

# Machine learning in female urinary incontinence: A scoping review

DIGITAL HEALTH  
Volume 10: 1–15  
© The Author(s) 2024  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20552076241281450  
journals.sagepub.com/home/dhj



Qi Wang<sup>1,2</sup>, Xiaoxiao Wang<sup>1,2</sup>, Xiaoxiang Jiang<sup>3</sup> and Chaoqin Lin<sup>1,2</sup> 

## Abstract

**Introduction and Hypothesis:** The aim was to conduct a scoping review of the literature on the use of machine learning (ML) in female urinary incontinence (UI) over the last decade.

**Methods:** A systematic search was performed among the Medline, Google Scholar, PubMed, and Web of Science databases using the following keywords: [Urinary incontinence] and [(Machine learning) or (Predict) or (Prediction model)]. Eligible studies were considered to have applied ML model to explore different management processes of female UI. Data analyzed included the field of application, type of ML, input variables, and results of model validation.

**Results:** A total of 798 papers were identified while 23 finally met the inclusion criteria. The vast majority of studies applied logistic regression to establish models (91.3%, 21/23). Most frequently ML was applied to predict postpartum UI (39.1%, 9/23), followed by de novo incontinence after pelvic floor surgery (34.8%, 8/23). There are also three papers using ML models to predict treatment outcomes and three papers using ML models to assist in diagnosis. Variables for modeling included demographic characteristics, clinical data, pelvic floor ultrasound, and urodynamic parameters. The area under receiver operating characteristic curve of these models fluctuated from 0.56 to 0.95, and only 11 studies reported sensitivity and specificity, with sensitivity ranging from 20% to 96.2% and specificity from 59.8% to 94.5%.

**Conclusion:** Machine learning modeling demonstrated good predictive and diagnostic abilities in some aspects of female UI, showing its promising prospects in near future. However, the lack of standardization and transparency in the validation and evaluation of the models, and the insufficient external validation greatly diminished the applicability and reproducibility, thus a focus on filling this gap is strongly recommended for future research.

## Keywords

Machine learning, prediction model, female urinary incontinence, artificial intelligence

Submission date: 3 June 2024; Acceptance date: 20 August 2024

## Introduction

Urinary incontinence (UI) is a relatively common condition that is gaining wider attention because of its potential for significant negative physical and psychological effects on women.<sup>1,2</sup> According to an authoritative review of the literature published in recent years, the results of worldwide epidemiologic surveys suggested a prevalence of UI ranging from 5% to 72%, with large fluctuations due to differences in the populations interviewed, survey methodologies, and diagnostic criteria.<sup>3</sup> A recent large-scale epidemiological survey from China showed that 24.8% of females aged 20–70 suffers from UI.<sup>4</sup> In reality, however, the true prevalence of UI has been difficult to assess correctly because the condition is often underestimated and undiagnosed, and a significant proportion of women hold

<sup>1</sup>College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

<sup>2</sup>Department of Gynecology, Fujian Maternity and Child Health Hospital, Fuzhou, China

<sup>3</sup>Fujian Provincial Key Laboratory of Women and Children's Critical Diseases Research, Fuzhou, China

Xiaoxiao Wang contributed equally to this work.

### Corresponding authors:

Chaoqin Lin, Department of Gynecology, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, No.18 Dao-shan Road, Gu-lou District, Fuzhou 350005, China.

Email: lcqfjsfy@126.com

Xiaoxiang Jiang, Fujian Provincial Key Laboratory of Women and Children's Critical Diseases Research, Fuzhou, China, No. 568 Banzhong Road, Jin'an District, Fuzhou 350012, China.

Email: 19959179535@163.com



the incorrect view that urinary leakage is a natural process that accompanies aging rather than a disease, suggesting that the number of women in the unreported group afflicted with UI may be considerable.<sup>5</sup> How to help these women out is an urgent and challenging task for urogynecologists, and supervised pelvic floor muscle training (PFMT) has been reported to potentially reduce the prevalence of UI in specific populations.<sup>6,7</sup> A variety of risk factors, such as obesity, multiple vaginal deliveries, advanced age, and instrumented delivery, are closely related to the occurrence of UI.<sup>8</sup> However, it is difficult to accurately assess the risk of UI by only obtaining rough information on such risk factors, and blindly implementing early prevention measures without accurately identifying high-risk individuals may lead to a waste of healthcare resources and unnecessary treatment costs, which ultimately leads to the ineffectiveness of early prevention and intervention strategies.<sup>9,10</sup>

Based on the abovementioned, the application of technological tools capable of recognizing the risk of urine leakage in women is constructive and practical. In the past decade, artificial intelligence (AI) has been increasingly applied to various scientific fields, and healthcare is one of the important branches. Such emerging technologies are centered on machine learning (ML) algorithms, which play an increasingly important role in the prediction, prevention, diagnosis, and decision-making of clinical treatment programs for diseases. More importantly, the increasing electronic, digital, and automation degree of healthcare system provides a broader potential application scenario for ML. There is a large volume of literature and reviews on the application of ML in the field of healthcare, covering a wide range of medical specialties from urology, cardiovascular, and geriatrics, suggesting a broad and growing interest in the application of ML in the medical community.<sup>11–15</sup>

Notably, the application of ML in the field of obstetrics and gynecology is quite extensive, covering health care during pregnancy to decision-making on surgical protocols and prognosis, indicating its potential for significant predictive power.<sup>16–19</sup> Despite the high frequency of AI and ML in the literature, it is worth noting that research on their application to female UI is still uncommon. Theoretically, ML has tremendous application prospects in the field of UI prevention and treatment, through the extraction of women's individual characteristics, clinical data and follow-up results, analysis and screening of high-risk factors for diseases, increasing the accuracy of prediction of the occurrence and prognosis of diseases, and assisting the doctors and patients to make clinical decisions together with a view to obtaining better preventive and therapeutic effects. However, clinicians still face non-negligible challenges in applying ML to solve UI problems, which resulted from their lack of systematic understanding of ML technologies and the imperfections of existing ML technologies.

The purpose of this scoping review is to comprehensively map the literature from the last decade (2013–2023) concerning the application of ML techniques in addressing female UI. By cataloging and assessing the breadth of existing research, this review aims to identify key strengths and limitations within the field, highlight existing knowledge gaps, and offer informed insights to guide future research endeavors and clinical applications.

## Methods

Based on the significant heterogeneity of current published research in this field, this study adopted a scoping review, which is a powerful tool to better articulate the current state of research in the field. This study was conducted following the specification requirements of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Review (PRISMA-ScR).<sup>20</sup>

### Search strategy and eligibility

Based on the following keywords we assembled a specialized search character set: [Urinary incontinence] and [(Machine learning) or (Predict) or (Prediction model)]. Subsequently, the Medline, Google Scholar, PubMed, and Web of Science databases were sequentially searched for all relevant papers published in the decade from 2013 to 2023. Inclusion criteria included (1) It is necessary to use ML algorithms in the intervention process of female UI, such as prevention, diagnosis, or prognosis; (2) These studies need to evaluate the models developed and report the corresponding metrics. Exclusion criteria included: (1) Reviews, commentaries, abstracts, book chapters, and animal experimental studies were excluded from this study; (2) The study subjects were not adult females or the topic of the study was not UI; (3) Indicators of model assessment were not reported; and (4) Literature that was not in English or for which the full text was not available.

### Study selection and data extraction

Two authors (Q.W. and X.X.W.) conducted a literature search and review based on the inclusion and exclusion criteria described above: first, the titles, abstracts, keywords, and conclusions of the articles were read to screen for potentially eligible studies. Each potentially eligible study was thoroughly assessed by full-text reading to determine that it met the inclusion criteria. Disagreements regarding whether or not they qualified were considered by additional reviewers (X.X.J and C.Q.L) and resulted in a consensus opinion. References of included studies were also manually screened to avoid the presence of omitted potentially eligible studies. The search for this scoping review was completed in March 2024.

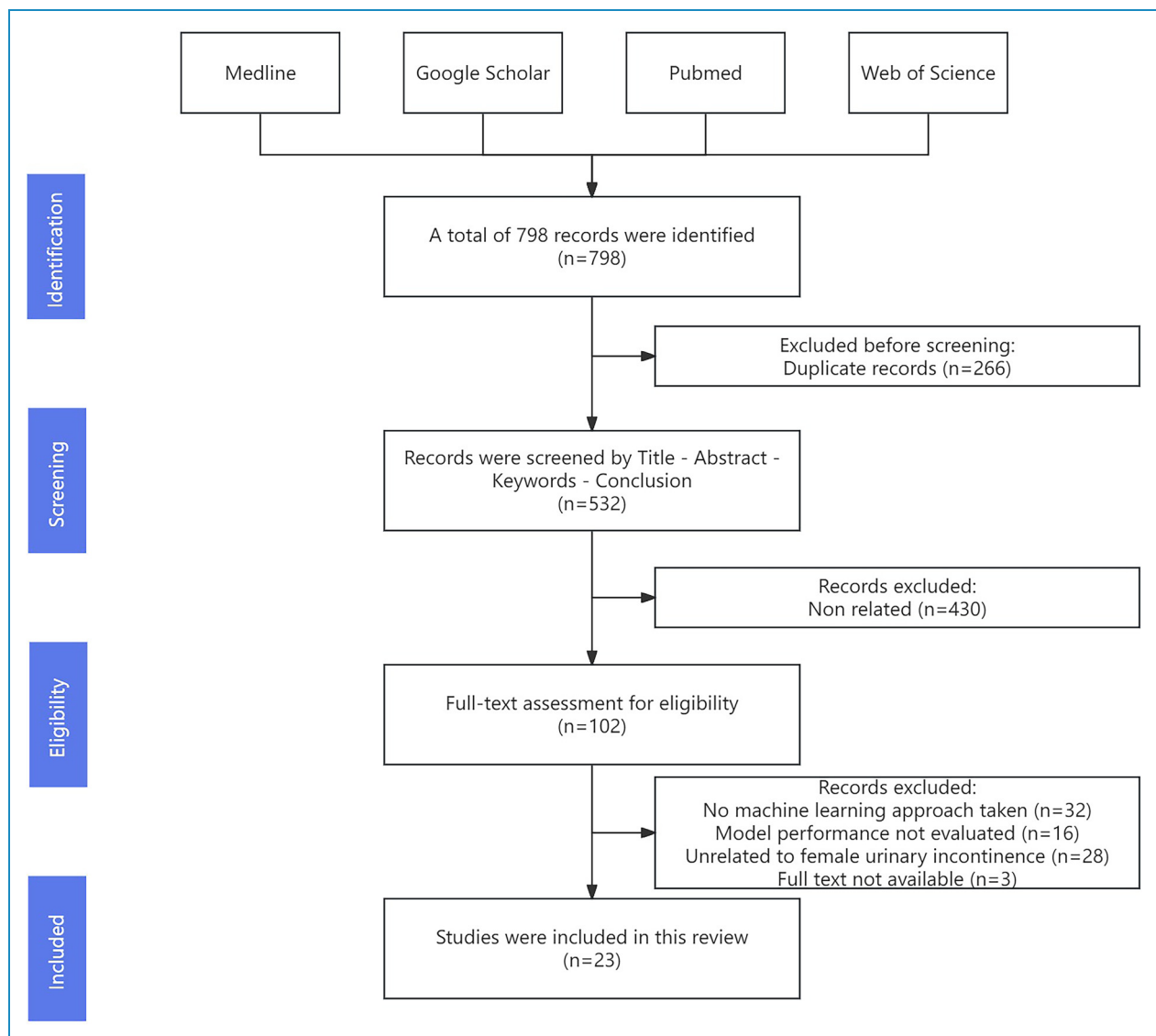
We extracted the following data from the finalized included studies: year of publication, type of study, purpose of study, sample size, and type of ML algorithm. Depending on the purpose of the study, we will further detail the following relevant information: type of incontinence, input variables, model establishment and validation methods, evaluation metrics, and the method of model visualization. It is worth noting that if the study builds multiple models, the effectiveness metrics of the best performing model will be included in the final table.

## Results

The review team found a total of 798 relevant records from Medline, Google Scholar, PubMed, and Web of Science, with 102 papers remaining after removal of duplicate

documents and initial screening. After full-text review, the review team finally included 23 literatures. The process of searching, identifying, and screening the literature is shown in Figure 1. Over the time period of this scoping study (2013–2023), the volume of published literature demonstrated a significant upward trend, with 87.0% (20/23) of the literature published in 2018–2023, indicating the rapidly rising interest in this topic among the scientific community. Table 1 presented information specific to the included studies, including year of publication, type of study, purpose, sample size, type of ML algorithm, study objectives, and assessment of the performance of the model created.

Based on the purpose of the study, these studies can be categorized into four categories: predicting postpartum and pregnancy UI (9, 39.1%), predicting postoperative de novo



**Figure 1.** Flowchart of study identification and inclusion.

**Table 1.** Overview of studies on applications of machine learning in female urinary incontinence.

Year	Reference	Study design	Purpose	Sample size	Machine learning tool	Objective	Accuracy for best model
2013	Jelovsek et al. <sup>21</sup>	Prospective	Prediction	759 patients	Logistic Regression	Predicting the risk of postpartum urinary incontinence in primiparous women based on perinatal clinical data	AUC = 0.69 SE = NR SP = NR
2014	Jelovsek et al. <sup>22</sup>	Prospective	Prediction	773 patients (457 training set, 316 validation set)	Logistic Regression	Predicting de novo incontinence after pelvic organ prolapse surgery by applying 12 preoperative clinical features	AUC = 0.73 SE = NR SP = NR
2016	Jelovsek et al. <sup>23</sup>	Prospective	Prediction	1499 patients (597 training set, 902 validation set)	Logistic Regression	Predicting recurrence of urinary incontinence 12 months after midurethral sling surgery based on patient characteristics and urodynamic parameters	AUC = 0.73 SE = NR SP = NR
2018	Van der Ploeg et al. <sup>24</sup>	Prospective	Prediction	356 patients	Logistic Regression	Predicting the likelihood of de novo incontinence after vaginal prolapse repair surgery based on clinical data	AUC = 0.79 SE = NR SP = NR
2018	Sabadell et al. <sup>25</sup>	Retrospective	Prediction	169 patients	Logistic Regression	External validation of established models to predict de novo incontinence after pelvic organ prolapse surgery	AUC = 0.69 SE = 20.0% SP = 92.6%
2018	Chen et al. <sup>26</sup>	Retrospective	Prediction	521 patients (830 training set, 356 validation set)	Logistic Regression	Predicting the presence of stress urinary incontinence during pregnancy based on pelvic floor ultrasound parameters and clinical data	AUC = 0.79 SE = 78.7% SP = 69.3%
2019	Xiao et al. <sup>27</sup>	Prospective	Diagnosis	337 patients	Fisher linear discriminant analysis	Diagnosis of stress urinary incontinence based on 3D pelvic floor ultrasound parameters	AUC = 0.82 SE = 60.5% SP = 94.5%
2019	Jelovsek et al. <sup>28</sup>	Prospective	Prediction	152 patients	Logistic Regression	External validation of established models using international cohort data to predict de novo incontinence after pelvic organ prolapse surgery	AUC = 0.63 SE = NR SP = NR

(continued)

Table 1. Continued.

Year	Reference	Study design	Purpose	Sample size	Machine learning tool	Objective	Accuracy for best model
2020	Keshavarz et al. <sup>29</sup>	Prospective case-control	Diagnosis	88 patients (44 controls, 44 cases)	Logistic Regression	Diagnosis of stress urinary incontinence in married women based on pelvic floor ultrasound data	AUC = 0.89 SE = 89% SP = 79%
2020	Chen et al. <sup>30</sup>	Prospective	Prediction	727 patients	Logistic Regression	Predicting the likelihood of postpartum stress urinary incontinence in primiparous and multiparous women	AUC = 0.78 SE = NR SP = NR
2020	Yasa et al. <sup>31</sup>	Retrospective	Prediction	225 patients	Logistic Regression	External validation of established models to predict de novo incontinence after pelvic organ prolapse surgery	AUC = 0.56 SE = 41.7% SP = 65.4%
2020	Zhong et al. <sup>32</sup>	Retrospective	Prediction	327 patients	Logistic Regression	Predicting the prognosis of patients with stress urinary incontinence after undergoing TVT surgery	AUC = 0.69 SE = NR SP = NR
2021	Brooks et al. <sup>33</sup>	Prospective	Prediction	77 patients	Logistic Regression	Predicting the efficacy of pelvic floor muscle training in females with stress urinary incontinence	AUC = 0.63 SE = 70% SP = 75%
2021	Zhang et al. <sup>34</sup>	Retrospective	Diagnosis	400 patients (320 training set, 80 test set)	Convolutional neural network	Diagnosis of stress urinary incontinence based on 2D transperineal ultrasound parameters	AUC = 0.92 SE = 75.0% SP = 92.3%
2021	Oh et al. <sup>35</sup>	Retrospective	Prediction	1142 patients (915 training set, 227 validation set)	Logistic Regression	Predicting the likelihood of bothersome stress urinary incontinence one year after pelvic organ prolapse surgery	AUC = 0.78 SE = NR SP = NR
2022	Liu et al. <sup>36</sup>	Prospective	Prediction	255 patients	Logistic Regression	Application of clinical and ultrasound data to predict the risk of postpartum stress urinary incontinence	AUC = 0.88 SE = NR SP = NR
2022	Wang et al. <sup>37</sup>	Prospective	Prediction	1186 patients (830 training set, 356 validation set)	Logistic Regression	Predicting the odds of postpartum urinary incontinence based on patient characteristics	AUC = 0.76 SE = NR SP = NR

(continued)

Table 1. Continued.

Year	Reference	Study design	Purpose	Sample size	Machine learning tool	Objective	Accuracy for best model
2022	Cheng et al. <sup>38</sup>	Retrospective	Prediction	360 patients	Logistic Regression	Predicting the presence of postpartum stress urinary incontinence in primiparous women	AUC = 0.80 SE = NR SP = NR
2022	Van Doorn et al. <sup>39</sup>	Retrospective	Prediction	512 patients	Logistic Regression	Predicting treatment outcomes for pure or predominant urge urinary incontinence based on patient characteristics, disease history, and treatment modalities	AUC = 0.70 SE = NR SP = NR
2023	You et al. <sup>40</sup>	Prospective case-control	Prediction	133 patients (81 controls, 52 cases)	Logistic Regression	Predicting the risk of postpartum stress urinary incontinence in primiparous women based on postpartum MRI parameters	AUC = 0.95 SE = 96.2% SP = 86.4%
2023	Xu et al. <sup>41</sup>	Retrospective	Prediction	3051 patients (2441 training set, 610 validation set)	Logistic Regression	Predicting the likelihood of stress urinary incontinence after vaginal delivery based on clinical data	AUC = 0.85 SE = 85% SP = 72%
2023	Liang et al. <sup>42</sup>	Retrospective	Prediction	660 patients	Logistic Regression	Predicting the presence of stress urinary incontinence during pregnancy by applying patient characteristics	AUC = 0.79 SE = 79.9% SP = 65.9%
2023	Fu et al. <sup>43</sup>	Retrospective	Prediction	555 patients (445 training set, 110 validation set)	Logistic regression, random forest, XGBoost	Predicting the likelihood of stress urinary incontinence 1 year after undergoing prolapse surgery based on patient characteristics and clinical data	AUC = 0.70 SE = 78.3% SP = 59.8%

Abbreviations: NR: not reported; AUC: areas under the curve; SE: sensitivity; SP: specificity; RKKS: reproducing kernel Krein space; TVT: tension free vaginal tape; MRI: magnetic resonance imaging; XGBoost: extreme gradient boosting.

UI (8, 34.8%), predicting the outcome of UI treatment (3, 13.0%) and assisted Diagnostics of UI (3, 13.0%). Retrospective and prospective studies were included in roughly equal proportions (47.8% vs. 52.2%), and the sample sizes included in these studies ranged from 77 to 3051, with a mean sample size of 620. The ML algorithms adopted included Fisher linear discriminant analysis, random forests, logistic regression, convolutional neural network, and XGBoost (eXtreme gradient boosting), but

logistic regression clearly dominated among them as 91.3% (21/23) of the studies selected to apply logistic regression. It is particularly noteworthy that the two studies which did not use logistic regression employed ultrasound parameters and ML algorithms to aid in the diagnosis of UI.

The input variables varied greatly among these models, depending on the target and purpose of the prediction, and could be broadly categorized into patient characteristics,

clinical data, ultrasound and urodynamic parameters, and so on. These studies took a variety of methods to evaluate the model prediction efficacy, including setting up the validation set independently, K-fold cross-validation, and bootstrapping method. All of these models reported area under receiver operating characteristic curve (AUC) with values fluctuating from 0.59 to 0.95, unfortunately more than half of the studies (12/23, 52.2%) did not disclose further details of model efficacy such as sensitivity and specificity, the remaining models reported sensitivity fluctuating from 20% to 96.2% and specificity fluctuating from 59.8% to 94.5%, it is worth noting that the unsatisfactory data were all from external validation studies applying additional cohorts to the original models. The following is a detailed analysis of each of the four categories of models according to research intent:

### Predicting postpartum and pregnancy UI

A total of nine papers<sup>21,26,30,36–38,40–42</sup> applied ML to develop models to predict the occurrence of UI during pregnancy and postpartum. As demonstrated in Table 2, all of these literatures employed logistic regression, and the predictors were mostly stress UI (SUI), but they had a variety of different input predictors, which can be summarized and categorized into the following categories, such as basic characteristics including (age, body mass index [BMI], race, level of education, and income), previous maternal history (parity, mode of delivery, infant weight, forceps deliveries, and the presence of urinary leakage during pregnancy), and auxiliary findings (ultrasound and magnetic resonance imaging [MRI]). Chen et al.<sup>26</sup> established a prediction model by combining clinical data and ultrasound parameters, and the final variables included in the model included bladder neck (BN) funneling and  $\beta$  angle at rest, in addition to BMI gain, constipation, previous delivery mode, and the model achieved an AUC of 0.79 with a sensitivity of 78.7% and a specificity of 69.3%.

As the only study that adopted MRI parameters, You et al.<sup>40</sup> developed a model using MRI measurements of retrovesicourethral angle during straining, functional urethral length during straining, bladder funnel. After cross-validation, the model reached an AUC of 0.95, a sensitivity of 96.2%, and a specificity of 86.4%, which were the best among the models in this series, suggesting that the application of MRI imaging may have potential advantages and prospects for the prediction of postpartum UI, but unfortunately, this study did not give a visual and interpretable approach, whereas all other models in the series used a nomogram approach to make the model interpretable. It is also difficult to compare the predictive efficacy of the models because of the great variety of variables included in each model.

### Predicting postoperative de novo UI

Eight studies were conducted on predicting de novo UI after pelvic floor repair surgery,<sup>22–25,28,31,35,43</sup> most of which

focused on SUI, and the variables included in the models can be categorized into two categories: basic characteristics (including age, BMI, and number of vaginal deliveries), perioperative clinical data (e.g., preoperative stress test and leakage, type of prolapse surgery, and concomitant anti-incontinence surgery), and detailed data are summarized in Table 3. Notably, in 2014, Jelovsek et al.<sup>22</sup> first successfully constructed a model to predict de novo SUI 12 months after pelvic floor repair using the aforementioned variables and logistic regression, and its AUC reached 0.73 and 0.62 after internal and external validation, respectively. The researchers visualized this model in the form of a nomogram, since then, scholars have carried out external validation of this model based on multiple cohorts in different locations,<sup>25,28,31,35</sup> and the AUC fluctuated from 0.56 to 0.69, making this model the most frequently externally validated model in the field. In addition to this, Oh et al.<sup>35</sup> validated the original model using a prospective cohort and developed a novel model based on this cohort, which showed a significant increase in AUC value compared to the original model (0.74 vs. 0.63). In other independent studies, age, BMI, and parity were the most frequent predictors, while logistic regression remained the most common algorithm, with AUCs for these models fluctuating from 0.70 to 0.79.

### Predicting the outcome of UI treatment

As demonstrated in Table 4, only three studies<sup>32,33,39</sup> were relevant for predicting the effect of UI treatment, but they all focused on different branches. Zhong et al.<sup>32</sup> incorporated urodynamic examination parameters such as maximal urethral closure pressure and Valsalva leak point pressure in establishing a model to predict the efficacy of anti-incontinence surgery, and the resulting model had an AUC of 0.69. Another study predicted the therapeutic effect of PFMT, and the final inclusion in the model was the following four predictors: the International Consultation on Urinary Incontinence Questionnaire Urinary Incontinence Short Form, pelvic floor muscle tone, BN height during quiet standing, and BN height during standing cough, with the model having an AUC of 0.80, a sensitivity of 70%, and a specificity of 75%.<sup>33</sup> The last study predicted the corresponding efficacy of responding to urgency UI (UUI) with pharmacological, conservative, and invasive treatments, respectively, with a model AUC of 0.70, which may be helpful in the choice of treatment for UUI, but unfortunately this study did not provide a visualization tool.

### Assisted diagnostics of UI

Only three studies<sup>27,29,34</sup> explored the possibility of applying ML to assist in the diagnosis of UI, and there are some commonalities in these studies, such as the predicted events all assisting in the diagnosis of SUI, and it is clear that

**Table 2.** Predicting postpartum and pregnancy UI.

Year	Reference	Sample size	ML tool	Type of UI	Input variables	Validation methods	Accuracy for best model	Interpretation methods
2013	Jelovsek et al. <sup>21</sup>	759 patients	LR	UI	Race, UI before pregnancy, UI during pregnancy, prepregnancy BMI, predelivery BMI, maternal age, planned mode of delivery	Bootstrapping and cross-validation	AUC = 0.69 SE = NR SP = NR	Nomogram
2018	Chen et al. <sup>26</sup>	521 patients (830 training set, 356 validation set)	LR	SUI	BMI gain, constipation, previous delivery mode, BN funneling, $\beta$ angle at rest	Bootstrapping and independent validation set	AUC = 0.79 SE = 78.7% SP = 69.3%	Nomogram
2020	Chen et al. <sup>30</sup>	727 patients	LR	SUI	Age, SUI during pregnancy, abortion history, mode of delivery	Bootstrapping	AUC = 0.78 SE = NR SP = NR	Nomogram
2022	Liu et al. <sup>36</sup>	255 patients	LR	SUI	Age, parity, angle of internal urethral orifice funnel, BND, mode of delivery	Independent validation set	AUC = 0.88 SE = NR SP = NR	Nomogram
2022	Wang et al. <sup>37</sup>	1186 patients (830 training set, 356 validation set)	LR	UI	Residence, mode of delivery, age at first birth, parity, UI before pregnancy, UI during pregnancy, feeding pattern	Independent validation set	AUC = 0.76 SE = NR SP = NR	Nomogram
2022	Cheng et al. <sup>38</sup>	360 patients	LR	SUI	Gravidity, residence, occupation, education, monthly income, mode of delivery, oxytocin	Cross-validation	AUC = 0.80 SE = NR SP = NR	Nomogram
2023	You et al. <sup>40</sup>	133 patients (81 controls, 52 cases)	LR	SUI	Retrovesicourethral angle during straining, functional urethral length during straining, bladder funnel measured by MRI	Cross-validation	AUC = 0.95 SE = 96.2% SP = 86.4%	NR
2023	Xu et al. <sup>41</sup>	3051 patients (2441 training set, 610 validation set)	LR	SUI	Age, parity, infant weight, duration of second stage of labor, forceps delivery	Cross-validation And independent validation set	AUC = 0.85 SE = 85% SP = 72%	Nomogram
2023	Liang et al. <sup>42</sup>	660 patients	LR	SUI	Constipation, education level, previous delivery mode	Bootstrapping	AUC = 0.79 SE = 79.9% SP = 65.9%	Nomogram

Abbreviations: UI: urinary incontinence; SUI: stress urinary incontinence; ML: machine learning; NR: not reported; AUC: areas under the curve; SE: sensitivity; SP: specificity; LR: logistic regression; SUI: stress urinary incontinence; BMI: body mass index; BN: bladder neck; BND: bladder neck descent; MRI: magnetic resonance imaging.

ultrasound parameters were quite important for these studies, details are shown in Table 5. Xiao et al.<sup>27</sup> developed several prediction models by utilizing different combinations of four ultrasound parameters with good results,

namely, BN position on maximal Valsalva maneuver, levator hiatus area on maximal Valsalva maneuver, BN descent, urethral rotation angle, and finally the model with all variables was validated to have the best predictive



**Table 3.** Predicting postoperative de novo UI.

Year	Reference	Sample size	ML tool	Type of UI	Input variables	Validation methods	Accuracy for best model	Interpretation methods
2014	Jelovsek et al. <sup>22</sup>	773 patients (457 training set, 316 validation set)	LR	SUI	Age, number of vaginal births, BMI, preoperative stress test, continence procedure, urine leakage associated with a feeling of urgency, diabetes	Bootstrapping and independent validation set	AUC = 0.73 SE = NR SP = NR	Nomogram
2016	Jelovsek et al. <sup>23</sup>	1499 patients (597 training set, 902 validation set)	LR	UI	Age, race, parity, menopausal status, Urinary Distress Inventory irritative, Urinary Distress Inventory stress, Urinary Distress Inventory obstructive, feeling of incomplete emptying, leaking related to urgency, and limitations in entertainment activities	Bootstrapping and independent validation set	AUC = 0.73 SE = NR SP = NR	NR
2018	Van der Ploeg et al. <sup>24</sup>	356 patients	LR	SUI	Age, vaginal parity, point Ba of the POP-Q, subjective UI, vaginal prolapse repair without MUS	Bootstrapping	AUC = 0.79 SE = NR SP = NR	Score chart
2018	Sabadell et al. <sup>25</sup>	169 patients	LR	SUI	Age at surgery, parity, BMI, preoperative stress test, presence of urge UI, association with an MUS procedure	Independent validation set	AUC = 0.69 SE = 20.0% SP = 92.6%	NR
2019	Jelovsek et al. <sup>28</sup>	152 patients	LR	SUI	Age, number of vaginal births, BMI, preoperative stress test, continence procedure, urine leakage	Independent validation set	AUC = 0.63 SE = NR SP = NR	NR

(continued)

Table 3. Continued.

Year	Reference	Sample size	ML tool	Type of UI	Input variables	Validation methods	Accuracy for best model	Interpretation methods
					associated with urgency, diabetes			
2020	Yasa et al. <sup>31</sup>	225 patients	LR	SUI	Age, number of vaginal births, BMI, preoperative stress test, continence procedure, urine leakage associated with a feeling of urgency, diabetes	Independent validation set	AUC = 0.56 SE = 41.7% SP = 65.4%	NR
2021	Oh et al. <sup>35</sup>	1142 patients (915 training set, 227 validation set)	LR	SUI	Age, diabetes, subjective UI, prolapse reduction stress test, type of prolapse surgery, concomitant MUS	Cross-validation and independent validation set	AUC = 0.78 SE = NR SP = NR	Nomogram
2023	Fu et al. <sup>43</sup>	555 patients (445 training set, 110 validation set)	LR RF XGBoost	SUI	Age, BMI, parity, point C, Aa and Ba of the POP-Q, preoperative UI	Cross-validation and independent validation set	AUC = 0.70 SE = 78.3% SP = 59.8%	NR

Abbreviations: UI: urinary incontinence; ML: machine learning; NR: not reported; AUC: areas under the curve; SE: sensitivity; SP: specificity; LR: logistic regression; RF: random forest; XGBoost: extreme gradient boosting; POP-Q: pelvic organ prolapse quantification; MUS: midurethral sling.

efficacy, with an AUC of 0.82, a sensitivity of 60.5% and a specificity of 94.5%. Keshavarz et al.<sup>29</sup> also utilized ultrasound parameters for prediction, with the difference that their model was simpler, and their results found that a  $\beta$  angle higher than  $127^\circ$  with the Valsalva maneuver, was a strong predictor with an AUC of 0.89, with 89% sensitivity and 79% specificity. The last study<sup>34</sup> applied convolutional neural network algorithm to build an AI image recognition system to predict the occurrence of SUI by recognizing ultrasound images, this simple and pioneering method has good predictive efficacy with AUC of 0.92, sensitivity of 75.0% and specificity of 92.3%.

## Discussion

The pursuit of human beings to predict the occurrence and prognosis of diseases has a long history, from the ancient times when phenomena were summarized as experiences, and the development of statistics in the last century which provided more effective methods for this purpose, until in the last decade, the high-speed development of AI and big data systems has greatly assisted health professionals in exploring the intrinsic developmental patterns of specific

diseases. Since then, an increasing number of predictive models based on complex databases and diverse information emerged in anticipation of providing more accurate predictions of disease prevention, treatment, prognosis, and follow-up.<sup>44–48</sup>

Since the latest review on predicting female UI was published nearly a decade ago,<sup>49,50</sup> an update on this is necessary, and this study provides an up-to-date overview of the application of ML algorithms and techniques to predicting female UI. Because there are multiple domains of female UI, with marked differences in the purpose, applicable populations, and methodologies of the corresponding models, as well as significant heterogeneity in the variables entered, it is difficult to conduct a systematic review and meta-analysis of the literature for this category. In order to provide a more in-depth and intuitive understanding of these literatures, this study recorded a series of detailed information about these models: year of publication, type of study, purpose, sample size, type of ML algorithm, objectives of the study, and an assessment of the performance of the constructed models. Based on the purpose of the predictive models, this review classified the included literature into four categories, which were further analyzed in

**Table 4.** Predicting the outcome of UI treatment.

Year	Reference	Sample size	ML tool	Type of UI	Input variables	Validation methods	Accuracy for best model	Interpretation methods
2020	Zhong et al. <sup>32</sup>	327 patients	LR	SUI	Total cholesterol, maximal urethral closure pressure	Cross-validation	AUC = 0.69 SE = NR SP = NR	Nomogram
2021	Brooks et al. <sup>33</sup>	77 patients	LR	SUI	Bladder neck height in a quiet standing position, bladder neck height during a cough in a standing position, PFM tone, and ICIQ-UI-SF score	Bootstrapping	AUC = 0.63 SE = 70% SP = 75%	Logistic regression equation
2022	Van Doorn et al. <sup>39</sup>	512 patients	LR	UUI	Number of incontinence episodes per day, voiding frequency during the day, subjective quantity of UI, coexistence of SUI, night incontinence, and bladder capacity	Bootstrapping	AUC = 0.70 SE = NR SP = NR	NR

Abbreviations: UI: urinary incontinence; SUI: stress urinary incontinence; UUI: urgency urinary incontinence; ML: machine learning; NR: not reported; AUC: areas under the curve; SE: sensitivity; SP: specificity; LR: logistic regression; PFM: pelvic floor muscle; ICIQ-UI-SF: the International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form.

**Table 5.** Assisted diagnostics of UI.

Year	Reference	Sample size	ML tool	Type of UI	Input variables	Validation methods	Accuracy for best model	Interpretation methods
2019	Xiao et al. <sup>27</sup>	337 patients	FLDA	SUI	BNP, LHA, BND, URA	Cross-validation	AUC = 0.82 SE = 60.5% SP = 94.5%	NR
2020	Keshavarz et al. <sup>29</sup>	88 patients (44 controls, 44 cases)	LR	SUI	BND, $\beta$ angles with and without the Valsalva Maneuver	Cross-validation	AUC = 0.89 SE = 89% SP = 79%	NR
2021	Zhang et al. <sup>34</sup>	400 patients (320 training set, 80 test set)	CNN	SUI	Transperineal ultrasound images	Cross-validation	AUC = 0.92 SE = 75.0% SP = 92.3%	NR

Abbreviations: UI: urinary incontinence; SUI: stress urinary incontinence; ML: machine learning; NR: not reported; AUC: areas under the curve; SE: sensitivity; SP: specificity; LR: logistic regression; FLDA: Fisher linear discriminant analysis; CNN: convolutional neural network; BNP: bladder neck position on maximal Valsalva maneuver; LHA: levator hiatus area on maximal Valsalva maneuver; BND: bladder neck descent; URA: urethral rotation angle.

detail for their subtypes of predicted UI, input variables, validation methods, and visualization pathways.

Through this scoping review, we found that sociodemographic factors such as age, BMI, education, and income level proved to be important in predicting de novo UI in the postpartum period and after pelvic floor repair surgery, and that equally important predictors of postpartum UI

were obstetrical-related factors such as mode of delivery, infant weight, and instrumented or noninstrumented delivery, and the presence of urinary leakage during pregnancy. Ultrasound parameters could also assist in predicting the development of postpartum UI. Although there was less literature on assisting in the diagnosis of UI, it was also possible to intuitively realize the crucial role that ultrasound

and MRI examinations play in this program, and such studies can intelligently analyze ultrasound and MRI images to get an accurate prediction. In addition to this, some studies have combined clinical data and ultrasound parameters to create models that can be used to predict the outcome of UI patients after undergoing surgery, drugs, and PFMT, respectively.

This is a timely scoping review that provides clinicians and analysts with a broader and deeper understanding of the application of ML to female UI, and the fact that the majority of studies (87.0%) were published in the last five years is indicative of the growing interest in this topic in the medical community. Currently, ML-based UI prediction studies have focused on the assessment of UI risk in specific populations (e.g., postpartum and after pelvic floor repair surgeries), and these ML-based prediction models are very interesting for both pregnant women (patients) and physicians, as they have the opportunity to improve the prognosis of the patients by identifying the corresponding high-risk populations and delivering early interventions to improve their quality of life as well as to reduce the associated healthcare costs. For health service providers, empowering them with simple and more accurate primary screening tools can enable precise management of high-risk populations, undoubtedly contributing to the saving of medical resources while improving the effectiveness of interventions. At the same time, the development of telemedicine assessments, telemedicine, and the proliferation of wearable devices may facilitate the collection and processing of medical data, and such solutions may be of great importance to women living in developing countries and in rural areas, where access to health care may be limited mainly by socioeconomic factors.

Although the application of ML modeling for UI diagnosis and efficacy assessment is still relatively limited, it is not surprising that this topic will be the most promising development in the field in the coming years. The clinical decision-making process of humans can often be accompanied by errors, biases, or shackled by personal experience paths,<sup>51</sup> and assisted decision-making systems based on predictive modeling can help clinicians reduce the risk of misdiagnosis and incorrect treatment. From the patient's perspective, when consulting with a physician about treatment options, compared with simply being informed of possible risk factors and prognosis in general, predictive modeling can provide individualized risk and prognosis information based on the patient's situation, which is undoubtedly more conducive to the patient's understanding of his or her own situation and decision-making together with the physician.

The application of ML to female UI is still in its infancy, although some of its current clinical applications show its future potential to revolutionize the prevention, diagnosis, treatment, and prognosis of UI in women. In order for the corresponding predictive model to be suitable for wider clinical implementation, it must be accurate and

generalizable. There are several important shortcomings that should not be overlooked when using ML techniques to achieve these goals.

The first is that the process of model development, validation, and evaluation should follow appropriate standards. As more and more predictive models emerged, the need to develop standards to increase the accuracy and credibility of research results became more and more urgent. Because some ML algorithms are often considered a black box because it is difficult to explain how a prediction is derived,<sup>52</sup> the lack of transparency in the process of model development and validation will undoubtedly weaken the credibility of the model's prediction results and reduce clinicians' willingness to apply the model. An important milestone was the publication of the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines in 2015,<sup>53</sup> a guideline that sought to ensure that researchers provide sufficient information when reporting on prediction modeling studies to enable others to understand, evaluate, and reproduce the results. By following the TRIPOD guidelines, researchers can improve the reliability and transparency of predictive models and promote their effective application in healthcare and other fields. Unfortunately, there are still many studies that lack sufficiently transparent descriptions of model development and the necessary evaluation metrics in the model evaluation stage. More than half of the literature in this review (12/23, 52.2%) only reported the AUC of the model and lacked further details about the model performance, such as sensitivity, specificity, accuracy, positive predictive value, and negative predictive value, which undoubtedly weakened the credibility of the model.

Secondly, although a few studies have combined multiple ML algorithms, overall, the current ML algorithms applied to the field of female UI are very homogeneous, and the proportion of studies that applied logistic regression in this review was as high as 91.3% (21/23). Although this approach has been widely used, it has certain deficiencies in dealing with multidimensional data and feature interactions, and this deficiency cannot be ignored in the era of big data. In fact, emerging ML algorithms in recent years, such as XGBoost, support vector machines, and convolutional neural network, have demonstrated excellent performance in prediction tasks in other medical scenarios.<sup>34,54–56</sup> Applying multiple algorithms to model separately and then selecting the best performer from them seems to be an ideal solution.

By screening the existing literature, we found that many models (43.5%, 10/23) on female UI prediction lacked effective visualization tools. This phenomenon limits the generalization and application of predictive models in clinical practice. Often, the use of visualization tools such as nomogram or web calculators can make complex statistical models intuitive and understandable, thus facilitating their use by clinicians and researchers.

In addition, the visualization of the model helps other researchers to externally validate and improve the original model. External validation is an indispensable step in determining the reliability and applicability of a model, which means validating the predictive model using at least one other dataset separate from the development dataset. While the methods of internal validation are well established, this does not replace the role of external validation. While the population for which a predictive model is applicable should be explicitly characterized, ideally, predictive models should be applicable to patients from a wide range of races, ethnicities, and backgrounds that are common in clinical practice. The predictive model for post-operative de novo UI developed by Jelovsek et al.<sup>22</sup> in 2014 was validated by several external cohorts worldwide,<sup>25,28,31,35</sup> the key reason for which is that the availability of a nomogram in this study allowed for a significant enhancement of the model's usability. Therefore, we suggest that future studies should emphasize and incorporate effective visualization methods when developing predictive models in order to improve the usability and interpretability of the models and facilitate their widespread use and validation in clinical work.

## Conclusion

This review provides a timely summary of the current status of the development of applying ML to female UI. The increasing interest of the scientific community in applying ML techniques to the prevention, diagnosis, treatment, and prognosis phases of female UI suggests that ML may have a meaningful impact on the field of female UI. Future research in this field should employ more diverse ML algorithms while providing clearer and more transparent descriptions of the model building process and validated results. These models also need to provide effective visualization tools to facilitate large-scale external validation to ensure the applicability of the models. To quote a popular saying, the future mass adoption of AI and predictive modeling will not replace the role of the specialist, but those who understand and can use AI and predictive modeling techniques may replace those who cannot.

**Acknowledgements:** The authors are grateful for the fund support provided by Fujian Provincial Health Commission.

**Contributorship:** Qi Wang: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft. Xiaoxiao Wang: Data Curation, Formal Analysis, Investigation, Methodology. Xiaoxiang Jiang: Formal Analysis, Investigation, Writing – review & editing. ChaoQin Lin: Formal Analysis, Investigation, Writing – review & editing. All authors approved the final article.

**Data availability:** The data of this study are available from the corresponding author upon reasonable request.

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval:** This observational study was conducted in full compliance with the principles of the Declaration of Helsinki. Since this study was a scoping review and did not incorporate any new data, a waiver was obtained from the institutional ethics review board.

**Funding:** The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Joint Funds for the innovation of science and Technology, Fujian province (Grant number: 2023Y9367).

**Guarantor:** XJ

**ORCID ID:** Chaoqin Lin  <https://orcid.org/0000-0003-4631-0218>

**Supplemental material:** Supplemental material for this article is available online.

## References

1. Favre-Inhofer A, Dewaele P, Millet P, et al. Systematic review of guidelines for urinary incontinence in women. *J Gynecol Obstet Hum* 2020; 49: 101842.
2. Lukacz ES, Santiago-Lastra Y, Albo ME, et al. Urinary incontinence in women: a review. *JAMA* 2017; 318: 1592–1604.
3. Aoki Y, Brown HW, Brubaker L, et al. Urinary incontinence in women. *Nat Rev Dis Primers* 2017; 3: 1–20.
4. Wang Q, Que YZ, Wan XY, et al. Prevalence, risk factors, and impact on life of female urinary incontinence: an epidemiological survey of 9584 women in a region of southeastern China. *Risk Manag Healthc Policy* 2023; 16: 1477–1487.
5. Wang Q, Que Y, Yang Y, et al. A population-based cross-sectional survey on the prevalence, severity, risk factors, and self-perception of female urinary incontinence in rural Fujian, China. *Int Urogynecol J* 2023; 34: 2089–2097.
6. Ko PC, Liang CC, Chang SD, et al. A randomized controlled trial of antenatal pelvic floor exercises to prevent and treat urinary incontinence. *Int Urogynecol J* 2011; 22: 17–22.
7. Wesnes SL and Lose G. Preventing urinary incontinence during pregnancy and postpartum: a review. *Int Urogynecol J* 2013; 24: 889–899.
8. Wang Q, Jiang XX, Que YZ, et al. Development and validation of a risk prediction model for female stress urinary incontinence in rural Fujian, China. *Risk Manag Healthc Policy* 2024; 17: 1101–1112.
9. Dufour S and Wu M. No. 397–conservative care of urinary incontinence in women. *J Obstet Gynaecol Can* 2020; 42: 510–522.
10. Nambiar AK, Bosch R, Cruz F, et al. EAU Guidelines on assessment and nonsurgical management of urinary incontinence. *Eur Urol* 2018; 73: 596–609.

11. Haug CJ and Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med* 2023; 388: 1201–1208.
12. Wehbe RM, Katsaggleos AK, Hammond KJ, et al. Deep learning for cardiovascular imaging: a review. *JAMA Cardiol* 2023; 8: 1089–1098.
13. Salem H, Soria D, Lund JN, et al. A systematic review of the applications of expert systems (ES) and machine learning (ML) in clinical urology. *BMC Med Inform Decis Mak* 2021; 21: 1–36.
14. Checcucci E, De Cillis S, Granato S, et al. Applications of neural networks in urology: a systematic review. *Curr Opin Urol* 2020; 30: 788–807.
15. Shiwani T, Relton S, Evans R, et al. New horizons in artificial intelligence in the healthcare of older people. *Age Ageing* 2023; 52: afad219.
16. Oprescu AM, Miro-Amarante G, García-Díaz L, et al. Artificial intelligence in pregnancy: a scoping review. *IEEE Access* 2020; 8: 181450–181484.
17. Jeong GH. Artificial intelligence, machine learning, and deep learning in women's health nursing. *Korean J Women Health Nurs* 2020; 26: 5–9.
18. Ramakrishnan R, Rao S and He JR. Perinatal health predictors using artificial intelligence: a review. *Womens Health* 2021; 17: 1–7.
19. Kannaiyan A, Bagchi S, Vijayan V, et al. Revolutionizing women's health: artificial intelligence's impact on obstetrics and gynecology. *J South Asian Fed Obstet Gynecol* 2024; 16: 161–168.
20. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473.
21. Jelovsek JE, Piccorelli A, Barber MD, et al. Prediction models for postpartum urinary and fecal incontinence in primiparous women. *Urogynecology* 2013; 19: 110–118.
22. Jelovsek JE, Chagin K, Brubaker L, et al. A model for predicting the risk of de novo stress urinary incontinence in women undergoing pelvic organ prolapse surgery. *Obstet Gynecol* 2014; 123: 279–287.
23. Jelovsek JE, Hill AJ, Chagin KM, et al. Predicting risk of urinary incontinence and adverse events after midurethral sling surgery in women. *Obstet Gynecol* 2016; 127: 330–340.
24. van der Ploeg JM, Steyerberg EW, Zwolsman SE, et al. Stress urinary incontinence after vaginal prolapse repair: development and internal validation of a prediction model with and without the stress test. *Neurourol Urodyn* 2019; 38: 1086–1092.
25. Sabadell J, Salicrú S, Montero-Armengol A, et al. External validation of de novo stress urinary incontinence prediction model after vaginal prolapse surgery. *Int Urogynecol J* 2019; 30: 1719–1723.
26. Chen L, Luo D, Yu X, et al. Predicting stress urinary incontinence during pregnancy: combination of pelvic floor ultrasound parameters and clinical factors. *Acta Obstet Gynecol Scand* 2018; 97: 966–975.
27. Xiao T, Xiao T, Chen Y, et al. Can stress urinary incontinence be predicted by ultrasound? *Am J Roentgenol* 2019; 213: 1163–1169.
28. Jelovsek JE, van der Ploeg JM, Roovers JP, et al. Validation of a model predicting de novo stress urinary incontinence in women undergoing pelvic organ prolapse surgery. *Obstet Gynecol* 2019; 133: 683–690.
29. Keshavarz E, Pouya EK, Rahimi M, et al. Prediction of stress urinary incontinence using the retrovesical ( $\beta$ ) angle in transperineal ultrasound. *J Ultrasound Med* 2021; 40: 1485–1493.
30. Chen L, Luo D, Chen X, et al. Development of predictive risk models of postpartum stress urinary incontinence for primiparous and multiparous women. *Urol Int* 2020; 104: 824–832.
31. Yasa C, Gungor Ugurlucan F, Dural O, et al. External validation of a model predicting de novo stress urinary incontinence after pelvic organ prolapse surgery. *Neurourol Urodyn* 2021; 40: 688–694.
32. Zhong C, Pan L, Deng Z, et al. Nomogram for preoperative estimation of prognosis after retropubic tension free vaginal tape in female patients with stress urinary incontinence. *Ann Palliat Med* 2021; 10: 3684–3691.
33. Brooks KCL, Varette K, Harvey MA, et al. A model identifying characteristics predictive of successful pelvic floor muscle training outcomes among women with stress urinary incontinence. *Int Urogynecol J* 2021; 32: 719–728.
34. Zhang M, Lin X, Zheng Z, et al. Artificial intelligence models derived from 2D transperineal ultrasound images in the clinical diagnosis of stress urinary incontinence. *Int Urogynecol J* 2022; 33: 1179–1185.
35. Oh S, Lee S, Hwang WY, et al. Development and validation of a prediction model for bothersome stress urinary incontinence after prolapse surgery: a retrospective cohort study. *BJOG* 2022; 129: 1158–1164.
36. Liu W and Qian L. Establishment and validation of a risk prediction model for postpartum stress urinary incontinence based on pelvic floor ultrasound and clinical data. *Int Urogynecol J* 2022; 33: 3491–3497.
37. Wang X, Jin Y, Xu X, et al. Development and validation of a predictive model for urinary incontinence postpartum: a prospective longitudinal study. *Int Urogynecol J* 2022; 33: 1609–1615.
38. Cheng H, Gong F, Shen Y, et al. A nomogram model predicting the risk of postpartum stress urinary incontinence in primiparas: a multicenter study. *Taiwan J Obstet Gynecol* 2022; 61: 580–584.
39. van Doorn T, Reuvers SHM, Roobol MJ, et al. Development of a prediction model in female pure or predominant urge urinary incontinence: a retrospective cohort study. *Ther Adv Urol* 2022; 14: 1–15.
40. You C, Zhao Y, Zhang C, et al. Pelvic floor parameters predict postpartum stress urinary incontinence: a prospective MRI study. *Insights Imaging* 2023; 14: 160.
41. Xu C, Guo Y, Chi X, et al. Establishment and validation of a simple nomogram for predicting early postpartum stress urinary incontinence among women with vaginal delivery: a retrospective study. *BMC Public Health* 2023; 23: 8.
42. Liang S, Huang S, Andarini E, et al. Development and internal validation of a risk prediction model for stress urinary incontinence throughout pregnancy: a multicenter retrospective longitudinal study in Indonesia. *Neurourol Urodyn* 2024; 43: 354–363.
43. Fu L, Huang G, Sun Z, et al. Predicting the occurrence of stress urinary incontinence after prolapse surgery: a machine learning-based model. *Ann Transl Med* 2023; 11: 251.
44. Niigaki DI, Silva RSP, Bortolini MAT, et al. Predictors for long-term adherence to vaginal pessary in pelvic organ prolapse: a prospective study. *Int Urogynecol J* 2022; 33: 3237–3246.

45. Labrie J, Lagro-Janssen ALM, Fischer K, et al. Predicting who will undergo surgery after physiotherapy for female stress urinary incontinence. *Int Urogynecol J* 2015; 26: 329–334.
  46. Chen Z, Mikhail SM, Buttini M, et al. Online prediction tool for female pelvic floor dysfunction: development and validation. *Int Urogynecol J* 2021; 33: 1–9.
  47. Goldstein A and Cohen S. Self-report symptom-based endometriosis prediction using machine learning. *Sci Rep* 2023; 13: 5499.
  48. Salz T, Baxi SS, Raghunathan N, et al. Are we ready to predict late effects? A systematic review of clinically useful prediction models. *Eur J Cancer* 2015; 51: 758–766.
  49. Troko J, Bach F and Tooze-Hobson P. Predicting urinary incontinence in women in later life: a systematic review. *Maturitas* 2016; 94: 110–116.
  50. Jelovsek E. Predicting urinary incontinence after surgery for pelvic organ prolapse. *Curr Opin Obstet Gynecol* 2016; 28: 399–406.
  51. Beam AL and Kohane IS. Translating artificial intelligence into clinical care. *JAMA* 2016; 316: 2368–2369.
  52. Diprose WK, Buist N, Hua N, et al. Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *J Am Med Inform* 2020; 27: 592–600.
  53. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; 162: W1–W73.
  54. Zhu J, Zheng J, Li L, et al. Application of machine learning algorithms to predict central lymph node metastasis in T1-T2, non-invasive, and clinically node negative papillary thyroid carcinoma. *Front Med* 2021; 8: 635771.
  55. Zhang Y, Wan DH, Chen M, et al. Automated machine learning-based model for the prediction of delirium in patients after surgery for degenerative spinal disease. *CNS Neurosci Ther* 2023; 29: 282–295.
  56. Liu H, Li J, Leng J, et al. Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. *Diabetes Metab Res Rev* 2021; 37: e3397.
-