Fetal fibronectin in the prediction of preterm birth
Bruijn, M.M.C.

Citation for published version (APA):
Bruijn, M. M. C. (2016). Fetal fibronectin in the prediction of preterm birth

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary and general discussion
Accurate prediction of preterm birth is a big clinical challenge in obstetrics. Most of the women presenting with symptoms of preterm labour will not deliver within one week and the majority will even deliver at term. Correct discrimination between women with a high and a low risk to deliver on short term would, on the one hand, enable targeted interventions and transfer to an appropriate referral centre to improve neonatal outcomes, and, on the other hand, reduce overtreatment and unnecessary health care costs. Fetal fibronectin (fFN) is the most commonly used biomarker in the prediction of preterm birth. The aim of the research presented in this thesis was to evaluate the use of the fFN test in women with symptoms of preterm labour and to obtain the best approach for individualised risk assessment. First, we focussed on the most commonly used qualitative fFN test, which gives a positive or negative result based on a 50 ng/mL threshold. Subsequently, we evaluated the novel quantitative fFN test in combination with cervical length. Finally, we gave an overview of the current implementation of the fFN test in the Netherlands. Below we summarise the main findings of each chapter.

Concerns have been raised that the validity of the fFN test may be influenced by external factors. However, in some situations it is inevitable to avoid this. In chapter 2 we assessed the effect of external factors, including the use of vaginal soap, digital examination, transvaginal ultrasound, sexual intercourse (all within 24 hours prior to testing), and vaginal bleeding during testing, on false-positive, false-negative and invalid fFN test results, using spontaneous delivery within seven days as the outcome. Out of 708 women, 237 (33%) had a false-positive result. Vaginal bleeding increased the proportion of positive fFN results, but was associated with a lower risk of false-positive test results. The other factors did not show any association. Ten women (1%) had a false-negative result. None of the investigated factors was associated with a higher risk of false-negative results. Twenty-one tests (3%) were invalid; only vaginal bleeding was associated with a higher risk of invalid test results. Based on these results, we agree with the manufacturer's recommendation not to use fFN testing in women with vaginal bleeding. In situations where transvaginal ultrasound or digital examination prior to fFN testing are inevitable, for example after transfer from another hospital, the fFN test result can still be relied on.

It was already known that qualitative fFN and cervical length is a useful combination in the prediction of preterm birth in symptomatic women. There are many other risk factors for preterm birth that have been identified as well, but it is unclear whether these risk factors further improve the prediction in addition to fFN and cervical length. In chapter 3 we compared two prediction models to predict spontaneous preterm birth within seven days in 600 women with symptoms of preterm labour and intact membranes between 24 and 34 weeks of gestation. In total, 73 women (12%) delivered spontaneously within seven days after presentation. The first prediction model included only qualitative fFN and cervical length as predictors. In the second model we added other patient characteristics to these predictors, including maternal age, parity, previous preterm birth, vaginal bleeding, C-reactive protein and dilatation. Both prediction models had a high discriminative capacity, with an AUC of 0.92 (95% CI 0.88 – 0.95) for the first and 0.95 (95% CI 0.92 – 0.97) for the second model. However, the extended model did not result in a more precise identification of women who are at low risk of spontaneous preterm birth on short term
compared to qualitative fFN and cervical length alone.

Many studies focused on the prediction of preterm birth on short term (i.e. within 48 hours or seven days), but it is also interesting to know what the risk is after that timeframe. In chapter 4 we described the predictive value of the qualitative fFN test in combination with cervical length in the prediction of spontaneous preterm birth in women with threatened preterm labour who remained pregnant after seven days. We found that the risk of delivery between seven and fourteen days after initial presentation was increased for women with a cervical length below 15 mm, or a cervical length between 15 and 30 mm and a positive fFN test, compared to women with a cervical length above 30 mm. For spontaneous preterm birth before 34 weeks the risk was increased for women with a cervical length below 15 mm, or with a cervical length between 15 and 30 mm with either a positive or a negative fFN test, compared to women with a cervical length above 30 mm. Thus, risk stratification based on qualitative fFN and cervical length could be used to send women, in whom delivery between seven and fourteen days is highly unlikely, home safely.

In chapter 5 we described the comparison of the qualitative fFN test with another biochemical test, the Actim Partus test, in combination with cervical length measurement in the prediction of spontaneous preterm birth within seven days. The Actim Partus test can be done without the need of an analyser and is therefore cheaper than the fFN test. We tested 350 frozen samples from women with signs of preterm labour between 24 and 34 weeks of gestation and a cervical length below 30 mm. Predictive values of the combination of the Actim Partus test and cervical length were comparable to those of the qualitative fFN test and cervical length. This information is particularly important for settings where the fFN test is not available at all, for example because of the costs of the test. In that case, the Actim Partus test could be considered as an alternative, as it is the cheaper option.

Recently, a new bedside test has been developed, measuring the fFN concentration in the cervicovaginal fluid within ten minutes. Studies evaluating this quantitative test demonstrated enhanced prediction compared to the traditional qualitative fFN test (positive/negative with a threshold of 50 ng/mL) in symptomatic women, but the test has not been evaluated in combination with cervical length.

In chapter 6 we used the same 350 frozen samples as in chapter 5 to evaluate the use of the novel quantitative fFN bedside test to predict preterm birth in symptomatic women undergoing cervical length measurement. Quantification of fFN showed to have added value across the risk range, as the risk of preterm birth within seven days increased with rising fFN concentrations. However, we were not able to demonstrate an improved identification of women with a low risk (below 5%) to deliver on short term.

In chapter 7 we evaluated the predictive value of the quantitative fFN test in combination with cervical length in a European multicentre cohort study. We studied 455 women with symptoms of preterm labour and intact membranes between 24 and 34 weeks of gestation, of whom 48 (11%) delivered spontaneously within seven days. The risk of preterm birth within seven days increased with rising fFN concentrations. We found that in women undergoing cervical length measurement, the quantitative fFN
improves the differentiation between high- and low-risk women (5% risk threshold), compared to the qualitative fFN test. On the one hand, women with a short cervix (below 15 mm) and a very low fFN concentration (below 10 ng/mL) were identified as low risk and could be prevented from overtreatment and unnecessary transfers. On the other hand, women with a long cervix (above 30 mm) and a very high fFN concentration (above 500 ng/mL) were identified as high risk and could benefit from targeted interventions.

The implementation of an effective innovation into routine daily practice is a complex process and often, it takes up to ten years before the innovation is used. The fFN test is a good example of such an innovation with a suboptimal implementation. In chapter 8 we gave an overview of the current implementation of the fFN test in the Netherlands. We sent an online questionnaire to gynaecologists in 86 hospitals to inquire about the use of the fFN test and the reasons in case they did not use the test. The test was used in 34% of the hospitals, and another 16% reported that they expected to start using the test soon. The most reported reason not to use the test was financial; almost half of the respondents that did not use the test found it too expensive. In addition, we performed a retrospective cohort study to measure the actual implementation of the fFN test. We collected data from 96 women with symptoms of preterm labour and intact membranes between 24 and 34 weeks of gestation, who were transferred from general hospitals to the Academic Medical Centre (AMC) in Amsterdam during one year. In 52 women with a cervical length between 15 and 30 mm, an additional fFN test was performed in the AMC; 17 (33%) had a positive result and 35 (67%) had a negative result. In our cohort, 36% (35/96) of the transfers could have been prevented by implementation of the fFN test in the general hospitals, which comes down to a potential cost saving of more than 250,000 euros per year.

IMPLICATIONS FOR CLINICAL PRACTICE AND FUTURE RESEARCH

The work presented in this thesis demonstrated that the traditional qualitative fFN test is a reliable and accurate predictor of preterm birth in women with symptoms of preterm labour, especially in combination with cervical length. Not only for prediction of preterm birth on short term, but also for prediction beyond seven days this combination is useful. Moreover, the novel quantitative bedside fFN test seems to further improve the prediction of preterm birth within seven days in women undergoing cervical length measurement.

However, despite the evidence showing the effectiveness of fFN, we must also conclude that the positive predictive values are still not optimal and yet too many women are exposed to unnecessary interventions, hospital admissions and in utero transfers. Therefore, there is an ongoing need for improved personalised risk assessment. Here, we propose some implications for clinical practice and future research.
Alternative biochemical markers

In chapter 8 we found that the most common reasons not to use the fFN test were the costs of the test and the complexity to perform the test. Therefore, there may be a need for other diagnostic tests that are cheaper and easier to use. A number of other commercially available biochemical tests are available for the prediction of preterm birth in symptomatic women. In chapter 5 we evaluated the bedside Actim Partus test, which measures phosphorylated insulin-like growth factor binding protein 1 (phIGFBP-1) in the cervicovaginal fluid. We showed that with comparable accuracy and lower costs, this test may be an alternative for the qualitative fFN tests in settings where this test is not available or if it is considered to be too expensive. However, future research should focus on prospective evaluation of the Actim Partus test in combination with cervical length and a direct comparison with the fFN test in studies with adequate sample sizes.

Another biochemical marker of interest is placental alpha macroglobulin 1 (PAMG-1). PAMG-1 is a glycoprotein synthesised by the decidua, and is found in high concentrations in amniotic fluid, but in low concentrations in the cervicovaginal fluid. Because of these characteristics, PAMG-1 is useful in the diagnosis of premature rupture of fetal membranes, and the use of the Amnisure ROM test, which is developed to detect the protein, is increasingly accepted as diagnostic tool. In 2012, Lee et al. suggested that the detection of PAMG-1 in cervicovaginal fluid of women with symptoms of preterm labour and clinically intact membranes might represent evidence of micro leakage of amniotic fluid, indicating impending delivery. Accordingly, a device that is more sensitive in the detection of PAMG-1 has been developed in order to detect more women at risk for imminent delivery: the PartoSure™ time-to-delivery (TTD) test (AmniSure International, Boston, MA, USA). Nikolova et al. were the first to evaluate this novel bedside test and demonstrated promising results, with a negative predictive value (NPV) and positive predictive value (PPV) for preterm birth within seven days of 97% and 78%, respectively (n=101). In a subsequent study, they compared the PartoSure test with the fFN test and concluded that the former was superior, especially with respect to the PPV (76% and 29%, PartoSure and fFN, respectively). However, the fFN test was only performed in a small subgroup of the study (n=66, 33%). In addition, they evaluated the combination of the PartoSure test with cervical length, and stated that the PartoSure test was of most added value when the cervical length was between 15 and 35 mm. Although these results show promising short-term prediction potential, the PartoSure test requires further evaluation in studies with larger sample sizes and studies that compare the different preterm birth biomarker tests directly with one another, before it is used as an alternative.

Prediction models

Clinical prediction models combine multiple predictors, such as biochemical markers, biophysical markers and patient characteristics, to accurately predict an individual’s risk of a certain outcome. The popularity of such models has increased greatly over the past years. As mathematical tools, they intend to guide clinicians in their everyday decision-making and aid patient counselling, determining an
individual risk instead of a calculated summary value.\textsuperscript{5} In chapter 2 we developed a prediction model including qualitative fFN, cervical length and other patient characteristics. The results showed that the model did not improve the prediction of preterm birth within seven days, compared to a model with only qualitative fFN and cervical length as predictors. However, this may be caused by a lack of power due to the relatively small number of events in comparison to the number of predictors included in the model. Moreover, we were not able to perform an external validation of the model.

In the United Kingdom, a mobile app has recently been developed as a tool for determining the risk of spontaneous preterm birth in symptomatic women, based on a prediction model incorporating quantitative fFN and previous preterm birth or preterm pre-labour rupture of membranes (PPROM) as predictors.\textsuperscript{6} The model was developed in a training set and validated in a validation set obtained from the same study population. The model performed well when tested on the validation set, but it is unclear whether this model remains accurate in different populations. Future research should focus on external validation of prediction models to evaluate the generalizability. In the UK, cervical length measurement is not part of routine daily practice and is therefore not included in this tool. For the Netherlands, where cervical length measurement is common practice, it would be interesting to extend this model with cervical length as a predictor.

\textbf{Individual Participant Data (IPD) meta-analysis}

With the increasing popularity of prediction research over the last decades, international collaboration has become more commonplace. This has led to an increased sharing of datasets, and to appropriately synthesise these data to develop a single prediction model, the need for IPD meta-analyses has been exposed.\textsuperscript{7} In IPD meta-analyses, the data from multiple studies are analysed all together in one dataset, accounting for the clustering among participants in the same study. Prediction models resulting from IPD meta-analyses are more generalizable than models developed on a single dataset, as the inclusion of different studies results in a more varied study population with more variation in the characteristics of the included participants.\textsuperscript{8}

A good example of an IPD meta-analysis in the prediction of preterm birth is the QUIDS study (Quantitative Fibronectin to help Decision-making in women with Symptoms of Preterm Labour), which is currently conducted in the United Kingdom.\textsuperscript{9} This study aims to develop a decision support tool for the management of women with symptoms and signs of preterm labour, based on a validated prediction model including quantitative fFN and other risk factors. With an IPD meta-analysis, the prediction model will be developed based on five existing European datasets, including the APOSTEL-I and Eufis study (chapter 7), on which this thesis was mainly based.\textsuperscript{10} Subsequently, the model will be externally validated in a prospective cohort study in at least eight centres in the United Kingdom, and a mobile app decision support tool will be developed afterwards. In addition, an economic analysis will be undertaken to assess potential cost-effectiveness of the prognostic model.
Impact analyses

The abovementioned prediction models are also referred to as ‘prediction rules’, providing diagnostic probabilities using a score. We speak of ‘decision rules’ or ‘decision tools’ if clinicians use the probabilities obtained from the prediction rules to help make decisions for patients. After successful development and validation of decision tools, their impact on behaviour in daily practice should be studied in so called impact studies. Impact studies quantify whether the use of such tools improves doctors’ decision-making and consequently patient outcome, as a result of better targeted interventions to women with the highest risk to deliver preterm. Ideally, this is studied in a randomised trial, comparing the use of the tool with a control group of health care professionals who provide usual care.

Economic evaluations

Another important aspect in the evaluation of diagnostic tests or decision tools are economic analyses, as controlling the increase of healthcare costs is a continuing and growing concern. Besides the evidence on clinical effectiveness of available diagnostic strategies, evidence on costs is of importance for both health care professionals and policy makers as well. The results of an original trial could be suggestive that a diagnostic strategy may be worthwhile to introduce into clinical practice, while an economic evaluation may lead to a different conclusion.

A previous cost-effectiveness analysis showed that an additional qualitative fFN test in women with an intermediate cervical length could save €800 per patient, compared to cervical length measurement only, without significant additional health loss. In the Netherlands, this could lead to a cost saving of €4.0 million annually. In chapter 7 we demonstrated that the novel quantitative fFN test may improve the prediction of preterm birth within seven days. However, before the use of this new test is introduced, an economic analysis should indicate whether quantitative fFN testing in women undergoing cervical length measurement is cost-effective. As compared to the additional qualitative fFN testing in women with a cervical length between 15 and 30 mm only, the quantitative fFN strategy requires all women to be tested, which implies a substantial increase in costs of the test.

Implementation of diagnostic tests and decision support tools

Despite continuous development of healthcare innovations with promising results of more effective and safe patients’ care, studies have suggested that about 30-40% of patients do not receive care that is, according to the available evidence, best for them. The implementation of innovations into routine daily practice is a complex process and often, it takes up to ten years until an effective innovation is introduced. In chapter 8 we demonstrated that several years after the first presentation of the results of the APOSTEL-I study, which demonstrated the additional value of the fFN test in combination with cervical length, the fFN test is used in less than 50% of all hospitals in the Netherlands.
Changing clinical practice is not only within doctors’ control; effective change often also requires organisational, economic and political approaches to overcome obstacles in different levels. However, health care professionals have the responsibility to take their role in the process of a successful implementation.

The development of evidence-based guidelines appears to be one of the most effective tools for improving the quality of care. They provide doctors with a summary and translation of the best available evidence and support them to make better choices in daily practice. The aforementioned impact analyses and economic evaluations can be of support in this process. In the Netherlands, the use of fFN is not yet included in the national guideline on threatened preterm labour of the Dutch Association of Obstetrics and Gynaecology (NVOG). Incorporating this effective diagnostic innovation in the national guideline would be a first step to effectuate a change in daily practice, reducing costs and improving health outcome.
REFERENCES


