Seminal significance: the forgotten father in recurrent pregnancy loss

Fossé, N.A. du

CHAPTER 5

Toward more accurate prediction of future pregnancy outcome in couples with unexplained recurrent pregnancy loss: taking both partners into account

N.A. du Fossé
M.L.P. van der Hoorn
R. de Koning
A.G.M.G.J. Mulders
J.M.M. van Lith
S. le Cessie
E.E.L.O. Lashley

Fertility and Sterility 2021;117(1):144-152

ABSTRACT

Objective

To identify, besides maternal age and the number of previous pregnancy losses, additional characteristics of couples with unexplained recurrent pregnancy loss (RPL) that improve the prediction of an ongoing pregnancy.

Design

Hospital-based cohort study in couples who visited specialised RPL units of two academic centres between 2012-2020.

Setting

Two academic centres in the Netherlands.

Patients

Clinical data from 526 couples with unexplained RPL were used in this study.

Intervention(s)

None.

Main Outcome Measure(s)

The final model to estimate the chance of a subsequent ongoing pregnancy was determined with a backward selection process and internally validated using bootstrapping. Model performance was assessed in terms of calibration and discrimination (area under the ROC curve; AUC).

Results

Subsequent ongoing pregnancy was achieved in 345/526 couples (66%). Number of previous pregnancy losses, maternal age, paternal age, maternal body mass index (BMI), paternal BMI, maternal smoking status and previous IVF/ICSI treatment were predictive for the outcome. The optimism corrected AUC was 0.63, compared to 0.57 when using only the number of previous pregnancy losses and maternal age.

Conclusion

The identification of additional predictors for a subsequent ongoing pregnancy after RPL, including male characteristics, is important for both clinicians and couples with RPL. At the same time we showed that the predictive ability of the current model is still limited and more research is warranted to develop a model that can be used in clinical practice.

INTRODUCTION

Recurrent pregnancy loss (RPL) is a condition characterised by the spontaneous loss of two or more pregnancies before 24 weeks of gestation, affecting 2-3% of couples of reproductive age.(1, 2) Over time, various risk factors for RPL have been identified and several diagnostic investigations are recommended by international guidelines, including screening for uterine anomalies, acquired thrombophilia, thyroid abnormalities and parental chromosomal translocations.(2) Despite the extensive diagnostic work-up being offered to couples with RPL, no underlying condition can be identified in 60-70% of cases.(3) For these unexplained cases, no evidence-based therapeutic options are available, which adds to the frustrating nature of this condition.(2) Indeed, multiple studies have shown that couples with RPL are more likely to deal with depression and anxiety.(4) It is considered important to offer supportive care to couples with RPL, consisting of intensive monitoring and care during early pregnancy as well as psychological support.(5, 6) Moreover, supportive care should certainly include reliable counselling regarding prognosis.

For couples with recurrent pregnancy loss (RPL) one question is vital: what is the chance of a future successful pregnancy? Even when aetiological mechanisms are not fully elucidated, well-developed and validated prediction models may provide adequate estimates of future pregnancy outcomes.(7) Currently, two prognostic tools are recommended by the ESHRE guideline on RPL.(2) Both models base their predictions on two factors: the number of preceding pregnancy losses and maternal age. Brigham et al.(8) predicted the chance of a subsequent ongoing pregnancy with fetal survival beyond 24 weeks of gestation, while Lund et al.(9) predicted pregnancy success rates at five, ten and fifteen years after referral. Yet, some important limitations must be kept in mind when using these prediction models.

First of all, as neither performance measures nor validation procedures were described for both models, their predictive performance remains unknown. Second, as these models were developed 21 and nine years ago, changing definitions and diagnostic investigations for RPL have most probably affected the reliability of the models in today's clinical practice. In addition, a limited number of candidate predictors were examined in both studies. Although it is indisputable that maternal age and previous number of losses are important predictors for future pregnancy outcome(2), it is likely that inclusion of other factors may improve accuracy of prediction. Lifestyle factors such as cigarette smoking have been associated with pregnancy loss in previous studies and may thus influence future pregnancy outcome.(10, 11) Moreover, although the focus has been on the female partner for many years, evidence is emerging that characteristics of the male partner also contribute to (recurrent) pregnancy loss.(12, 13)

The aim of this study was to explore whether predicting the chance of a subsequent ongoing pregnancy in couples with unexplained RPL could be improved by taking, besides maternal age and the number of previous pregnancy losses, additional candidate predictors into account. To the best of our knowledge, this is the first time that the predictive potential of both maternal and paternal factors was evaluated in this context.

MATERIAL AND METHODS

This study was conducted following the recommendations of the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement.(14) This study was approved by the Medical Research Ethics Committee of the Leiden University Medical Center (reference number P19.014).

Source of data

In this hospital-based cohort study, data from two specialised RPL units located in two Dutch academic hospitals (Erasmus MC, University Medical Center Rotterdam and Leiden University Medical Center) was obtained, covering the period between January 2012 and December 2019. Couples with RPL were referred to these clinics for diagnostic investigations, counselling, supportive care and/or intensive monitoring during the first trimester of a subsequent pregnancy. Baseline characteristics (described in more detail in the paragraph Candidate predictors) of all couples that visited the RPL clinics were registered in electronic patient records during the intake consultation, using a standardised template. Data on baseline characteristics and subsequent pregnancy outcome were extracted from the hospital database systems and entered in a study database, using a standardised template.

Eligibility criteria

Couples were included in the study database with at least two pregnancy losses before 24 weeks of gestation (following the definition of the ESHRE guideline on RPL) in the current relationship. Couples with pregnancy losses following oocyte or sperm donation and couples with an identified underlying condition for RPL (specified in the next paragraph) were excluded.

Diagnostic investigations for RPL

Diagnostic investigations considered for this study were based on recommendations of the current ESHRE guideline on RPL(2) and included screening for uterine anomalies, thyroid abnormalities (anti-thyroid peroxidase (TPO) and thyroid-stimulating hormone (TSH) levels), acquired thrombophilia (antiphospholipid antibodies(15)) and parental chromosomal translocations. Parental karyotyping was only performed in case of increased risk of abnormalities, following the risk table of Franssen et al.(16)

Outcome

We estimated the chance of a subsequent ongoing pregnancy, defined as fetal survival beyond 24 weeks of gestation(2) in the first pregnancy after intake consultation at the RPL clinic. All first pregnancy outcomes that occurred after intake consultation and before January 2021 were analysed. Pregnancies conceived by a new male partner (i.e.

a different partner than during the intake consultation) or conceived following oocyte or sperm donation were excluded from the analysis. Also, couples with no further pregnancy or with an unknown pregnancy outcome after intake consultation were also excluded from the present analysis.

Sample size calculation

For sample size considerations, we followed the recommendations as published by van Smeden et al.(17) An established rule of thumb for the required sample size to develop a prediction model is to ensure at least 10 events per candidate predictor parameter. However, van Smeden et al. stated this rule is insufficient to minimise the risk of model overfitting and to target precise model predictions. For binary outcomes, they showed that the number of candidate prediction parameters, the total sample size and the outcome proportion are the main drivers of the mean predictive accuracy of a prediction model. Therefore a sample size formula was presented, that aims to ensure that a new prediction model will on average have a small prediction error in the estimated outcome probabilities, as measured by the mean absolute prediction error (MAPE). An interactive calculation tool is available online and was used for this study: https://mvansmeden.shinyapps.io/BeyondEPV/. Before performing the present study, the number of available patients and predictors was known. For this situation, the calculation tool could be used to identify the maximum number of candidate predictors to be considered. With an anticipated outcome proportion of 70% couples with an ongoing pregnancy (8, 9, 18), a sample size of 526 (the number of couples available in our database) and a MAPE of 0.05 between observed and true outcome probabilities (as recommended by van Smeden et al.), the maximum number of candidate prediction parameters was determined a priori as 12.

Candidate predictors

The following candidate predictors were considered based on theoretical plausibility following previous research, expert opinion and availability: the number of previous pregnancy losses, primary or secondary RPL (with primary RPL being defined as no live birth in the current relationship), previous pregnancies conceived by *in vitro* fertilisation (IVF) or intracytoplasmic sperm injection (ICSI),maternal and paternal age, maternal and paternal body mass index (BMI) and maternal and paternal smoking status. All candidate predictor variables were collected during the intake consultation. The number of previous pregnancy losses, maternal and paternal ages were treated as continuous variables. Previous IVF or ICSI treatment and maternal and paternal smoking status were treated as dichotomous variables.

Statistical analysis

All analyses were performed in R studio version 1.3.9.50 and R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria).

Handling of missing data

To avoid a decrease of statistical power and selection bias, missing values were imputed. We assumed that missing values were missing at random. Based on the amount of missing data, missing values were imputed 30 times using multiple imputation with chained equations (MICE) with predictive mean matching.(19, 20) All candidate predictors and the outcome variable were included in the imputation model.(19) Rubin's rules were applied for pooling estimates across the imputed datasets.(21)

Model development

Initially we fitted univariable logistic regression models to assess the effect of individual predictors. Possible non-linearity in the associations between continuous predictors and the outcome were examined using the R studio package 'rcspline.plot'. Maternal age had a significant non-linear relation with the probability of a subsequent ongoing pregnancy and was modelled using a restricted cubic spline. For model development we used the R studio package 'pfmsi' which provides functions to apply pooling and variable selection in multiple imputed datasets. We performed multivariable logistic regression analysis with ongoing pregnancy as binary outcome. A backward selection process was used to determine the final multivariable logistic regression model, using the Akaike Information Criteria (AIC) as a stopping rule (corresponding to a *p*-value of 0.157).(20, 22) To assess the added value of additional predictors, we fitted smaller models including only a subset of the predictors derived from the backward selection.

Model performance

The resulting final model was internally validated using bootstrapping with 250 bootstrap samples, yielding estimates for the optimism in the performance for discrimination and calibration. The bootstrapping procedure was performed in combination with backward selection, as it is known that variable selection is a major reason for model overfitting.(20) Model calibration was ascertained by visual inspection of a calibration plot. Receiver operator characteristic (ROC) curve analysis was used as a measure for discrimination. Discrimination refers to the ability of a model to correctly assign higher probabilities to subjects with the outcome (ongoing pregnancy) compared to subjects without the outcome. An area under the ROC curve (AUC) of 0.5 indicates no discrimination and is comparable with tossing a coin: the ability of the model to assign a higher probability to a couple with ongoing pregnancy than to a couple without ongoing pregnancy is 50%. An AUC of 1.0 indicates perfected discrimination. The explained variance was described in terms of the Nagelkerke R^2. To prevent the model from overfitting, the calibration slope from the bootstrapping procedure was used to shrink the pooled regression coefficients and to determine a new intercept, being aligned with the shrunken coefficients.(20) Performance measures of the final model and smaller models including fewer predictors were compared.

RESULTS

After exclusions, the dataset included 526 couples with unexplained RPL and a subsequent pregnancy outcome after intake consultation at one of the two participating clinics. The flow of participants through the study is shown in Supplemental Figure 1. All included couples were in follow-up for at least one year after intake consultation. In 345 couples (66%) the first pregnancy after intake consultation was an ongoing pregnancy beyond 24 weeks of gestation. Of the remaining 181 couples (34%) without an ongoing pregnancy, 168 (93%) had a spontaneous pregnancy loss, eight (4%) had an ectopic pregnancy and five (3%) had a termination of pregnancy due to fetal abnormalities. Fifty-six pregnancy outcomes occurred in 2020, during the Covid-19 pandemic. None of these women were known to have had a SARS-CoV-2 infection during their pregnancy. Table 1 shows the characteristics of the total cohort and of couples with and without ongoing pregnancy separately. Percentages of missing values ranged from 0 to 22.8% per candidate predictor.

Table 1. Cohort characteristics

Characteristics	All couples (*n* = 526)	Ongoing pregnancy* (*n* = 345)	No ongoing pregnancy (*n* = 181)	Missing data *n* (%)
Mean age (SD), range				
Women	33.58 (4.67), 20-45	33.28 (4.42) 20-43	34.14 (5.08), 21-45	0 (0)
Men	35.50 (6.11), 20-67	35.10 (5.79) 20-67	36.28 (6.63), 21-55	26 (4.9)
Median number of pregnancy losses (IQR), range	3 (2-4), 2-11	3 (2-3), 2-10	3 (3-4), 2-11	0 (0)
Primary RPL, *n* (%)	308 (58.6)	202 (58.6)	106 (58.6)	0 (0)
History of IVF/ICSI treatment, *n* (%)	72 (13.7)	39 (11.3)	33 (18.2)	0 (0)
Mean BMI (SD), range				
Women	24.55 (4.59), 16.18-44.98	24.71 (4.83), 17.71-44.98	24.24 (4.08), 16.18-42.91	24 (4.6)
Men	25.51 (3.60), 18.26-41.77	25.36 (3.50), 18.26-41.77	25.79 (3.77), 19.27-40.75	120 (22.8)
Smoking, *n* (%)				
Women	65 (12.4)	37 (10.7)	28 (15.4)	6 (1.1)
Men	133 (25.3)	83 (24.1)	50 (27.6)	61 (11.6)

SD = standard deviation; IQR = interquartile range; RPL = recurrent pregnancy loss; IVF = in vitro fertilisation; ICSI = intracytoplasmic sperm injection; BMI = body mass index.
*Ongoing pregnancy defined as fetal survival beyond 24 weeks of gestation.

Predicting the chance of ongoing pregnancy

The number of previous pregnancy losses, maternal and paternal age and previous conceptions by IVF/ICSI treatment had statistically significant univariable associations with an ongoing pregnancy (Supplemental Table 1). Figure 1 shows the unadjusted relations between the predicted probability of an ongoing pregnancy and the continuous predictors

number of previous pregnancy losses, maternal and paternal age and maternal and paternal BMI. The probability of an ongoing pregnancy gradually declined with increasing number of previous pregnancy losses and increasing paternal age and sharply declined starting from maternal age 35. Although parental BMI effects were small, we observed a negative association between increasing paternal BMI and an ongoing pregnancy, while increasing maternal BMI slightly improved the chance of an ongoing pregnancy.

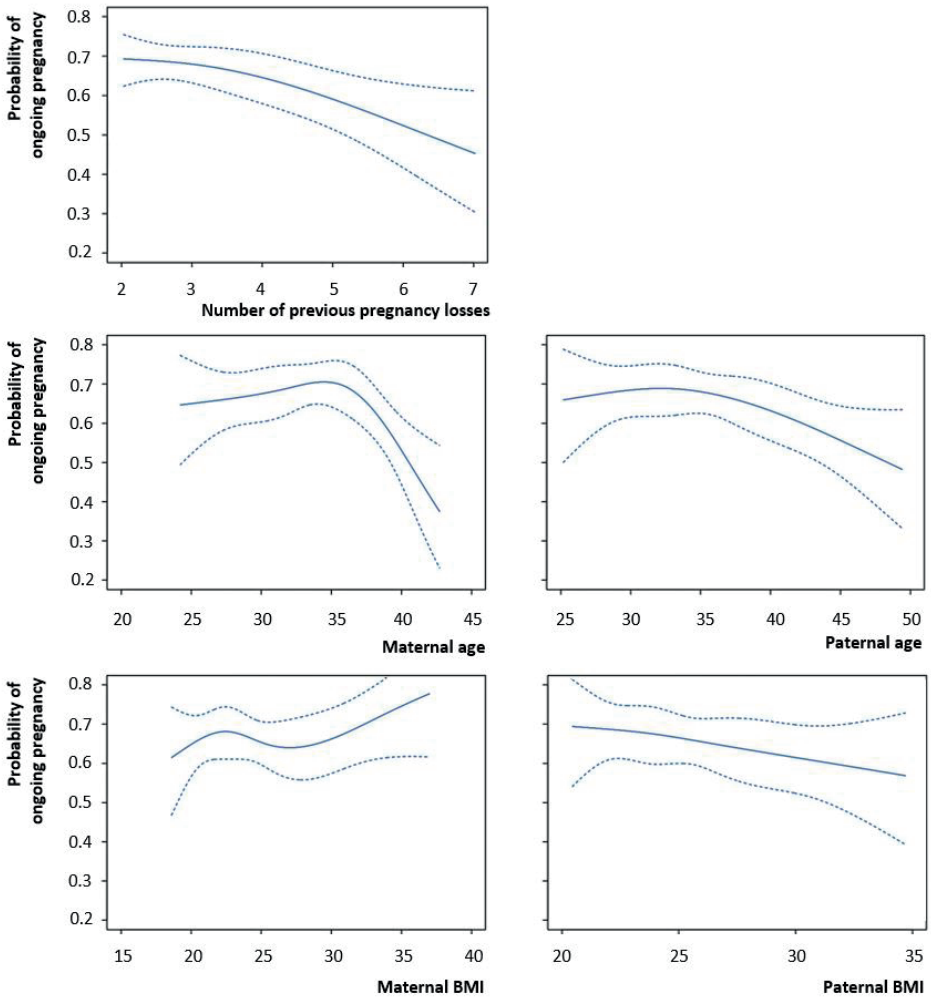

Figure 1. Univariable relations between continuous baseline variables and ongoing pregnancy

BMI = body mass index. Each panel depicts the probability of ongoing pregnancy (solid curve) with 95% confidence bands (dashed curves) as function of the baseline variable. Relations were characterised by restricted cubic spline functions. Only maternal age had a significant non-linear relation with the outcome.

The factors in the final multivariable model (Table 2) to predict the probability of having a subsequent ongoing pregnancy were the number of previous pregnancy losses, maternal age, paternal age, maternal BMI, paternal BMI, maternal smoking status and mode of conception (with or without history of IVF/ICSI treatment). The bootstrapping procedure yielded an adjusted calibration slope of 0.77, which was applied as a shrinkage factor to the intercept and coefficients of the final model. The odds of a subsequent ongoing pregnancy decreased with every increasing previous pregnancy loss. For example, the odds of an ongoing pregnancy after three pregnancy losses were 19% lower than the odds of an ongoing pregnancy after two pregnancy losses, and the odds after six pregnancy losses were 47% less than after three losses. A smoking woman had 38% lower odds of an ongoing pregnancy compared to a non-smoking woman. Couples with a history of IVF/ICSI treatment had a 46% reduced odds of an ongoing pregnancy compared to couples with spontaneous conceptions.

Table 2. Final logistic regression model for ongoing pregnancy

Intercept and predictors	β coefficient[a]	Odds ratio (95% CI)	*P*-value
Intercept	0.53		
Number of previous pregnancy losses	-0.16	0.81 (0.70-0.93)	0.004
Maternal age as restricted cubic spline[b]			
Maternal age	0.06	1.08 (0.92-1.25)	0.34
Maternal age'	-0.01	0.98 (0.71-1.38)	0.94
Maternal age''	-0.46	0.55 (0.12-2.46)	0.43
Maternal smoking	-0.36	0.62 (0.36-1.07)	0.09
Maternal BMI	0.03	1.04 (0.99-1.09)	0.09
Paternal age	-0.02	0.97 (0.93-1.01)	0.15
Paternal BMI	-0.04	0.95 (0.89-1.01)	0.11
History of IVF/ICSI treatment	-0.47	0.54 (0.312-0.92)	0.02

BMI = body mass index; 95% CI = 95% confidence interval; IVF = in vitro fertilisation; ICSI = intracytoplasmic sperm injection. *The predicted probability of a subsequent ongoing pregnancy can be calculated for individual couples using the formula shown in the Supplemental data. [a]Regression coefficients were multiplied with a shrinkage factor of 0.77 that was obtained from the bootstrapping procedure (described in Methods). β-values are expressed per 1-unit increase for continuous predictors and for the condition present (prediction value = 1) for dichotomous predictors. [b]Maternal age was fitted using a restricted cubic spline function with four knots placed at 25.27, 31.84, 35.94 and 40.53 years. The age variables with tick-marks (', '') represent the new variables created to allow for non-linear contributions from maternal age. These coefficients cannot be interpreted on their own; the partial effect plot for maternal age is shown in Supplemental Figure 3.

Model performance

The calibration plot of the final multivariable model indicated overall good calibration (Supplemental Figure 2). We compared the discrimination of the final model to that of smaller models including only a subset of the predictors. The optimism corrected AUCs ranged from 0.57 for a model only including the predictors maternal age (fitted as a linear variable) and number of previous pregnancy losses, to 0.63 for the final model including all predictors derived from the backward selection procedure. Performance measures for all models are shown in Supplemental Table 2.

Predicting ongoing pregnancy for specific couples

Figure 2 shows four couples with their respective characteristics and predicted chances of a subsequent ongoing pregnancy according to our final multivariable prediction model, including the number of previous pregnancy losses, maternal and paternal age, maternal and paternal BMI, maternal smoking status and mode of conception (with or without a history of IVF/ICSI treatment). We compared the predicted probabilities of our model with those provided by the commonly used prediction model of Brigham et al.(8), including only the number of previous pregnancy losses and maternal age fitted as a linear variable.

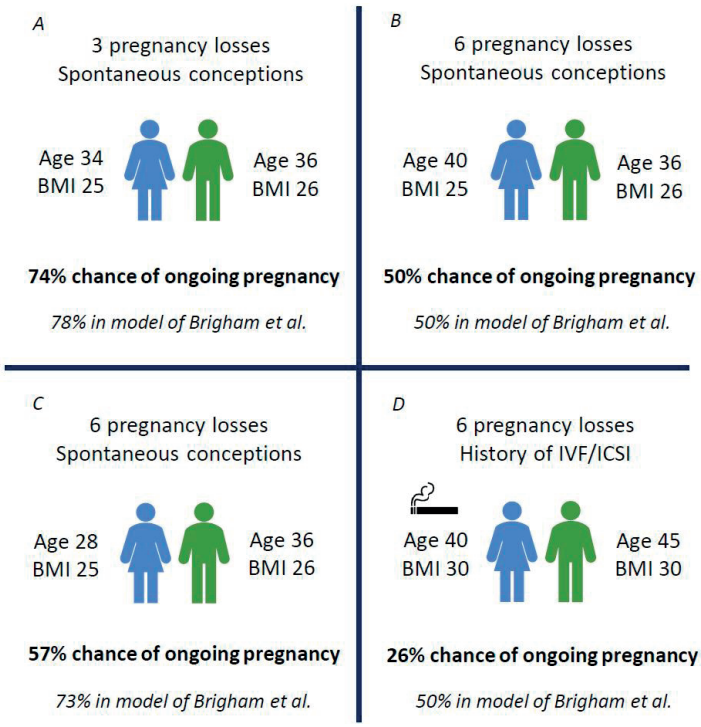

Figure 2. Predicting ongoing pregnancy: four scenarios

BMI = body mass index; IVF = in vitro fertilisation; ICSI = intracytoplasmic sperm injection. Chances of an ongoing pregnancy >24 weeks' gestation based on our final prediction model, including the following variables: number of previous pregnancy losses, maternal age (fitted as restricted cubic spline with four knots), paternal age, maternal BMI, paternal BMI, maternal smoking status and mode of conception. Predicted probabilities are shown for four couples and compared to the model of Brigham et al. Scenario A shows a couple with average characteristics based on our population statistics, i.e. with the median number of pregnancy losses, mean ages and BMIs as shown in Table 1. In scenario B, the number of previous pregnancy losses and maternal age are higher, while other characteristics are unchanged. Scenario C is similar to scenario B, except for a younger maternal age. In scenario D the number of pregnancy losses and the woman's age are similar to scenario B, but here the male partner is also of advanced age, the couple has a history of fertility treatment (IVF/ICSI), they are obese and the woman smokes.

For scenario A and B, the predicted chances of a subsequent ongoing pregnancy calculated with our model and with the model of Brigham et al. were similar (74% vs. 78% for scenario A and 50% both for scenario B). In scenario C our model provided a lower chance of an ongoing pregnancy compared to the model of Brigham et al. (57% vs. 73%). In scenario D the predicted probabilities resulting from both models were even more deviating. The estimate of our model was a 26% chance of an ongoing pregnancy, almost half the probability as calculated for scenario B. However, the model of Brigham et al. still estimated a 50% chance of an ongoing pregnancy, since this model is only based on the number of previous pregnancy losses and maternal age, being equal in scenarios B and D

DISCUSSION

We showed that predicting the chance of a subsequent ongoing pregnancy beyond 24 weeks of gestation in couples with RPL becomes more accurate when, besides the conventional predictors maternal age and the number of previous pregnancy losses, more variables are incorporated into the model. The additional predicting variables include both male and female characteristics, advocating a couple-focused rather than a female-focused approach in RPL. Still, the predictive ability of the current model remains limited and we emphasize that more research is needed in order to develop a model that can be used in clinical practice.

The apparent predictive performance of our final multivariable model in terms of the AUC was 0.66 (0.63 after internal validation with bootstrapping), compared to 0.57 for a model restricted to the conventional predictors maternal age and number of previous pregnancy losses. Although showing an improvement in predictive ability, an AUC between 0.60-0.70 is still considered as poor to moderate performance and indicates that the model will not successfully predict outcomes for many couples.(20) As Brigham et al.(8) and Lund et al.(9) did not mention any performance measures, it was not possible to make a direct comparison with their models. A recently published nationwide Danish cohort study that aimed to predict the chance of subsequent live birth in the general population based on maternal age and prior pregnancy events, reported an AUC of 0.60. Both this Danish cohort study and our study illustrate the difficulty of predicting future ongoing pregnancy. This may be due to the complex and largely unexplained multifactorial aetiology of (recurrent) pregnancy loss.

While we confirmed earlier findings showing that the number of previous pregnancy losses and woman's age are prognostic variables of great importance(2, 8, 9, 18), we also found that additional maternal variables (smoking status, BMI) as well as paternal parameters (age, BMI) increased predictive performance. Furthermore, we observed that previous IVF/ICSI treatment lowers the predicted chance of a subsequent ongoing pregnancy in couples with RPL. Our candidate predictors were chosen based on previous epidemiologic and basis research and although one should be cautious with interpreting the results of a prediction study aetiologically(7), it is likely that some of the predictors have a causal relation with the outcome.

Maternal age is strongly associated with a higher risk of fetal aneuploidy, an established cause of pregnancy loss.(23) Advanced paternal age has been linked to increased levels of sperm DNA fragmentation, which is associated with (recurrent) pregnancy loss.(13, 24, 25) Likewise, paternal obesity may cause excessive oxidative stress and affect pregnancy outcome by damaging DNA integrity of the spermatozoa.(26) Maternal smoking is well-

known to increase the risk of pregnancy complications, including pregnancy loss.(10) On the other hand, the relation between assisted reproductive techniques, including IVF/ICSI treatment, and an increased risk of pregnancy loss is less straight-forward. It is complex to determine whether this increased risk can be attributed to the treatment itself, whether it is a proxy for underlying (unidentified) patient characteristics, or whether it is due to the fact that ART pregnancies are closely monitored and subsequent (early) pregnancy loss is more often detected compared to couples who conceived naturally.(27) Furthermore, we observed a positive association between increasing maternal BMI and the chance of an ongoing pregnancy in our cohort. A previous study in couples with unexplained RPL demonstrated a U-shaped relationship between miscarriage rate in the subsequent pregnancy and maternal pre-pregnancy BMI, with the highest risk of miscarriage in underweight women, followed by obese women (BMI >30 kg/m^2).(28) Although we observed similar high risks of pregnancy loss in underweight women with BMI <20 kg/m^2, in our population the highest chance of an ongoing pregnancy was found in obese women. However, it should be noted that the number of obese women in our sample was limited and the observed BMI effect was relatively weak and uncertain.

When developing a prediction model it is important to assess the presence of non-linear patterns between continuous predictors and the outcome of interest.(29) We found that maternal age had a non-linear relationship with the chance of an ongoing pregnancy, with a negative effect starting around 35 years and we estimated this relationship using a restricted cubic spline. A similar pattern for the maternal age effect was observed in two prior studies(30, 31) predicting chances of live birth in other (large) populations, not restricted to RPL patients; these studies also fitted maternal age as restricted cubic spline in their models. However, previous prediction models for RPL handled maternal age as a linear term, which probably differs substantially from the “true” predictor-outcome relationship, as it assumes that the effect is the same at each part of the range of maternal age.

We believe that our study holds several strengths compared to other prediction studies on unexplained RPL. We followed TRIPOD recommendations for model development and reporting.(14) To prevent overfitting, we determined the maximum number of candidate predictors a priori.(17, 32) Furthermore, we selected candidate predictors based on theoretical plausibility instead of choosing predictors on the basis of the strength of their unadjusted univariable associations with the outcome. The last strategy is undesired as this most often leads to substantial uncertainty in model structure and important predictors may be rejected because of nuances in the study data.(29, 33, 34) We used backward elimination with AIC for predictor selection, being a preferred method, especially in smaller data sets.(20) In addition, we performed internal bootstrap validation and used the shrinkage factor to adjust the regression

coefficients and apparent performance for optimism(20), which was not done in any of the previously published prediction models for RPL. Besides these methodological assets, we used data of a strictly defined population of couples with unexplained RPL, containing information on both partners, being systematically collected during intake consultations. Still, some missing data existed, mainly on paternal variables. However, it was possible to impute these data using multiple imputation. This technique takes into account statistical uncertainty in the imputed values and, if data are missing at random, provides less biased results compared to complete case analysis.

The aim of this study was to identify predictors for a subsequent ongoing pregnancy beyond 24 weeks of gestation, after referral to the clinic. This outcome was available for the vast majority of couples in our database, while the outcome of a subsequent live birth as well as outcomes of later occurring pregnancies were more often missing (due to the fact that many women were referred back to their local hospital or midwifery practice). Ideally, patients would like to know their overall chances of having a future live birth. Therefore, the ultimate model should predict the cumulative chances of live birth within a certain time period, for instance within five years after referral. This would require a prospective cohort study with structural follow-up of couples with RPL for at least five years after first consultation. Furthermore, in future research, the effects of more potential predictors such as alcohol consumption of both partners and level of sperm DNA fragmentation should be evaluated, which have previously been associated with pregnancy loss but were unavailable in this study. In a sufficiently large cohort including couples with both explained and unexplained RPL, it may also be considered to assess identified risk factors (for instance presence of anti-TPO antibodies or APS) as predicting variables and to assess meaningful interactions between different predictors.

Conclusions

Couples with RPL need something to hold on to, that helps to shape their expectations and assists in making decisions regarding new pregnancy attempts. In addition, stratification of couples into risk groups can be used for further in-depth personalised research, for instance on interventions. To facilitate this, an accurate well-developed and validated prediction model is needed. To date, such a model is not yet within reach. Although we showed in this study that we should look beyond the number of previous pregnancy losses and maternal age and we should also consider additional predictors including male factors and lifestyle factors, the predictive ability - and therefore the clinical applicability- of the model is still insufficient. However, our findings serve as an important starting point for the development of a new prediction tool to use in clinical practice.

REFERENCES

1. Rai R, Regan L. Recurrent miscarriage. Lancet 2006;368:601-11.
2. Bender Atik R, Christiansen OB, Elson J, Kolte AM, Lewis S, Middeldorp S et al. ESHRE guideline: recurrent pregnancy loss. Hum Reprod Open 2018;2018.
3. Jaslow CR, Carney JL, Kutteh WH. Diagnostic factors identified in 1020 women with two versus three or more recurrent pregnancy losses. Fertil Steril. 2010;93:1234-43.
4. Voss P, Schick M, Langer L, Ainsworth A, Ditzen B, Strowitzki T et al. Recurrent pregnancy loss: a shared stressor---couple-orientated psychological research findings. Fertil Steril. 2020;114:1288-96.
5. Musters AM, Taminiau-Bloem EF, van den Boogaard E, van der Veen F, Goddijn M. Supportive care for women with unexplained recurrent miscarriage: patients' perspectives. Hum Reprod. 2011;26:873-7.
6. Musters AM, Koot YE, van den Boogaard NM, Kaaijk E, Macklon NS, van der Veen F et al. Supportive care for women with recurrent miscarriage: a survey to quantify women's preferences. Hum Reprod 2013;28:398-405.
7. van Diepen M, Ramspek CL, Jager KJ, Zoccali C, Dekker FW. Prediction versus aetiology: common pitfalls and how to avoid them. Nephrol Dial Transplant. 2017;32:ii1-ii5.
8. Brigham SA, Conlon C, Farquharson RG. A longitudinal study of pregnancy outcome following idiopathic recurrent miscarriage. Hum Reprod. 1999;14:2868-71.
9. Lund M, Kamper-Jørgensen M, Nielsen HS, Lidegaard Ø, Andersen AM, Christiansen OB. Prognosis for live birth in women with recurrent miscarriage: what is the best measure of success? Obstet Gynecol 2012;119:37-43.
10. Pineles BL, Park E, Samet JM. Systematic review and meta-analysis of miscarriage and maternal exposure to tobacco smoke during pregnancy. Am J Epidemiol 2014;179:807-23.
11. Wang L, Yang Y, Liu F, Yang A, Xu Q, Wang Q et al. Paternal smoking and spontaneous abortion: a population-based retrospective cohort study among non-smoking women aged 20–49 years in rural China. J Epidemol Community Health 2018;72:783.
12. du Fossé NA, van der Hoorn MP, van Lith JMM, le Cessie S, Lashley E. Advanced paternal age is associated with an increased risk of spontaneous miscarriage: a systematic review and meta-analysis. Hum Reprod Update 2020;26:650-69.
13. McQueen DB, Zhang J, Robins JC. Sperm DNA fragmentation and recurrent pregnancy loss: a systematic review and meta-analysis. Fertil Steril. 2019;112:54-60.e3.
14. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162:55-63.
15. Miyakis S, Lockshin MD, Atsumi T, Branch DW, Brey RL, Cervera R et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). J Thromb Haemosat. 2006;4:295-306.
16. Franssen MT, Korevaar JC, Leschot NJ, Bossuyt PM, Knegt AC, Gerssen-Schoorl KB et al. Selective chromosome analysis in couples with two or more miscarriages: case-control study. BMJ 2005;331:137-41.
17. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ et al. Sample size for binary

logistic prediction models: Beyond events per variable criteria. Statistical Methods in Medical Research 2019;28:2455-74.

18. Sugiura-Ogasawara M, Ozaki Y, Kitaori T, Suzumori N, Obayashi S, Suzuki S. Live birth rate according to maternal age and previous number of recurrent miscarriages. Am J Reprod Immunol. 2009;62:314-9.
19. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30:377-99.
20. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162:W1-73.
21. Rubin DB. Inference and Missing Data. Biometrika 1976;63:581-92.
22. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. Med Decis Making 2001;21:45-56.
23. Grande M, Borrell A, Garcia-Posada R, Borobio V, Muñoz M, Creus M et al. The effect of maternal age on chromosomal anomaly rate and spectrum in recurrent miscarriage. Hum Reprod. 2012;27:3109-17.
24. Wyrobek AJ, Eskenazi B, Young S, Arnheim N, Tiemann-Boege I, Jabs EW et al. Advancing age has differential effects on DNA damage, chromatin integrity, gene mutations, and aneuploidies in sperm. Proc Natl Acad Sci. 2006;103:9601.
25. Schmid TE, Grant PG, Marchetti F, Weldon RH, Eskenazi B, Wyrobek AJ. Elemental composition of human semen is associated with motility and genomic sperm defects among older men. Hum Reprod. 2013;28:274-82.
26. Wright C, Milne S, Leeson H. Sperm DNA damage caused by oxidative stress: modifiable clinical, lifestyle and nutritional factors in male infertility. Reprod Biomed Online 2014;28:684-703.
27. Wang JX, Norman RJ, Wilcox AJ. Incidence of spontaneous abortion among pregnancies produced by assisted reproductive technology. Hum Reprod. 2004;19:272-7.
28. Metwally M, Saravelos SH, Ledger WL, Li TC. Body mass index and risk of miscarriage in women with recurrent miscarriage. Fertil Steril. 2010;94:290-5.
29. Harrell FE, Jr. Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression, and Survival Analysis. 1 ed: Springer, New York, NY, 2001.
30. McLernon DJ, Steyerberg EW, te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: population based study of linked cycle data from 113873 women. BMJ 2016;355:i5735.
31. Kolte AM, Westergaard D, Lidegaard Ø, Brunak S, Nielsen HS. Chance of live birth: a nationwide, registry-based cohort study. Hum Reprod. 2021;36(4):1065-1073.
32. Riley RD, Ensor J, Snell KIE, Harrell FE, Jr., Martin GP, Reitsma JB et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441.
33. Steyerberg EW. Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating. 1 ed: Springer-Verlag New York, 2009.
34. Sauerbrei W, Boulesteix A-L, Binder H. Stability Investigations of Multivariable Regression Models Derived from Low- and High-Dimensional Data. J Biopharm Stat. 2011;21:1206-31.