## Beyond successful external cephalic version

de Hundt, M.

2015

Chapter 2

# External validation of a prediction model for successful external cephalic version

M. de Hundt
F. Vlemmix
M. Kok
J.W. van der Steeg
J.M.J. Bais
B.W.J. Mol
J.A.M. van der Post

*American Journal of Perinatology 2012;29(3):231-6*

# Abstract

## OBJECTIVE

We sought external validation of a prediction model for the probability of successful external cephalic version (ECV).

## METHODS

We evaluated the performance of the prediction model with calibration and discrimination. For clinical practice, we developed a score chart to calculate the probability of a successful ECV.

## RESULTS

We studied 320 women undergoing ECV, of which 117 (37%) were successful. The model underestimated the success rate by 4 to 14%. The area under the receiver operating characteristic curve was moderate (0.66, 95% CI 0.60-0.72), but the model was able to make good distinction between women with a higher predicted probability of a successful ECV (more than 50%) compared with women with a lower predicted probability of a successful ECV (less than 20%).

## CONCLUSION

Our model to predict the outcome of ECV holds in external validation. It can be used to support patient counseling and decision making for ECV in singleton term breech presentations.

# Introduction

Breech delivery is associated with an increased risk of neonatal morbidity and mortality as compared with cephalic delivery. As a result more cesarean sections are performed in women who have a fetus in breech presentation at term,[1,2] with the consequence of additional risks in subsequent pregnancies due to complications such as uterine rupture and abnormal placental implantation.[3,4] External cephalic version (ECV) is a relatively simple and safe maneuver and is proven to be effective in the reduction of the frequency of breech presentation at term and consequently the number of cesarean deliveries for this condition. In the literature, the ECV success rates vary from 35 to 86%.[5]

There are many factors thought to be associated with the success or failure of an ECV attempt. A meta-analysis showed that clinical characteristics such as non-engagement of the breech, multiparity, relaxation of the uterus, a palpable fetal head, and maternal weight less than 65 kg are related to a successful ECV outcome.[6] Ultrasound characteristics related with a successful ECV outcome are complete breech presentation, posterior placental localization, and an amniotic fluid index (AFI) >10 cm.[7]

Based on this information, we have developed a model to predict the outcome of a successful ECV in women with a singleton breech presentation at term.[8] At internal validation the model showed a fair discriminative capacity and a good calibration, thus allowing to discriminate between women with a good probability of a successful ECV and women with a poor probability of a successful ECV. External validation is a crucial step before a predictive model can be used in daily practice, and thus far the model has not been externally validated.[9] The aim of this study was to validate the ECV prediction model with data from an external population.

# Materials and methods

## PATIENTS

From August 2002 to January 2009, data of consecutive women undergoing ECV at the Medical Centre Alkmaar were collected. ECV attempts performed in the Medical Centre Alkmaar between August 2004 and December 2006 were excluded because these data were used for development of the original prediction model. ECV was applied in women with a singleton fetus in breech presentation at term (36 0/7 weeks gestation and onward). Exclusion criteria for ECV were women with a contraindication for vaginal labor, placenta abruption in the history, uterus anomalies, suspicion of ruptured membranes, HELLP syndrome (hemolysis; elevated liver enzymes; low platelet count), preeclampsia or severe pregnancy-induced

hypertension, fetus with congenital abnormalities, and fetal growth restriction (estimated fetal weight (EFW) <5th percentile). Before and after every ECV attempt, a cardiotocography for half an hour was conducted. All ECV attempts were performed by either an experienced midwife or gynecologist, without use of uterine relaxants. After 30 minutes, the position of the fetus was confirmed by ultrasound.



We collected data on parity, maternal age, body mass index, ethnicity, gestational age, fetal position, placental localization (anterior, posterior, in fundo, lateral), EFW, and amniotic fluid. Parity was recorded as nulliparous or multiparous, where multiparity was defined as having delivered at least once abdominally or vaginally after 24 weeks of gestation. Fetal position was categorized as frank breech versus non frank breech. The AFI was registered by ultrasound and divided in two groups, normal amniotic fluid and decreased amniotic fluid (AFI <10 cm). EFW was calculated with the Hadlock formula, based on ultrasound measurements.[10]

## DATA ANALYSIS

The primary end point of the study was a successful ECV, defined as by ultrasound confirmed cephalic presentation 30 minutes after the ECV attempt. The probability of a successful ECV was calculated using the formula: probability = 1/[1+exp(-β)], where β = -5.1 + (multiparity x 1.05) + (EFW ≤ 3000 g x 0.13) + (anterior placental localization x -0.32) + (normal AFI x 0.82).

The model was evaluated with discriminative capacity and calibration. Missing data of the predicted variables were imputed (''filled in''), because deleting them would lead to a loss of statistical power in multivariable analysis and, more seriously, potentially biased results.[11] We generated a single imputed data set, using the first step of the ''aRegImpute'' multiple imputation function in Splus 8.0. This is an efficient implementation of Bayesian multiple imputation, a recommended state-of-the-art method.[12] Discriminative capacity of the model was assessed by receiver operating characteristics (ROC) analysis. Sensitivity was defined as the percentage of correctly predicted successful ECV attempts as compared with all successful ECV attempts, whereas specificity was defined as the percentage correctly predicted unsuccessful ECV attempts as compared with all unsuccessful ECV attempts. Calibration was evaluated by measuring the agreement between the predicted ECV success rates and the mean observed ECV outcomes. A calibration plot was constructed with the predicted and the observed successful ECV rate as calculated from the model in five different categories. Women were qualified for a category based on their prognosis: the first 20% with the lowest predicted chance in the first category, the subsequent 20% in the second category, and so on. In case of a perfect calibration, all predictions and observations would be on the line of equality (X=Y). Finally, we developed a prognostic score chart to make the prediction model useful in clinical practice. The scores were derived from the β coefficients of the initial model. The variable EFW was categorized into practical useful groups, based on the spline function that visualized the association of the probability of a successful ECV and

EFW.[8] By categorizing EFW, a more precise prediction with smaller confidence intervals can be calculated and this is preferable above a dichotomous approach of EFW. The score per group was calculated as the mean probability of success within this group.

# Results

We included all 320 women who underwent a first ECV attempt for breech position within the study period, of which 117 (37%) were successful as the fetus was in cephalic position after the procedure. Baseline characteristics are summarized in Table 1.

**Table 1** - Baseline Characteristics and successful outcome of ECV

| | Missing data | Presence of the characteristics (n=320) | Successful ECV (%) |
|---|---|---|---|
| Gestational age (mean) (range) | 0 | 36 6/7 (35 5/7 – 40 3/7) | 117 (37) |
| Gestational age <37 weeks | | 205 | 69 (34) |
| Gestational age >37 weeks | | 115 | 48 (42) |
| Parity (mean) (range) | 0 | 0.52 (0-5) | |
| Nulliparous | | 198 (62) | 48 (24) |
| Multiparous | | 122 (38) | 69 (57) |
| Maternal age (y) (mean) (range) | 0 | 31 (18-42) | |
| BMI (kg/m$^2$) (mean) (range) | 62 | 22 (17-41) | |
| EFW (g) (mean) (range) | 0 | 2702 (1550-4589) | |
| EFW ≤3000 (%) | | 266 (83) | 93 (35) |
| EFW >3000 (%) | | 54 (17) | 24 (44) |
| Gestational age in days (mean) (range) | 0 | 258 (250-283) | |
| Ethinicity (%) | 0 | | |
| Caucasian | | 308 (96) | 110 (36) |
| Non Caucasian | | 12 (4) | 7 (58) |
| Placental localization (%) | 0 | | |
| Posterior | | 103 (32) | 42 (41) |
| Anterior | | 92 (29) | 21 (23) |
| Fundal | | 72 (23) | 30 (42) |
| Lateral | | 53 (17) | 24 (45) |
| Fetal position (%) | 0 | | |
| Frank Breech | | 46 (14) | 20 (43) |
| Complete Breech | | 251 (78) | 90 (36) |
| Incomplete Breech | | 23 (7) | 7 (30) |
| Amniotic Fluid (%) | 1 | | |
| Decreased amniotic fluid | | 123 (38) | 35 (29) |
| Normal amniotic fluid | | 196 (62) | 82 (42) |

BMI; body mass index, ECV; external cephalic version, EFW; estimated fetal weight

All the ECV attempts performed within the study period were eligible for inclusion in the analysis. None of the ECV attempts were complicated by an adverse event, and no emergency cesarean sections were needed as all fetal heart rate recordings after the attempts were reassuring. Within this study period, the developed prediction model was not used in the study population, as it was not published yet at the end of the study period.
In total, 4.8% of the data points were missing and subsequently imputed. Of the four prognostic factors, 0.3% was missing and imputed. The predicted probability of a successful ECV ranged from 6 to 66%, with a median of 25%. The ROC curve is shown in Figure 1.

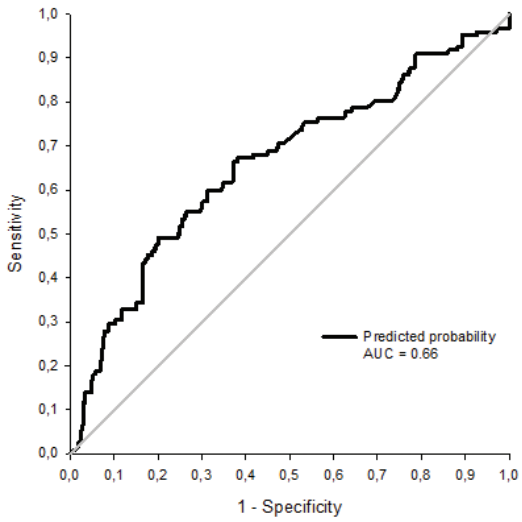


**Figure 1** - Receiver Operating Curve

The area under the ROC curve was 0.66 (95% CI 0.60-0.72). Calibration of the model is shown in Figure 2. The maximum difference between the predicted and observed probability was 14% (Table 2). For all categories, a slight underestimation of ECV success was seen. However, there was no overlap in the confidence intervals of the group with a poor predicted probability (less than 20%) and the group with a good predicted probability (more than 50%), thereby indicating a reliable distinction between these prognostic groups. The Hosmer-Lemeshow test resulted in a p value of 0.30, which points out that there was a good fit of the model.

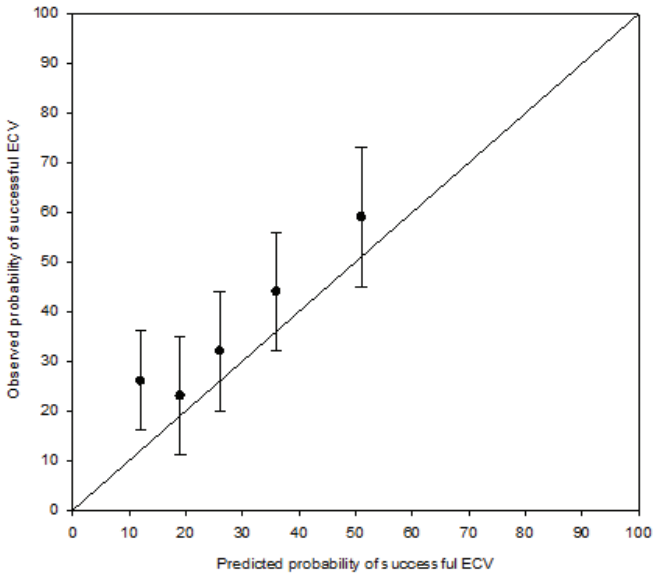


Figure 2 - Callibration plot demonstrating the association between the chance of a successful as predicted by the logistic model and the observed chance of a successful ECV.

Table 2 - Predicted chance of a successful ECV versus observed chance of a successful ECV: calibration.

| Predicted Chance (%) | No. of Patients Per Group | Mean Predicted Chance (%) | No. of Successful ECVs | Mean Observed Chance (%) | Predicted Minus Observed Chance (%) |
|---|---|---|---|---|---|
| 6-16 | 68 | 12 | 18 | 26 | -14 |
| 16-23 | 60 | 19 | 14 | 23 | -4 |
| 23-30 | 68 | 26 | 22 | 32 | -6 |
| 30-41 | 66 | 36 | 29 | 44 | -8 |
| 41-66 | 58 | 51 | 34 | 59 | -8 |

ECV; external cephalic version

# Discussion

We assessed the external validity of our previously developed prediction model of successful ECV for breech presentation at term in a retrospective cohort (n=320). In this study, we demonstrated that the predictive model by Kok et al had a fair discriminative ability.[8] The calibration was fairly good. In external validation, the calibration is almost always less perfect compared with the calibration of the model in the original population, which the model was based on, due to different intercorrelations in the study populations. Nevertheless, this model holds on external validation and it can make a good distinction between women with a poor probability of a successful ECV (less than 20%) and a good probability of a successful ECV (more than 50%). External validation of prognostic models is a vital step, which needs to be performed before the model can be used in clinical practice.[13] Thus far, four studies have developed a prediction model to predict a successful ECV outcome.[14-17] One study tested their model on a second data set, but this study did not address the calibration or the discriminative capacity of the model.[17]

The success rate of 37% is at the lower range of success rates reported in the literature, but similar to the success rate of the population on which the prognostic model was developed.[5,8] The centers in these studies use limited exclusion criteria for ECV. Even though the probability of a successful ECV is expected to be low, women are still offered an ECV, for instance in case of an engaged breech or an anterior located placenta.[18,19]

Our study has several limitations. First, we externally validated an already existing model. As a consequence, other factors known to be predictive of ECV success that were not included in the original model, such as engagement of the breech and palpation of the fetal head, could not be taken into account. Besides, the end point of the model is a cephalic presentation after ECV and not the ultimate goal: a successful vaginal delivery of a child in cephalic presentation. Future studies should evaluate if addition of these factors improve the performance of the model that we validated. Second, data for the external validation were acquired from just one hospital. Therefore, it can be debated if the model fits the overall population of women with a breech presentation. We do think this model can be used for the average woman with a breech presentation as the original model was developed on a population from multiple centers.

The area under the ROC curve of 0.66 indicates a low outcome in assessing the accuracy of a diagnostic test to accurately predict in whom ECV will be successful and in whom it will fail. However, this outcome is to be expected in the assessment of a prognostic test, such as our model. ROC analysis presumes to express the capacity to distinguish between a successful and a nonsuccessful ECV. However, the probability of a successful ECV is never 0% or 100% and even seldom above 60%, which inevitably will affect the area under the curve. In predicting the outcome of a successful ECV, it is more important to separate the patients

with a minimal probability of success from those with a good probability of a success. Therefore, the calibration is more important than the discrimination.[20] This is supported by a cost-effectiveness analysis of ECV by Tan et al., who calculated from a society's perspective that an estimated probability of successful ECV less than 32% costs more to society and has poorer quality-adjusted life years (QALYs) for the patient. With a probability of successful ECV between 32% and 63%, ECV costs more than cesarean delivery but with greater associated QALY such that the cost-effectiveness ratio was less than $50,000/ QALY. If the probability of successful ECV was greater than 63%, the computer modeling indicated that a trial of ECV is less expensive and with better QALYs than a scheduled cesarean.[21]

In clinical practice, this prediction model can also be used as a paper score worksheet (Table 3 and Figure 3). The scores were derived from the β coefficients (β) of multivariable analysis, in which $e^{\beta}$ = odds ratio. Variables were categorized into practical useful groups. The score per group was calculated as the product term of the group mean, the β coefficient, and divided by the factor 0.26. For example, the variable EFW was grouped per 100 g into EFW < 20 (< 20 x 100 = 2000 g), 20 to 25 (2000 to 2500 g), > 25 (> 2500 g). The mean EFW in the 20 to 25 group was 2350 or 23.5 x 100 g and the β was 0.13. All women in this group get an additional score of 12 points (23.5 x 0.13/0.26). The sum of all scores of the different variables resulted in a prognostic index. The prognostic index versus the probability of a successful ECV is plotted on a graph (Figure 3). Once the prognostic index is known, the probability of a successful ECV can be derived from this chart.

The question whether one should withhold ECV to women with a poor predicted probability of success (less than 30%) is still a question under debate. The model underestimates the probability of success in all five groups. This should be taken into account when deciding on a cutoff point to determine when to withhold an ECV attempt, so as to avoid exclusion of too many patients who would have otherwise had successful ECV. In view of the low complication rate of the procedure, one might advocate that even with low success rates, one should attempt ECV. However, it is unknown whether complication rates are similar throughout the entire population or maybe higher in patients with lower success rates. Further research is needed to detect such differences within patient groups to further validate decision making on offering ECV to a selected population.

In conclusion, in this study we performed the vital step in the development of a prediction model, namely external validation. This study demonstrates that the model can distinguish between women with a poor or good probability of successful ECV. It allows the clinician to identify women who could benefit from ECV and this model can support clinical decision making.

**Table 3** - Prognostic Score Chart

| Parameter | Score |
|---|---|
| Parity | |
| Nullipara | 0 |
| Multipara | 4 |
| Placenta localization | |
| Anterior | 0 |
| Non anterior | 1 |
| Amniotic fluid | |
| Normal | 3 |
| Decreased amniotic fluid (AFI <10) | 0 |
| Estimated fetal weight (g) | |
| <2000 | 9 |
| 2000-2500 | 12 |
| >2500 | 14 |
| Prognostic index (total)* | min 9 , max 22 |

*Note: Encircle the prognostic score for each parameter and total them in bottom row. Use the curve in figure 3 to estimate the chance of a successful ECV.
For example, a multipara with an posterior placenta localization, normal amniotic fluid and an estimated fetal weight of 3250 gram has a prognostic index of: 4+1+3+14 = 22 (Table 3). This score corresponds with a 60% chance of a successful ECV.

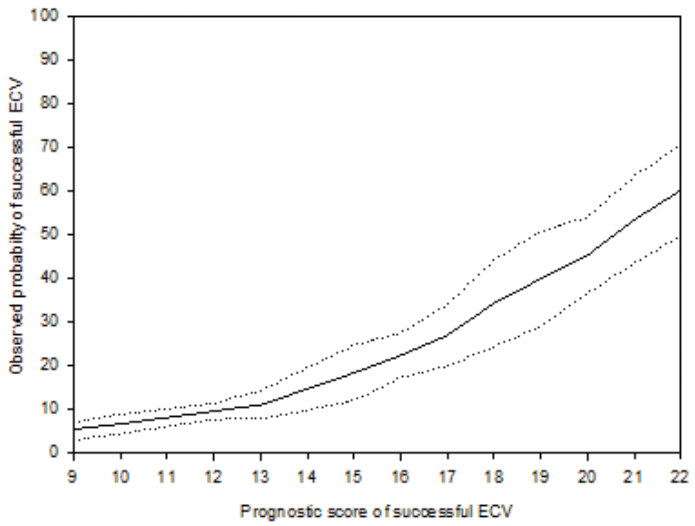


**Figure 3** - Prognostic score versus chance of successful ECV (95% CI)

# References


1. Hannah ME, Hannah WJ, Hewson SA, Hodnett ED, Saigal S, Willan AR; Term Breech Trial Collaborative Group. Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. Lancet 2000;356:1375–1383

2. Rietberg CC, Elferink-Stinkens PM, Brand R, van Loon AJ, Van Hemel OJ, Visser GH. Term breech presentation in The Netherlands from 1995 to 1999: mortality and morbidity in relation to the mode of delivery of 33824 infants. BJOG 2003;110:604–609

3. Sheiner E, Shoham-Vardi I, Hallak M, Hershkowitz R, Katz M, Mazor M. Placenta previa: obstetric risk factors and pregnancy outcome. J Matern Fetal Med 2001;10:414–419

4. Villar J, Valladares E, Wojdyla D, et al; WHO 2005 global survey on maternal and perinatal health research group. Caesarean delivery rates and pregnancy outcomes: the 2005 WHO global survey on maternal and perinatal health in Latin America. Lancet 2006;367:1819–1829

5. Hofmeyr GJ. Interventions to help external cephalic version for breech presentation at term. Cochrane Database Syst Rev 2004;(1):CD000184

6. Kok M, Cnossen J, Gravendeel L, van der Post J, Opmeer B, Mol BW. Clinical factors to predict the outcome of external cephalic version: a metaanalysis. Am J Obstet Gynecol 2008;199:630–637; e1–e7; discussion e1–e5

7. Kok M, Cnossen J, Gravendeel L, Van Der Post JA, Mol BW. Ultrasound factors to predict the outcome of external cephalic version: a meta-analysis. Ultrasound Obstet Gynecol 2009;33:76–84

8. Kok M, van der Steeg JW, van der Post JA, Mol BW. Prediction of success of external cephalic version after 36 weeks. Am J Perinatol 2011;28:103–110

9. Custers IM, Steures P, van der Steeg JW, et al. External validation of a prediction model for an ongoing pregnancy after intrauterine insemination. Fertil Steril 2007;88:425–431

10. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements—a prospective study. Am J Obstet Gynecol 1985;151:333–337

11. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002;7:147–177

12. Schafer JL. Analysis of Incomplete Multivariate Data. London: Chapman & Hall; 1997

13. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003;56:826–832

14. Aisenbrey GA, Catanzarite VA, Nelson C. External cephalic version: predictors of success. Obstet Gynecol 1999;94(5 Pt 1):783–786

15. Fortunato SJ, Mercer LJ, Guzick DS. External cephalic version with tocolysis: factors associated with success. Obstet Gynecol 1988;72:59–62

16. Lau TK, Lo KW, Wan D, Rogers MS. Predictors of successful external cephalic version at term: a prospective study. Br J Obstet Gynaecol 1997;104:798–802

17. Newman RB, Peacock BS, VanDorsten JP, Hunt HH. Predicting success of external cephalic version. Am J Obstet Gynecol 1993;169(2 Pt 1):245–249; discussion 249–250

18. Kok M, Cnossen J, Gravendeel L, van der Post J, Opmeer B, Mol BW. Clinical factors to predict the outcome of external cephalic version: a metaanalysis. Am J Obstet Gynecol 2008;199:630; e1–e7; discussion e1–e5

19. Kok M, Cnossen J, Gravendeel L, Van Der Post JA, Mol BW. Ultrasound factors to predict the outcome of external cephalic version: a meta-analysis. Ultrasound Obstet Gynecol 2009;33:76–84

20. Coppus SF, van der Veen F, Opmeer BC, Mol BW, Bossuyt PM. Evaluating prediction models in reproductive medicine. Hum Reprod 2009;24:1774–1778

21. Tan JM, Macario A, Carvalho B, Druzin ML, El-Sayed YY. Cost-effectiveness of external cephalic version for term breech presentation. BMC Pregnancy Childbirth 2010;10:3