GYNECOLOGY

Predicting risk of pelvic floor disorders 12 and 20 years after delivery

J. Eric Jelovsek, MD; Kevin Chagin, MS; Maria Gyhagen, MD, PhD; Suzanne Hagen, PhD; Don Wilson, MD; Michael W. Kattan, PhD; Andrew Elders, MSc; Matthew D. Barber, MD, MHS; Björn Areskoug, PhD; Christine MacArthur, PhD; Ian Milsom, MD, PhD

BACKGROUND: Little progress has been made in the prevention of pelvic floor disorders, despite their significant health and economic impact. The identification of women who are at risk remains a key element in targeting prevention and planning health resource allocation strategies. Although events around the time of childbirth are recognized clinically as important predictors, it is difficult to counsel women and to intervene around the time of childbirth because of an inability to convey a patient's risk accurately in the presence of multiple risk factors and the long time lapse, which is often decades, between obstetric events and the onset of pelvic floor disorders later in life. Prediction models and scoring systems have been used in other areas of medicine to identify patients who are at risk for chronic diseases. Models have been developed for use before delivery that predict short-term risk of pelvic floor disorders after childbirth, but no models that predict long-term risk exist.

OBJECTIVE: The purpose of this study was to use variables that are known before and during childbirth to develop and validate prognostic models that will estimate the risks of these disorders 12 and 20 years after delivery.

STUDY DESIGN: Obstetric variables were collected from 2 cohorts: (1) women who gave birth in the United Kingdom and New Zealand (n=3763) and (2) women from the Swedish Medical Birth Register (n=4991). Pelvic floor disorders were self-reported 12 years after childbirth in the United Kingdom/New Zealand cohort and 20 years after childbirth in the Swedish Register. The cohorts were split so that data during the first half of the cohort's time period were used to fit prediction models, and validation was performed from the second half (temporal validation). Because there is currently no consensus on how to best define pelvic floor disorders from a patient's perspective, we chose to fit the data for each model using multiple outcome definitions for prolapse, urinary incontinence, fecal incontinence, ≥1 pelvic floor disorder, and ≥2 pelvic floor disorders. Model accuracy was measured in the following manner: (1) by ranking an individual's risk among all subjects in the cohort (discrimination) with the use of a concordance index and (2) by observing whether the predicted probability was too high or low (calibration) at a range of predicted probabilities with the use of visual plots.

RESULTS: Models were able to discriminate between women who experienced bothersome symptoms or received treatment at 12 and 20 years, respectively, for pelvic organ prolapse (concordance indices, 0.570, 0.627), urinary incontinence (concordance indices, 0.653, 0.689), fecal incontinence (concordance indices, 0.618, 0.676), ≥1 pelvic floor disorders (concordance indices, 0.639, 0.675), and ≥2 pelvic floor disorders (concordance indices, 0.635, 0.619). Route of delivery and family history of each pelvic floor disorder were strong predictors in most models. Urinary incontinence before and during the index pregnancy was a strong predictor for the development of all pelvic floor disorders in most models 12 years after delivery. The 12- and 20-year bothersome symptoms or treatment for prolapse models were accurate when predictions were provided for risk from 0% to approximately 15%. The 12- and 20-year primiparous model began to over predict when risk rates reached 20%. When we predicted bothersome symptoms or treatment for urinary incontinence, the 12-year models were accurate when predictions ranged from approximately 5–60%; the 20-year primiparous models were accurate from 5% and 80%. For bothersome symptoms or treatment for fecal incontinence, the 12- and 20-year models were accurate from 1–15% risk and began to over predict at rates at >15% and 20%, respectively.

CONCLUSION: Models may provide an opportunity before birth to identify women who are at low risk of the development of pelvic floor disorders and may provide institute prevention strategies such as pelvic floor muscle training, weight control, or elective cesarean section for women who are at higher risk. Models are provided at http://riskcalc.org/UR_CHOICE/.

Key words: fecal incontinence, machine learning, pelvic floor disorder, pelvic organ prolapse, prediction model, urinary incontinence

Pelvic floor disorders such as pelvic organ prolapse, urinary incontinence, and fecal incontinence constitute a huge global health problem that affects millions of women throughout the world. The prevalence of pelvic floor disorders has been reported to be 46%, and many women have >1.[1] Pelvic floor disorders can have a negative influence on a woman's well-being, quality of life, body image, and sexual function and prevent many women from participating in recreational and sporting activities.[1,2] The global costs of pelvic floor disorders to healthcare systems and society are enormous.[1,3] Approximately, 1 of 5 women will undergo surgery for prolapse or urinary incontinence by age 85 years.[4,5] Current treatments, often surgical, carry risks and relatively high rates of recurrence.[6,7]

EDITORS' CHOICE

Little progress has been made in the prevention of pelvic floor disorders, despite their significant health and economic impact.[8] The identification of women who are at risk remains a key element in the targeting of prevention and planning health of resource

Cite this article as: Jelovsek JE, Chagin K, Gyhagen M, et al. Predicting risk of pelvic floor disorders 12 and 20 years after delivery. Am J Obstet Gynecol 2018;218:222.e1-19.

0002-9378/free

https://doi.org/10.1016/j.ajog.2017.10.014

allocation strategies. The cause of pelvic floor disorders is known to be multifactorial; obstetric trauma during childbirth is 1 of the most important identifiable risk factors.[1] Numerous epidemiologic studies indicate an increased prevalence of pelvic floor disorders with increasing parity with the greatest increase in risk that is attributed to the birth of the first child.[1] Although events around the time of childbirth are recognized clinically as important predictors, many women undergo the labor and delivery process and do not experience long-term pelvic floor dysfunction. At present, it is difficult to counsel women and intervene around the time of childbirth because of an inability to convey a patient's risk accurately in the presence of multiple risk factors and the long-time lapse, often decades, between obstetric events and the onset of pelvic floor disorders later in life.

Prediction models and scoring systems have been used in other areas of medicine to identify patients who are at risk for chronic diseases.[9,10] Models have been developed for use before delivery that predict short-term risk of pelvic floor disorders after childbirth; however, no models that predict long-term risk exist.[11,12] The aims of this study were to construct and validate models that are capable of predicting the development of pelvic floor disorders 12 and 20 years after delivery with the use of data from 2 large independent international cohort studies.[13,14] Such models have potential to provide individual women more accurate predictions than the current standard of care that is given: (1) the paucity of existing tools, (2) the large amount of variability in the predicted rates of pelvic floor disorders that are provided by clinicians in practice, and (3) the increasing evidence that clinical prediction models consistently show superiority over expert clinicians because they avoid common cognitive biases.[15,16]

Methods

This study used methods set forth in the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis: the TRIPOD statement.[17] The study population consisted of 2 longitudinal, prospective cohort studies. The Prolapse and Incontinence Long-term (ProLong) study aimed to determine whether delivery mode was predictive of pelvic floor disorders in 10,989 primiparous and multiparous women 12 years after the index birth.[14] The second cohort was the Swedish Pregnancy, Obesity and Pelvic Floor (SwePOP) study. The aim of SwePOP was to compare the prevalence of pelvic floor disorders in a cohort of 10,117 primiparous women identified from the Swedish Medical Birth Register 20 years after 1 delivery.[13,18] Both studies were designed to investigate delivery mode as a predictor of pelvic floor disorders and therefore captured key maternal, labor, and delivery variables that were known, at that time, to be potential risk factors of pelvic floor disorders. Study details have been published previously and are summarized in Figure 1.[14,18]

In the ProLong study, prolapse symptoms were measured with the use of the validated Pelvic Organ Prolapse Symptom Score.[19] Urinary and fecal incontinence questions were designed by the study team because, at the time of recruitment (1993/94), there were no suitable validated questionnaires on incontinence. Family history was measured with the use of a response of either "yes" or "no" to the following questions: "Have any of your blood relatives ever had a prolapse?" "If yes, how are they related to you (eg, mother, sister)?" In the SwePOP study, prolapse was defined with the use of the validated five-item questionnaire[20]; urinary incontinence was defined with the use of the Sandvik severity scale,[21] and fecal incontinence was defined with the use of the Wexner score.[22] Family history was measured with a response of either "yes" or "no" to each of the following questions: "Has your mother suffered from urinary leakage?" "Has your mother suffered from prolapse?" and "Has your mother suffered from leakage of flatus/gas or feces?" Each study received ethics committee approval at all centers. Written informed consent was obtained from participants in both studies.

To allow for temporal validation, each cohort was split temporally so that women who gave birth in the first half of the cohort's time period were considered for the training dataset and used to build each model. For the ProLong dataset, data from primiparous and multiparous women who gave birth between September 11, 1993, and May 1, 1994, and who responded at 12 years (n=2095) were used to build models to predict 12-year outcomes for women who gave birth between May 2, 1994, and November 11, 1994 (n=1668). Similarly, in the SwePOP dataset, data from primiparous women who gave birth between January 1, 1985, and June 30, 1987 (n=2607) were used to build models to predict 20-year outcomes for women who gave birth between July 1, 1987, and December 31, 1988 (n=2384). For each training dataset, the multiple imputation with the use of chained equations method was used to calculate missing values for predictors.[23] Predictors for the test dataset and outcomes for all models were based on actual, not imputed values.

Because there is currently no consensus on how to best define pelvic floor disorders from a patient's perspective, we chose to fit the data for each model using multiple outcome definitions for prolapse, urinary incontinence, fecal incontinence, $\geq$1 pelvic floor disorders and $\geq$2 pelvic floor disorders. We developed models in the following 4 categories to predict (1) the presence of "any symptoms" regardless of severity, (2) the presence of bothersome symptoms, (3) treatment for the disorder, or (4) the combination of either bothersome symptoms or receiving treatment for each disorder (prolapse, urinary incontinence, fecal incontinence) and their combination (any pelvic floor disorder or $\geq$2 pelvic floor disorders; Table 1). Data are presented only for category 4, and all remaining outcomes are available in the supplemental results.

Multiple logistic regression models were fit to the training data that consisted of the full set of candidate predictors and each outcome. Harrell's "Model Approximation" process of backwards elimination was used to rank the variables in order of importance,