46:1717–26. doi: 10.1016/j.ejso.2020.04.011
29. Smith D, Willan K, Prady SL, Dickerson J, Santorelli G, Tilling K, et al. Assessing and predicting adolescent and early adulthood common mental disorders using electronic primary care data: Analysis of a prospective cohort study (ALSPAC) in Southwest England. *BMJ Open.* (2021) 11:e53624. doi: 10.1136/bmjopen-2021-053624
30. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* Switzerland: Springer Nature Switzerland AG (2019). p. 558. doi: 10.1007/978-3-030-16399-0
31. Wynants L, Collins GS, Van Calster B. Key steps and common pitfalls in developing and validating risk models. *BJOG.* (2017) 124:423–32. doi: 10.1111/1471-0528.14170
32. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* (2015) 350:g7594. doi: 10.1136/bmj.g7594
33. Pena AS, Metz M. What is adolescent polycystic ovary syndrome? *J Paediatr Child Health.* (2018) 54:351–5. doi: 10.1111/jpc.13821
34. Branch EGAG. Chinese guidelines for diagnosis and treatment of polycystic ovary syndrome. *Chin J Obstet Gynecol.* (2018) 53:2–6. doi: 10.3760/cma.j.issn.0529-567X.2018.01.002
35. Group REAP. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod.* (2004) 19:41–7. doi: 10.1093/humrep/deh098
36. Prevention BODC. *Notice of the Ministry of Health on the Issuance of the "standards for the Management and Treatment of Major Mental Illness (2012 edition)".* (2012). Available online at: <http://www.nhc.gov.cn/jkj/s5888/201204/16ebc49bf504f979eb31070fc3ac5bf.shtml> (accessed August 1, 2022).
37. Biedermann F, Fleischhacker WW. Psychotic disorders in DSM-5 and ICD-11. *CNS Spectr.* (2016) 21:349–54. doi: 10.1017/S109285291600316
38. Liu Z, Li J, Wang Y, Miao M, Zhong J. Structural verification and measurement invariance of Chinese version of children's depression inventory. *Chin J Clin Psychol.* (2019) 27:1172–6. doi: 10.16128/j.cnki.1005-3611.2019.06.019
39. Wu WF, Lu YB, Tan FR, Yao SQ. Reliability and validity of the Chinese version of children's depression inventory. *Chin Ment Health J.* (2010) 24:775–9. doi: 10.3969/j.issn.1000-6729.2010.10.014
40. Hong X, Li JQ, Liang YQ, Wang ZY, Xu F. Reliability and validity of the Chinese version of children's depression inventory. *Chin J Sch Health.* (2012) 33:1182–5. doi: 10.16835/j.cnki.1000-9817.2012.10.011
41. Chen YL, Shen LJ, Gau SS. The Mandarin version of the Kiddie-schedule for affective disorders and schizophrenia-epidemiological version for DSM-5 - a psychometric study. *J Formos Med Assoc.* (2017) 116:671–8. doi: 10.1016/j.jfma.2017.06.013
42. Eysenck HJ, Eysenck S, Eysenck HJ, Eysenck S, Eysenck HJ, Eysenck SB, et al. Manual of the Eysenck personality questionnaire. *J Card Fail.* (1975) 20:S67. doi: 10.1037/t05462-000
43. Gullone E, Hughes EK, King NJ, Tonge B. The normative development of emotion regulation strategy use in children and adolescents: A 2-year follow-up study. *J Child Psychol Psychiatry.* (2010) 51:567–74. doi: 10.1111/j.1469-7610.2009.02183.x
44. Okajima I, Miyamoto T, Ubara A, Omichi C, Matsuda A, Sumi Y, et al. Evaluation of severity levels of the Athens insomnia scale based on the criterion of insomnia severity index. *Int J Environ Res Public Health.* (2020) 17:8789. doi: 10.3390/ijerph17238789
45. Zhou J, Yu J, Zhou Y, Qiu J. Study of item text in the Chinese Symptom Checklist-90. *Medicine.* (2021) 100:e24841. doi: 10.1097/MD.00000000000024841
46. Li D, Yu L, Zhang W, Wang J, Li L. Reliability and validity of medical outcomes study social support survey (Chinese version) in infertile female patients. *J Nurs.* (2020) 27:66–70. doi: 10.16460/j.issn1008-9969.2020.20.066
47. Hatch R, Rosenfield RL, Kim MH, Tredway D. Hirsutism: implications, etiology, and management. *Am J Obstet Gynecol.* (1981) 140:815–30. doi: 10.1016/0002-9378(81)90746-8
48. Cao Z. *Obstet Gynecol.* Beijing: People's Medical Publishing House (2014). p. 3180.
49. Chen H, Jackson T, Huang X. The negative physical self-scale: initial development and validation in samples of Chinese adolescents and young adults. *Body Image.* (2006) 3:401–12. doi: 10.1016/j.bodyim.2006.07.005
50. Basu S, Poole J. The brief illness perception questionnaire. *Occup Med.* (2016) 66:419–20. doi: 10.1093/occmed/kqv203
51. Riley RD, Ensor J, Snell K, Harrell FJ, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* (2020) 368:m441. doi: 10.1136/bmj.m441
52. Erensoy H, Niafar M, Ghafarzadeh S, Aghamohammadzadeh N, Nader ND. A pilot trial of metformin for insulin resistance and mood disturbances in adolescent and adult women with polycystic ovary syndrome. *Gynecol Endocrinol.* (2019) 35:72–5. doi: 10.1080/09513590.2018.1498476