IHE Report

Post policy implementation review (PPIR) of rapid fetal fibronectin testing for preterm labour in Alberta

March 2015 (updated July 7, 2015)

IHE
INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

INSTITUTE OF HEALTH ECONOMICS

The Institute of Health Economics (IHE) is an independent, not-for-profit organization that performs research in health economics and synthesizes evidence in health technology assessment to assist health policy making and best medical practices.

IHE BOARD OF DIRECTORS

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

IHE Report

Post policy implementation review (PPIR) of rapid fetal fibronectin testing for preterm labour in Alberta

IHE Report: Post Policy Implementation Review Report that examines the health system impact of technology adoption decision conducted by the Alberta Health Technologies Decision Process. The reports are written under contract with the Alberta Health Technologies Decision Process.

Acknowledgements

Post policy implementation review working group:

Dr. Rhada Chari	Ms. Selikke Janes-Kelley (Co-Chair)	Ms. Bridget Smith
Dr. James Wesenberg	Dr. Doug Wilson	Ms. Denise Brind
Dr. Thomas Corbett	Dr. Carolyn DeCoster	Ms. Margo Galvon
Ms. Heather Crosland		

CRM and Data Access, Alberta Health
The Alberta Perinatal Health Program
Lois Hole Hospital for Women – Royal Alexandra Hospital
Calgary Laboratory Services
Laboratory Services, Alberta Health Services
Institute of Health Economics – Dr. Ann Scott, Ms. Christa Harstall, Ms. Dagmara Chojecki, Ms. Carla Martin and Ms. Stefanie Kletke

Corresponding Author

Please direct any inquiries about this report to:
Dr. Anderson Chuck (Co-Chair, Working Group)
Director of Economic Evaluation and Analytics, Institute of Health Economics
achuck@ihe.ca
780-448-4881

Funding

This report was supported by a financial contribution from Alberta Health (AH) through the Alberta Health Technologies Decision Process, the Alberta model for health technology assessment and policy analysis. The views expressed herein do not necessarily represent the official policy of Alberta Health.

Declared Competing Interest of Authors

Competing interest is considered to be financial interest of non-financial interest, either direct or indirect, that would affect the research contained in this report or create a situation in which a person's judgement could be unduly influenced by a secondary interest such as personal advancement. The authors of this publication claim no competing interest.

Suggested Citation (ICMJE or Vancouver Style)

Institute of Health Economics. *Post policy implementation review (PPIR) of rapid fetal fibronectin testing for preterm labour in Alberta*. Edmonton AB: Institute of Health Economics. 2015.

Web Address

This publication is available for free download from the IHE website at http://www.ihe.ca.

Executive Summary

Background

The ability of health care providers to accurately assign risk of imminent PTD among symptomatic women with intact membranes is still limited, as early detection of PTL remains a diagnostic challenge. Therefore, the timely diagnosis of PTL and accurate prediction of risk for imminent PTD in symptomatic women presenting for care at rural or urban clinical settings is an ongoing and important goal for health care providers, in order to enable targeting of effective treatments and avoidance of unnecessary interventions.

The absence of fetal fibronectin (fFN) and of phosphorylated insulin-like growth factor binding protein-1 (*ph*IGFBP-1) in cervicovaginal secretions have shown potential to become clinically useful tests to aid in diagnosing PTL for symptomatic women with intact membranes. These tests are readily available in the form of commercial rapid response testing (point-of-care/bedside) kits (devices/systems).

Technology

Three testing options for ruling out PTL are available in North America. A qualitative rapid assay (the Rapid fFN® for the TLi_{IQ}® System), which gives a positive or negative result. A quantitative rapid assay (the Rapid fFN® 10Q System), which gives information about the fFN concentration level in the cervicovaginal secretion of the tested woman and a rapid response method for *ph*IGFBP-1 detection (the Actim™ Partus test) which is a qualitative assay (www.medixbiochemica.com).

Policy Decision

A review of the evidence on the testing options to rule out PTL was conducted under the auspices of the Alberta Health Technologies Decision Process in 2006 (others were conducted in 2008). The evidence suggested that fFN had the potential to reduce health care utilization and unnecessary treatment by more accurately identifying women who were experiencing false PTL.

On September 19, 2006, the Alberta Ministry of Health issued the following policy decision to the Chief Executive Officers of the Regional Health Authorities (RHAs) in Alberta:

- fFN should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008.
- Given the potential for better and more appropriate care, potential savings, and the modest costs involved, RHAs are asked to fund the introduction and on-going operating costs of fFN testing from within existing budget allocations. The target date is intended to allow sufficient time for health regions to address implementation and budget issues.
- Alberta Health and Wellness supports RHAs taking a provincial approach to developing appropriate quality assurance mechanisms alongside vigilant practice guidelines and standards for the introduction of this service. Regions may also wish to explore other opportunities for collaboration, for example, bulk purchasing arrangements.

The Ministry of Health issued a subsequent letter to the RHAs in March 2008, indicating that health regions could choose the service delivery option that would best serve the needs of its residents.

The rationale for issuing the policy was as follows:

- The evidence (at the time) suggested that fFN had the potential to reduce health care utilization and unnecessary treatment (ambulance transfers and/or length of hospital stay) by more accurately identifying women who were experiencing false PTL.
- Adopting fFN testing province wide would ensure women had equitable access to fFN testing.
- fFN testing would result in cost savings to the provincial health system through the avoidance of ambulance transfers and decreased length of stay.

Objective

All RHAs adopted fFN testing ($TLi_{IQ}^{®}$ System) into the health system between 2006 and 2008. We report on the results of a post-policy implementation review (PPIR) to determine whether the policy achieved its objectives.

Methods

The PPIR consisted of the following components:

- A literature review update was conducted to compare the effectiveness of the $TLi_{IQ}^{®}$ System (fFN), 10Q (fFN) system and he Actim™ Partus test. This was conducted to determine whether there was new evidence since the health technology assessment published in 2008 that more clearly identified differences in performance between the two systems.
- Key informant interviews were conducted with health system implementers from the former Alberta health regions to determine how implementation was conducted in terms of inputs and outputs, and to identify barriers, facilitators, and unintended consequences.
- Economic analysis of the impact on ambulance transfers, hospital admissions, hospital length of stay and health system costs.

Key Findings/Insights

Literature Review Update

Three diagnostic accuracy studies recently conducted in Canada compared the performance of the Actim™ Partus test to that of the $TLi_{IQ}^{®}$ System for predicting PTD in symptomatic women. The following results were reported by these studies:

- Specificity and negative predictive value estimates were high for both tests at most clinical endpoints of interest and did not differ greatly between the two tests, meaning they performed well in predicting the majority of women who were not at risk for PTD.
- Sensitivity and positive predictive values were poor for both tests at all clinical endpoints of interest meaning they did not perform well in predicting the majority of women who were at risk of PTD.
- Compared to the $TLi_{IQ}^{®}$ System, the performance of the Actim™ Partus test was associated with a greater number of false positive results for PTD before 35 or 37 weeks of gestation, and within 7 or 14 days from sampling/testing.

- The LR+ values for the $TLi_{IQ}^{®}$ System were greater than 6.0 for predicting risk of delivery before 35 or 37 weeks of gestation and within 7 days from testing, while the LR+ values for the Actim™ Partus test were lower than 3.0 for these clinical endpoints, meaning that the $TLi_{IQ}^{®}$ System was more accurate in predicting risk of PTD.

According to these results, the overall accuracy of the $TLi_{IQ}^{®}$ System in predicting PTD in symptomatic women appears to be higher than that of the Actim™ Partus test. Hence, in terms of diagnostic performance, there is no evidence to suggest that the system adopted in Alberta should be changed.

Key Informant/Stakeholder Interviews

- All the regional health authorities (RHAs) with representatives who participated in the study fully implemented testing for preterm labour policy within the timeframe indicated by Alberta Health (AH) (2006-2008). All RHAs chose fFN testing to manage patients with symptoms of preterm labour. Availability of fFN testing equipment (that is, specimen collection kits and analyzers) varied across RHAs.
- The majority of RHAs absorbed fFN equipment costs through the existing budgets of Women's Health, Obstetrical Programs, and Laboratory Services Program areas. One RHA allocated additional dollars to programs to purchase fFN test kits and analyzers.
- RHAs trained staff on fFN test collection and analysis using multiple lines of communication (for example, memos and rounds, vendor presentations, orientation session, clinical staff educators, video health teleconferencing, and e-learning opportunities). RHAs mainly drew from the Alberta Perinatal Health Program (APHP) and vendor fFN testing educational resources to train staff and develop fFN testing protocols.All RHAs provided training to obstetrical physicians and nurses and laboratory staff during policy implementation. Informants explained that family physicians may not have received training directly, but would have had access to APHP, vendor, and More^{OB} Program materials. After the policy implementation period, training for fFN testing mainly occurred through new staff orientations.
- Policy implementation was not formally monitored by the government or the APHP; however, some RHAs recorded information pertaining to fFN test usage for procurement purposes.
- A few factors that facilitated policy implementation in the RHAs included:
    - Existing research that demonstrated the efficacy of the fFN testing and encouraged perceptions of its legitimacy among health care providers;
    - Local proponents of fFN testing who championed the establishment of this preterm labour testing option;
    - Intensive staff education efforts;
    - The existence of different groups within RHAs that helped organize and communicate policy implementation;
    - Immediate availability of testing equipment after the policy directive was issued; and
    - An organizational culture that aimed to increase the efficiency of health care service delivery.

- Policy implementation barriers identified by informants included:
    - The costs of fFN testing equipment;
    - Lack of access to fFN analyzers in some of the facilities in the RHAs;
    - Determining how many fFN tests to order and process at the site level;
    - Training staff to use the fFN test only when appropriate; and
    - Ongoing staff education, particularly in sites that manage fewer births on average.
- According to key informants, obstetricians and family physicians generally trusted the fFN test and consider fFN test results as part of their routine for managing patients with symptoms of suspected preterm labour.
- While most informants did not believe that policy implementation resulted in unintended consequences, a few individuals suspected that the policy has not prevented all unnecessary hospital transfers, as not all hospital sites are equipped with an fFN test analyzer.

Economic Analysis

- Physicians placed greater significance on positive test results (inappropriate use of test) compared to negative test results (appropriate use of test), despite the fact that the clinical utility of fFN testing is predicated on a high specificity (approximately 90% reported in the literature, and 98% calculated from administrative databases). Note literature review results that test has low sensitivity and positive predictive values.
- fFN testing did not reduce the number of unnecessary ambulance transfers or admissions for preterm pregnancies in false labour. Unnecessary ambulance transfers increased due to the significance placed on a positive fFN test result. Considerations of other factors such as geographic distance to a level D facility may attenuate the utility of fFN testings because of the time needed to return to a level D hospital if sent home. Level D facilities are those with full obstetrical services and access to tertiary care.
- fFN testing increased the number of appropriate ambulance transfer and admissions for preterm pregnancies in true labour.
- fFN testing did not reduce health system costs by reducing unnecessary resource utilization. In some cases, unnecessary utilization increased with fFN testing which was an unintended consequence.
- Total health system costs increased due to the purchasing of the fFN test kits and analyzers, increased resource utilization associated with unanticipated increases in appropriate clinical care (for example, appropriate transfers and admissions), and increased resource utilization associated with increasing unnecessary health service use.
- The potential maximum cost savings resulting from fFN adoption were small at the outset.

Conclusion

The PPIR suggests that the policy decision to adopt fFN testing in Alberta did not achieve the intended aims of reducing unnecessary utilization of health services to achieve health system savings. Physicians placing greater significance on positive test results compared to negative test results resulted in the inadvertent increase in health care utilization. Hence, when factoring the costs of fFN

testing as well, the total cost for the health system increased. If access to fFN testing services is to continue, it is imperative that further education and training be provided to ordering physicians on how to interpret fFN test results along with a mechanism for ongoing management and assessment of fFN testing that can feed back to these clinicians as well as health system managers.

Abbreviations

All abbreviations that have been used in this report are listed here unless the abbreviation is well known, has been used only once, or has been used only in tables or appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

ACCS	Ambulance Care Classification System
AH	Alberta Health
AHS	Alberta Health Services
APHP	Alberta Perinatal Health Program
Aug	August
CI	confidence interval
95% CI	95% confidence interval
CRD	Centre for Reviews and Dissemination
d	day(s)
DAD	Discharge Abstract Database
DARE	Database of Abstracts of Reviews of Effect
FDA	U.S. Food and Drug Administration
fFN	fetal fibronectin
FN	false negative
FP	false positive
GA	estimated gestational age
h	hour(s)
HTA	Health Technology Assessment
ICD	International Classification of Disease
IHE	Institute of Health Economics
Jan	January
L&D	labour and delivery unit
LIS	Lab Information System
LR	likelihood ratio(s)
LR+	positive likelihood ratio (likelihood ratio for positive results)
LR-	negative likelihood ratio (likelihood ratio for negative results)
min	minute(s)
N	sample size

NCCHTA	National Coordinating Centre for Health Technology Assessment
NHS EED	NHS Economic Evaluation Database
NPV	negative predictive value
NR	not reported
Oct	October
PCD	Physician Claims Databases
*ph*IGFBP-1	phosphorylated insulin-like growth factor binding protein-1
PPV	positive predictive value
PROM	preterm premature rupture of membranes
PTD	spontaneous preterm birth
PTD	spontaneous preterm delivery
PTL	spontaneous preterm labour
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
QC	quality control
RCT	randomized controlled trial
RDS	respiratory distress syndrome
RHA	Regional Health Authorities
s	second(s)
SD	standard deviation
Sn	sensitivity
Sp	specificity
TN	true negative
TP	true positive
TVUS	transvaginal ultrasound (ultrasonography)
UK	United Kingdom
USA	United States of America
US	ultrasound (ultrasonography)
VE	vaginal examination
vs.	versus
wk	week(s)
y	year(s)

Glossary

The glossary terms listed below were obtained and adapted from the following sources:

- Medical Dictionary Online (www.online-medical-dictionary.org)
- fFNTest Glossary of Terms (http://www.ffntest.com/info/utilities/glossary.html)
- Genetics Home Reference Glossary (http://ghr.nlm.nih.gov/glossary=Glossary)
- Medical Glossary (www.medicalglossary.org)

Amnion – The extraembryonic membrane, which contains the embryo and amniotic fluid.

Amniotic fluid – A clear, slightly yellowish liquid (contained in the amniotic sac) that surrounds the unborn baby (fetus) during pregnancy.

Assisted reproductive technologies – All fertility treatments in which both eggs and sperm are handled.

Cervical cerclage – A surgical procedure during which the cervix is sewn closed with suture material during pregnancy.

Cervical length (CL) – Length of the cervix is a measure inversely related to the risk of preterm labour. It is assessed manually or with transvaginal ultrasonography.

Cervical ripening – A complex process that results in physical softening and distensability of the cervix, ultimately leading to partial cervical effacement and dilatation.

Cervix – The lower, narrow end of the uterus where it joins with the top end of the vagina.

Chorion – The outermost extraembryonic membrane.

Contractions – When the muscles of the uterus get tight and then relax. Contractions help push the baby out of the uterus.

Decidua – The term for the uterine lining (endometrium) during a pregnancy, which forms the maternal part of the placenta. This inner layer of the wall of the uterus, which envelops the embryo, is discharged with the placenta.

Effacement – The thinning of the cervix, which occurs before and while it dilates.

Extremely low birth weight – An infant that weighs less than 1000 g at delivery, regardless of the gestational age at birth.

False negative – A negative test result when the condition is present.

False positive – A positive test result when the condition is not present.

Fetal fibronectin – A "glue-like" protein that bonds the baby to the uterus. When the body is getting ready for delivery, this glue breaks down and leaks from the uterus.

Fetal fibronectin test – A test that measures the level of fetal fibronectin in cervicovaginal secretions of a pregnant woman and indicates whether she may be at increased risk for premature labour.

Fetal membranes – Thin layers of tissue which surround the embryo or fetus and provide for its nutrition, respiration, excretion, and protection; they are the yolk sac, allantois, amnion, and chorion.

Foetus/fetus – The unborn offspring of any viviparous mammals, in the postembryonic period, after the major structures have been outlined.

Full term pregnancy – A pregnancy that lasts from 38 to 42 weeks.

Glycoprotein – Conjugated protein-carbohydrate compounds including mucins, mucoid, and amyloid glycoproteins.

Iatrogenic – In a preterm birth context, the term means that the physician decides that the baby needs to be delivered preterm, due to serious maternal or fetal complications.

Index test – The test whose performance is being evaluated.

Insulin-like growth factors (IGF) – Hormones that stimulate protein synthesis and sulfation. IGF I and II play a role in uterine and placental growth and early fetal growth during pregnancy.

Insulin-like growth factor binding protein-1 – One of the six homologous proteins that specifically bind insulin-like growth factors (somatomedins) and modulate their mitogenic and metabolic actions.

Likelihood ratio – A measure of the increase or decrease of the odds of the presence of a disease based on the results of a test. Positive likelihood ratio (LR+) is calculated by sensitivity/1-specificity. Negative likelihood ratio (LR-) is calculated by 1-sentivity/specificity.

Low-birth weight – An infant that weighs less than 2500 grams at delivery, regardless of the gestational age at birth.

Negative predictive value – The probability that a person does not have the target disorder when a negative test result is observed.

Phosphorylation – A chemical reaction resulting in the addition of a phosphate group to a protein or other organic molecule.

Phosphorylated insulin like growth factor binding protein-1 – A carrier protein for insulin-like growth factor 1. The highly phosphorylated isoform (*ph*IGFBP-1) is produced by the decidua. It is an indicator of tissue damage at the choriodecidual interface in pregnant women and a marker of increased risk of infectious complications such as bacterial vaginosis.

Placental abruption – Separation of the placenta from the wall of the uterus.

Premature baby – A baby born before 37 completed weeks of pregnancy.

Preterm birth – Delivery that occurs before 37 completed weeks of pregnancy.

Preterm premature rupture of membranes – Rupture of fetal membranes prior to 37 weeks gestation.

Preterm labour – Labour that occurs before 37 completed weeks of gestation.

Prevalence – The number of events in a given population at a designated time (point prevalence) or during a specified period (period prevalence).

Positive predictive value – The probability that a person has the target disorder when a positive test result is observed.

Respiratory distress syndrome – A syndrome caused in premature infants by developmental insufficiency of surfactant production and structural immaturity in the lungs.

Sensitivity – Probability that the test result will be positive when the disease is present.

Specificity – Probability that the test result will be negative when the disease is absent.

Spontaneous delivery – When a baby is born without being induced.

Swab – A small piece of absorbent material attached to the end of a stick used for obtaining a specimen.

Tocolytic – Medications used to suppress or slow preterm premature labour.

Transvaginal ultrasound (ultrasonography) – Ultrasound (ultrasonography) examination in which a probe is inserted into the vagina.

True negative - A negative test result when the condition is not present.

True positive – A positive test result when the condition is present.

Very low birth weight – An infant that weighs less than 1500 g at delivery, regardless of the gestational age at birth.

Table of Contents

SECTION ONE: Report Introduction

Preface

In the fall of 2012, the Alberta Health Technologies Decision Process (AHTDP)[a] asked the Institute of Health Economics (IHE) to conduct a post policy implementation review (PPIR) of their policy decisions (regarding technology adoption) that could be evaluated. An evaluability assessment of their previous policy decisions was conducted using the tools outlined in the PPIR framework developed by the IHE (www.ihe.ca/publications). Based on the evaluability assessment of previous policy decisions, rapid fFN testing for the detection of false PTL was determined be the best suited to undergo a PPIR.

The PPIR of fFN testing was based on the PPIR framework developed by IHE. The framework provides a theoretical construct and practical guidance and tools for evaluating policies developed within the AHTDP, in order to close the feedback loop from post-policy review to policy development and, thereby, strengthen the link between policy and evidence within the AHTDP. The framework is an attempt to apply scientific methods to the complex and sometimes "unscientific" process of policy evaluation.

Condition

The ability of health care providers to accurately assign risk of imminent PTD among symptomatic women with intact membranes is still limited, as early detection of PTL remains a diagnostic challenge. Therefore, the timely diagnosis of PTL and accurate prediction of risk for imminent PTD in symptomatic women presenting for care at rural or urban clinical settings is an ongoing and important goal for health care providers, in order to enable targeting of effective treatments and avoidance of unnecessary interventions.

The absence of fetal fibronectin (fFN) and of phosphorylated insulin-like growth factor binding protein-1 (*ph*IGFBP-1) in cervicovaginal secretions have shown potential to become clinically useful tests to aid in diagnosing PTL for symptomatic women with intact membranes. These tests are readily available in the form of commercial rapid response testing (point-of-care/bedside) kits (devices/systems).

Technology

Three testing options for ruling out PTL are available in North America. A qualitative rapid assay (the Rapid fFN® for the TLi$_{IQ}$® System) (www.hologicworldwide.com; www.ffntest.com), which gives a positive or negative result. A quantitative rapid assay (the Rapid fFN® 10Q System), which gives information about the fFN concentration level in the cervicovaginal secretion of the tested woman and a rapid response method for *ph*IGFBP-1 detection (the Actim™ Partus test) which is a qualitative assay (www.medixbiochemica.com).

Policy Decision

A review of the published evidence on the testing options to rule out PTL was conducted under the auspices of the Alberta Health Technologies Decision Process in 2008. The evidence suggested that

[a] The AHTDP is an AH initiative that aims to identify and review health technologies to improve health outcomes among Albertans and create a more sustainable health system.

fFN had the potential to reduce health care utilization and unnecessary treatment by more accurately identifying women who were experiencing false PTL.

On September 19, 2006, the Alberta Ministry of Health issued the following policy decision to the Chief Executive Officers of the Regional Health Authorities (RHAs) in Alberta:

- fFN testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008.
- Each RHA may determine the service delivery option that would best serve the needs of its residents.
- Given the potential for better and more appropriate care, potential savings, and the modest costs involved, RHAs are asked to fund the introduction and on-going operating costs of fFN testing from within existing budget allocations. The target date is intended to allow sufficient time for health regions to address implementation and budget issues.
- Alberta Health and Wellness supports RHAs taking a provincial approach to developing appropriate quality assurance mechanisms alongside vigilant practice guidelines and standards for the introduction of this service. Regions may also wish to explore other opportunities for collaboration, for example, bulk purchasing arrangements.

The Ministry of Health issued a subsequent letter to the RHAs in March 2008, indicating that health regions could choose to implement either the Actim™ Partus or the the Rapid fFN® 10Q Syste.

The rationale for issuing the policy was as follows:

- Based on evidence reviews conducted under the auspices of the AHTDP at the time, it was suggested that fFN had the potential to reduce health care utilization and unnecessary treatment (ambulance transfers and/or hospital length of stay) by more accurately identifying women who were experiencing false PTL.
- Adopting fFN testing province wide would ensure women had equitable access to fFN testing. There would not be a disproportionate risk for unnecessary medical care and ambulance transfer, depending on where they resided in Alberta.
- fFN testing would result in cost savings to the provincial health system through the avoidance of ambulance transfers and decreased length of stay.

Objective

All RHAs adopted fFN testing (TLi_{IQ}® System) into the health system between 2006 and 2008. We report on the results of a post-policy implementation review (PPIR) to determine whether the policy achieved its objectives.

The PPIR consisted of the following components:

- A literature review update was conducted to compare the effectiveness of the TLi_{IQ}® System (fFN), 10Q (fFN) system and the Actim™ Partus test. This was conducted to determine whether there was new evidence since the health technology assessment published in 2008 that more clearly identified differences in performance between the two systems.

- Key informant interviews were conducted with health system implementers from the former Alberta health regions to determine how implementation was conducted in terms of inputs and outputs, and to identify barriers, facilitators, and unintended consequences.
- Economic analysis of the impact on ambulance transfers, hospital admissions, hospital length of stay, and health system costs.

The PPIR was coordinated by the IHE with guidance from the fFN working group. The fFN working group was comprised of clinical and health system content experts (see Acknowledgements).

SECTION TWO: Literature Review Update

Paula Corabian, BSc, MPH; Bing Guo, MD, MSc

Introduction

This evidence review updates a rapid review[1] published in 2008, with the intent to summarize the results from the published research that reported on the safety, diagnostic accuracy, clinical effectiveness (in terms of patient and resource usage outcomes), and costs of using the Actim™ Partus test when compared with the Rapid fFN® for the TLi_{IQ}® System (referred to here as the TLi_{IQ}® System). The Actim™ Partus test and the TLi_{IQ}® System were the only rapid response tests licensed in Canada in 2007 to aid in diagnosing PTL and predicting PTD in symptomatic women. Since 2007, Health Canada has also licensed the Rapid fFN® 10Q System (referred to here as the 10Q System) for the same indication.

The Actim™ Partus test has been advocated as a cheaper alternative to the TLi_{IQ}® System, without its limitations.[1] The 10Q System has been introduced as an improvement to the TLi_{IQ}® System because it provides fFN concentration thresholds to potentially help in discriminating the risk of imminent PTD in symptomatic women.[2] Therefore, the aim for the current evidence review was to answer the following question: "How does the Actim™ Partus test or the 10Q System compare to the TLi_{IQ}® System in terms of diagnostic accuracy, clinical outcomes (patients and resource usage outcomes), cost per test, and economic utility when used to aid in diagnosing 'true' PTL in symptomatic women presenting for care with intact membranes?"

The scope of this evidence review was defined as follows:

- ***Population:*** pregnant women (all ages, all ethnic groups, with single or multiple gestations) presenting for care with symptoms and signs of PTL and intact membranes at inpatient or outpatient settings (rural or urban)
- ***Intervention (index test):*** use of the Actim™ Partus test or use of the 10Q System to aid in diagnosing PTL and predicting the risk of imminent PTD in this population
- ***Comparators:*** use of the TLi_{IQ}® System to aid in diagnosing PTL and predicting the risk of imminent PTD in this population
- ***Reference standard:*** a diagnosis based on subsequent course of pregnancy (PTD before 37 weeks of gestation or within a defined period from sampling/testing) in this population
- ***Outcomes:*** diagnostic accuracy (sensitivity, specificity, positive and negative predictive values, and/or likelihood ratios for positive and negative results); clinical outcomes (patient and resource usage outcomes in terms of impact on gestation age at delivery, maternal anxiety/stress, and need for woman's removal from her home support; rates of PTB/PTD, maternal transfers, and hospital admissions; and impact on assessment time, length of hospital stay, use of other diagnostic tests, and use of therapeutic interventions); and costs associated with adding the test for PTL management.

This evidence review included the following elements of assessment:

- a systematic and comprehensive search for scientific literature published on the use of the rapid response testing devices/systems of interest to aid in the diagnosis of PTL and prediction of PTB/PTD in symptomatic women presenting for care with intact membranes at rural or urban clinical settings; and

- a systematic review and critical appraisal of published research studies reporting on the comparative efficacy/effectiveness (diagnostic accuracy and clinical outcomes), and/or costs of using the rapid response testing devices/systems of interest to aid in diagnosing PTL or predicting PTB/PTD for symptomatic women presenting for care with intact membranes.

This evidence review does not cover the use of the rapid response testing devices/systems of interest (alone or in conjunction with other diagnostic tests) for other categories of pregnant women, including asymptomatic women.

More details on the methodology and results for this evidence review are provided in Appendices 2.A-D. Appendix 2.A describes the literature search strategy and summarizes the methodological approach used for study selection, data extraction, quality assessment of selected studies, and data analysis. Appendix 2.B lists the excluded research studies and the reasons for their exclusion. Appendix 2.C summarizes in a tabulated format the characteristics of and the results reported by the included research studies, and Appendix 2.D presents the results for assessment of methodological quality for included studies.

Clinical Condition: Spontaneous Preterm Labour

Spontaneous preterm labour (PTL) is defined as the demonstrated progressive cervical change in the presence of uterine activity (contractions) between 20 and 37 weeks of gestation.[3-14] PTL includes cases with intact fetal membranes and cases with preterm premature rupture of membranes (PROM). Although the pathogenesis of PTL is currently not well understood, its etiology is likely to be complex and influenced by a variety of maternal and fetal causative factors, including cervical infection, cervical inflammation, placental abruption or decidual hemorrhage, uterine stretch or overdistension (for example, resulting from multiple gestation), cervical incompetence (for example, resulting from trauma), and hormonal changes (for example, mediated by maternal or fetal stress).

Recognized risk factors for PTL include a history of previous PTD, multiple gestations, infection (such as chlamydia, gonorrhea, and bacterial vaginosis) and inflammation during pregnancy, as well as maternal nutritional and psychological status.[4-7;9;10;12-22] Maternal ethnicity, extremes of maternal weight (underweight or obesity), extremes of maternal age (< 16 or > 35 years of age), drug abuse, smoking, low socio-economic status, and various diseases during pregnancy (such as heart disease, thyroid disease, diabetes mellitus, and periodontal disease) also contribute to the risk for PTL. Over half of PTL cases resulting in PTD occur in women with no known risk factors.[3;5;6;13;20]

Women with signs and symptoms suggestive of PTL may be at high risk for PTD.[4;5;7;8;10-12;18;19;21;23-25] Although all deliveries or births before 37 weeks of gestation are defined as preterm, prematurity is often divided into subgroups such as extremely preterm (before 28 weeks of gestation), very preterm or early preterm (before 32 weeks of gestation), moderately preterm (at 32-34 weeks of gestation), and late preterm (at 34-36 completed weeks of gestation).[5;12;18;23] PTL in symptomatic women with intact membranes is responsible for up to 50% of PTD cases.[1;3-5;7;9;12]

Prevalence and Incidence

Worldwide, an estimated 13 million infants are born before 37 weeks of gestation annually.[11;23] While the rates of PTD vary between countries, it is estimated that they account for 5% to 15% of all pregnancies.[3;5;7;9;11;12;14;18-20;22-24;26-31] Spontaneous PTD cases, which exclude iatrogenic PTD (indicated for maternal or fetal conditions), represent 66% to 80% of all PTD cases, of which up to a third are associated with PROM.[4;8;9;11;12;12;18;19;27]

In Canada, the PTD rate was approximately 8.1% in 2006-2007, accounting for almost 29,000 births.[18] In 2010-2011, the Canadian PTD rate was 7.9%, and across the country it ranged from 5.2% in the Northwest Territories to 10.7% in Nunavut. Among the Canadian provinces and territories, the PTD rate remained fairly stable from 2006-2007 to 2010-2011.[18;24;32;33]

In Alberta, the PTD rate is significantly higher than the Canadian rate, and has varied over time.[18;24;32;33] After increasing from 7.5% in 1998 to a peak of 9.1% in 2004 and 2005, the rate decreased in 2006 and 2007 to 8.4%. In 2010, the PTD rate in Alberta increased to 8.7% (4,353 live preterm births).[33]

Burden

PTD is a leading cause of neonatal mortality and morbidity in countries with developed market economies.[4;5;7;8;10;11;13;15;18;20;22;34;35] It has been associated with 60% to 80% of deaths in infants without congenital anomalies, it accounts for up to 75% of neonatal morbidity, and it contributes to neuro-developmental problems, respiratory/pulmonary dysfunction, hearing and visual impairment, and other long-term health problems.[1;4;5;7;8;11;15;20;28;31;36] Most morbidity and mortality occur in infants that are delivered or born before 34 weeks of gestation.[4;5;7-12;18;19;37] Spontaneous PTD before 34 weeks' gestation affects about 4% of pregnancies.[8;13;31]

The morbidity and mortality arising from PTD are gestational age-dependent and impose a significant burden on families as well as on health care, education, and social services.[4;5;7-11;13;15;18;20] Neonatal morbidity arising from PTD leads to increased health care costs due to the use of specialized equipment, longer length of hospital stay, and increased health care personnel resources. Some health problems can persist for years, leading to long-lasting use of health care, education, and social services, including special education and rehabilitation for those with physical and/or mental disabilities.

Diagnosis and Management of PTL

There is considerable variation in the way PTL is diagnosed and managed worldwide.[4-12;14;15;19-21;38-41] The goals of clinical management for women presenting for care with symptoms suggesting PTL are the identification of "true" PTL (that is, resulting in PTD) at an early stage and the appropriate application of effective prenatal interventions to reduce perinatal morbidity and mortality. Usually, PTL is diagnosed by clinical history (including assessment of obstetric history, symptoms, and epidemiological risk factors) and physical examination (which usually begins with digital examination of the cervix). The hallmarks of PTL are uterine activity (contractions) and cervical change; however, uniformly accepted standards for diagnosing PTL do not exist.[15;21;38-42]

Up to 75% of the women who present for care with symptoms suggestive of PTL will deliver or give birth at term,[4;5;9-11;13;19;21;22;26;40;43] approximately 10% will deliver within the next 7 days,[11;21;22] and as few as 5% will deliver within the next 14 days.[28] Early identification of "true" PTL and prediction of risk for imminent PTD is challenging and with limited accuracy because early symptoms are often mild, non-specific, and may occur in normal pregnancies.[4;5;7-9;9-11;19;21;22;38-40] When a woman presents with intact membranes and her cervix is dilated less than 3 cm, early diagnosis of "true" PTL is even more difficult to establish.

For women with "true" PTL, there are some potentially beneficial interventions.[4;5;9;10;14;17;19;21;40] These include the initiation of tocolytic therapy when indicated, maternal transfer to a facility with a neonatal intensive care unit, the use of prophylactic interventions (including prenatal steroids to improve fetal outcomes, magnesium sulphate for fetal neuroprotection, and antibiotics for

intrauterine bacterial infection), as well as the use of intensive prenatal surveillance and monitoring. However, the available therapeutic interventions are associated with various side effects and risks of complications for mother and fetus. Remote maternal transfer, when necessary, is disruptive to the woman and her family. All these interventions are costly to the health care system. For these reasons, if possible, any of these interventions should be directed at those women who are most likely to benefit from prevention of PTD.

Harm to the woman and/or fetus can be caused by unnecessary treatments that may follow a false positive diagnosis of PTL.[1;4;5;8-10;14;17;19;28;31] The added psychological stress for the woman and her family, the inconvenience and costs associated with unnecessary maternal transfers, and the costs associated with unnecessary treatment, additional testing, and the use of additional resources to monitor a falsely predicted development of PTL are also unwanted outcomes. Another significant risk is withholding appropriate and effective interventions because of a false negative diagnosis of PTL, which can lead to excessive and unjustified morbidity, inconvenience, and expense as a consequence of PTD.

Given the consequences of false positive and false negative diagnoses of PTL, various diagnostic and predictive markers have been explored as adjunct tools to help clinicians in timely identification of symptomatic women who are in "true" PTL with increased certainty.[1;4;5;8-11;13;15;17;19;20;28;31;34;39] Because complications associated with prematurity are significantly reduced after 34-35 weeks of gestation, most efforts have focused on developing rapid response tests to aid in diagnosing early PTL and identifying imminent risk for PTD before this gestational age.

Rapid Response Tests to Aid in Diagnosing PTL

Over the last 15 years or so, various biochemical markers in different body fluids have been investigated as potential diagnostic markers for PTL and predictive markers for risk of PTD in symptomatic women presenting for care before 34-35 weeks of gestation.[4;5;8-10;13;15;17;19;20;30;31;39;44;45] The detection of proteins such as fFN and *ph*IGFBP-1 in cervicovaginal secretions has shown potential to become a clinically useful rapid response testing method to aid in diagnosing PTL and predicting PTD for symptomatic women with intact membranes.

Because the exact mechanism(s) underlying the onset of labour in humans are unknown,[3-5;7;9;10;12;16;17;20;24] the mechanisms by which fFN and *ph*IGFBP-1 are released in the cervicovaginal secretions remain unknown.[29;37] The release of these proteins is likely attributable to various processes associated with the onset of labour. fFN is a glycoprotein produced by the fetal amnion cells that is found in high concentrations in amniotic fluid and between the chorion and decidua.[37] Cervicovaginal secretions have detectable (high) levels of fFN early in gestation (during the first 24 weeks) and again just before term delivery. The presence of fFN at detectable levels in cervicovaginal secretions between 24 and 34 completed weeks of gestation may indicate disruption of the amnion-chorion interface, and have been associated with an increased risk of PTL and imminent PTD. *ph*IGFBP-1 is a protein produced by placental decidual cells.[29] The process of labour, whether term or preterm, is hypothesized to disrupt the chorio-decidual interface, releasing *ph*IGFBP-1 into cervical secretions. The identification of *ph*IGFBP-1 in cervical secretions would thus be indicative of onset of PTL and predictive of imminent PTD.

The main advantage of *ph*IGFBP-1 and fFN as diagnostic for PTL and predictive markers for PTD is that they can be detected on cervicovaginal secretion samples/specimens that are easy and safe to collect, with minimal risk and discomfort for the fetus and expectant mother.[1;4;5;9;10;13;19;20;28]

Advancements in biomedical engineering have allowed the development of non-invasive rapid response (point-of-care/bedside) test kits (devices or systems) that are commercially available for the detection of these markers. This evidence review concentrates on the use of the rapid response tests that are currently licensed in Canada to aid in diagnosing PTL and predicting the risk of imminent PTD in symptomatic women presenting for care with intact fetal membranes. These tests include the TLi$_{IQ}$® System and the 10Q System for detection of fFN, and the Actim™ Partus test for detection of *ph*IGFBP-1.

The TLi$_{IQ}$® System and the 10Q System

The TLi$_{IQ}$® and 10Q Systems are *in vitro* diagnostic devices (manufactured by Hologic Inc., United States) intended for rapid detection of fFN in cervicovaginal secretions collected from women with single and multiple gestations who present for care with signs and symptoms of PTL, intact fetal membranes, and minimal cervical dilatation (< 3 cm).[2;37;46-48] The TLi$_{IQ}$® System is indicated for qualitative detection of fFN to aid in assessing the risk of PTD within 7 and 14 days from testing for symptomatic women sampled between 24 weeks, 0 days, and 34 weeks, 6 days gestation. The 10Q System is indicated for quantitative detection of fFN to aid in assessing the risk of PTD within 7 and 14 days for symptomatic women sampled between 22 weeks, 0 days, and 35 weeks, 6 days of gestation. It is further indicated for use as an aid to rapidly assess the risk of PTD in less than 34 weeks, 0 days of gestation in women with signs and symptoms of PTL with intact amniotic membranes, and minimal cervical dilatation (< 3 cm) (sampled between 22 weeks, 0 days, and 33 weeks, 0 days of gestation).

Both the TLi$_{IQ}$® and 10Q Systems are lateral-flow, solid-phase immunosorbent assay devices that use the same patented monoclonal antibody for detection of fFN, and are based on the same principle.[2;37;46;46;48;49] Although they both use the same Rapid fFN® Test Specimen Collection Kit,[47] each uses its own special equipment (which includes the specimen cassette kit, the automated analyzer, and the control kit for daily quality control via the QCette). After the test specimen is collected during a sterile speculum examination, the sample at room temperature is added to a solid single-use cassette device (patient specimen cassette), which is placed into an automated analyzer (a hardware device with printer). The automated analyzer uses optical reflectance technology to create a digitized format of the reacted patient specimen cassette. The data are analyzed using multiple parameters, including a comparison of sample data to calibration data.

After a reaction time (20 minutes with the TLi$_{IQ}$® System, or 7 minutes with the 10Q System), the analyzer reads the test specimen cassette and interprets the test result based upon unique test characteristics that must be met (which are pre-programmed in the hardware device).[2;37;46;48;49] For the TLi$_{IQ}$® System, the test result is displayed as positive or negative, based on determining the intensity of the signal derived from the patient sample and ascertaining whether it is greater than, equal to, or less than the signal intensity specified by the reference calibration value (0.50 µg/mL fFN). If inadequate sample is added to the patient specimen cassette, the analyzer will display an invalid result. The 10Q System reports the fFN concentration (a quantitative result, in ng/mL) that indicates the level of fFN in the test specimen.[2;49] The 10Q analyzer reports fFN concentrations ranging from 0 to 500 ng/mL and concentrations greater than 500 ng/mL. The result from the 10Q analyzer is reported as invalid if specific internal test criteria have not been met.

The test result is displayed within 23-25 minutes from specimen collection when using the TLi$_{IQ}$® System, and within 10 minutes when using the 10Q System.[2;37;46;48;49] Total time (from specimen collection to reporting results to the health care provider) depends on the location of the analyzer or

testing site. Quality control (QC) is built into the entire system; both the patient specimen cassette and the analyzer. Upon completion, the analyzer automatically prints and displays the result(s) and the QC information.

Specimens not tested within 8 hours from collection must be stored at 2° to 8°C and assayed within 3 days of collection, or frozen and assayed within 3 months.[2;37;46;47] Specimens can be transported at 2° to 25°C, or frozen. The specimens should not be exposed to temperatures above 25°C. The cassettes and QCettes can be stored at room temperature (15° to 30°C). The shelf life of unopened cassettes is 18 months from the date of manufacture.

The TLi_{IQ}® System or the 10Q System can be performed in a central or hospital laboratory (as a laboratory test), as well as in a labour and delivery unit (L&D) at the bedside (as a point-of-care test) by a laboratory technician or a health care provider (Hologic Inc., personal communication, May 2013). Collection of test specimen is simple to perform, and it appears that there is no risk to the mother or fetus from performing the test itself when following the procedure recommended by the manufacturer.[1;28;50] Minimal training is required to perform each test, and the necessary training on specimen collection and running the analyzer is provided free of charge by the manufacturer (Hologic Inc., personal communication, May 2013).

Limitations

There are some limitations associated with the use of the TLi_{IQ}® or 10Q Systems, due to several factors that can confound the interpretation of their results.[2;37;46;47] The fFN concentration may be elevated by cervical disruption caused by, but not limited to, events such as sexual intercourse, digital cervical examination, or vaginal probe ultrasound examination. Manipulations of the cervix may lead to false positive test results. Therefore, specimens should be collected prior to digital and/or transvaginal ultrasound examinations, and not within 24 hours after cervical manipulation or after sexual intercourse. However, according to the manufacturer, when a woman reports having had sexual intercourse within the previous 24 hours, a negative result (from the TLi_{IQ}® System) or an fFN concentration of less than 10 ng/ml (from the 10Q System) can be considered valid.[2;37;46]

These two tests are not intended for women with advanced cervical dilatation (≥ 3 centimeters), cervical cerclage, suspected or known placental abruption, placenta previa, or visual evidence of moderate or gross vaginal bleeding.[2;3;37;46;47] False positive test results can occur in the presence of vaginal bleeding and cervical cerclage. Other contraindications include sampling before 22 weeks of gestation, PROM, and abnormal vaginal flora. fFN is found in amniotic fluid, so the test is irrelevant once the fetal membranes rupture. Contamination of the collection swab with lubricants, soaps, disinfectants, or creams may also interfere with the test results.[2;37;46;47]

Test results from the TLi_{IQ}® and 10Q Systems cannot be interpreted visually, and must be interpreted by the automated analyzer.[2;46;48;49]

Clinical use

The current clinical use of the TLi_{IQ}® System remains defined by its strong negative predictive value (NPV), while the clinical importance of a positive test result remains unclear.[3;5;7-11;17;22;26;28;39;43;51;52] The NPV estimates for the TLi_{IQ}® System reported by the available research vary, depending on the gestational age at onset of PTL symptoms and the time between the onset of symptoms and delivery. Its high NPV (over 90%), in conjunction with clinical assessment, has been shown to rule out "true" PTL, and it is deemed as a potent predictor of low risk of imminent PTD in symptomatic

women with singleton or multiple gestations (tested between 24 and 34 weeks of gestation, with intact membranes and minimal dilatation of < 3 cm) within the next 7 to 14 days from testing.

The clinical and economic utility of adding the TLi$_{IQ}$® System to the management of PTL in symptomatic women have been evaluated in several prospective cohort studies and randomized controlled trials (RCTs).[3;17;26;28;36;38;43;53;54] However, its clinical and economic impact when used for this indication is still uncertain.[26;28;36]

Deshpande et al.[55] recently published the results from a systematic review and cost analysis, which assessed the accuracy, clinical effectiveness, and cost-effectiveness of rapid fFN testing (using the TLi™ System or the TLi$_{IQ}$® System, also known as FullTerm™) in predicting PTD in symptomatic women. The comparator was usual care. Evidence from 15 diagnostic accuracy studies suggest that fFN testing has a moderate accuracy for predicting PTD (within 7-10 days of testing, <34 weeks' gestation, or <37 weeks' gestation), and may be most sensitive for predicting PTD within 7-10 days of testing. Evidence from RCTs suggests that fFN does not increase adverse outcomes and may reduce resource use. The studies included in this systematic review did not provide information on the effect of fFN testing on clinical decision-making, and no RCT reported significant effects of fFN testing on maternal or neonatal outcomes.

The base-case cost analysis conducted by Deshpande et al.[55] showed a modest cost difference in favour of fFN testing, which is largely dependent on whether or not fFN testing indeed reduces hospital admission. According to Deshpande et al., "[t]his depends on precisely the place of fFN testing in the care pathway (i.e. essentially the weight placed on the fFN test results in conjunction with or as opposed to other information such as signs, symptoms and physical examination). When fFN testing reduces admissions testing will be very likely to save costs. When it does not, there obviously is only a very limited possibility that fFN testing will save costs; given the assumption that testing will not impact on the delivery and subsequent events."[55]

Berghella et al.[36] recently conducted a Cochrane systematic review and meta-analysis to assess the effectiveness of PTL management based on knowledge of fFN results for preventing PTD. They reviewed five RCTs that included 474 symptomatic women (most with singletons), of which 235 were randomized to PTL management with knowledge of fFN results, and 249 to PTL management without knowledge of fFN results. Although they found that PTD before 37 weeks of gestation was significantly decreased with management based on knowledge of fFN results (15.6% when using fFN versus 28.6 % when not using fFN), the reviewers concluded there was not enough evidence to recommend its use, and they encouraged further research.

One multicentre observational study[2;50] was recently conducted to evaluate the use of the 10Q System in 300 symptomatic women with singletons (between 22 and 35 weeks of gestation). The purpose of this first and only study conducted to date on the use of the 10Q system in symptomatic women was to determine whether quantification of cervicovaginal fFN concentration improves predictive accuracy for spontaneous PTD. In this study, all participating clinicians were trained in the use of the 10Q and TLi$_{IQ}$® Systems, and the two tests were performed concurrently. Clinicians were blinded to the result from the 10Q System until after the delivery, while the result from the qualitative TLi$_{IQ}$® System was made available. Overall, there was a rate of 8.7% for spontaneous PTD before 34 weeks of gestation, and a rate of 12% for spontaneous PTD before 37 weeks of gestation. The reported NPV estimates for PTD within 7 or 14 days from sample collection and before 34 weeks of gestation were over 95% for all fFN concentrations (10, 50, 200, and 500 ng/mL).[50] According to the reported results, the increasing concentration of fFN correlated with

increased risk of spontaneous PTD within 14 days from sample collection and before 34 weeks of gestation.[2;50] The results suggest that the quantitative measurement of fFN concentration using the 10Q has the potential to more accurately identify symptomatic women who are most likely to deliver preterm, compared to qualitative detection of fFN (by using the TLi$_{IQ}$® System).

To date, no study has evaluated the clinical and/or economic utility of adding the 10Q System to the management of PTL in symptomatic women.

Costs

The cost for the TLi$_{IQ}$® System is approximately $2,400 CAD, and the cost per patient/test is about $100 (Hologic Inc., personal communication, May 2013). The cost per test is for the disposables. There is no other laboratory equipment or upkeep that is above and beyond standard laboratory supplies, practices, or procedures. Service, replacement, and training is provided free of charge by Hologic Inc. The cost of the TLi$_{IQ}$® System may vary somewhat by the volume of testing being done at an account. The cost for a 10Q System and the cost per patient/test are comparable to those for the TLi$_{IQ}$® System (Hologic Inc., personal communication, May 2013).

Regulatory status in North America

In Canada, the TLi$_{IQ}$® and 10Q Systems are currently licensed as Class III devices to aid in rapidly assessing the risk of PTD within 7 and 14 days from sample collection in symptomatic pregnant women with intact membranes and minimal dilatation (< 3 cm) (Hologic Inc., personal communication, May 2013).[1] The TLi$_{IQ}$® System is also approved by the Food and Drug Administration (FDA) in the United States for the same indication. The 10Q System is not approved by the FDA at this time (Hologic Inc., personal communication, May 2013).

Utilization in Canada

Currently, there are more than 200 TLi$_{IQ}$® System units in use in Canada, of which 50% are installed in rural settings and 50% in urban settings (Hologic Inc., personal communication, May 2013). About 20 TLi$_{IQ}$® System units are installed in Alberta, and the majority of the other units are in British Columbia, Ontario, Quebec, and Nova Scotia. fFN testing with the TLi$_{IQ}$® System is also available in Nunavut and the Yukon. Hologic Inc. has not sold any 10Q System units in Canada.

The TLi$_{IQ}$® System is used as a laboratory test in Canada and requires a technician or health care provider to run the test (Hologic Inc., personal communication, May 2013). However, it can also be performed in a L&D unit at bedside, if preferred.

Several Canadian studies have evaluated the clinical application of the TLi$_{IQ}$® System.[1;28;53;56-58] These studies reported that knowledge of a negative test result used to complement clinical diagnosis had a significant impact on the evaluation of risk for PTB/PTD in Level 3 health care centres, and also in Level 1 and Level 2 centres that lack the resources for intensive care of the preterm newborns. The impact was reported in terms of reducing the rates of and costs associated with unnecessary maternal transfers and hospital admissions, as well as reducing the use of unnecessary therapeutic interventions and the indirect costs associated with displacement of the mother from her family and community.

The Actim™ Partus Test

The Actim™ Partus test (manufactured by Medix Biochemica, Finland) is a qualitative immunochromatographic dipstick test based on monoclonal antibodies for detecting the presence of

*ph*IGFBP-1 in cervical secretions during pregnancy (www.medixbiochemica.com).[29;59;60] Medix Biochemica markets the Actim™ Partus test as a laboratory or a rapid bedside (point-of-care) test kit. Alere Canada (www.alere.ca) is currently the only distributor of the Actim™ Partus test kits in Canada (Alere Canada, personal communication, April 2013).

According to the manufacturer, the Actim™ Partus test is most useful for pregnant women with singletons and a gestational age of 22 weeks until term who present for care with signs and symptoms of PTL (www.medixbiochemica.com).[16] The test is intended for professional use to aid in predicting the risk of imminent PTD when fetal membranes are intact. A negative test result is considered a clear indication that the symptomatic woman will not deliver within the next 7 to 14 days from testing. The results are available within 5 minutes from specimen collection, and no reader device is required to interpret them (test results are visually interpreted by user).

The Actim™ Partus test is intended for *in vitro* diagnostic use only. The test is simple to perform. and it appears that there is no risk to the mother and fetus from performing the test itself when following the procedure recommended by the manufacturer (www.medixbiochemica.com).[1;29;59-61] During a sterile speculum examination, cervical secretion is collected from the endocervix with a swab provided in the test kit. The swab is inserted and swirled vigorously for 10-15 seconds in an extraction solution, in which the dipstick is dipped after the swab is discarded. The dipstick is kept in the extraction solution until the liquid front becomes visible in the result area (window). Then it is removed and let to develop for 5 minutes in a horizontal position. A blue line (sample line) will appear in the result area (window) if the concentration of *ph*IGFBP-1 in the sample exceeds the cut-off value for the test (10 µg/l). A second blue line (control line) confirms correct performance of the test.

The result can be interpreted as positive (indicating elevated risk of imminent PTD) as soon as the two blue lines (sample and control lines) become visible in the result area (www.medixbiochemica.com).[29;59-61] A negative test result (indicating that imminent PTD within the next 2 weeks is highly unlikely) must be confirmed at 5 minutes. If only the control line has appeared after 5 minutes, the test result is interpreted as negative. If a control line does not appear, the test is considered invalid. The test requires at least 150 µl of extracted cervical fluid to perform correctly.

The Actim™ Partus test has a shelf life of 24 months from the manufacture date, and can be stored at 2° to 25°C (www.medixbiochemica.com).[59] The manufacturer recommends testing the specimen immediately. If necessary, the specimen can also be stored for up to 4 hours after collection. If a specimen cannot be tested within this time, it should be frozen. After thawing, the specimens should be mixed and tested as described in the pack insert.

The Actim™ Partus test can be performed in the hospital laboratory or at the bedside in both rural and urban health care settings, either by a laboratory technician, a physician, or a nurse (Alere Canada, personal communication, April 2013). The performance of the Actim™ Partus test is not dependent on outside laboratory personnel and equipment. If required, Alere Canada will provide any training support documents and education.

Limitations

Because phIGFBP-1 is undetectable in urine or seminal fluid, the mechanical stress caused by recent intercourse does not affect Actim™ Partus test results (www.medixbiochemica.com).[29;59;60] However, specimens for the Actim™ Partus test should be collected prior to digital and/or transvaginal

ultrasound examinations (www.medixbiochemica.com).[29;59;60] Before performing the Actim™ Partus test, the manufacturer recommends ensuring that the fetal membranes are intact by performing a test to detect PROM. With ruptured fetal membranes, the test will give a positive result. The manufacturer also recommends not testing women with moderate or heavy vaginal bleeding. The specimens for the test should not be contaminated with blood to avoid false positive results. The test result indicates the risk at the time of sampling, and changes in the woman's condition may later affect the final outcome of the pregnancy.

Clinical use

Available clinical data regarding the validity of *ph*IGFBP-1 as a diagnostic and predictive marker for PTL and risk of imminent PTD and the efficacy of the Actim™ Partus test when used to aid in predicting risk of PTD in symptomatic women is limited to several observational test-accuracy studies published since 2002 (www.medixbiochemica.com).[1;5;8;20;31;39;53;61;62] The reported NPV estimates suggest that the Actim™ Partus test can be a clinically useful tool for ruling out "true" PTL in symptomatic women with singletons presenting for care between 20 and 37 weeks of gestation with intact membranes. Its high NPV (over 90%) may be a reassuring sign that the likelihood of imminent PTD in women sampled between 22 and 36 weeks of gestation is low within the next 7 to 14 days from testing. However, the clinical importance of a positive test result remains unclear.

In the available observational studies, the clinical staff in charge was blinded to the Actim™ Partus test results, and managed the suspected PTL cases according to the standard protocols in their clinical settings. None of these studies reported on the clinical and/or economic utility of adding the test to PTL management, in terms of improved patient outcomes and reduced resource usage and associated costs.

Costs

According to the manufacturer, everything required for running the Actim™ Partus test is included in an individually packaged kit, with no additional capital equipment required to perform or read the test (www.medixbiochemica.com).[61] The Actim™ Partus products are available in kits of 10 tests. The cost is $25 CAD per test ($250 CAD/kit) (Alere Canada, personal communication, April 2013).

Regulatory status in North America

The Actim™ Partus test is currently licensed in Canada as a one-step dipstick test (Class III device) for detecting the presence of *ph*IGFBP-1 in cervical secretions to predict PTD or susceptibility to deliver at term when fetal membranes are intact (Alere Canada, personal communication, April 2013).[1] The test is not currently approved for marketing by the FDA; however, it is undergoing clinical evaluation in the United States (Alere Canada, personal communication, April 2013). The timeline for its clearance by the FDA and its availability on the market is not yet known.

Utilization in Canada

The promotion of the Actim™ Partus test to Canadian health care settings is still in its early stages. Currently in Alberta, there are no health care centres using the Actim™ Partus test (Alere Canada, personal communication, May 2013). However, in other parts of Canada, the test is used by physicians in rural and urban settings (at the bedside or in the laboratory).

To date, there is no published Canadian study that has investigated the clinical and/or economic utility of adding the Actim™ Partus test to the management of PTL in symptomatic women.

Summary of the Available Rapid Response Tests' Characteristics

Table 2.1 summarizes the characteristics of the $TLi_{IQ}^{®}$ and 10Q Systems and the Actim ™ Partus test, as they are presented by their manufacturers/distributers.[2;29;37;46-49;59-61]

Table 2.1: Characteristics of the Actim™ Partus test and $TLi_{IQ}^{®}$ and 10Q Systems

Characteristics	Actim™ Partus test	TLi_{IQ}® System	10Q System
Intended use	To aid in predicting PTD in symptomatic women (singleton gestation, intact membranes confirmed by first performing a test to detect PROM) sampled between 22^{+0} wk GA until term	To aid in predicting PTD in ≤ 7 or ≤ 14 d from testing in symptomatic women (singleton/multiple gestation, intact membranes, cervical dilation < 3 cm) sampled between 24^{+0} and 34^{+6} wk GA	To aid in predicting PTD in ≤ 7 or ≤ 14 d from testing in symptomatic women (singleton/multiple gestation, intact membranes, cervical dilation < 3 cm), sampled between 22^{+0} and 35^{+6} wk GA To aid in predicting PTD before 34 wk in symptomatic women (intact membranes, cervical dilation < 3 cm), sampled between 22^{+0} and 33^{+0} wk GA
Contraindication	Moderate to heavy bleeding, PROM, sampling at < 22 wk GA	Excessive blood, sexual intercourse or digital exam within previous 24 h, PROM, sampling at < 22 wk GA	Excessive blood, sexual intercourse or digital exam within previous 24 h, PROM, sampling at < 22 wk GA
Safety	No risk to woman or fetus from performing test itself following manufacturer recommended procedure	No risk to woman or fetus from performing test itself following manufacturer recommended procedure	No risk to woman or fetus from performing test itself following manufacturer recommended procedure
Clinical use	Knowledge of a negative test result may supplement clinical judgment to rule out PTL and predict low risk of PTD in symptomatic women within 7 to 14 d from testing	Knowledge of a negative test result may supplement clinical judgment to rule out PTL and predict low risk of PTD in symptomatic women within 7 to 14 d from testing	Knowledge of quantitative fFN levels may supplement clinical judgment in discriminating risk of PTD in symptomatic women within 7 to 14 d from testing and at < 34 wk GA
Assay format	Immunochromatographic dipstick test	Lateral-flow cassette, optical reader	Lateral-flow cassette, optical reader
Classification	Qualitative *in vitro* diagnostic device	Qualitative *in vitro* diagnostic device	Quantitative *in vitro* diagnostic device
Specimen collection	Speculum exam, specific swab and collection tube (supplied); collect specimen before digital exam and/or TVUS from endocervix; confirm membranes are intact by first performing a test to detect PROM	Speculum exam, specific swab and collection tube (supplied); collect specimen before digital exam and/or TVUS from posterior fornix of vagina; discard specimen if > 3 cm dilated	Speculum exam, specific swab and collection tube (supplied); collect specimen before digital exam and/or TVUS from posterior fornix of vagina; discard specimen if > 3 cm dilated
Specimen stability	Once collected, < 4 h at RT, then should be frozen	Once collected, 8h at RT, 3 d at 2° to 8°C, or frozen and assayed within 3 mo	Once collected, 8h at RT, 3 d at 2° to 8°C, or frozen and assayed within 3 mo

Storage requirements	Test kit at 2° to 25°C; test kit and packs at 2° to 30°C (for 2 mo)	Cassettes and QCettes at 15° to 30°C	Cassettes and QCettes at 15° to 30°C
Assay timing	Controlled by user	Controlled by analyzer	Controlled by analyzer
Test interpretation	Based on visual reading by user	Based on reader device (analyzer)	Based on reader device (analyzer)
Time to result	5 minutes from specimen collection to reporting result	23-25 minutes from specimen collection to reporting result	10 minutes from specimen collection to reporting result
Test result	Positive/negative	Positive/negative	Quantification of fFN concentration
Patient record result	None	Print label and display	Print label and display
Traceability	None	Result, time, date, patient ID, user ID, cassette lot, calibration code, analyzer ID, internal controls	Result, time, date, patient ID, user ID, cassette lot, calibration code, analyzer ID, internal controls
Patient record storage	None	Results stored and retrievable	Results stored and retrievable
Performance requirements	Hospital laboratory or at bedside in rural and urban settings No reader device required	Central or hospital laboratory or at bedside in rural and urban settings Reader device required	Central or hospital laboratory or at bedside in rural and urban settings Reader device required
Qualified personnel to perform test	Physicians, nurses, or laboratory technicians	Physicians, nurses, or laboratory technicians	Physicians, nurses, or laboratory technicians
Training requirements	Minimal training (provided free of charge by Alere Canada)	Minimal training (provided free of charge by Hologic Inc.)	Minimal training (provided free of charge by Hologic Inc.)
Costs*	$25 CAD per test; sold in kits of 10 tests ($250 CAD per kit)	~$2,400 CAD per system; ~ $ 100 CAD per patient result	~$2,400 CAD per system; ~ $ 100 CAD per patient result
Regulatory status*	Health Canada licensed Not FDA approved	Approved by Health Canada and FDA	Health Canada licensed Not FDA approved

°C = degrees Celsius; CAD = Canadian dollars; d = day(s); FDA = Food and Drugs Administration in the United States; GA = gestational age; h = hour(s); mo = month(s); QC = quality control; PROM = preterm premature rupture of membranes; PTD = spontaneous preterm delivery; PTL = spontaneous preterm labour; RT = room temperature; TVUS = transvaginal ultrasound; wk = week(s)

* This information was obtained from personal communication with the Canadian distributor of the Actim™ Partus test (Alere Canada) and the manufacturer of the $TLi_{IQ}^{®}$ and 10Q Systems (Hologic Inc., United States)

Guidelines and Patient Test Protocols

Some of the published clinical practice guidelines on the diagnosis and management of PTL and assessment of risk for imminent PTD recommend only the use of rapid fFN testing to complement clinical assessment for diagnosing PTL in symptomatic women (when clinical diagnosis is doubtful, to identify women at low risk for imminent PTD).[11;15;21;41;42;63] Other guidelines recommend the use of rapid fFN testing and also mention rapid *ph*IGFBP-1 testing as an alternative to rapid fFN testing for this indication.[39;40] Follow-up on positive fFN test results is unclear, and no specific algorithms exist at this time.

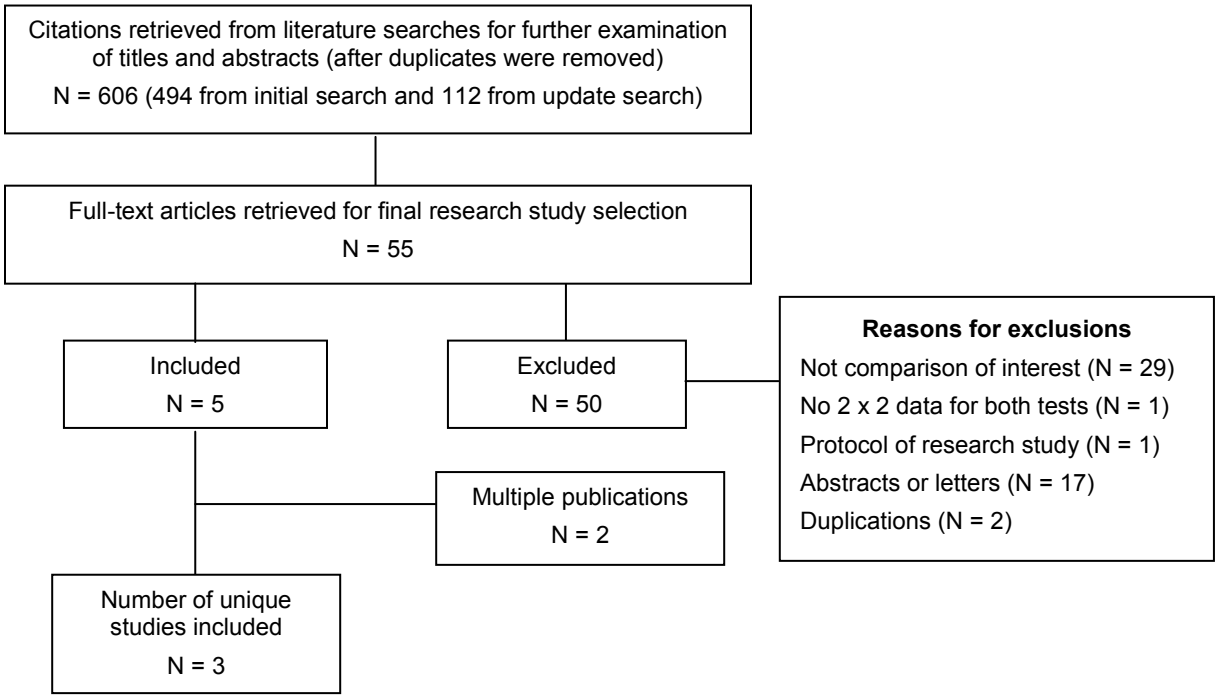
Patient test protocols and guidelines for using the TLi$_{IQ}$® System were developed by the Canadian Perinatal Partnerships Coalition,[57] and by several reproductive care and perinatal programs in Alberta,[64] British Columbia,[1] Ontario,[65;66] and Nova Scotia (http://rcp.nshealth.ca/resources-reports/fetal-fibronectin-working-group).[1]

The literature searches conducted for this evidence review did not identify any guidelines or patient test protocols specifically developed for the use of the 10Q System or the Actim™ Partus test to aid in diagnosing PTL in symptomatic women.

Available Research Evidence

The initial comprehensive literature search conducted for this evidence review between 20 April and 20 June 2013 identified 494 citations. The update search conducted on 11 February 2014 yielded another 112 citations. The application of the selection criteria to 55 full-text articles retrieved as potentially relevant research studies resulted in 50 being excluded (the main reasons for their exclusions are listed in Table 2.B.1, Appendix 2.B). Figure 2.1 below outlines the research study selection process for this review.

Figure 2.1: Research study selection process



The literature searches conducted for this evidence review did not identify any systematic reviews of clinical studies that directly compared the use of the rapid response tests of interest in terms of any of the outcomes of interest. The literature searches also did not identify any primary research studies that specifically compared the clinical and/or economic utility of adding the rapid response tests of interest to the management of PTL.

Three diagnostic accuracy studies met the selection criteria developed for this evidence review and were included for data extraction and analysis. [35;67-70] None of these studies compared the use of the

10Q System versus the use of the TLi$_{IQ}$® System. The only study conducted to date on the use of the 10Q System in symptomatic women aimed to determine whether quantification of cervicovaginal fFN concentration improves predictive accuracy for spontaneous PTD.[2;50] Although in this study both the 10Q System and the TLi$_{IQ}$® System were performed in the same population, the objective was not to compare the performance of the two tests for the prediction of spontaneous PTD. This study was not included in this evidence review for data extraction and analysis because it did not report 2 x 2 data for both tests.

Description of Selected Studies

All three studies selected for this evidence review were conducted in large Canadian urban cities: Victoria, Calgary, and Montreal (see Table 2.2 and Appendix 2.C, Table 2.C.1). Two studies[35;67] were published in peer-reviewed journals in 2010 and 2012. The evidence obtained from the third study was reported in 2008 in a poster presentation at the 60th Annual Meeting of the American Association for Clinical Chemistry,[68;69] and in an oral presentation at the 64th Annual Clinical Meeting of the Society of Obstetricians and Gynaecologists of Canada.[70] Table 2.C.1 in Appendix 2.C provides information about the individual studies (study's and participants' characteristics, index test, comparator test, reference standard, and outcomes) and summarizes their findings.

All selected studies[35;67-70] directly compared the performance of the Actim™ Partus test (index test) versus the TLi$_{IQ}$® System (comparator test) for predicting PTD in women who presented for care between 22 and 34 weeks of gestation with symptoms and signs of PTL and intact membranes (see Table 2.2 and Appendix 2.C, Table 2.C.1). In each study, the reference standard was the outcome of pregnancy: spontaneous PTD reported at different cut offs of gestational age (before 34 weeks, 35 weeks, and/or 37 weeks of gestation) and/or within the next 7 to 14 days (or within 2 weeks) from sample collection (testing). None of the selected studies reported on clinical outcomes of interest (patient and resource usage outcomes), or on costs associated with adding the tests to management of PTL.

Table 2.2: Summary of selected research studies

Study	Study characteristics and authors' conclusions
Cooper et al. (2012)[35] **Type:** cohort study, prospective enrollment (NR if consecutive or random) **Objective:** to examine performance of *ph*IGFBP-1 test in predicting PTD in symptomatic women and to compare characteristics of *ph*IGFBP-1 and fFN tests **Duration:** Oct 2005 to May 2009 **Setting:** tertiary care centre **Province (city):** Alberta (Calgary)	**Sample*:** 288 symptomatic women (singletons and multiple pregnancies; nulliparous and multiparous; with and without previous PTD) **Testing week*:** 24-34 wk GA **Index test*:** Actim™ Partus test (commercially available kit) **Comparator test*:** TLi$_{IQ}$® System (commercially available kit) **Reference standard*:** PTD **Clinical endpoints*:** PTD < 37 wk GA **PTD prevalence*:** 16% for PTD < 37 wk GA **Authors' conclusion**:** "NPV did not differ between *ph*IGFBP-1 and fFN for delivery < 37 weeks. Neither test improves on pretest probability of delivery < 37 weeks, so clinicians must decide whether the use of either test is justified."
Audibert et al. (2011)[67] **Type:** cohort study, prospective non consecutive enrollment **Objective:** to validate use of *ph*IGFBP-1 as a predictor of PTD	**Sample*:** 62 symptomatic women (singletons and twins; nulliparous and multiparous; with and without previous PTD) **Testing week*:** 24-34 wk GA **Index test*:** Actim™ Partus test (commercially available kit) **Comparator test*:** TLi$_{IQ}$® System (commercially available kit)

Duration: Jan 2006 to Jan 2007 **Setting:** tertiary care centre **Province (city):** Quebec (Montreal)	**Reference standard*:** PTD **Clinical endpoints*:** PTD < 34 wk GA; PTD < 37 wk GA; PTD within 2 wk **PTD prevalence*:** 22.6% (PTD < 34 wk GA); 37.1% (PTD < 37 wk GA); 9.7% (PTD within 2 wk) **Authors' conclusion**:** "In this study, IGFBP-1 screening did not predict preterm delivery and fFN screening provided the best predictive capacity."
Dansereau et al. (2008)[68-70] **Type:** cohort study, prospective consecutive enrollment **Objective:** to compare performance of $TLi_{IQ}^{®}$ System with that of Actim™ Partus test **Duration:** Oct 2004 to Aug 2007 **Setting:** tertiary care centre **Province (city):** British Columbia (Victoria)	**Sample*:** 361 symptomatic women (singletons and twins) **Testing week*:** 22-34 wk GA **Index test*:** Actim™ Partus test (commercially available kit) **Comparator test*:** $TLi_{IQ}^{®}$ System (commercially available kit) **Reference standard*:** PTD **Clinical endpoints*:** PTD < 35 wk GA; PTD < 37 wk GA; PTD within 7d; PTD within 14 d **PTD prevalence*:** 10% (PTD < 35 wk GA); 23% (PTD < 37 wk GA); 2.8% (PTD within 7 d); 3.9% (PTD within 14 d) **Authors' conclusion**:** "For all outcomes, AP was more (or equally) sensitive than FFN. However, AP was also less specific than FFN, resulting in worse PPV, with only a marginal improvement in NPV. Both tests individually proved similarly inadequate to 'rule out' PTD. A combination or sequential approach may prove the best strategy. Semi-quantitative results would be preferable to a dichotomous one. Other tests should be assessed and combined with FFN and AP."[70]

AP = Actim™ Partus test; d = day(s); fFN/FFN = fetal fibronectin; GA = gestational age; IGFBP-1 = insulin-like growth factor binding protein-1; NR = not reported; NPV = negative predictive value; *ph*IGFBP-1 = phosphorylated insulin-like growth factor binding protein-1; PPV = positive predictive values; PTD = spontaneous preterm delivery; wk = week(s)

* Information for the cohort of women who had both index and comparator tests

** Conclusions stated by the author(s) of the selected studies and quoted directly from the published report

Participants in the selected studies[35;67-70] were recruited prospectively as single clinical cohorts that included a total of 711 symptomatic women who attended major teaching hospitals and referral centres and were tested with both rapid response tests before the reference standard (see Table 2.2 and Appendix 2.C, Table 2.C.1). Although the majority of participants were women with singletons, each study also included symptomatic women across other clinical risk spectrums. Cooper et al.[35] included 20 women with multiple pregnancies (6%) in their study; however, it is unclear how many twin or higher order pregnancies were in the cohort of 288 women who had both index and comparator tests. Audibert et al.[67] included seven women with twins (11%) and Dansereau et al.[68-70] included 19 women with twins (5%) in their cohorts of women who had both tests. The prevalence of spontaneous PTD varied in the selected studies (see Table 2.2).

One study[67] reported the mean maternal age (27.6 years) for the cohort of 62 women who had both index and comparator tests (see Appendix 2.C, Table 2.C.1). Audibert et al.[67] also reported that 47% of these women were nulliparous. The other two studies[35,68-70] did not report on these characteristics for the cohort of women who had both tests. None of the selected studies[35;67-70] reported on the number of women who experienced previous PTD in the cohort that had both tests. The selected studies[35;67-70] used similar methods for estimating gestational age at testing.

Cervicovaginal specimens were collected for both index and comparator tests at recruitment time,[35;67-70] or within 24 hours of admission if a digital examination had been performed within 24 hours before the participant's inclusion in the study.[67] Test specimens were analyzed and test results

were interpreted at various times in each study. Cooper et al.[35] used an fFN protocol according to which the test specimen was to be held for 1 hour after it was collected, and then sent to the laboratory for analysis or discarded in cases of unequivocal progression to spontaneous PTD. No protocol for the index or comparator test was used in the other two studies.[67-70]

In all studies[35;67-70] fFN was measured and the fFN results were read by laboratory personnel who were blinded to clinical information and outcomes in two studies.[67-70] The *ph*IGFBP-1 test specimens were tested and the results were interpreted by the study research nurse (blinded to fFN results) in one study,[35] by trained clinical staff (obstetrician, resident, or research nurse, blinded to fFN results) in another study,[67] and by laboratory technician (not blinded to fFN results) in the third study.[68-70] In all studies,[35;67-70] the results from the TLi_{IQ}® System were reported to the clinical team involved in the care of the symptomatic women, while the Actim™ Partus test results were not disclosed to the clinical or nursing staff in charge.

Methodological Quality of Selected Studies

The methodological quality of the included test accuracy studies was appraised using a modified version of the QUADAS-2 tool, as described in Appendix 2.D. The risk of bias results for each study are presented in Table 2.D.2 (Appendix 2.D), and Table 2.3 (below) summarizes the judgments for risk of bias within the four domains for the selected studies.

Table 2.3: Overall risk of bias results

Domain	Judgments for risk of bias		
	Low	High	Unclear
Patient selection	1 study[68-70]	1 study[67]	1 study[35]
Index test	1 study[67]	1 study[68-70]	1 study[35]
Reference standard	3 studies[35;67;68]		
Patient flow and timing		3 studies[35;67-70]	

Some of the methodological issues that have been shown to potentially bias test accuracy results, such as the absence of test descriptions, the use of different test thresholds, the use of inappropriate reference standards, or the use of different reference tests,[71-77] were generally not applicable to the selected studies. Although not all studies provided a complete description of the index and comparator tests and their performance, they all used commercially available kits that have pre-specified thresholds and did not change their characteristics. Also, it appears that, in all studies, the index and comparator tests were performed according to the manufacturer instructions. None of the studies were potentially affected by partial verification bias, because all participants included in the final analyses received confirmation of the diagnosis by the reference standard (spontaneous PTD). In addition, no study was potentially affected by differential verification bias, since the index and comparator test results were verified by the same reference standard. The reference standard was independent of both the index and comparator tests, thus avoiding the incorporation of bias.

However, the interpretation of test accuracy data from the selected studies was limited by other potential threats to validity (see Appendix 2.C, Table 2.C.1 and Appendix 2.D, Table 2.D.2). Particularly, blinding and consecutive enrolment were either not part of the study design or were not reported. Comparator review bias might have occurred in the study conducted by Dansereau and colleagues,[68-70] in which the same laboratory technician interpreted both tests and, hence, the

evaluator was not blinded to the results of the comparator test when interpreting the results of the index test. The same type of bias might have occurred in the study conducted by Cooper et al.,[35] in which it was not reported if the results of the reference standard were known when interpreting results of the index test. The extent to which comparator review bias may affect test results is related to the degree of subjectiveness in the interpretation of the test result.[74] In the study conducted by Cooper et al.,[35] it is not clear if the interpreter of the index test was blinded or not to the available clinical data (such as patient and pregnancy history, symptoms, and obstetric risk profile). This study might have been affected by problems related to clinical review bias, where the availability of clinical data during the interpretation of the index test results might have affected estimates of test performance. Knowledge of such data can influence the result if the test involves an interpretative component.[74]

The lack of consecutive enrolment might have resulted in a differing clinical spectrum of women being included by the selected studies, leading to a spectrum bias that potentially may influence test accuracy results.[8;31;74-77] Spectrum bias refers to the variation in test performance across subgroups of symptomatic women with different risks of imminent PTD (nulliparous and multiparous, with and without previous PTD, with singletons and with twins or higher order multiple pregnancy, etc.). It is typically thought to occur when a study does not adequately represent all subgroups. As an example, none of the selected studies purposely recruited only symptomatic women with singletons, and the inclusion of women with multiple pregnancies and the lack of stratified reporting of performance measures for the clinical endpoints of interest can cause uncertainties about the reported predictive ability of each test. Similarly, none of the selected studies stratified reporting for nulliparous and multiparous women and/or for women with and without previous PTD.

Treatment paradox may bias the estimates of test performance if treatment is started based on the knowledge of the results of the index or comparator test and the reference standard is applied after treatment has started.[8;31;74;77] In all selected studies, there is a time interval between testing the participants with both index and comparator tests and the occurrence of spontaneous PTD (the reference standard). During this time, the clinicians providing care for the participants were not blinded to the results of the $TLi_{IQ}^{®}$ System, and this might have led to changes in prenatal care that could affect the outcome of pregnancy, which in turn could influence the final accuracy estimates. This is the case of women with fFN positive results who might have received effective treatment(s) leading to prevention of PTD, which could have made the $TLi_{IQ}^{®}$ System appear inaccurate.

Absence of primary data in areas such as description of participants and study protocol limited the ability to extract and explore the data as completely as would have been desired. Attempts to minimize this problem by writing to the corresponding authors of the selected studies for the required information met with variable results. At the time this report was completed, additional information was received only for two of the selected studies.[67-70]

All studies[35;67-70] included in their final analyses less participants than they initially recruited, and did not provide details about those women who were excluded or lost to follow-up. Exclusion of these women from analysis could have biased the results.

The following commentary summarizes the findings reported by the selected studies.[35;67-70] The test performance measures reported by each study are summarized in Table 2.C.1 (Appendix 2.C). For two of the selected studies,[67-70] some of the information presented in Table 2.C.1 (Appendix 2.C) and in the commentary below was obtained from their lead authors.

Results

In 2012, Cooper et al.[35] published the results from a prospective cohort study that compared the performance of the Actim™ Partus test to that of the $TLi_{IQ}^{®}$ System in predicting PTD for women with symptoms of PTL (symptoms of uterine activity judged by the assessing physician to be indicative of PTL) at 24 to 34 weeks of gestation who attended labour and delivery units (L&D) in two tertiary care centres in Calgary. Women were excluded if they had ruptured membranes, antepartum hemorrhage, active labour, and suspected chorioamnionitis. Women who could not have an fFN test because they had a digital exam or sexual intercourse within the past 24 hours were eligible to join the study and had an *ph*IGFBP-1 test, as described by the manufacturer.

The study initially recruited 366 consenting women, and then excluded 15 because their swabs were collected outside the eligible gestational age and two because they were mistakenly entered twice in the study.[35] Results for the Actim™ Partus test were available for all 349 women included in the final analysis, while results for the $TLi_{IQ}^{®}$ System were available only for 288 women. The fFN swabs for the remaining 61 women (15.7%) were not tested either because of ineligibility for swab, or because the swab was discarded following clarification of clinical status. The primary outcome was PTD before 37 weeks of gestation.

The protocol for fFN testing was to hold the swab for one hour and, in cases of unequivocal progression to labour, the swab was discarded.[35] $TLi_{IQ}^{®}$ System results were read by laboratory personnel and were reported on the woman's chart. The *ph*IGFBP-1 swabs were tested and the results were read by the study research nurse (not clear if trained for interpreting the test) blinded to fFN results (not clear if blinded to clinical data and/or to clinical outcome).[35] The results of the Actim™ Partus test were unknown to the clinical or nursing staff in charge.

For the comparison between the Actim™ Partus test and the $TLi_{IQ}^{®}$ System, data were analyzed only for the 288 women (with singletons and multiple pregnancies; not clear how many were twins) who had both tests done.[35] In this subgroup, 46 women (16%) delivered before 37 weeks of gestation. When the performance measures for predicting PTD before 37 weeks were compared, both tests had comparably high NPV and low positive predictive value (PPV) estimates (see Table 2.4).

Table 2.4: Summary of results reported by Cooper et al.

Clinical endpoint	Sn (%)	Sp (%)	PPV (%)	NPV (%)	LR+	LR-
PTD before 37 wk GA	AP: 39 TLi: 33	AP: 74 TLi: 95	AP: 22 TLi: 54	AP: 86 TLi: 88	AP: 1.50 TLi: 6.07	AP: 0.82 TLi : 0.71

AP = Actim™ Partus test; GA = gestational age; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; Sn = sensitivity; Sp = specificity; NPV = negative predictive value; PPV = positive predictive value; PTD = spontaneous preterm delivery; TLi = $TLi_{IQ}^{®}$ System; wk = weeks

Source: [35]

However, the greatest problem with positive and negative predictive values is that they are dependent on the prevalence of the target condition in the population in which the measurements are performed. Therefore, they have limited clinical applicability outside of the study population. More useful for describing the diagnostic performance of a test is the likelihood ratio (LR), which combines sensitivity and specificity estimates into a single value that indicates by how much the test result will reduce the uncertainty of a given diagnosis. LR values are considered clinically meaningful measures of a test's diagnostic accuracy because they are less likely to change with the prevalence of

the target condition, and can be used to calculate the post-test probability for the target condition.[3;8;31;45;52;58;77-79] In general, the higher the positive likelihood ratio (LR+) value is above 1, the more accurate the test is in ruling in the target condition; the lower the negative likelihood ratio (LR-) value is below 1, the more accurate the test is in ruling out the target condition.

According to Honest and colleagues,[8;31] in the setting of pregnant women who present for care with symptoms of PTL, diagnostic tests with LR+ values from 2 to 5 "may be useful" (likely to generate small but sometimes important changes from the pre- to post-test probabilities of having PTD), those with LR+ values from 5 to 10 are "useful" (likely to generate moderate changes) and those with LR+ values above 10 are "very useful" (likely to generate large and conclusive changes). Similarly, tests with LR- values from 0.2 to 0.5 "may be useful" (likely to generate small but sometimes important changes), those with LR- values from 0.1 to 0.2 are "useful" (likely to generate moderate changes), and those with LR- values lower than 0.1 are "very useful" (likely to generate large and conclusive changes).[8;31] Tests with LR+ value from 1 to 2 and those with LR- value from 0.5 to 1 are considered "not useful" (may alter pre- to post-test probabilities to a small, and rarely important, degree).[8;31]

The results reported by Cooper et al.[35] suggest that, of the two tests, only the TLi_{IQ}® System had a clinically useful value for LR+ (see Table 2.4). However, the LR- value reported for the TLi_{IQ}® System was not useful.

The investigators also evaluated the agreement between *ph*IGFBP-1 and fFN test results within subjects, and found "that for individual patients the *ph*IGFBP-1 swab was more likely to give a positive result than fFN swab: 28.1% versus 9.7%, P<0.001."[35]

Based on their analyses, Cooper et al.[35] concluded that "NPV did not differ between *ph*IGFBP-1 and fFN for delivery < 37 weeks. Neither test improves on pretest probability of delivery < 37 weeks, so clinicians must decide whether the use of either test is justified." According to Cooper et al., "[f]urther research is clearly needed to identify a test that would be more effective than either *ph*IGFBP-1 or fFN in correctly identifying women who will not deliver preterm or within a defined period. Until the time when a better test is available, institutions and clinicians must decide whether the use of either test is clinically justified in women with symptoms of preterm labor."[35]

In 2010, Audibert et al.[67] published the results obtained from a smaller prospective cohort study that compared the performance of the Actim™ Partus test and that of the TLi_{IQ}® System for predicting PTD in women admitted with a clinical diagnosis of PTL between 24 and 34 weeks of gestation at the Centre Hospitalier Universitaire Sainte-Justine in Montreal. They recruited 71 women between January 2006 and January 2007, and tested them for the presence of fFN and *ph*IGFBP-1 in the cervicovaginal specimens collected before cervical length (CL) measurement by transvaginal ultrasound. PTL was defined by the presence of regular uterine contractions, lasting at least 30 seconds and occurring at least four times per 30 minutes, and significant cervical changes on digital examination. Women were excluded if they had confirmed or suspected rupture of membranes, cervical dilatation > 3 cm, cervical cerclage, vaginal bleeding, placenta previa, placental abruption, severe intrauterine growth restriction, preeclampsia, or medically indicated preterm delivery before 34 weeks.

The investigations were carried out either on admission, or within 24 hours of admission if a digital examination had been performed in the 24 hours before the woman's inclusion in the study.[67] The TLi_{IQ}® System was performed in the laboratory by a technician, who was not aware of the clinical situation. The Actim™ Partus test was performed and read by an obstetrician, resident, or research

nurse (trained for interpreting the test) blinded to fFN results and to clinical outcome. The clinical team members were not blinded to the results of fFN testing and CL measurement, and these results were available in the medical record. However, the results of *ph*IGFBP-1 testing were not disclosed to the clinical team and were not reported in the medical record. The standard local management protocol for PTL was applied, according to which administration of corticosteroids and tocolytics and bed rest were prescribed by the attending physician, depending on clinical evaluation and on the results of investigations, including CL measurement and fFN testing. The clinical endpoints were PTD within 2 weeks of admission to the study, before 34 weeks of gestation, and before 37 weeks of gestation.

Only 62 women (55 with singletons and seven with twins) were included in the final analysis because the outcome of pregnancy could not be determined for five women (who had been discharged and delivered in another centre), and four women were excluded from the final analysis (two had fFN testing but no CL measurement, and two had CL measured but no fFN testing).[67] Six women (9.6%) delivered within 2 weeks from testing, 14 women (22.5%) delivered before 34 weeks of gestation, and 23 women (37.1%) delivered before 37 weeks of gestation. According to the reported results, although both tests were less sensitive for predicting PTD before 34 weeks and 37 weeks of gestation than for predicting PTD within 2 weeks from testing, the sensitivity of the Actim™ Partus test was extremely low (below 20%) at all clinical endpoints (see Table 2.5 and Appendix 2.C, Table 2.C.1).

Table 2.5: Summary of results reported by Audibert et al.

Clinical endpoints	Sn (%)	Sp (%)	PPV (%)	NPV (%)	LR+	LR-
PTD before 34 wk GA	AP: 14 TLi: 50	AP: 94 TLi: 85	AP: 40 TLi: 50	AP: 79 TLi: 85	AP: 2.30 TLi: 3.40	AP: 0.90 TLi: 0.60
PTD before 37 wk GA	AP: 13 TLi: 48	AP: 95 TLi: 92	AP: 60 TLi: 79	AP: 65 TLi: 75	AP: 2.50 TLi: 6.20	AP: 0.90 TLi: 0.50
PTD within 14 days (2 wk)	AP: 17 TLi: 83	AP: 93 TLi: 84	AP: 20 TLi: 36	AP: 91 TLi: 98	AP: 2.30 TLi: 5.20	AP: 0.90 TLi: 0.20

AP = Actim™ Partus test; FN = false negative; FP = false positive; GA = gestational age; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; Sn = sensitivity; Sp = specificity; NPV = negative predictive value; PPV = positive predictive value; PTD = spontaneous preterm delivery; TLi = TLi_{IQ}® System; TN = true negative; TP = true positive; wk = weeks

Source: [67]

Comparable high NPV and specificity estimates were reported for both tests at all clinical endpoints.[67] However, at all endpoints, the TLi_{IQ}® System had higher LR+ values and lower LR- values than the Actim™ Partus test. The TLi_{IQ}® System was the best test for predicting PTD within 2 weeks (sensitivity of 83%, specificity of 84%, LR+ of 5.2, and LR- of 0.2).

Based on their analyses and results, Audibert et al. concluded that "[i]n this study, IGFBP-1 screening did not predict preterm delivery and fFN screening provided the best predictive capacity."[67]

In 2008, Dansereau and colleagues[68-70] reported the results from a prospective cohort study that compared the performance of the Actim™ Partus test with that of the TLi_{IQ}® System for predicting PTD in women with symptoms of PTL (symptoms of uterine activity and cervical changes judged by the assessing physician or midwife to be indicative of PTL) who presented for care at the Victoria

General Hospital in Victoria, British Columbia, between October 2004 and August 2007. The study included symptomatic women between 22 and 34 weeks of gestation, with intact membranes, cervical dilatation of < 3 cm, and no presence of blood on speculum. Women with cerclage, women who had vaginal examination or sexual intercourse within past 24 hours, or those who had vaginal probe ultrasound examination within past 24 hours were excluded from the study.

Dansereau and colleagues[68-70] initially recruited 407 women, and then excluded 41 because one of the two tests was not done or was sampled inadequately and deemed inappropriate for processing. Five women were lost to follow-up and were also excluded from the final analysis (none of these women delivered during their admission to the hospital). Results for both tests were available for 361 symptomatic women (342 with singletons and 19 with twins) who were included in the final analysis. The clinical endpoints were PTD within 7 days from testing, within 14 days from testing, before 35 weeks of gestation, and before 37 weeks of gestation.

The concentration levels of fFN and *ph*IGFBP-1 were concurrently measured in cervicovaginal specimens collected during the physical examination of the included women.[68-70] The attending physician or midwife collected the samples for both tests, which were sent to the laboratory and analyzed and interpreted by the laboratory technician. The evaluators of the two tests were blinded to the clinical information and outcomes of the pregnancy. However, the same laboratory technician interpreted both tests and, hence, was not blinded to the result of the other test. The attending physician was aware of the fFN results; however he/she was kept blinded to the *ph*IGFBP-1 results. Dansereau and colleagues[68-70] conducted a chart review to determine the four clinical end-points of interest for all included women.

Of all 361 women included in the final analysis, 10 (2.77%), 14 (3.88%), 36 (9.97%), and 83 (22.99%) women delivered within 7 days from testing, within 14 days from testing, before 35 weeks of gestation, and before 37 weeks of gestation, respectively.[68-70] The results reported at all clinical end-points showed that the Actim™ Partus test was as or more sensitive than the $TLi_{IQ}^{®}$ System, while the $TLi_{IQ}^{®}$ System was more specific than the Actim™ Partus test (see Table 2.6 and Appendix 2.C, Table 2.C.1). Overall, the agreement of *ph*IGFBP-1 and fFN test results for the 361 women was 79% (that is, 79% of the time, both tests gave the same result), largely due to the discrepancy between the two tests in regards to the lesser specificity of the *ph*IGFBP-1 test.

Table 2.6: Summary of results reported by Dansereau et al.

Clinical endpoints	Sn (%)	Sp (%)	PPV (%)	NPV (%)	LR+	LR-
PTD before 35 wk GA	AP: 61 TLi: 39	AP: 79 TLi: 95	AP: 24 TLi: 48	AP: 95 TLi: 93	AP: 2.80 TLi: 8.50	AP: 0.50 TLi: 0.60
PTD before 37 wk GA	AP: 40 TLi: 23	AP: 79 TLi: 97	AP: 37 TLi: 69	AP: 81 TLi: 80	AP: 1.90 TLi: 7.00	AP: 0.50 TLi: 0.80
PTD within 7 days	AP: 60 TLi: 60	AP: 76 TLI: 93	AP: 7 TLi: 21	AP: 99 TLi: 99	AP: 2.40 TLi: 9.10	AP: 0.50 TLi: 0.40
PTD within 14 days	AP: 71 TLi: 64	AP: 76 TLi: 94	AP: 11 TLi: 31	AP: 99 TLi: 99	AP: 3.00 TLi: 11.1	AP: 0.40 TLi: 0.40

AP = Actim™ Partus test; FN = false negative; FP = false positive; GA = gestational age; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; Sn = sensitivity; Sp = specificity; NPV = negative predictive value; PPV = positive predictive value; PTD = spontaneous preterm delivery; TLi = TLi$_{IQ}$® System; TN = true negative; TP = true positive; wk = weeks
Source: [68-70]

Although both tests had weak PPV estimates, those reported for the TLi$_{IQ}$® System were higher at all clinical endpoints (see Table 2.6). Equally high NPV estimates were reported at all clinical endpoints for both tests in the cohort of 361 that included both singletons and twins. The TLi$_{IQ}$® System had higher LR+ values at all endpoints and lower LR- values than the Actim™ Partus test at most endpoints in this cohort.

Dansereau and colleagues[68-70] also compared the performance of the Actim™ Partus test to that of the TLi$_{IQ}$® System in a small subgroup that included 17 symptomatic women with twins (the other two women with twins delivered electively at 32 weeks, and they were excluded from this subgroup analysis). Nine of the 17 symptomatic women delivered before 35 weeks of gestation (PTD prevalence of 53%). The specificity and NPV estimates for PTD before 35 weeks of gestation were strongly diminished for both tests (see Table 2.7). Neither test had useful LR+ and LR- values (likely to generate moderate changes in pre- to post-test probabilities of having PTD).

Table 2.7: Summary of results reported by Desjardins et al. for twin pregnancies

Clinical endpoint	Sn (%)	Sp (%)	PPV (%)	NPV (%)	LR+	LR-
PTD before 35 wk GA	AP: 78 TLi: 44	AP: 63 TLi: 75	AP: 70 TLi: 67	AP: 71 TLi: 55	AP: 2.10 TLi: 1.75	AP: 0.34 TLi: 0.75

AP = Actim™ Partus test; FN = false negative; FP = false positive; GA – gestational age; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; Sn = sensitivity; Sp = specificity; NPV = negative predictive value; PPV = positive predictive value; PTD = spontaneous preterm delivery; TLi = TLi$_{IQ}$® System; TN = true negative; TP = true positive; wk = weeks
Source: [68-70]

Discussion

The clinical diagnosis of PTL is complicated by its uncertain and multi-factorial etiology. A tool that can aid in timely identification of "true" PTL and prediction of whether a symptomatic woman presenting for care with intact membranes is at risk for imminent PTD would be valuable in enabling the choice of the most appropriate effective interventions. Such an adjunct test would also help identify those symptomatic women who are not in "true" PTL and are unlikely to benefit from the available effective interventions, and could therefore avoid the associated side effects and complications. Three rapid response tests are currently available in Canada for this indication: the Actim™ Partus test and the TLi$_{IQ}$® and 10Q Systems.

All three tests are relatively safe, simple to perform, can be run at bedside (as point-of-care tests) in urban and rural settings, and require minimal personnel training (Table 2.1). Any of the tests has the potential to reduce utilization of unnecessary interventions by helping clinicians to identify with increased certainty the symptomatic women who are not in "true" PTL. However, a positive result from the qualitative tests (the Actim™ Partus test or the TLi$_{IQ}$® System) is not predictive of spontaneous PTD. In contrast, quantitative fFN testing results quickly produced by the 10Q System (within 10 minutes) may be better at discriminating the risk of imminent PTD, and may ultimately lead to greater confidence in interpreting positive results. When compared to the TLi$_{IQ}$® System, the

attracting features of the Actim™ Partus test include availability of results in less time (5 minutes versus 25 minutes), lower cost per test ($25 versus $100), the possibility to use it after recent sexual intercourse or in the presence of urine, and independence from reader device (automated analyzer). However, because the Actim™ Partus test is a visual read test, the interpretation of its results is subjective.

Based of the available evidence, no conclusions could be drawn on whether the 10Q System has clear advantages over the TLi_{IQ}® System in terms of diagnostic performance and clinical and economic impact when added to PTL management in symptomatic women with intact membranes.

The value of the Actim™ Partus test as an alternative to the TLi_{IQ}® System to complement clinical examination for diagnosing PTL suspected in symptomatic women with intact membranes remains unclear. The studies selected for this evidence review[35;67-70] compared the performance of the two tests for predicting the risk of PTD in 711 symptomatic women (with singletons and multiple gestations) presenting for care (between 22 and 34 weeks) with intact membranes at tertiary centres in three large Canadian urban cities. However, their findings did not clarify whether the Actim™ Partus test has clear advantages over the TLi_{IQ}® System in terms of diagnostic performance for this indication. The studies vary in terms of sample size (from 62 to 361 participants), the definition of PTL, the prevalence of PTD at specific cut offs of gestational age, and blinding of test interpreters, making it difficult to comment on differences and similarities between studies. The interpretation of their findings was limited by potential threats to validity identified in the assessment of study quality, and by poor reporting. All studies demonstrated low risk of bias in the reference standard domain, and high risk of bias in the patient flow and timing domain. Only one study[68-70] demonstrated low risk of bias in the patient selection domain, and only one study[67] demonstrated low risk of bias in the index test domain.

The selected studies[35;67-70] gave insights regarding the comparative diagnostic performance of the evaluated tests when used to predict imminent PTD at different gestational time points and/or within a certain period of time from testing. However, none of these studies compared the clinical and/or economic utility of adding the tests to PTL management in terms of improved patient outcomes and reduced resource usage and associated costs.

Diagnostic Performance

Despite differences in design, the selected studies[35;67-70] are consistent in reporting sensitivity values lower than specificity values and NPVs higher than PPVs for both the Actim™ Partus test and TLi_{IQ}® System at all clinical endpoints (see Tables 2.4-7). Specificity and NPV estimates are high for both tests and do not differ greatly between the two tests, while sensitivity and PPV estimates are poor for both tests at all clinical endpoints. In general, both tests were highly specific and predicted the majority of women who would not go on to deliver preterm, although their negative results did not exclude all imminent PTD cases (see Table 2.8). A greater number of false positive results were reported for the Actim™ Partus test than for the TLi_{IQ}® System at most clinical endpoints (see Table 2.8).

Table 2.8: Comparison of performance between the Actim™ Partus test and the TLi_{IQ}® System

Study	TP	TN	FP	FN	LR+	LR-
PTD before 34 weeks of gestation						

Audibert (2010)[67] *PTD prevalence: 22.6%*	AP: 2 TLi: 7	AP: 45 TLi: 41	AP: 3 TLi: 7	AP: 12 TLi: 7	AP: 2.3 TLi: 3.4	AP: 0.9 TLi: 0.6
PTD before 35 weeks of gestation						
Dansereau (2008)[68-70] *PTD prevalence: 10%*	AP: 22 TLi:14	AP: 255 TLi: 310	AP: 70 TLi: 15	AP: 14 TLi: 22	AP: 2.8 TLi: 8.5	AP: 0.5 TLi: 0.6
PTD before 37 weeks of gestation						
Cooper (2012)[35] *PTD prevalence: 16%*	AP: 18 TLi: 15	AP: 179 TLi: 229	AP: 63 TLi: 13	AP: 28 TLi: 31	AP: 1.5 TLi: 6.1	AP: 0.8 TLi: 0.7
Audibert (2010)[67] *Prevalence: 37.1%*	AP: 3 TLi: 11	AP: 37 TLi: 36	AP: 2 TLi: 3	AP: 20 TLi: 12	AP: 2.5 TLi: 6.2	AP: 0.9 TLi: 0.5
Dansereau (2008)[68-70] *PTD prevalence: 23%*	AP: 33 TLi: 19	AP: 220 TLi: 270	AP: 58 TLi: 8	AP: 50 TLi: 64	AP: 1.9 TLi: 7.0	AP: 0.8 TLi: 0.5
PTD within 7 days from testing						
Dansereau (2008)[68-70] *PTD prevalence: 2.8%*	AP: 6 TLi: 6	AP: 265 TLi: 328	AP: 86 TLi: 23	AP: 4 TLi: 4	AP: 2.4 TLi: 9.1	AP: 0.5 TLi: 0.4
PTD within 14 days (2 weeks) from testing						
Audibert (2010)[67] *PTD prevalence: 9.7%*	AP: 1 TLi: 5	AP: 52 TLi: 47	AP: 4 TLi: 9	AP: 5 TLi: 1	AP: 2.3 TLi: 5.2	AP: 0.9 TLi: 0.2
Dansereau (2008)[68-70] *PTD prevalence: 3.9%*	AP: 10 TLi: 9	AP: 265 TLi: 327	AP: 82 TLI: 20	AP: 4 TLi: 5	AP: 3.0 TLi: 11.1	AP: 0.4 TLi: 0.4

AP = Actim™ Partus test; FN = false negative; FP = false positive; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; PTD = spontaneous preterm delivery; TLi = TLi_{IQ}® System; TN = true negative; TP = true positive

Because predictive values are dependent on the prevalence of PTD (which varied in the selected studies[35;67-70]), the best single measure for these tests' diagnostic performance is the likelihood ratio (LR). The LR values are less likely to change with the prevalence of PTD, and can be used to calculate the post-test probability for the target condition. In the setting of pregnant women presenting for care with symptoms of PTL and intact membranes, a LR value indicates how much a given test result will increase or decrease the pre-test probability of having PTD at various gestational ages, or within a defined period from testing. The further LR values are from 1, the stronger the evidence for the presence or absence of "true" PTL.

According to the LR+ and LR- values summarized in Table 2.8, of the two tests, the TLi_{IQ}® System appears to be more clinically useful because it has LR+ values greater than 5.0 at most clinical endpoints of interest, and even greater than 10.0 for one clinical endpoint. This means that the use of the TLi_{IQ}® System results is likely to generate moderate to large conclusive changes in the pre-test probability of having PTD before 35 weeks, before 37 weeks, and within 7 to 14 days from testing.

However, the TLi_{IQ}® System tended to have a greater usefulness for LR+ than for LR- value. According to Honest and colleagues,[8;31] given this trade off, the balance between LR+ and LR- values for a good adjunct test depends largely on the outcomes of PTL and the costs associated with the available effective intervention(s) (including potential mortality and morbidity). Due to the consequences of false-positive results, it is important that the LR+ value is suitably high, because providing unnecessary interventions leads to unjustifiable morbidity, inconvenience, and expense.[8;31] Given the consequences of false-negative results, it is important that the LR- value be suitably low,

because withholding necessary effective interventions leads to excessive morbidity and expense as a consequence of spontaneous PTD.[8;31]

Considering both LR+ and LR- values, the TLi$_{IQ}$® System appears to be more clinically useful than the Actim™ Partus test for predicting PTD within 14 days from testing (see Table 2.8).

Traditionally, PTD has been perceived as a problem before 37 weeks of gestation. However, because morbidity and mortality associated with prematurity are reduced significantly after 34 weeks, this gestational age has become a clinically and economically relevant threshold.[4;5;7-12;18;19;31;34;80] The performance of the Actim™ Partus test versus that of the TLi$_{IQ}$® System for predicting PTD before 34 weeks of gestation was compared only in one of the selected studies (see Table 2.8). The Audibert et al. study[67] reported very poor sensitivity for the Actim™ Partus test compared with the TLi$_{IQ}$® System. Although the PPV and NPV estimates were lower for the Actim™ Partus test compared with the TLi$_{IQ}$® System, they did not differ greatly, and specificity estimates were similar for the two tests. More false positive results and fewer false negative results were reported for the TLi$_{IQ}$® System than for the Actim™ Partus test. None of the two tests had useful LR+ and LR- values. The interpretation of these results is limited by the small sample size and the exclusion from final analysis of nine participants after recruitment because of delivery in other hospitals. There is a need for adequately powered prospective studies to investigate which of these tests is more accurate in predicting PTD before 34 weeks of gestation.

There is also a need for adequately powered prospective studies to investigate whether the Actim™ Partus test or the TLi$_{IQ}$® System is more accurate in predicting imminent risk of spontaneous PTD for symptomatic women with twin and higher order multiple pregnancies, which carry a substantial risk of PTD.[13;39] Although all selected studies included twins and higher order multiple pregnancies, only one[68-70] reported separate results for the prediction of PTD before 35 weeks of gestation in a small subgroup of 17 symptomatic women with twin pregnancies (see Table 2.9). Although the reported specificity was higher for the TLi$_{IQ}$® System, the Actim™ Partus test had higher sensitivity and NPV estimates in this subgroup. Reported LR+ and LR- values suggest that neither test is a clinically useful test for this indication.

Table 2.9: Performance of Actim™ Partus test and TLi$_{IQ}$® System for twin pregnancies

Study	TP	TN	FP	FN	LR+	LR-
PTD before 34 weeks of gestation						
Dansereau (2008)[68-70] *PTD prevalence: 53%*	AP: 7 TLi: 4	AP: 5 TLi: 6	AP: 3 TLi: 2	AP: 2 TLi: 5	AP: 2.10 TLi: 1.75	AP: 0.34 TLi: 0.75

AP = Actim™ Partus test; FN = false negative; FP = false positive; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; PTD = spontaneous preterm delivery; TLi = TLi$_{IQ}$® System; TN = true negative; TP = true positive

Clinical and Economic Utility

According to the results reported by the selected studies,[35;67-70] the clinical utility of both the Actim™ Partus test and the TLi$_{IQ}$® System for PTL lies in being able to identify symptomatic women who will not go on to deliver preterm. Both are bedside tests that require minimal training of health care professionals. The Actim™ Partus test has been introduced in Canada as a cheaper alternative to the TLi$_{IQ}$® System, without its limitations. According to Cooper et al., in the Calgary health zone "where at present ~720 fFN tests are processed by the laboratory service each year, the use of *ph*IGFBP-1 bedside testing 870 women (substituting for fFN, plus use in additional women who had a recent

vaginal exam or sexual intercourse as recommended by the manufacturer), would save approximately $70,000 per year in testing costs."[35] However, the characteristics of either test do not guarantee that its adoption into a clinical setting will achieve the hypothesized clinical and economic benefits.

The high NPV of the $TLi_{IQ}^{®}$ System was confirmed in recently published RCTs which evaluated the use of the $TLi_{IQ}^{®}$ System as an adjunct diagnostic tool in the management of PTL in symptomatic women with intact membranes.[28] However, according to a recently conducted Cochrane systematic review and meta-analysis of RCT data, there is insufficient evidence to support or refute its use for the management of PTL in symptomatic women.[36] Although PTD before 37 weeks of gestation was significantly decreased with PTL management based on knowledge of fFN results versus management without such knowledge, all other outcomes for which data were available (including PTD before 34 weeks of gestation) were similar in the two groups.[36] The reviewers found insufficient data for women with multiple gestations to make meaningful comparisons.[36]

The question as to whether adding the Actim™ Partus test or the 10Q System to the management of PTL for symptomatic women would change clinical practice and affect patient outcome, resource usage, and associated costs also remains unanswered. The use of either of these tests has not been evaluated in RCTs and/or in non-randomized controlled studies to determine whether clinicians can use the additional information provided by the test results to improve clinical practice and patient outcomes, and to reduce resource usage and costs associated with PTL management.

None of the available research studies directly compared the impact of using the Actim™ Partus test or the 10Q System to that of using the $TLi_{IQ}^{®}$ System in terms of improved patient outcomes and reduced resource usage and associated costs when added to PTL management in symptomatic women.[1;35;67-70]

Both the Actim™ Partus test and the $TLi_{IQ}^{®}$ System can be of particular utility in rural and remote areas to prevent the unnecessary and distressing transportation of pregnant women away from their homes. They both seem to provide useful information when there is uncertainty about whether to transport a symptomatic woman for PTL from a Level 1 or Level 2 health care centre to a Level 2 or Level 3 health care centre. However, the value of using either of these tests to complement clinical assessment of symptomatic women presenting for care in rural settings and remote areas is yet to be determined. All studies directly comparing the performance of the Actim™ Partus test to that of the $TLi_{IQ}^{®}$ System for predicting PTD in symptomatic women were conducted in Level 3 health care centers.[35;67-70] Therefore, it is yet to be determined which of the two tests is likely to have a greater impact on the decision regarding maternal transfers from rural and remote areas.

Health Canada has approved the Actim™ Partus test, the 10Q System, and the $TLi_{IQ}^{®}$ System to aid in diagnosing PTL and predicting PTD in symptomatic women. Canadian clinical practice guidelines recommend only the use of rapid fFN testing to complement clinical assessment for diagnosing PTL in symptomatic women. Patient test protocols for using the $TLi_{IQ}^{®}$ System have been developed by several reproductive care and perinatal programs in Canada. No guidelines or patient test protocols specifically developed for the use of the 10Q System and its results or for the use of the Actim™ Partus test and its results to aid in diagnosing PTL in symptomatic women are currently available in Canada.

Strengths and Limitations

The strengths of this evidence review pertain to the comprehensiveness of the literature searches, the criterion-based selection of relevant evidence, the rigorous appraisal of study validity, and the evidence-based inferences.

The present review also has several limitations. The literature review was confined to published reports of controlled or comparative studies that were written in English.

Only full text articles were included for data extraction and analysis, because abstracts provide insufficient details to allow an accurate, unbiased assessment and comparison of the study results. However, for the purpose of this review, the information contained in abstract publications reporting Canadian research results was summarized to inform the above "Available Research Evidence" section.

The selected studies were assessed using a quality tool, with the expectation that this would aid in identifying the studies that should be given more weight in the overall synthesis. However, the findings of the selected studies were not directly comparable, as their authors took different approaches and reported on different clinical endpoints, and none of them met all the criteria used to judge their methodological quality. Although the original aim of quality assessment became redundant because of these factors, it still had value in highlighting the study design and execution flaws.

Planned meta-analyses of results reported for the prediction of PTD before 34, 35, or 37 weeks of gestation, and prediction of PTD within 7 and 14 days (2 weeks) from testing could not be conducted because the number of studies per meta-analysis was small (≤ two studies).

The present review only summarizes the recommendations from reports of relevant clinical practice guidelines, position statements, and consensus documents, and does not appraise their scientific foundations.

Qualitative research literature, which provides information about the benefits, limitations, and utility (from physicians' and women's perspective) of using the tests of interest to aid in diagnosing PTL that may develop in imminent PTD in symptomatic women was not included.

The extent of publication bias was not assessed.

Conclusions

Due to the lack of evidence, no conclusions could be drawn on whether the 10Q System is superior to the $TLi_{IQ}^{®}$ System in terms of diagnostic performance and clinical and economic impact when added to the management of spontaneous preterm labour in symptomatic women presenting for care with intact membranes.

The value of using the Actim™ Partus test as an alternative to the $TLi_{IQ}^{®}$ System to complement clinical examination for diagnosing spontaneous preterm labour suspected in this population remains unclear. This evidence review confirms the findings from the earlier rapid review that both tests are relatively safe and simple to perform and have the potential to reduce unnecessary treatment and health care utilization by identifying women who are not in "true" spontaneous preterm labour. When compared to the $TLi_{IQ}^{®}$ System, the attractive features of the Actim™ Partus test include the possibility to use it after recent sexual intercourse, the availability of results in less time, lower cost per test, and independence from a reader device. However, according to the

findings reported by the three recently completed Canadian comparative studies that were included in this evidence review, the overall accuracy of the Actim™ Partus test in predicting spontaneous preterm delivery in symptomatic women appears to be lower in comparison to the $TLi_{IQ}^{®}$ System.

The findings reported by the three recently completed Canadian comparative studies that were included in this evidence review did not clarify the roles of the Actim™ Partus test and the $TLi_{IQ}^{®}$ System as adjunct tools to clinical examination for diagnosing spontaneous preterm labour suspected in symptomatic women presenting for care with intact membranes. Their findings helped in gaining more understanding regarding the comparative diagnostic performance of the two tests; however, they did not help to inform the decision on which test under which circumstances should be adopted for this indication. Clinicians and institutions deciding to use the Actim™ Partus test or the $TLi_{IQ}^{®}$ System for this indication should be aware of the following:

- Any modifications to the assay protocol as described by the manufacturer of each test may yield erroneous results.
- Both tests may produce false positive and false negative results; however, the performance of the Actim™ Partus test is associated with more false positive results.
- Although the negative predictive values do not differ greatly between the two tests, the $TLi_{IQ}^{®}$ System has more useful positive likelihood ratio values for most clinical endpoints of interest and appears to be more accurate in predicting spontaneous preterm delivery.
- Considering the estimated positive and negative likelihood ratio values, the $TLi_{IQ}^{®}$ System appears to be more clinically useful than the Actim™ Partus test in predicting preterm delivery within 14 days from testing.
- There is insufficient evidence to determine whether one test is more accurate than the other in predicting imminent risk of spontaneous delivery before 34 weeks of gestation.
- There is insufficient evidence to determine which test is most accurate in ruling in or out spontaneous preterm labour in symptomatic women at high risk, including those with twin or higher order multiple pregnancies.
- Neither of these tests has been shown to be superior to the other as a definitive adjunct tool that is likely to have a direct impact on the decisions regarding maternal transfers from rural and remote areas.
- The clinical and economic implications of using the Actim™ Partus test as an alternative to the $TLi_{IQ}^{®}$ System to aid in diagnosing spontaneous preterm labour in symptomatic women have yet to be determined.

Depending on the setting, resources, number of tests conducted, and available expertise, one test may be more appropriate than the other. Well-designed research studies conducted in other settings besides large urban hospitals to measure their diagnostic and predictive performance, as well as resource utilization related to better outcomes for mothers and newborns, may address these issues.

Appendix 2.A: Methodology

Literature Search

A comprehensive literature search was conducted by an Information Specialist from the Institute of Health Economics (IHE) between 20 April and 20 June 2013. An update search was conducted by the same Information Specialist on 11 February 2014. Major electronic databases used include: The Cochrane Library, CRD Databases: (NHS EED, HTA, DARE), PubMed, EMBASE, CINAHL and Web of Science. In addition, relevant library collections, websites of practice guidelines, regulatory agencies, evidence-based resources, and other HTA-related agency resources (AETMIS, CADTH, ICES) were searched. Internet search engines were also used to locate grey literature.

The search was developed and carried out prior to the study selection process, and was limited to English language publications and human studies published from January 2008 onwards. The search was further limited to systematic reviews, controlled or comparative studies, health technology assessments, economic evaluations, and clinical practice guidelines (CPGs).

Medical Subject Headings (MeSH) terms relevant to this topic include: Premature birth; Predictive Value of Tests; Insulin-Like Growth Factor Binding Proteins; Fibronectins.

In addition to the search strategy outlined in Table 2.A.1, the bibliographies and reference lists of all retrieved articles were examined and internet searches were conducted to retrieve grey literature. Grey literature searches were conducted to identify literature from non-indexed sources, health technology assessment reports, guidelines, government documents, and regulatory status information (that is, National Guidelines Clearinghouse, Health Canada, and Google).

Table 2.A.1: Search strategy

Database	Edition or date searched	Search terms[††]
Core databases		
MEDLINE (includes in process and other non-indexed citation) OVID Licensed Resource	20 April 2013 Update: 11 February 2014 Results: 21	1 obstetric labor, premature/ or premature birth/ 2 prenatal care/ 3 ((premature or preterm) adj2 (labor or labour or deliver$ or birth$)).mp. 4 1 or 2 or 3 5 exp "Sensitivity and Specificity"/ 6 (sensitivity or specificity).tw. 7 exp Diagnosis/ 8 Prenatal Diagnosis/ 9 exp "Predictive Value of Tests"/ 10 Point of Care Systems/ 11 exp Risk Factors/ 12 exp Risk assessment/ 13 (diagnos* or manag* or assess* or risk* or predict* or test* or detect* or screen* or point of care or bedside or rapid).tw. 14 exp Insulin-Like Growth Factor Binding Proteins/ 15 (actim partus or igfbp or phigfbp or somatomedin-binding protein$).mp. 16 Fibronectins/

		17 Fetal Proteins/ 18 (fetal fibronectin or foetal fibronectin or ffn or tliq).mp. 19 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 20 14 or 15 or 16 or 17 or 18 21 4 and 19 and 20 22 limit 21 to yr="2008 -Current" **(128 results)**
Embase	20 April 2013 Update: 11 February 2014 Results: 89	1 prematurity/ or "immature and premature labor"/ or premature labor/ 2 prenatal care/ 3 ((premature or preterm) adj2 (labor or labour or deliver$ or birth$)).mp. 4 1 or 2 or 3 5 "sensitivity and specificity"/ or sensitivity analysis/ 6 (sensitivity or specificity).tw. 7 exp diagnosis/ 8 prenatal screening/ 9 predictive value/ 10 exp "point of care testing"/ 11 risk factor 12 risk assessment/ 13 (diagnos* or manag* or assess* or risk* or predict* or test* or detect* or screen* or point of care or bedside or rapid).tw. 14 somatomedin binding protein/ 15 (actim partus or igfbp or phigfbp or somatomedin-binding protein* or growth factor binding protein*).tw. 16 fibronectin/ 17 fetoprotein/ 18 (fetal fibronectin or foetal fibronectin or ffn or tliq).mp. 19 5 or 6 or 7 or 9 or 10 or 11 or 12 or 13 20 14 or 15 or 16 or 17 or 18 21 4 and 19 and 20 22 limit 21 to yr="2008 -Current" **(283 results)**
Cochrane	22 April 2013 Update: 11 February 2014 Results: 8	1 MeSH descriptor: [Obstetric Labor, Premature] this term only 2 MeSH descriptor: [Premature Birth] this term only 3 MeSH descriptor: [Prenatal Care] this term only 4 (premature or preterm) next/2 (labor or labour or deliver* or birth*):ti,ab,kw (Word variations have been searched) 5 #1 or #2 or #3 or #4 6 MeSH descriptor: [Sensitivity and Specificity] explode all trees 7 (sensitivity or specificity):ti,ab,kw (Word variations have been searched) 8 MeSH descriptor: [Diagnosis] explode all trees 9 MeSH descriptor: [Prenatal Diagnosis] explode all trees 10 MeSH descriptor: [Predictive Value of Tests] explode all trees 11 MeSH descriptor: [Point-of-Care Systems] this term only 12 MeSH descriptor: [Risk Factors] this term only 13 MeSH descriptor: [Risk Assessment] explode all trees 14 (diagnos* or manag* or assess* or risk* or predict* or test* or detect* or

		screen* or point of care or bedside or rapid):ti,ab,kw (Word variations have been searched) 15 MeSH descriptor: [Insulin-Like Growth Factor Binding Proteins] explode all trees 16 (actim partus or igfbp or phigfbp or somatomedin-binding protein*):ti,ab,kw (Word variations have been searched) 17 MeSH descriptor: [Fibronectins] this term only 18 MeSH descriptor: [Fetal Proteins] this term only 19 (fetal fibronectin or foetal fibronectin or ffn or tliq):ti,ab,kw (Word variations have been searched) 20 #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 21 #15 or #16 or #17 or #18 or #19 22 #5 and #20 and #21 from 2008 to 2013 **(19 results)**
Web of Science	22 April 2013 Update: 11 February 2014 Results: 40	Topic=(((premature or preterm) NEAR/2 (labor or labour or deliver* or birth*))) AND Topic=(sensitivity or specificity or diagnos* or manag* or assess* or risk* or predict* or test* or detect* or screen* or point of care or bedside or rapid) AND Topic=(actim partus or igfbp or phigfbp or somatomedin-binding protein* or growth factor binding protein* or fetal fibronectin or foetal fibronectin or ffn or tliq) Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=2008-01-01 - 2013-04-23 **(187 results)**
Grey literature		
Dissertations and theses		
Proquest Dissertations and Theses	20 June 2013	all((premature or preterm) NEAR/2 (labor or labour or deliver* or birth*)) AND all((sensitivity or specificity or diagnos* or manag* or assess* or risk* or predict* or test* or detect* or screen* or point of care or bedside or rapid)) AND all((actim partus or igfbp or phigfbp or somatomedin-binding protein* or growth factor binding protein* or fetal fibronectin or foetal fibronectin or ffn or tliq)) **(1 result)**
Clinical practice guidelines		
AMA Clinical Practice Guidelines *http://www.topalbertadoctors.org*	29 May 2013	Browsed list of guidelines **(0 results)**
CMA Infobase *http://mdm.ca/cpgsnew/cpgs/index.asp*	29 May 2013	Premature; Preterm **(0 results)**
National Guideline Clearinghouse *http://www.ngc.gov*	29 May 2013	"preterm labor" or "preterm birth" or "preterm delivery" or "premature birth" or "premature labor" or "premature delivery" **(1 result)**
NICE *http://www.nice.org.uk/*	20 June 2013	Insulin like growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq **(0 results)**

Clinical trials		
ClinicalTrials.gov (US) http://clinicaltrials.gov	29 May 2013	growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq or "preterm labor" or "preterm birth" or "preterm delivery" or "premature birth" or "premature labor" or "premature delivery" **(10 results)**
metaRegister of Controlled Trials (mRCT) http://www.controlled-trials.com/mrct/	17 June 2013	growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq **(4 results)**
International Clinical Trials Registry Platform http://apps.who.int/trialsearch/	17 June 2013	growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq **(4 results)**
HTA resources		
INESS http://www.inesss.qc.ca/index.php?id=49&L=1&code=RECHERCHE	20 June 2013	growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq or "preterm labor" or "preterm birth" or "preterm delivery" or "premature birth" or "premature labor" or "premature delivery" **(0 results)**
CADTH http://www.cadth.ca/	20 June 2013	growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq or "preterm labor" or "preterm birth" or "preterm delivery" or "premature birth" or "premature labor" or "premature delivery" **(3 results)**
ICES www.ices.on.ca	20 June 2013	Browsed list of publications **(0 results)**
Health Technology Assessment Unit At McGill http://www.mcgill.ca/tau/	20 June 2013	Browsed list of publications **(0 results)**
Medical Advisory Secretariat http://www.hqontario.ca/evidence/publications-and-ohtac-recommendations	20 June 2013	Browsed list of publications **(0 results)**
Search engines		
NHS Evidence https://www.evidence.nhs.uk/	20 June 2013	Insulin like growth factor binding protein* or fetal fibronectin or ffn or actim partus or igfbp or tliq **(7 results)**
Google www.google.com	20 June 2013	preterm labor or preterm birth or preterm delivery or premature birth or premature labor or premature delivery fetal fibronectin OR growth factor binding protein OR ffn OR igfbp OR tliq filetype:pdf **(15 results)**

Other sites		
Society of Gynecologists and Obstetricians of Canada *www.sogc.org*	20 June 2013	Browsed publications **(0 results)**
American Congress of Obstetricians and Gynecologists *http://www.acog.org/Resources_And_Publications*	20 June 2013	Browsed publications **(0 results)**

Note: †† "*", "# ", and "?" are truncation characters that retrieve all possible suffix variations of the root word, e.g. surg* retrieves surgery, surgical, surgeon, etc.

The Canadian distributor of the Actim™ Partus test (Alere Canada) and the manufacturer of the $TLi_{IQ}^{®}$ and 10Q Systems (Hologic Inc., United States) were contacted for information on these tests' characteristics and their effectiveness when used to aid in diagnosing PTL, as well as their regulatory status, availability, and cost in Canada. They were also contacted for information regarding ongoing or completed research studies directly comparing the value of adding these tests in diagnosing PTL suspected in symptomatic women presenting for care with intact membranes.

Study selection

Titles and abstracts of the citations retrieved by the electronic search conducted for this evidence review were screened independently by two reviewers (PC and BG) using the predefined inclusion/exclusion criteria that are listed below. Full-text publications of potentially relevant articles were retrieved, and their eligibility was determined by the same two reviewers according to the same predefined inclusion/exclusion criteria. Disagreements were resolved through discussion and consensus.

Inclusion criteria

Research studies were included if they met the following criteria:

- **Population:** pregnant women (all ages, all ethnic groups, with multiple or single gestations) presenting for care with symptoms and signs of PTL and intact membranes at inpatient or outpatient settings (rural or urban)
- **Index test:** use of the Actim™ Partus test or use of the 10Q System as an adjunct rapid response diagnostic test for PTL that develops in imminent PTD for the target population
- **Comparator:** use of the $TLi_{IQ}^{®}$ System as an adjunct rapid response diagnostic test for PTL that develops in imminent PTD for the target population
- **Reference standard:** a diagnosis based on subsequent course of pregnancy (spontaneous PTD before 37 weeks of gestation or within a defined period from sampling/testing) for the target population
- **Outcomes:** diagnostic accuracy (sensitivity, specificity, positive and negative predictive values, and/or likelihood ratios for positive and negative results); clinical outcomes (patient

and resource usage outcomes in terms of impact on gestation age at delivery, maternal anxiety/stress, and need for woman's removal from her home support; rates of spontaneous PTB/PTD, maternal transfers and hospital admissions; and impact on assessment time, length of hospital stay, use of other diagnostic tests, and use of therapeutic interventions); risks and complications to mother and/or fetus associated with performing the rapid diagnostic test itself; and costs associated with adding the test for PTL management

- **Study design:**
    - clinical studies that directly compared the use of the Actim™ Partus test or the use of the 10Q System to the use of the $TLi_{IQ}^{®}$ System to aid in diagnosing PTL and predicting PTD for symptomatic women presenting for care with intact membranes, and reported on these tests' efficacy (in terms of their diagnostic accuracy and patient and resource usage outcomes) and associated costs; and/or
    - systematic reviews (quantitative and/or qualitative), health technology assessment studies, and economic analyses reporting on the effectiveness, and/or costs and cost-effectiveness of using the Actim™ Partus test or the 10Q System versus the $TLi_{IQ}^{®}$ System to aid in diagnosing PTL and predicting PTD for symptomatic women presenting for care with intact membranes
        - *Note:* A review is considered to be systematic if it meets all of the following criteria:[81]
            - focused clinical question;
            - explicit search strategy;
            - use of explicit, reproducible, and uniformly applied criteria for article selection;
            - critical appraisal of the included studies; and
            - qualitative or quantitative data synthesis.
- **Time frame:** published from January 2008 onwards

Research studies were included if the published report was publicly available, and only full, peer-reviewed articles were included because abstracts do not provide adequate detail on the review methodology. However, for the purpose of this evidence review, information contained in abstract publications of relevant comparative studies conducted in Canada was summarized to inform the "Available Research Evidence" section (Section Two), even if these were available only in poster or oral abstract presentation format. The authors of the abstract-only publications were contacted by e-mail for details of their studies. In the case of multiple publications, the most recent and complete version was included.

Only articles reporting on research studies conducted in countries with developed market economies were considered, since the health status and disease burden of individuals, cultural and legal norms, and access to health care in countries with another status are likely to be too different from those of Canada to be clinically relevant. Countries deemed to have developed market economies, as defined by the United Nations, include Australia, Canada, Japan, New Zealand, the United States of America, and European countries (except for countries with market economies in transition) (http://unpan1.un.org/intradoc/groups/public/documents/un/unpan008092.pdf).

To inform the "Guidelines and Patient Test Protocols" section (Section Two) and the other sections of the review, the following publicly available published reports were also considered for inclusion:

Relevant clinical practice guidelines (CPGs), position papers, and consensus statements issued on the diagnosis and/or management of PTL, if developed by national bodies in Canada and other countries with developed market economies;

- Patient test protocols developed specifically for using the $TLi_{IQ}^{®}$ System, the 10Q System, or the Actim™ Partus test to aid in diagnosing suspected PTL and predicting PTD in symptomatic women presenting for care with intact membranes; and
- clinical reviews, overview articles, commentaries and discussion papers, narrative and descriptive reviews, letters, conference material, commentaries, discussion papers, editorials, and abstracts presenting background information on the use of the rapid response tests of interest to aid in diagnosing PTL or predicting PTD for symptomatic women presenting for care with intact membranes.

Exclusion criteria

This review does not cover the use of the rapid response tests of interest, alone or in conjunction with other diagnostic tests, for other categories of pregnant women, such as symptomatic women with premature rupture of membranes, asymptomatic women, or for other indications (for example, for prediction of pre-eclampsia or postterm delivery, or for selection of the most suitable PTL induction methods).

Excluded from data extraction were published reports of research studies that:

- did not provide the necessary data to complete 2x2 tables to compute indices of test accuracy, or did not report the calculated performance measures for both the index and comparator tests;
- involved both symptomatic and asymptomatic women and did not report separately the results for the symptomatic women; or
- included women who experienced preterm rupture of membranes and/or medically indicated PTL and did not separately report on these subjects.

Published reports of narrative and descriptive reviews that summarized the research on the topic but lacked an explicit description of a systematic approach to the identification and interpretation of evidence were also excluded. They were considered only as a source of background information, where appropriate.

Editorials, letters, and technical reports were excluded.

Data extraction

Two reviewers (PC and BG) independently abstracted the following data from the published reports of the selected research studies using data extraction tabulated forms developed *a priori*. Study profile information and outcome data from the selected research studies were summarized in Appendix 2.C, Table 2.C.1.

- **Study:** author(s), year of publication, country (number of centres), objective(s), recruitment method, setting, duration, and source of funding

- **Study's and women's characteristics:** sample (number of women included, number of women who had both tests), inclusion and exclusion criteria, protocol used for *ph*IGFBP-1 or fFN test, method used to estimate gestational age at testing and at delivery, definition used for PTL, and women's characteristics at baseline (age, parity, previous PTD, pregnancy type – singletons, twins, higher order pregnancy)
- **Diagnostic interventions and outcomes:** the rapid response tests (index and comparator tests) used (device/system, time of swab collection, time of test performance, test interpreter, blinding), reference standard used, and information on primary/secondary outcomes
- **Reported results of interest:** diagnostic accuracy, clinical outcomes, and costs

For studies in which the reporting of the study methodology, outcomes, and/or results was unclear, one reviewer (PC) contacted their lead authors by email for further information.

Methodological quality assessment

Two reviewers (PC and BG) independently assessed methodological quality of included test accuracy studies using a modified version of the QUADAS-2 tool. The reviewers were not blinded to any aspects of the research studies being evaluated. Quality assessment results from the two reviewers were compared and disagreements were resolved by discussion and consensus. Appendix D presents the modified QUADAS-2 tool, the guidance developed and piloted by the two reviewers, and the quality assessment results. The quality assessment results are incorporated into the review by investigating whether a relationship exists between quality concerns and study findings. Quality assessment results were not used to include or exclude research studies.

No attempt was made to appraise the scientific foundations of the selected CPGs.

Data analysis and synthesis

A narrative approach was used to summarize the research findings from the included studies. The performance measures for individual studies have been summarized and displayed in tabulated format; however, these results were combined statistically to provide a "pooled" measure for each outcome of pregnancy. This approach was considered reasonable given the small number of included studies (n = 3) and the limited number of cases with PTD for the different cut-offs of gestational age, because the small absolute numbers of affected cases would have introduced imprecision.

External review

The draft report was reviewed by the members of the Post Policy Implementation Review of fFN Testing in Alberta Working Group assembled for this project.

Appendix 2.B: Excluded Research Studies

The application of the selection criteria for research studies described in Appendix 2.A resulted in 50 full text articles being excluded from data extraction and synthesis. Table 2.B.1 lists the excluded full text reports of the retrieved research studies and the main reasons for their exclusion.

Table 2.B.1: Excluded full text reports of retrieved research studies

Main reason for exclusion: *No comparison of interest* (n = 28)
Deshpande et al. Rapid fetal fibronectin testing to predict preterm birth in women with symptoms of premature labour: a systematic review and cost analysis. *Health Technology Assessment* (Winchester, England) 17(40), 1
Lee et al. Does the use of fetal fibronectin in an algorithm for preterm labor reduce triage evaluation times? *Journal of Maternal-Fetal & Neonatal Medicine* 2013;26(7):706-709
Kallioniemi et al. Early pregnancy vaginal fluid phosphorylated insulin-like growth factor binding protein-1 predicts preterm delivery. *Prenatal Diagnosis* 2013;33(4):378-383
Laudanski et al. Assessment of the selected biochemical markers in predicting preterm labour. *Journal of Maternal-Fetal & Neonatal Medicine* 2012;25(12):2696-2699
Dutta D, Norman E. Pilot study into the efficacy of foetal fibronectin testing in minimising hospital admissions in women presenting with symptoms of preterm labour: a randomised controlled trial of obstetric and neonatal outcomes. *Archives of Gynecology & Obstetrics* 2011;284(3):559-565
Bogavac et al. The role of insulin-like growth factor in prediction and prevention of preterm delivery. *Vojnosanitetski Pregled* 2010;67(11):883-886
Conde-Agudelo A, Romero R. Cervicovaginal fetal fibronectin for the prediction of spontaneous preterm birth in multiple pregnancies: a systematic review and meta-analysis. *Journal of Maternal-Fetal & Neonatal Medicine* 2010;23(12):1365-1376
Adeyemi O, Osoba L. The role of phosphorylated insulin-like growth factor binding protein-1 in predicting pre-term labour in twin pregnancies. *Journal of Obstetrics & Gynaecology* 2010;30(6):571-573
Honest et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technology Assessment* (Winchester, England) 2009;13(43):1-627
Wilms et al. Predicating imminent preterm labour based on a determination of foetal fibronectin in a vaginal smear. *Nederlands Tijdschrift voor Geneeskunde* 2009;153:B398
Sanchez-Ramos et al. Fetal fibronectin as a short-term predictor of preterm birth in symptomatic patients: a meta-analysis. *Obstetrics & Gynecology* 2009;114(3):631-640
Altinkaya et al. Cervical phosphorylated insulin-like growth factor binding protein-1 in prediction of preterm delivery. *Archives of Gynecology & Obstetrics* 2009;279(3):279-283
Pelaez et al. Negative fetal fibronectin: who is still treating for threatened preterm labor and does it help? *Journal of Perinatal Medicine* 2008;36(3):202-205
Balic et al. Insulin-like growth factor-binding protein-1 (IGFBP-1) in cervical secretions as a predictor of preterm delivery. *Journal of Maternal-Fetal & Neonatal Medicine* 2008;21(5):297-300
Roman et al. Vaginal fetal fibronectin as a predictor of spontaneous preterm delivery in triplet gestations. *Journal of Maternal-Fetal and Neonatal Medicine* 2012;10:1921-1923
Riboni et al. Biochemical markers predicting pre-term delivery in symptomatic patients: Phosphorylated insulin-like growth factor binding protein-1 and fetal fibronectin. *Archives of Gynecology and Obstetrics* 2011;6:1325-1329
Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta Obstetricia et Gynecologica Scandinavica* 2011;11:1189-1199
Raminez et al. Analysis of two strategies for the management of threatened preterm labor. *Progresos de Obstetricia y Ginecologia* 2012;7:261-266
Dutta et al. The efficacy of fetal fibronectin testing in minimising hospital admissions, length of hospital stay and cost savings in women presenting with symptoms of pre-term labour. *Journal of Obstetrics and Gynaecology* 2010;30(8):768-773

Wiwanitkit V. Diagnostic property of cervical phosphorylated insulin-like growth factor binding protein-1 in the prediction of preterm labor in symptomatic patients. *Archives of Gynecology and Obstetrics* 2010;1;175-176
Rahkonen et al. Cervical length measurement and cervical phosphorylated insulin-like growth factor binding protein-1 testing in prediction of preterm birth in patients reporting uterine contractions. *Acta Obstetricia et Gynecologica Scandinavica* 2009;8:901-908
Tsourapas et al. An economic evaluation of alternative test-intervention strategies to prevent spontaneous pre-term birth in singleton pregnancies. *Acta Obstetricia et Gynecologica Scandinavica* 2009;8:1319-1330
Wilms et al. Prediction of threatened premature birth by determination of fetal fibronectin in vaginal fluid. *Nederlands Tijdschrift voor Geneeskunde* 2009;153(31):1514-1520
Berghella et al. Fetal fibronectin testing for reducing the risk of preterm birth. *Cochrane Database of Systematic Reviews* 2008;4: CD006843
Fuchs I, Dudenhausen JW. Use of foetal fibronectin for prediction of a premature birth. *Gynakologische Praxis* 2008;32(4):646-648
Honest et al. Prediction of spontaneous preterm birth: no good test for predicting a spontaneous preterm birth. *Current Opinion in Obstetrics & Gynecology* 2012;24(6):422-433
Menon et al. Biomarkers of spontaneous preterm birth: An overview of the literature in the last four decades. *Reproductive Sciences* 2012;18(11):1046-1070
Tanir et al. Cervical phosphorylated insulin-like growth factor binding protein-1 for the prediction of preterm delivery in symptomatic cases with intact membranes. *Journal of Obstetrics and Gynaecology Research* 2009;35(1):66-72
Main reason for exclusion: *No 2x2 data for both index and comparator tests (n = 1)*
Abbott et al. Evaluation of a quantitative fetal fibronectin test for spontaneous preterm birth in symptomatic women. *American Journal of Obstetrics & Gynecology* 2013;208(2):122-126
Main reason for exclusion: *Protocol of a research study (n = 1)*
Rapid fetal fibronectin (fFN) testing to predict pre-term birth in women with symptoms of premature labour: A systematic review and cost-effectiveness analysis (Project record). Health Technology Assessment Database 2012. Available: *http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32012000531/frame.html*
Main reason for exclusion: *Abstracts (or letters) of research studies (n = 17)*
Bruijn et al. Does quantitative fetal fibronectin testing improve the prediction of spontaneous preterm delivery as compared to qualitative fetal fibronectin testing in symptomatic women: A post-hoc analysis. *American Journal of Obstetrics and Gynecology*. Conference: 34th Annual Meeting of the Society for Maternal-Fetal Medicine: The Pregnancy Meeting, February 3-8, 2014, New Orleans, United States, 2014:S364
Abbott et al. Quantifiable fetal fibronectin (FFN) using 10q to accurately predict preterm birth in twin pregnancies. *Twin Research and Human Genetics*. Conference: 14th International Congress on Twin Studies and The 2nd World Congress on Twin Pregnancy, Florence, Italy, 2012:15(2)
Siddiqa M, Haloob R. An audit of fetal fibronectin test. *International Journal of Gynecology and Obstetrics*. Conference: 20th FIGO World Congress of Gynecology and Obstetrics, Rome, Italy, 2012:S752
Basak S, Babbur V. Audit of in-utero transfers following introduction of rapid fetal fibronectin test for assessment of preterm labour in a district general hospital in United Kingdom. *International Journal of Gynecology and Obstetrics*. Conference: 20th FIGO World Congress of Gynecology and Obstetrics, Rome, Italy, 2012;S291
Thandayathany et al. Fetal fibronectin rapid test versus phosphorylated insulin-like growth factor-1 (phIGFBP-1) as bedside test kits for prediction of preterm delivery in the clinical setting. *BJOG: An International Journal of Obstetrics and Gynaecology*. Conference: 10th International Scientific Congress of the Royal College of Obstetricians and Gynaecologists, 2012
Myint et al. Introducing fetal fibronectin testing to a District General Hospital in women with symptoms of preterm labour. *Archives of Disease in Childhood: Fetal and Neonatal Edition*. Conference: Perinatal Medicine 2011, Harrogate, United Kingdom, 2011:Fa83
Northridge et al. Predicting pre-term birth using fetal fibronectin (FFN) in Ninewells Hospital, Dundee. *Archives of Disease in Childhood: Fetal and Neonatal Edition*. Conference: Perinatal Medicine 2011, Harrogate, United Kingdom, 2011:Fa69

Chandiramani et al. The affirm study: Assessment of fetal fibronectin testing to improve preterm management. *Archives of Disease in Childhood: Fetal and Neonatal Edition*. Conference: Perinatal Medicine 2011, Harrogate, United Kingdom 2011:FA65-FA66
Pelletier et al. *Actim Partus* Test as therapeutic guide in patients with preterm contractions. *Archives of Gynecology and Obstetrics.* Conference: 58th Congress of the German Society for Gynecology and Obstetrics (Deutsche Gesellschaft fur Gynakologie und Geburtshilfe, DGGG), Munich, Germany, 2010
Osorio et al. The impact of fetal fibronectin assay on preterm labor management: A randomized controlled trial. *Journal of Maternal-Fetal and Neonatal Medicine*. Conference: 22nd European Congress of Perinatal Medicine, 2010, Granada, Spain, 2010:304-305
Randhawa TS, Kwek K. Assessment of the predictive value of phosphorylated insulinlike growth factor binding protein-1 for preterm delivery. *Journal of Maternal-Fetal and Neonatal Medicine*. Conference: 22nd European Congress of Perinatal Medicine, 2010, Granada, Spain, 2010:294
Benson et al. Fetal fibronectin for evaluation of preterm labor in the setting of cervical cerclage. *American Journal of Obstetrics and Gynecology*. Conference: 2010 30th Annual Meeting of the Society for Maternal-Fetal Medicine (SMFM), Chicago, United States, 2010
Abbott et al. The first report of a quantitative fetal fibronectin bedside analyser in the prediction of preterm birth in symptomatic women. *BJOG-An International Journal of Obstetrics and Gynaecology* 2012;119(6):e1-e2
Roman et al. Fetal fibronectin as a predictor of spontaneous preterm birth in triplet gestations. *American Journal of Obstetrics and Gynecology* 2011;204:S186
Adhikari et al. Cervical length compared to Bishop's score for prediction of pre-term birth in women with pre-term labour. *Journal of Obstetrics and Gynaecology* 2011;31(3):213-216
Conde-Agudelo A, Romero R. Fetal fibronectin as a predictor of spontaneous preterm delivery in multiple gestations: a systematic review and meta-analysis. *American Journal of Obstetrics and Gynecology* 2009;201(6):S196
yala-Mendez et al. Fetal fibronectin status markedly modifies the risk of preterm delivery in low risk patients with symptomatic preterm labor and cervical shortening. *American Journal of Obstetrics and Gynecology* 2008;199(6):S235
Alere Canada. *Comparison of the clinical performance of Actim Partus to fetal fibronectin*. Pamphlet from Alere Canada, 2012. Available from: *www.alere.ca*
Main reason for exclusion: *Duplications (n = 2)*
Berghella et al. Fetal fibronectin testing for reducing the risk of preterm birth. *Cochrane Database of Systematic Reviews*, 2008(4)
Conde-Agudelo A, Romero R. Cervicovaginal fetal fibronectin for the prediction of spontaneous preterm birth in multiple pregnancies: a systematic review and meta-analysis (Provisional abstract). *Journal of Maternal-Fetal and Neonatal Medicine* 2010;23(12):1365-1376

Appendix 2.C: Results from Selected Research Studies

Table 2.C.1: Studies comparing performance of the Actim™ Partus test versus theTLi$_{IQ}$® System for predicting PTD in symptomatic women

<table>
<tr><th>Study</th><th>Study profile</th><th>Tests and outcomes</th><th>Diagnostic accuracy results</th></tr>
<tr>
<td>Cooper et al. 2012[35]

Country (# centres): Canada (2)

Objective(s): to examine performance (Sn, Sp, PPV, NPV) of phIGFBP-1 test in predicting PTD in symptomatic women and to compare characteristics of phIGFBP-1 and fFN tests

Recruitment: cohort; prospective enrollment; NR if random or consecutive enrollment

Setting: tertiary care centre (L&D, Foothills Medical Centre and Peter Lougheed Centre, Calgary, Alberta)

Duration: Oct. 2005 to May 2009

Funding: peer-reviewed funding received from Calgary Health Region Perinatal Funding Competition; $300 from Somagen Diagnostics for staff prize draws</td>
<td>Sample: 349 women (mean age of 29 y; 151 nulliparous and 198 multiparous; 327 singletons and 20 multiple pregnancies; 56 with pervious PTD and 293 without previous PTD); only 288 women had both phIGFBP-1 and fFN tests

Inclusion: women presenting for care with PTL symptoms at 24 to 34 wk GA

Exclusion: ruptured membrane, antepartum hemorrhage, active labour, and suspected chorioamnionitis

Protocol for phIGFB-1 or fFN test: used hospital fFN protocol; NR for phIGFBP-1 test

Determination of GA:
At testing: last menstrual period and/or first trimester US
At delivery: premature (< 37 wk) vs. not premature (≥ 37 wk) based on assessment of GA at testing (baseline)

PTL definition: symptoms of uterine activity judged by the assessing physician to be indicative of PTL

Women’s characteristics at baseline*:
Age*: NR
Parity*: NR
Previous PTD*: NR
Pregnancy*: NR</td>
<td>Index test: Actim™ Partus
Performance: swabs collected and prepared at recruitment, then frozen for later testing (NR) following manufacturer instructions
Time to result: 5 min
Interpreter: study research nurse (no information on training)
Blinding: test interpreter blinded to fFN results; NR if test interpreter blinded to clinical information and outcome;clinical team in charge blinded to test results

Comparator test: TLi$_{IQ}$® System
Performance: swabs collected at recruitment were held for 1 hour and then tested (according to hospital fFN protocol)
Time to result: NR
Interpreter: lab personnel
Blinding: NR if lab technician blinded to clinical information and outcome; fFN results reported in patient chart (available to clinical team)

Reference standard: outcome of pregnancy (PTD)

Outcomes: PTD before 37 wk GA (primary) and PTD within 7 and 14 d from testing (secondary)</td>
<td>Delivery before 37 wk GA (prevalence of 16%) (n = 288 women)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>39 (25-53)</td><td>33 (19-46)</td></tr>
<tr><td>Sp % (95% CI)</td><td>74 (68-80)</td><td>95 (92-97)</td></tr>
<tr><td>PPV % (95% CI)</td><td>22 (13-31)</td><td>0.54 (35-72)</td></tr>
<tr><td>NPV % (95% CI)</td><td>86 (82-91)</td><td>0.88 (84-92)</td></tr>
<tr><td>LR+ (95% CI)</td><td>1.50 (0.99-2.28)</td><td>6.07 (3.10-11.89)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.82 (0.65-1.05)</td><td>0.71 (0.58-0.87)</td></tr>
</table>
</td>
</tr>
</table>

* Information for the cohort of 288 women who had both index and comparator tests

<table>
<tr><th>Study</th><th>Study profile</th><th>Tests and outcomes</th><th>Diagnostic accuracy results</th></tr>
<tr>
<td>Audibert et al. 2010[67]

Country (# centres): Canada (1)

Objective(s): to validate the use of phIGFBP-1 as a predictor of PTD (secondary objective)

Recruitment: cohort; prospective, not consecutive enrollment**

Setting: tertiary care center (L&D, Centre Hospitalier Universitaire Sainte-Justine, Université de Montréal, Montréal, Quebec)

Duration: Jan 2006 to Jan 2007

Funding: Research grant ($20,000) from the Sainte-Justine, Research Center, Montreal, Quebec**</td>
<td>Sample: 62 women (all had both phIGFBP-1 and fFN tests)

Inclusion: women admitted with clinical diagnosis of PTL between 24 to 34 wk GA

Exclusion: confirmed or suspected rupture of membrane, cervical dilation > 3 cm, cervical cerclage, vaginal bleeding, placenta previa, placental abruption, severe intrauterine growth restriction, preeclampsia, or medially indicated PTD before 34 wk GA

Protocol for phIGFBP-1 or fFN test: no protocol was used

Determination of GA:
<u>At testing:</u> last menstrual period or first trimester US (if discrepancy of > 5 d before 14 wk and > 7 d before 20 wk)**
<u>At delivery:</u> same dating as for GA at testing**

PTL definition: the presence of regular uterine contraction, lasting ≥ 30 s and occurring ≥ 4 times/30 min, and significant cervical changes on digital examination

Women’s characteristics at baseline*:
<u>Age (mean±SD)*:</u> 27.6 ± 6.2 y
<u>Parity*:</u> 29 (47%) nulliparous women
<u>Previous PTD*:</u> NR
<u>Pregnancy*:</u> 55 (88.7%) singletons; 7 (11.3%) twins</td>
<td>Index test: Actim™ Partus

<u>Performance:</u> swab collected, prepared and tested either on admission or 24 h of admission (if digital examination done in 24 h before woman’s inclusion in study)
<u>Time to results:</u> 5 min
<u>Interpreter:</u> clinical staff (obstetrician, resident, or research nurse, trained for interpreting the test)**

<u>Blinding:</u> test interpreters blinded to fFN results and to clinical information and outcome**; clinical team in charge blinded to phIGFBP-1 test results

Comparator test: TLi_{IQ}® System
<u>Performance:</u> swab collected and sent to laboratory for analysis either on admission or 24 h of admission (if a digital examination performed 24 h before woman’s inclusion in study)
<u>Time to result:</u> NR
<u>Interpreter:</u> lab technician**

<u>Blinding:</u> lab technician blinded to clinical information and outcome**; clinical team in charge not blinded to fFN results (reported in medical chart)

Reference standard: outcome of pregnancy (PTD)

Outcomes: PTD before 34 and 37 wk GA and PTD withing 2 wk from testing</td>
<td>
Delivery before 34 wk GA (prevalence of 22.6%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>14 (2-43)</td><td>50 (23-77)</td></tr>
<tr><td>Sp % (95% CI)</td><td>94 (83-99)</td><td>85 (72-94)</td></tr>
<tr><td>PPV % (95% CI)</td><td>40 (5-85)</td><td>50 (23-77)</td></tr>
<tr><td>NPV % (95% CI)</td><td>79 (66-89)</td><td>85 (72-94)</td></tr>
<tr><td>LR+ (95% CI)</td><td>2.3 (0.4-12.4)</td><td>3.4 (1.4-8.1)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.9 (0.7-1.1)</td><td>0.6 (0.3-1.0)</td></tr>
</table>
Delivery before 37 wk GA (prevalence of 37.1%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>13 (5-21)</td><td>48 (35-60)</td></tr>
<tr><td>Sp % (95% CI)</td><td>95 (89-100)</td><td>92 (86-99)</td></tr>
<tr><td>PPV % (95% CI)</td><td>60 (48-72)</td><td>79 (68-89)</td></tr>
<tr><td>NPV % (95% CI)</td><td>65 (53-77)</td><td>75 (64-86)</td></tr>
<tr><td>LR+ (95% CI)</td><td>2.5 (0.5-14)</td><td>6.2 (1.9-20.0)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.9 (0.8-1.1)</td><td>0.5 (0.4-0.8)</td></tr>
</table>
Delivery within 2 wk of testing (prevalence of 9.7%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>17 (0-64)</td><td>83 (36-100)</td></tr>
<tr><td>Sp % (95% CI)</td><td>93 (83-98)</td><td>84 (72-92)</td></tr>
<tr><td>PPV % (95% CI)</td><td>20 (0.5-72)</td><td>36 (13-65)</td></tr>
<tr><td>NPV % (95% CI)</td><td>91 (81-97)</td><td>98 (89-100)</td></tr>
<tr><td>LR+ (95% CI)</td><td>2.3 (0.3-17.6)</td><td>5.2 (2.6-10.4)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.9 (0.6-1.3)</td><td>0.2 (0.0-1.2)</td></tr>
</table>
</td>
</tr>
</table>

* Information for the cohort of 62 women who had both index and comparator tests
** Information obtained from personal communication with the lead author of the study

<table>
<tr><th>Study</th><th>Study profile</th><th>Tests and outcomes</th><th>Reported test performance measures</th></tr>
<tr>
<td>Dansereau et al.[68-70]

Country (No. of centres): Canada (1)

Objective(s): to compare performance of TLi_{IQ}® System with that of Actim™ Partus test

Recruitment: cohort; prospective consecutive enrollment**

Setting: tertiary care center (L&D, Victoria General Hospital, Victoria, British Columbia)

Duration: Oct 2004 to Aug 2007

Funding: the funding was internally provided from the laboratory budget**</td>
<td>Sample: 361 women (all had both phIGFBP-1 and fFN tests)

Inclusion: women presenting with symptoms of PTL between 24 and 34 wk GA

Exclusion: ruptured membrane, cervical dilation ≥ 3 cm, cervical cerclage, cervical exam or sexual intercourse or use of lubricant gel within 24 hours, presence of blood

Protocol for phIGFBP-1 or fFN: no protocol was used**

Determination of GA: last menstrual period if consistent with US +/- 7 d (if discrepancy of ≥ 8 d existed, the US dating was used)**

PTL definition: symptoms of uterine activity and cervical changes judged by the attending physician or midwife to be indicative of PTL**

Women’s characteristics at baseline*:
Age*: NR
Parity*: NR
Previous PTD*: NR
Pregnancy*: 342 (94.7%) singletons; 19 (5.3%) twins</td>
<td>Index test: Actim™ Partus
Performance: swab collected, prepared and tested at recruitment
Time to result: 5 min
Interpreter: lab technician
Blinding: lab technician blinded to clinical information and outcome but not blinded to fFN results**; clinical team blinded to phIGFBP-1 test results

Comparator test: TLi_{IQ}® System
Performance: swab collected, prepared and tested at recruitment
Time to result: over 20 min
Interpreter: NR
Blinding: lab technician blinded to clinical information and outcome but not blinded to phIGFBP-1 results**; clinical team not blinded to fFN results

Reference standard: outcome of pregnancy (PTD)

Outcomes: PTD before 35 wk and 37 wk GA and PTD within 7 and 14 d from testing</td>
<td>
Delivery before 35 wk GA (prevalence of 10%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>61 (NR)</td><td>39 (NR)</td></tr>
<tr><td>Sp % (95% CI)</td><td>79 (NR)</td><td>95 (NR)</td></tr>
<tr><td>PPV % (95% CI)</td><td>24 (NR)</td><td>48 (NR)</td></tr>
<tr><td>NPV % (95% CI)</td><td>94.8 (NR)</td><td>93.4 (NR)</td></tr>
<tr><td>LR+ (95% CI)</td><td>2.8 (NR)</td><td>8.5 (NR)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.5 (NR)</td><td>0.6 (NR)</td></tr>
</table>
Delivery before 37 wk GA (prevalence of 23%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>40 (NR)</td><td>23 (NR)</td></tr>
<tr><td>Sp % (95% CI)</td><td>79 (NR)</td><td>97 (NR)</td></tr>
<tr><td>PPV % (95% CI)</td><td>37 (NR)</td><td>69 (NR)</td></tr>
<tr><td>NPV % (95% CI)</td><td>81 (NR)</td><td>80 (NR)</td></tr>
<tr><td>LR+ (95% CI)</td><td>1.9 (NR)</td><td>7.0 (NR)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.5 (NR)</td><td>0.8 (NR)</td></tr>
</table>
Delivery within 7 d (prevalence of 2.8%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>60 (NR)</td><td>60 (NR)</td></tr>
<tr><td>Sp % (95% CI)</td><td>76 (NR)</td><td>93 (NR)</td></tr>
<tr><td>PPV % (95% CI)</td><td>7 (NR)</td><td>21 (NR)</td></tr>
<tr><td>NPV % (95% CI)</td><td>98.5 (NR)</td><td>98.8 (NR)</td></tr>
<tr><td>LR+ (95% CI)</td><td>2.4 (NR)</td><td>9.1 (NR)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.5 (NR)</td><td>0.4 (NR)</td></tr>
</table>
Delivery within 14 d (prevalence of 3.9%)
<table>
<tr><th>Accuracy</th><th>phIGFBP-1</th><th>fFN</th></tr>
<tr><td>Sn % (95% CI)</td><td>71 (NR)</td><td>64 (NR)</td></tr>
<tr><td>Sp % (95% CI)</td><td>76 (NR)</td><td>94 (NR)</td></tr>
<tr><td>PPV % (95% CI)</td><td>11 (NR)</td><td>31 (NR)</td></tr>
<tr><td>NPV % (95% CI)</td><td>98.5 (NR)</td><td>98.5 (NR)</td></tr>
<tr><td>LR+ (95% CI)</td><td>3.0 (NR)</td><td>11.1 (NR)</td></tr>
<tr><td>LR- (95% CI)</td><td>0.4 (NR)</td><td>0.4 (NR)</td></tr>
</table>
</td>
</tr>
</table>

* Information for the cohort of 361 women who had both index and comparator tests
** Information obtained from personal communication with the lead author of the study

Appendix 2.D: Quality Assessment of the Key Studies

Overview of the quality assessment tools

The recently published QUADAS-2 quality assessment tool[71] was used to assess the methodological quality of included test accuracy studies. QUADAS-2 tool consists of four key domains: 1) patient selection; 2) index test; 3) reference standard; and 4) patient flow and timing. Each domain is assessed in terms of risk of bias, and the first three domains are also assessed in terms of applicability concerns.

In the development of QUADAS-2, the main decision was to separate "quality" into "risk of bias" and "concerns regarding applicability".[71;73] "Quality" was defined as both the "risk of bias" (the degree to which estimates of diagnostic accuracy avoided risk of bias) and the "applicability" (the extent to which primary studies are applicable to the review's research question) of a study. There are 11 signaling questions for assessing risk of bias (Table 2.D.1).

QUADAS-2 tool is applied in four phases:

1) Summarizing the review questions:
    - Target condition: PTL that may lead to imminent PTD
    - Target population: pregnant women (all ages, all ethnicities, with singletons or multiple gestations) presenting for care with symptoms and signs of PTL and intact membranes
    - Index test: Actim™ Partus test or 10Q System (currently on the Canadian market)
    - Comparator test: $TLi_{IQ}^{®}$ System (currently on the Canadian market)
    - Reference standard: subsequent course of pregnancy (spontaneous PTD before 37 weeks or within a defined period from testing)
    - Setting: hospital or community
    - Intended use of the index and comparator tests: as an adjunct test for diagnosing PTL and predicting PTD in the target population
2) Tailoring the tool to the review (by adding or omitting signaling questions), producing review-specific guidance, and piloting the tool by two independent raters.
3) Constructing a flow diagram for the primary study.
4) Assessing risk of bias and concerns regarding applicability.

Judgment on risk of bias: Risk of bias is judged as "high", "low", or "unclear".

High risk of bias: any signaling question for a domain is answered "No".

Low risk of bias: all signaling questions for a domain are answered "Yes".

Unclear: any signaling question for a domain is answered "Unclear" (used only when insufficient data are reported to permit a judgment).

Table 2.D.1: Risk of bias and applicability judgments in QUADAS-2

Domain	Patient selection	Index test	Reference standard	Flow and timing
Description	Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting).	Describe the index test and how it was conducted and interpreted.	Describe the reference standard and how it was conducted and interpreted.	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s)/comparator test(s) and reference standard.
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled?	Were the index test results interpreted without knowledge of the results of the reference standard?	Is the reference standard likely to correctly classify the target condition?	Was there an appropriate interval between index test(s) and reference standard?
	Was a case-control design avoided?	If a threshold was used, was it pre-specified?	Were the reference standard results interpreted without knowledge of the results of the index test?	Did all patients receive a reference standard?
	Did the study avoid inappropriate exclusions?			Did all patients receive the same reference standard?
				Were all patients included in the analysis?
Risk of bias (high/low/unclear)	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability (high/low/unclear)	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	

C/A – cannot answer; N/A – not applicable

Source: [71]

Adaptation of QUADAS-2 tool

One reviewer (PC) drafted a modified version of the QUADAS-2 tool and the associated guidance for answering the signaling questions for risk of bias, which were review-specific (tailored for this evidence review). The modified tool excluded one signaling question in the "Index test" domain ("If a threshold was used, was it pre-specified?"), because both the index and comparator tests are commercially available point-of-care tests (with package insert or brochure describing the test and how to perform it) that have a pre-specified threshold. Three new signaling questions were added to the tool, as it was thought that the QUADAS-2 tool did not adequately cover all the important aspects of this review which assessed comparative tests. One was added under the "Index test" domain ("Were the results of the index test interpreted without knowledge of the results of the comparator test?"), and the other two were added under the "Patient flow and timing" domain ("Is the time period between the index test and the comparator test short enough to be reasonably sure that the target condition did not change between the two tests?" and "Were the results of both the index test and the comparator test verified with the same reference standard?"). These signaling questions are all based on existing QUADAS items, and were suggested by the QUADAS Steering Group members to be added for reviews assessing comparative tests.[73]

Both the modified QUADAS-2 tool and the associated guidance for the signaling questions for risk of bias were discussed with another reviewer (BG). After discussions, one signaling question in the "Patient flow and timing" domain ("Was there an appropriate interval between index test(s) and reference standard?") was removed from the modified tool. This is because, for both index and comparator tests, the swabs are to be collected always when the woman presents for care, which is before any treatment starts and before delivery occurs. Therefore, the review-specific QUADAS-2 tool consists of 12 signaling questions.

The same two reviewers (PC and BG) piloted the modified QUADAS-2 tool using the associated draft guidance in one of the three selected test accuracy studies. Discrepancies between the reviewers and the issues associated with the use of the modified tool and the associated draft guidance were discussed. The modified tool and the rating guidance were refined and finalized based on discussions between the two reviewers.

Quality assessment results

The quality assessment results of each of the included accuracy studies are summarized in Table 2.D.2. A "Yes" response indicates that the specific quality item in the signaling question was met and the study was deemed to have been conducted in such a way as to minimize the bias associated with the particular domain. As recommended by the developers of the QUADAS-2,[71;73] rating results of QUADAS-2 were not used to generate a summary "quality score". Instead, the quality assessment results were incorporated into the systematic review through the investigation of the association of individual quality items with estimates of test accuracy.[82;83]

Table 2.D.2: Quality assessment of test accuracy studies comparing the Actim™ Partus test and the TLi$_{IQ}$® System

Domain	Questions for risk of bias	Cooper[35]	Audibert[67]	Dansereau[68-70]
Patient selection	Was a consecutive or random sample of patients enrolled?	?	X	√
	Was a case-control design avoided?	√	√	√
	Did the study avoid inappropriate exclusions?	√	√	√
	Risk of bias	unclear	high	low
Index test	Were the index test results interpreted without knowledge of the results of the reference standard?	?	√	√
	Were the results of the index test interpreted without knowledge of the results of the comparator test?	√	√	X
	Risk of bias	unclear	low	high
Reference standard	Is the reference standard likely to correctly classify the target condition?	√	√	√
	Was the reference standard independent of the comparator test (i.e., the comparator test did not form part of the reference standard)?	√	√	√
	Risk of bias	low	low	low
Patient flow and timing	Is the time period between the index test and the comparator test short enough to be reasonably sure that the target condition did not change between the two tests?	√	√	√
	Did all patients receive a reference standard?	X	X	X
	Did all patients receive the same reference standard?	√	√	√
	Were the results of both the index test and the comparator test verified with the same reference standard?	√	√	√
	Were all patients included in the analysis?	X	X	X
	Risk of bias	high	high	high

Yes = √; No = X; Unclear = ?

References

(1) Corabian C. *The Actim Partus versus the TLiIQ system as rapid response tests in aid in diagnosing preterm labour in symptomatic women*. Edmonton AB, Institute of Health Economics, 2008.

(2) Hologic I. Rapid fFN 10Q cassette kit. Sunnyvale CA: Hologic, Inc., 2009.

(3) Woodworth A. *Fetal fibronection. A marker for predicting preterm birth*. 2008. Available from: www.aacc.org/publications/cln/2008/october/Pages/series_1008.aspx (accessed 29 January 2013).

(4) Hamilton SA, Tower CL. Management of preterm labour. *Obstetrics, Gynaecology and Reproductive Medicine* 2010;20(8):235-240.

(5) PerkinElmer Life and Analytical Sciences. *Preterm birth - challenges and opportunities in prediction and prevention*. Waltham MA: Perkin Elmer, Inc., 2009; 1-80.

(6) HealthLink BC. Preterm labour. 2011. Available from: www.healthlinkbc.ca/kb/content/major/hw222237.html (accessed 25 March 2013).

(7) Ross MG, Smith CV. Preterm labour.2011. Available from: http://emedicine.medscape.com/article/260998-overview (accessed 18 March 2013).

(8) Honest H, Forbes CA, Duree KH, et al. Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling. *Health Technology Assessment* (Winchester, England) 2009;13(43):1-627.

(9) Sayres J. Preterm labor. *American Family Physician* 2010;81(4):477-484.

(10) Wax JR, Cartin A, Pinette MG. Biophysical and biochemical screening for the risk of preterm labor. *Clinics in Laboratory Medicine* 2010;30(3):693-707.

(11) Chandiramani M, Di Renzo GC, Gottschalk E, et al. Fetal fibronectin as a predictor of spontaneous preterm birth: A European perspective. *Journal of Maternal-Fetal & Neonatal Medicine* 2011;24(2):330-336.

(12) Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet* 2008;371(9606):75-84.

(13) Rahkonen L. *Preterm delivery and selected biomarkers - phosphorylated insulin-like growth factor-binding protein-1 and matrix metalloproteinase-8 - in cervical fluid*. Helsinki Finland: University of Helsinki, 2010.

(14) Schleubner E. The prevention, diagnosis and treatment of premature labor. *Deutsches Arzteblatt International* 2013;110(13):227-236.

(15) Department of Veterans Affairs, Department of Defense, The Pregnancy Management Working Group. *Va/DoD clinical practice guideline: Management of pregnancy*. Version 2.0. Washington DC: United States Department of Veterans Affairs, Department of Defense, 2009.

(16) Rigby FB, Isaacs C. Fetal fibronectin tests. 2012. Available from: http://emedicine.medscape.com/article/2035269-overview (accessed 13 March 2013).

(17) Nash P. Fetal fibronectin: what does it mean? *Neonatal Network - Journal of Neonatal Nursing* 2010;29(6):385-387.

(18) Canadian Institute for Health Information. *Too early, too small: a profile of small babies across Canada.* Ottawa ON: CIHI, 2009. Available from: https://wecure.cihi.ca/free_products/too_early_too_small_en.pdf (accessed 13 March 2013).

(19) Petraglia F, Visser GHA. Prevention and management of preterm labour. *Journal of Maternal-Fetal and Neonatal Medicine* 2009;22 (Suppl 2):24-30).

(20) Rahkonen L. Prediction of pre-term delivery with phosphorylated insulin-like growth factor-binding protein-1. *European Obstetrics & Gynaecology* 2011;6(2):95-99.

(21) The American College of Obstetricians and Gynecologists.Women's Health Care Physicians. Management of preterm labor. *Obstetrics & Gynecology* 2012;119(6):1308-1317.

(22) Rincon M, Pereira LMR. Ambulatory Management of Preterm Labor. *Clinical Obstetrics and Gynecology* 2012;55(3):756-764.

(23) Lawn JE, Gravett MG, Nunes TM, et al. Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. *Pregnancy and Childbirth* 2010;10(Suppl 1):51.

(24) Canadian Institute for Health Information. *Highlights of 2010-2011 selected indicators describing the birthing process in Canada.* Ottawa ON: Canadian Institute for Health Information, 2012. Available from: https://secure.cihi.ca/free_products/Childbirth_Highlights_2010-11_EN.pdf (accessed 13 March 2013).

(25) Vis J, Mol B-W, Wilms F, et al. Risk of preterm delivery in pregnant women who have been discharged after an episode of threatened preterm labor. Conference: 31st Annual Meeting of the Society for Maternal-Fetal Medicine: The Pregnancy Meeting, San Francisco, CA United States, February 7-12, 2011. *American Journal of Obstetrics and Gynecology* 2011:S192-S193.

(26) Vis JY, Wilms FF, Oudijk MA, et al. Why were the results of randomized trials on the clinical utility of fetal fibronectin negative? A systematic review of their study designs. American Journal of Perinatology 2011; 28(2):145-150.

(27) Royal College of Obstetricians and Gynaecologists. Introduction to preterm labour. 2012. Available from: www.rcog.org.uk/stratog/page/introduction-preterm-labour.

(28) Corabian P, Harstall C. *The role of rapid fetal fibronection assay in the management of spontaneous preterm labour.* Edmonton AB: Institute of Health Economics, 2008.

(29) Alere Canada. *Alere Actim Partus. Summary of clinical evidence.* Ottawa ON: Alere Canada, 2012.

(30) Norman JE, Shennan AH. Prevention of preterm birth-why can't we do any better? *Lancet* 2013;381(9862):184-185.

(31) Honest H. *Accuracy of tests for predicting spontaneous preterm birth: Systematic reviews of diagnostic research.* Birmingham UK: Unviersity of Birmingham, 2009.

(32) Reproductive Health Working Group. *Alberta reproductive health: Pregnancies & births. Surveillance report 2009.* Edmonton AB: Alberta Health and Wellness, 2009. Available from: www.health.alberta.ca/documents/Reproductive-Health-2009.pdf (accessed 18 March 2013).

(33) Alberta Reproductive Health Report Working Group. *Alberta reproductive health: Pregnancies and births. Table update 2011.* Edmonton AB: Alberta Health and Wellness, 2011. Available from: www.health.alberta.ca/documents/Reproductive-Health-2011.pdf (accessed 18 March 2013).

(34) Tsourapas A, Roberts TE, Barton PM, et al. An economic evaluation of alternative test-intervention strategies to prevent spontaneous pre-term birth in singleton pregnancies. *Acta Obstetricia et Gynecologica Scandinavica* 2009;88(12):1319-1330.

(35) Cooper S, Lange I, Wood S, et al. Diagnostic accuracy of rapid phIGFBP-I assay for predicting preterm labor in symptomatic patients. *Journal of Perinatology* 2012;32(6):460-465.

(36) Berghella V, Hayes E, Visintine J, et al. Fetal fibronectin testing for reducing the risk of preterm birth. *Cochrane Database of Systematic Reviews* (Online) 2008;(4):CD006843.

(37) Hologic Inc. *Fetal fibronectin enzyme immunoassay and fapid fFN for the TliIQ system*. Marlborough MA: Holigic, Inc., 2009.

(38) BMJ Evidence Centre. Premature labour. 2011. Available from: http://bestpractice.bmj.com/best-practice/monograph/1002/diagnosis-step-by-step.html (accessed 18 March 2013).

(39) Di Renzo GC, Roura LC, Facchinetti F, et al. Guidelines for the management of spontaneous preterm labor: Identification of spontaneous preterm labor, diagnosis of preterm premature rupture of membranes, and preventive tools for preterm birth. *Journal of Maternal-Fetal and Neonatal Medicine* 2011;24(5):659-667.

(40) The Royal Australian and New Zealand College of Obstetricians and Gynaecologists. *Use of cervical fetal fibronectin and phosphorylated insulin-like growth factor binding protein 1 as screening tests for preterm birth*. C-Obs 26. 2011. Melborne, Australia: The Royal Australian and New Zealand College of Obstetricians and Gynaecologists. Ranzcog College Statement.

(41) National Institute for Health and Clinical Excellence (NHS). *Multiple pregnancy. The management of twin and triplet pregnancies in the antenatal period*. NICE clinical guideline 129. Manchester UK: National Institute for Health and Clinical Excellence, 2011.

(42) Institute for Clinical Systems Improvement. *Health Care Guideline: Management of labour*. Fourth Edition. 2011.

(43) Sanchez-Ramos L, Delke I, Zamora J, et al. Fetal fibronectin as a short-term predictor of preterm birth in symptomatic patients: a meta-analysis. *Obstetrics & Gynecology* 2009;114(3):631-640.

(44) Menon R, Torloni MR, Voltolini C et al. Biomarkers of spontaneous preterm birth: An overview of the literature in the last four decades. *Reproductive Sciences* 2011;18(11):1046-1070.

(45) Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta Obstetricia et Gynecologica Scandinavica* 2011;90(11):1189-1199.

(46) Hologic Inc. Rapid fFN cassette kit. Sunnydale CA: Hologic, Inc., 2009.

(47) Hologic Inc. Rapid fFN test. Specimen collection kit. Sunnyvale CA: Hologic, Inc., 2009.

(48) Cytyc Corporation. *TLi System. User manual*. Marlborough MA: Cytyc Corporation, 2008.

(49) Hologic Inc. *Rapid fFN 10Q system. User manual*. Marlborough MA: Hologic, Inc., 2013.

(50) Abbott DS, Radford SK, Seed PT, et al. Evaluation of a quantitative fetal fibronectin test for spontaneous preterm birth in symptomatic women. *American Journal of Obstetrics & Gynecology* 2013;208(2):122-126.

(51) Conde-Agudelo A, Romero R. Cervicovaginal fetal fibronectin for the prediction of spontaneous preterm birth in multiple pregnancies: A systematic review and meta-analysis. *Journal of Maternal-Fetal & Neonatal Medicine* 2010;23(12):1365-1376.

(52) Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: No good test for predicting a spontaneous preterm birth. *Current Opinion in Obstetrics & Gynecology* 2012;24(6):422-433.

(53) Canadian Agency for Drugs and Technologies in Health. *Fetal fibronectin testing for pre-term labour put to the test in Canada's Arctic*. Health Technology Update 2009.

(54) Dutta D, Norman JE. The efficacy of fetal fibronectin testing in minimising hospital admissions, length of hospital stay and cost savings in women presenting with symptoms of pre-term labour. *Journal of Obstetrics and Gynaecology* 2010;30(8):768-773.

(55) Deshpande SN, van Asselt AD, Tomini F, et al. Rapid fetal fibronectin testing to predict preterm birth in women with symptoms of premature labour: A systematic review and cost analysis. *Health Technology Assessment* (Winchester, England) 2013;17(40):1-138.

(56) Hepburn CM, Booth M. Fetal fibronectin testing in Ontario: Successful government-sector collaboration to achieve high-quality and sustainable system change. *Healthcare Quarterly* 2011;14(Spec No 3):90-94.

(57) Canadian Perinatal Partnerships Coalition. *Fetal fibronectin. Guideline for use in the management of preterm labour*. Practice guideline. 2007.

(58) Sprague A. Impact of fetal fibronectin test implementation on health services utilization in the Province of Ontario: an interrupted time series analysis. *Journal of Obstetrics and Gynaecology Canada* 2012;34(6(Suppl 1)).

(59) Alere. *Quickly identify or rule out the risk of preterm delivery*. Ottawa ON: Alere Canada, 2012.

(60) Medix Biochemica (ed.). *Actim Partus. Qualitative test for estimation of cervical ripeness*. Kauniainen Finland: Medix Biochemica, 2009.

(61) Medix Biochemica. *Actim Partus - a rapid test to estimate the risk of preterm delivery*. Kauniainen Finland: Medix Biochemica, 2009.

(62) Wiwanitkit V. Diagnostic property of cervical phosphorylated insulin-like growth factor binding protein-1 in the prediction of preterm labor in symptomatic patients. *Archives of Gynecology and Obstetrics* 2010;281(1):175-176.

(63) British Association of Perinatal Medicine. *Management of acute in-utero transfers: A framework for practice*. London UK: British Association of Perinatal Medicine, 2011.

(64) Alberta Perinatal Health Program ESC. *Educational materials for the introduction of fetal fibronectin testing*. Edmonton AB: Alberta Perinatal Health Program, 2007.

(65) Provincial Council for Maternal and Child Health. *Draft fetal fibronectin (fFN) testing policy and procedure*. British Columbia: Canadian Perinatal Partnerships Coalition, 2008.

(66) Ontario Provincial Maternal-Newborn Advisory Committee. *Fetal fibronectin: Guideline for use in the management of preterm labour*. Toronto ON: Ontario Provincial Maternal-Newborn Advisory Committee, 2008.

(67) Audibert F, Fortin S, Delvin E, et al. Contingent use of fetal fibronectin testing and cervical length measurement in women with preterm labour. *Journal of Obstetrics & Gynaecology Canada:*2010; 32(4):307-312.

(68) Desjardins PR, Dansereau J, Hoag GN. Comparing the clinical effectiveness of fetal fibronectin and IGFBP-1 measurements in cervico-vaginal secretions, in predicting preterm deliveries. *Clinical Chemistry* 2008;54(6):A39-A40.

(69) Desjardins PR, Dansereau J, Hoag GN. *Comparing the clinical effectiveness of fetal fibronectin (FFN) and IGFBP-1 (IGFBP) measurements in cervico-vaginal secretions, in predicting preterm deliveries.* Conference: 60th Annual meeting of the American Association for Clinical Chemistry, Washington DC, July 2008.

(70) Dansereau J, Desjardins P, Hoag G. *Comparison between fetal fibronection and actim partus (pIGFBP-1) in the prediction of preterm delilvery.* Slide presentation at SOGC Clinical Annual Meeting, Calgary AB. Victoria BC: Vancouver Island Health Authority, 2008.

(71) Whiting PF, Rutjes AWS, Westwood ME, et al. A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;155(8):529-536.

(72) Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003;3:25.

(73) Whiting P, Rutjes AW, Westwood ME, et al. *Updating QUADAS: evidence to inform the development of QUADAS-2.* Available from: www.bris.ac.uk/quadas/resources/quadas2reporty4.pdf (accessed 28 May 2013).

(74) Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technology Assessment* 2004 8:25.

(75) Elie C, Coste J, French Society of Clinical Cytology Study Group. A methodological framework to distinguish spectrum effects from spectrum biases and to assess diagnostic and screening test accuracy for patient populations: Application to the Papanicolaou cervical cancer smear test. *BMC Medical Research Methodology* 2008;8(7).

(76) Willis BH. Spectrum bias - why clinicians need to be cautious when applying diagnostic tests studies. *Family Practice* 2008;25:390-396.

(77) Agency for Health Research and Quality (AHRQ). *Methods guide for medical test reviews.* AHRQ Publication No. 12-EC017. Rockville MC: Agency for Health Research and Quality, 2012.

(78) Jasechke R, Guyatt GH, Sackett DL. Users' guide to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-707.

(79) Grimes DA, Schulz KF. Refining diagnosis with likelehood ratios. *Lancet* 2005;365:1500-1505.

(80) Di Renzo GC, Roura LC, the European Association of Perinatal Medicine-Study Group on "Preterm Birth". Guidelines for the management of spontaneous preterm labour. *Archives of Perinatal Medicine* 2007;13(4):29-35.

(81) Cook DJ, Mulrow CD, Haynes RB. Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine* 1997;126(5):376-380.

(82) Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology* 2005;5:19.

(83) Westwood ME, Whiting P, Kleijnen J. How does study quality affect the results of a diagnostic meta-analysis? *BMC Medical Research Methodology* 2004;5:20.

SECTION THREE: Key Informant and Stakeholder Interviews

Charis Management Consulting - Lindsay Wodinski, MA; Margaret Wanke, BSc, MHSA

Methodology

The qualitative component of the fFN PPIR project was developed in consultation with the Post Policy Implementation Review of fFN Testing in Alberta Working Group (PPIR Working Group). The PPIR Working Group consists of representatives from Women's Health, Obstetrics, and Laboratory Services. The key informant interview strategy and data collection instruments closely aligned with the PPIR framework, work plan, and logic model (data collection templates and interview guides are included in Appendix 3.A). Interviewing took place over the phone between the months of May and July, 2013.

Participant Selection and Recruitment

On May 15, 2008 the AHS provincial health authority was established and the nine existing RHAs were reconfigured into five health zones.[b] Given that fFN testing policy implementation occurred alongside major restructuring of the health system, it was necessary to include representatives from both the former RHAs and current health zones in the data collection process. PPIR Working Group members and interview participants identified several individuals as possible interview contacts. The PPIR Research Team developed a comprehensive list of key informants who had direct involvement with fFN testing policy implementation in the 2006-2008 period. This group included:

- former RHA Women's Health Directors or designates;
- former RHA Obstetrical Leads or designates;
- former RHA Laboratory Services Directors or designates; and,
- a former Covenant Health representative.

In addition, the PPIR Research Team contacted individuals who currently use the fFN testing policy for managing patients who present with symptoms of preterm labour. These informants included:

- current Zone Women's Health Administrative Leads or designates;
- current Zone Women's Health Medical (Obstetrical) Leads or designates;
- current Laboratory Services Leads or designates; and,
- Covenant Health (Edmonton facilities) representatives.

In total, 24 participants were interviewed over 21 sessions (Table 3.1). This included two group interviews with a total of five laboratory services representatives. Of the 24 participants, eight represented Women's Health, three represented Obstetrics, and six represented laboratory services in the former RHAs and/or current health zones. Four individuals represented Covenant Health in the Edmonton/Capital area. Additionally, one representative was interviewed for each of the following provincial areas: AH, the APHP, and the Provincial Medical/Scientific Director of AHS Laboratory Services.

[b] Alberta Health. (2013). Alberta Health Services. Retrieved from http://www.health.alberta.ca/services/Alberta-Health-Services.html.

Table 3.1. Number of participants and interview sessions by representation and timeframe

Representative group	# of participants (n = 24)	# of interview sessions (n = 21*)	
		Policy implementation (pre-2008)	Current policy status (2013)
Alberta Health	1	1	-
Alberta Perinatal Health Program	1	1	-
AHS laboratory services (provincial)	1	1	1
Women's Health (RHA & Zone)	8	6	4
WH – Covenant Health	1	1	1
Obstetrics	3	2	2
Ob – Covenant Health	0	-	-
Laboratory services	6	2	4
Lab – Covenant Health	3	-	2
Total	**24**	**14**	**14**

* In seven instances, the same interview covered both timeframes. Single sessions involving group or multiple individuals were counted only once.

Complete or partial data was obtained for seven of the nine former RHAs (no representatives were available to speak for the Peace Country and Chinook areas) and four of five current AHS health zones (no representative was available to speak for the South zone). Among Covenant Health regions, 2006-2008 and 2013 representation was obtained for Edmonton and area facilities only.[c] Representation by RHA and zone may be found in Appendix 3.B.

Data Collection and Analysis

Key informants responded to descriptive and opinion-based questions about fFN testing policy implementation and current status. Five unique data collection templates were developed. One template was administered to former RHA contacts to determine how the fFN testing policy was implemented in each region between 2006 and 2008. This template asked about the timing, funding, and extent of policy implementation, availability of training, educational resources, and protocols, as well as whether the policy was formally monitored or resulted in any specific outcomes. The second template was administered to current AHS representatives. It asked a similar set of questions, but focused on current policy status, opportunities for training, availability of protocols, policy monitoring, and current outcome trends associated with the fFN testing policy. Current AHS representatives were asked to confirm AHS Provincial Laboratory Services 2013 survey data of fFN test kit and analyzer availability in each zone.

Three other key informant interview guides were developed for conversations with AH, AHS Provincial Laboratory Services, and APHP contacts. The AH representative was asked to discuss the political context leading up to the fFN testing policy letters, as well as how the policy directive was communicated and monitored by AH. The Provincial Laboratory Services AHS informant was

[c] Only Edmonton-based representatives of Covenant Health agreed to participate in the study. In this report, Covenant Health responses are presented alongside Capital Health or Edmonton zone respondent feedback.

asked to comment about lab data collection and involvement during policy implementation. The APHP contact was asked to describe the role of the APHP in supporting RHAs to implement the fFN testing policy. All three contacts were asked to comment on policy implementation challenges and facilitators, as well as policy implementation outcomes.

Interview data were recorded, transcribed, and compiled. A thematic content analysis technique was used to analyze key informant responses. Descriptive results are presented at the RHA and zone level, while opinion-based information is reported at the provincial level.

Research Limitations

While this PPIR qualitative study included multiple key informant perspectives to construct an in-depth account of fFN testing policy implementation across health regions in Alberta, it is important to note a few limitations:

- While the PPIR Research Team made every effort to reach key informants from all former RHAs and current zones, not all participants were available to participate in the study. This was largely due to changes in personnel since the time of policy implementation, as well as major restructuring of the provincial health system in 2008. Consequently, the findings provided in this report do not include the views of all possible informants involved at the time of policy implementation. Still, given changes in the health system, we were still able to obtain responses from 80% of both former RHAs and current health zones.
- As this study is retrospective in nature, the results of the policy review draw from institutional memory and the ability of key informants to recall and describe existing environmental and contextual factors at the time of implementation. As such, the information provided in this report should be interpreted with a degree of caution.

Results

This section describes the context of the fFN testing policy according to the perspectives of key informants. The observed inputs, outputs, and outcomes of policy implementation are presented in relation to the key review questions and indicators outlined in the PPIR work plan.

Policy Implementation Context

Understanding the policy context is important for making meaningful connections between policy objectives and outcomes. Representatives of AH, AHS, and the APHP were asked to describe the environmental factors that contributed to the 2006 policy decision to implement fFN testing. Participants noted that there was an existing knowledge base concerning the use of fFN for treating women with suspected preterm labour. Interest and activities about the use of this test were occurring in Calgary and Edmonton. Key individuals from each of these regions influenced the APHP to submit an application to consider fFN testing in the AHTDP. The AHTDP accepted fFN testing as one of the first health technologies to be reviewed. AH subsequently asked the IHE to conduct a review of the effectiveness of this technology. The IHE report, *Using Fetal Fibronectin to Diagnose Pre-term Labour*,[d] highlighted fFN as a valuable clinical tool with the potential to improve patient outcomes and improve health care efficiencies. AH also asked the University of Calgary to carry out an economic analysis of fFN testing. The combined findings of these two reports, as well

[d] The Institute of Health Economics. (2008). Using Fetal Fibronectin to Diagnose Pre-term Labour. Retrieved from: http://www.ihe.ca/documents/FetalFibronectin.pdf

as a policy analysis conducted by Charis, resulted in a synthesis report used to consult with RHA stakeholders about whether or not fFN testing should be provided. The majority of stakeholders agreed with introducing fFN as a publicly funded service in Alberta. The synthesis report and consultation findings informed AH's decision to implement the preterm labour testing policy.

The AH representative who participated in this study noted that the purpose of the policy directive was to ensure equal access to a diagnostic test that would improve health care for pregnant women while enhancing the overall efficiency and effectiveness of the health care system. In 2006, RHAs received the high level request from AH to introduce testing for preterm labour in their regions. Each RHA was asked to draw from their existing operating budgets and decide which facilities would have testing supplies and equipment. No additional funds were provided by the government as RHAs had recently received an increase in their annual budgets, and it was believed that the cost savings of using preterm labour testing would eventually outweigh the costs of implementation.

During the implementation period, AH asked the IHE to conduct a follow-up analysis of Actim™ Partus, after a representative of this product questioned why the province was advocating for a specific testing option, when two were available at the time. The results of the subsequent IHE study, comparing fFN to the Actim™ Partus test, resulted in an amendment to the original policy directive. In 2008, the policy request was broadened to include both fFN and Actim™ Partus testing as possible options for preterm labour testing. According to the AH representative, the AHTDP was also aiming to expand the scope of preterm labour testing technology as its relationships with researchers and experts (that is, the IHE, University of Alberta, and University of Calgary) expanded to include manufacturers of preterm labour testing.

According to informants, two important environmental factors contributed to the preterm labour testing policy in Alberta. The policy initiative was precipitated by growing interest among health professionals and organizations to implement a test for managing patients with suspected preterm labour. This interest paired well with AH's movement toward reviewing health technologies through the AHTDP, a new initiative that aimed to benefit health care for Albertans and the health system as a whole.

Policy Implementation Inputs

Policy implementation inputs refer to financial, human, and material resources used to achieve the intended policy outcomes. In the case of the fFN testing policy directive, implementation inputs include purchased testing supplies and equipment, training opportunities for clinical, technical, and laboratory service staff, as well as any established processes for monitoring policy implementation. Interviews with key informants focused on RHA policy implementation, communication, and monitoring strategies.

Was the policy decision implemented as planned and how was it implemented?

Number of RHAs that implemented the policy, by type of test

Representatives from seven former health regions indicated that their RHA had fully implemented the preterm labour testing policy using the fFN testing option. One RHA reported implementation of fFN testing well before policy implementation (Palliser), two RHAs had testing in place by 2006 (Capital Health/Covenant Health Edmonton facilities and Calgary), followed by three RHAs in 2007 (Aspen, Northern Lights, East Central) and one RHA in 2008 (David Thompson) (see Table 3.2).

Table 3.2. Reported fFN testing policy implementation across RHAs

RHAs	Time of policy implementation				
	Before 2006	2006	2007	2008	Unknown
Palliser	✓				
Capital Health Covenant Health Edmonton facilities		✓			
Calgary		✓			
Aspen			✓		
Northern Lights			✓		
East Central			✓		
David Thompson				✓	
Chinook					✓
Peace Country					✓
Total	**1**	**2**	**3**	**1**	**2**

RHA implementation plan, actual implementation, and current implementation status

RHA representatives were asked about their region's policy implementation plan and roll-out activities. In four RHAs, direction to implement the preterm labour policy was reported to come from Women's Health and Obstetrical Program areas. To a lesser extent, direction was also received from senior RHA executive levels or other groups such as Laboratory Services or Maternal Child Steering Committee (three RHAs). In some regions, policy direction was delivered through both channels (Edmonton, North, and Central areas).

Respondents were asked to confirm current fFN test kit and analyzer availability in 2013 based on survey data provided by AHS Provincial Laboratory Services (for site-level fFN test kit and analyzer availability, see Appendix 3.C). Table 3.3 presents participant verified fFN testing equipment availability in 2013.

Table 3.3. 2013 fFN test kit and analyzer availability per zone

Zone	Number and percent of obstetrical sites[e] with fFN test kits	Number and percent of obstetrical sites with fFN test kits and analyzers
North[f]	16 of 19 (84%)	13 of 19 (68%)
Edmonton	5 of 5 (100%)	4 of 5 (80%)
Central[g]	16 of 20 (80%)	8 of 20 (40%)
Calgary	5 of 5 (100%)	5 of 5 (100%)
South	4 of 8 (50%)	4 of 8 (50%)
All zones	**46 of 57 (81%)**	**34 of 57 (60%)**

[e] Sites with obstetrics are based on the APHP's Hospital of Birth and Community of Residence Matrix 2011.
[f] The 2013 number of fFN test kits and analyzers is unknown for three sites in the North zone.
[g] The 2013 number of fFN test kits and analyzers is unknown for one site in the Central zone.

Zone representatives were also asked to report whether they thought women's access to fFN testing had increased, decreased, or stayed the same after the policy implementation period. Of the four zones represented in this study, women's access to preterm labour testing was reported to increase in Edmonton after policy implementation (one Edmonton representative could not provide an estimate). Representatives in the Calgary, Central, and North zones felt that access has either increased or stayed the same (a few representatives from these areas could not provide an estimate). From a provincial standpoint, half of AHS zone informants (n = 16) believed that the percentage of women with access to preterm labour testing in their regions had increased (8), while others felt that access had decreased (2) or remained unchanged (3) (three participants (3) could not provide an estimate).

RHA allocation or reallocation of funding to support policy implementation

AH did not provide RHAs with additional funding for the implementation of the preterm labour testing policy. Study participants were asked how funds were allocated or reallocated in their regions to allow for the purchasing of fFN supplies and equipment. In five regions (Aspen, Northern Lights, Capital Health/Covenant Health, David Thompson, and Calgary), the costs associated with implementation were absorbed within the existing budgets of Women's Health, Obstetrical Programs, and Laboratory Services Program areas. In the Palliser RHA, program areas received additional dollars from the RHA to support policy implementation.[h] These additional funds were only used for purchasing fFN test kits and analyzers.

Were clinicians and support staff educated in the use of tests?

Investment in staff training and education

RHAs reported that they invested in resources to facilitate clinician and staff training and education of fFN testing, mainly through e-mailed memos, fFN testing equipment displays, physician rounds, new staff orientations, clinical staff educators, video telehealth conferencing, and e-learning opportunities. RHAs drew from resources developed and provided by the APHP, vendor (Adeza) and More^{OB} Program.

The most widely used educational resources for fFN training originated from the APHP. According to informants, the APHP's educational materials document, PowerPoint slides, and fFN Clinical Care Map were used in at least seven RHAs (Aspen, Northern Lights, Capital Health/Covenant Health, Calgary, David Thompson, East Central, and Palliser). These resources were developed after policy implementation by the APHP's Education Standing Committee, and in collaboration with other Canadian perinatal programs.

The APHP's *Educational Materials for The Introduction of Fetal Fibronectin Testing* document and PowerPoint presentation describe the:

- purpose of fFN testing;
- values and (contra) indicators of using an fFN test;
- equipment required to complete the test;
- procedure for specimen collection; and,

[h] The representative for the East Central RHA could not comment on how policy implementation was funded in the region.

- analysis and interpretation of fFN test results.

In addition, the APHP fFN Clinical Care Map was developed to guide health practitioners in using fFN to manage patients with symptoms of preterm labour.

Key informants from four of the former health regions (Calgary, David Thompson, Capital Health/Covenant Health, and Aspen) also mentioned using vendor materials to educate staff. Adeza representatives provided on-site training, as well as a 20-minute training CD (which was provided to RHAs with their initial purchase of fFN testing equipment). A third source of training and education during policy implementation was the More^{OB} Program, which offered further clinical information and workshops about fFN testing. Two RHA interviewees highlighted the importance of educating multiple staff groups on using the fFN test. For example, increased knowledge of testing criteria (for example, the need to avoid a vaginal examination prior to fFN testing) resulted in enhanced physician, resident, and nurse communication about a patient's previous assessment and eligibility for the fFN test. Laboratory services departments in some RHAs (Aspen and Northern Lights) developed a training checklist to educate laboratory staff about running analyses on collected fFN test specimens.

Of the four health zones represented in this study, respondents from Calgary and the Edmonton/Covenant Health Edmonton areas noted that, since the official roll-out of the preterm labour testing policy, testing for fFN has become standard practice, and thus training and education in the use of fFN tests only occurs during new staff orientations in current health zones.

Was policy implementation monitored?

AH and the APHP did not formally monitor how RHAs implemented preterm labour testing after the policy directive was issued. It was reported that AHS Women's Health and Obstetrical Programs and Laboratory Service leaders had enough expertise and capability to establish a diagnostic test in sites. AH offered informal support by way of answering questions and attending meetings. The APHP also supported policy implementation by providing educational materials for staff training purposes.

Some RHAs monitored policy implementation in their facilities by completing a quality assurance study (Calgary) and tracking the number of patients tested, testing date, testing result, length of hospital stay, and date of delivery (Capital Health/Covenant Health Edmonton facilities, Palliser); this information was not available for inclusion in this report. Four of the RHAs did not report any monitoring of policy implementation. The David Thompson RHA, for example, did not see a need for monitoring, due to low test volumes in its facilities. Information about policy implementation monitoring was not available for two RHAs (Peace Country and Chinook).

In the current AHS health zones, the North zone laboratory services tracks fFN test usage (information was not available to share in this report). The Edmonton zone/Covenant Health Edmonton facilities, Central, and Calgary zones do not formally monitor fFN test usage at present (monitoring practices are unknown for the South zone).

Policy Implementation Outputs

In this study, policy implementation outputs refer to the mechanisms used to put fFN testing into effect. Outputs include the number of practitioners trained in the use of preterm labour testing, the extent to which protocols were made available, and whether RHAs experienced any barriers and facilitators related to policy implementation. These areas are explored in the following sections.

Where implemented, what proportion of relevant practitioners received training and by what means?

Training and education investment across different staff groups

Informants were asked to estimate the extent to which obstetricians, family physicians, obstetrical nurses, and laboratory staff were trained in the use of fFN testing.[i] In some cases, participants could not estimate the percentage of staff trained across all groups in their region or zone. Tables 3.4 and 3.5 present reported RHA fFN staff training estimates during the policy implementation period, and by zone in 2013.

Table 3.4. Proportion of RHA staff that were trained during policy implementation

RHAs	Staff groups			
	Obstetricians	Family physicians	Obstetrical nurses	Laboratory services
Aspen	-	-	-	90-100%
Peace Country	-	-	-	-
Northern Lights	50-89%	50-89%	50-89%	-
Capital Health	90-100%	50-89%	-	90-100%
Covenant Health Edmonton facilities	90-100%	50-89%	50-89%	-
David Thompson	90-100%	90-100%	90-100%	-
East Central	-	-	50-89%	-
Calgary	50-89%	50-89%	50-89%	50-89%
Chinook	-	-	-	-
Palliser	90-100%	50-89%	50-89%	-

Table 3.5. Proportion of staff currently trained by zone (in 2013)

Zones	Staff groups			
	Obstetricians	Family physicians	Obstetrical nurses	Laboratory services
North	-	-	-	50-89%
Edmonton	90-100%	50-89%	90-100%	90-100%
Covenant Health Edmonton facilities	90-100%	50-89%	50-89%	90-100%
Central	90-100%	50-89%	50-89%	90-100%
Calgary	90-100%	90-100%	90-100%	90-100%
South	-	-	-	-

[i] Key informants were not always able to estimate the proportion of staff groups who were trained to carry out fFN testing in their area. In addition, no representatives were available to comment about staff training in the South zone.

What proportion of hospitals had testing/management protocols in place?

During policy implementation, availability of fFN testing/management protocols varied across the different staffing groups (that is, obstetricians/family physicians, obstetrical nurses, and laboratory staff). Many RHAs developed protocols directly from the training and education resources offered by the APHP, vendor, and More^{OB} Program, or used the training materials as their official protocols. Not all RHAs issued formal protocols for fFN testing. Table 3.6 describes the protocols used in each former RHA (according to whether the protocols originated at the RHA or hospital level) and current health zone (according to current protocol availability in obstetrical sites).

Table 3.6. RHA and zone fFN protocol use

RHA	Staff groups that have protocols (by level)	Zone	Staff groups that have protocols (with availability at obstetrical sites)
Aspen	• Laboratory services (RHA level)	**North**	• Laboratory services uses former Northern Lights policy (all sites) • No established fFN testing policy/procedure for North zone
Northern Lights	• Obstetricians/family physicians (RHA and hospital level) • Obstetrical nurses (RHA and hospital level)		
Capital Health	• Obstetricians/family physicians (RHA level) • Obstetrical nurses (RHA level) • Laboratory services (RHA level)	**Edmonton**	• No change from past protocols used • Obstetricians/family physicians (all sites) • Obstetrical nurses (all sites)
Covenant Health	• Obstetricians/family physicians (hospital level) • Obstetrical nurses (hospital level)		• Established updated fFN procedure documents • Obstetricians/family physicians (all sites)
Calgary	• Obstetricians/family physicians (RHA level) • Obstetrical nurses (RHA level) • Laboratory services (RHA level)	**Calgary**	• Policies have not been revised or updated since initial roll out • Obstetricians/family physicians (all sites) • Obstetrical nurses (all sites) • Laboratory services (4 of 7 sites)
David Thompson	• Obstetricians/family physicians (unknown level) • Obstetrical nurses (unknown level) • Laboratory services (unknown level)	**Central**	• All sites with fFN test kits/analyzers have Standard Operating Procedures in place • Obstetricians/family physicians (all sites) • Obstetrical nurses (all sites) • Laboratory services (all sites)
East Central	• Obstetrical nurses (unknown level)		
Palliser	• Obstetricians/family physicians (RHA level) • Obstetrical nurses (RHA level)	**South**	• Unknown

What implementation facilitators and challenges were identified?

All RHA representatives (n = 13) indicated that policy implementation occurred as planned in their regions. Key informants described a number of factors that facilitated policy implementation across the province. These factors included the following:

- **A successful pilot study of fFN testing (5):** Participants noted that the successful results from an early pilot study of fFN testing helped to increase legitimacy and credibility of this testing method among health providers.
- **Local champions of fFN testing (4):** In a few of the RHAs, obstetrical physicians encouraged adoption of fFN testing in their regions prior to policy implementation.
- **Efforts to educate clinical staff about fFN testing (4):** Clinical educators and obstetrical nurses played a key role in educating staff about fFN testing and patient eligibility requirements (2). In addition, availability of ready-made and researched APHP materials also contributed to staff training activities (2).
- **Support from key groups to adopt fFN testing (3):** According to a few participants, RHA Laboratory Services and RHA level committees played an instrumental role in supporting policy implementation and roll out in regions. For example, RHA committees communicated the policy directive to sites and arranged purchasing of fFN testing supplies and equipment.
- **Already existing infrastructure for fFN testing (2):** In some RHAs, fFN specimen collection kits and analyzers were established prior to AH issuing the 2006 policy letter.
- **Timeliness (2):** As fFN testing supplies and equipment were made available to RHAs soon after AH released the preterm labour testing policy, RHAs were able to quickly implement fFN tests in facilities.
- **Organizational environment (1):** One participant commented that working in an environment that values organizational efficiencies contributed to interest in and uptake of preterm labour policy.

Key informants from the various RHAs and current zones (n = 21) described a number of barriers associated with policy implementation. These challenges included the following:

- **Cost of fFN tests and analyzers (5):** Particularly in the early stages of policy implementation, RHAs were unsure as to how many fFN test kits and analyzers should be ordered and the amount of funding this would require from their internal budgets. In Laboratory Services departments, staff had to determine the appropriate frequency for running fFN test batches, as well as the manpower required to complete fFN test analyses. In addition, a laboratory services representative observed that more fFN test kits have been ordered in recent years, likely due to an increase in the number of pregnant women in Alberta. It was noted that fFN testing continues to have cost implications for Laboratory Services departments in the current health zones.
- **fFN testing in rural areas (4):** Key informants explained that not all rural sites had immediate access to an fFN analyzer and would have to send test samples offsite for analysis. This posed some patient management challenges for physicians who had to wait for test results before discharging patients or admitting/transferring them to another site for further care. Participants observed that it was not uncommon for physicians at non-obstetrical sites to transfer patients along with their (non-analyzed) fFN test specimens to

tertiary hospitals with obstetrics. In some health zones, physicians have requested on-site testing availability; however, this had not yet been granted due to cost.

- **fFN testing efficiency (3):** Participants noted challenges with ensuring that fFN testing was carried out in an efficient and cost effective manner. For instance, sites that saw few patients with suspected preterm labour had to monitor fFN testing supplies to avoid expiration of tests ordered. This challenge is ongoing in current health zones.
- **Appropriate use of fFN testing (2):** Some participants noted a few instances when the test was not used appropriately during the policy implementation period. For instance, there was a need to enhance staff awareness about when to perform the test. In some cases, fFN tests were completed and swabs were sent to lab for analysis even though patients were in labour. In addition, staff had to spend additional time educating others about correctly sealing specimen samples to ensure that test results are reliable and valid. While appropriate use of fFN testing has improved over time, health zone representatives suggest there is ongoing need to ensure that tests are used appropriately.
- **Staff education and expertise (2):** Participants indicated that a challenge for smaller regions was a lack of expertise on using fFN testing for managing patients with symptoms of preterm labour. In addition, a potential challenge for some regions was ongoing staff education about fFN testing, particularly in sites that saw few patients with symptoms of preterm labour.

Policy Implementation Outcomes

An important objective of the PPIR study is to determine whether the AH policy decision to adopt testing for preterm labour led to more accurate assessment and improved management of patients presenting with symptoms of preterm labour. To gain insight on this ultimate outcome, key informants were asked to comment or provide tracked data on whether obstetricians and family physicians changed their practice patterns as a result of preterm labour testing availability (that is, did obstetricians and family physicians alter their management of patients possibly experiencing preterm labour). Participants were also asked to convey any unintended consequences resulting from implementing the preterm labour testing policy. Of note is that this qualitative data on policy outcomes compliments quantitative data analysis being conducted through other activities in this PPIR.

Did obstetricians and family physicians change their practice patterns/trust the test?

RHA representatives (n = 13) generally agreed that obstetricians and family physicians trusted the test and altered their patient management practices once fFN testing became available in the 2006-2008 period (11 participants agreed, and 2 participants could not comment on this item). Participants noted that fFN became part of physicians' overall patient assessment procedures (for example, physicians avoided vaginal examination of patients prior to using the fFN test), and laboratory services fFN test analysis volumes were higher than initially expected. Physicians awaited fFN test results before deciding whether to admit or transfer a patient, but noted they sometimes kept the patient in care regardless of testing results, due to other health factors (for example, high blood pressure).

Among AHS health zones representatives (n = 16), most informants indicated that physicians have continued to use fFN testing as part of their overall assessment of patients (14 participants agreed, and 2 participants could not comment on this question). Participants commented that physicians' use of fFN testing became more efficient over time (that is, physicians avoided using the test if the patient was clearly in labour), and that the test is routinely used to make decisions about patient discharge, admission, or transfer.

Did policy implementation result in unintended consequences?

All key informants (n = 24) were asked to discuss whether the preterm labour policy resulted in any unintended consequences. The majority of informants did not identify any unexpected activities that occurred. Three participants suspected that the use of fFN testing did not result in fewer hospital transfers of women in suspected preterm labour. It was reported that patients with negative fFN results were still being transferred, and, in some cases, patients were transferred along with a test swab that was yet to be analyzed (generally due to lack of access to analyzers). A participant suggested that patient and specimen transfer could be viewed positively, as health care workers are provided with additional information about the patient during assessment. Another informant noted that, while the policy to introduce preterm labour testing has not resulted in any adverse consequences, it has not been an overly successful cost saving measure for the government and health care system as whole.

Additional Considerations

During interview conversations, key informants were given the opportunity to provide additional comments about RHA policy implementation and current zone policy status. Among RHA informants (n = 13), seven participants commented on the following:

- RHAs implemented fFN testing because more evidence-based information about this health technology was available at the time, compared to Actim™ Partus testing (2).
- Policy roll-out was facilitated by an already existing fFN testing knowledge base in Women's Health and Laboratory Services areas and a collective willingness to adopt the test (2).
- There is a need for ongoing quality assurance processes to ensure that health service providers are using the most effective and efficient technology for managing patients in suspected preterm labour (1).
- Greater organizational flexibility is needed to be able to adopt new testing strategies in light of emergent research and technologies (1).
- A provincial approach would have helped RHAs to implement the policy in a more similar and consistent manner (1).

In addition, five zone representatives (n = 16) indicated the following:

- The fFN test is simple and easy to access. Laboratory Services almost always receives valid specimens for analysis (1).
- Staff knowledge of fFN testing is greater in hospitals with higher numbers of births, compared with staff in hospitals that see fewer births (1).
- Women in Alberta have different levels of access to preterm labour testing depending on their location, as not all sites have immediate access to fFN analyzers. This can pose problems related to patient care, as well as time and resource efficiency (1).

- Review of the choice to use fFN testing should have occurred sooner after policy implementation, as new information about other (less expensive) technologies is now available (1).
- Current AHS evaluation processes make it difficult to review current testing practices and reallocate funding to implement alternative options (1).

Conclusion

According to key informants, the preterm labour policy was implemented to the fullest extent possible in regions across the province. Each RHA (with representatives who participated in the study) reportedly selected and introduced fFN testing for preterm labour prior to 2008, as requested by AH. RHA implementation strategies varied in relation to available staff training and education opportunities and fFN testing protocols used. Implementation strategies were facilitated by resources made available by the APHP and Adeza. Many RHAs drew from APHP learning materials, among other resources, to inform training practices and protocol development. A few key factors that positively influenced policy implementation included an already existing knowledge base and willingness to adopt fFN testing for managing patients in suspected preterm labour. In addition, participants mentioned that obstetrical experts and champions of fFN testing were instrumental in generating interest and support for the test prior to policy roll-out. In contrast, some of the barriers of policy implementation noted by informants concerned funding fFN testing kits and analyzers and establishing test analysis procedures in rural sites, particularly in areas with low numbers of births.

Key informants suggested that fFN testing has been incorporated into physician patient assessment practices and physicians appear to trust and use fFN test results to decide whether to admit, transfer, or discharge patients. While fFN testing and analysis has become standard practice in the work of obstetrical physicians and nurses, family physicians, and laboratory staff, informants suggested that this has not necessarily resulted in reduced usage of health system resources. One unintended consequence identified by informants is that patients are being admitted and transferred to hospitals regardless of fFN test results, and, in some cases, patients are transferred with unanalyzed specimens. Informants emphasized the value of reviewing health policies and new technologies and having the flexibility to adopt new technologies to enhance patient care and overall health system efficiency.

Appendix 3.A: Key Informant Interview Guides

Post-Policy Implementation Review (PPIR) of Preterm Labour Testing in Alberta

DATA COLLECTION TEMPLATE – 2006-08 POLICY IMPLEMENTATION BY REGIONAL HEALTH AUTHORITY (RHA)

Regional Health Authority:		
Key Informants:	Former RHA Women's Health Director or Designate:	Name: Current title: Former title:
	Former RHA Obstetrical Lead or Designate:	Name: Current title: Former title:
	Former RHA Director of Laboratory Services or Designate:	Name: Current title: Former title:

INTRODUCTION

In September 2006, Alberta Health and Wellness issued the following policy directive to all RHAs: *Fetal fibronectin testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008.* A subsequent letter was issued in March 2008, indicating that health regions could choose implementation of either fetal fibronectin or Actim™ Partus testing.

Alberta Health has now asked the Institute of Health Economics (IHE) to conduct a retrospective review of the implementation of this policy. This review is being guided by a provincial working group that includes representatives from Women's Health, Obstetrics and Laboratory Services. The goal of the review is to determine the impact of the preterm labour testing policy, as well as explore how it was implemented in the former health regions in Alberta.

The IHE has asked Charis Management Consulting Inc. (Charis) to support the retrospective review by contacting key informants from the former health regions. We will be interviewing representatives from Women's Health, Obstetrics and Laboratory Services to obtain a complete picture of how the policy was implemented in each former health region in 2008. We would like to speak with you about the questions outlined in this data collection template. You may choose to complete as much of the template as possible before the telephone interview. If you are willing to participate in the project, we will ask you to take part in a 20-30 minute telephone interview or assist us in collecting information.

The template entails both descriptive and opinion-based questions. In the final project write-up, descriptive information (Section A) will be presented in template form, while all opinion-based information (Section B) will be compiled and presented in aggregate form. Responses to opinion-based questions will not be presented in individual form, nor will former health regions be identified.

We may ask to audio-tape the interview conversation to better capture the discussion details. The audio file will be held confidential and will only be used for the purpose of this project. If you choose to not have the interview recorded, we will only take notes during the conversation. Your completion of the form or willingness to participate in the telephone interview will constitute your consent for taking part in the project.

SECTION A: Descriptive Questions

RHA Policy Implementation

1. (a) To what extent was the preterm labour policy (as described in the introduction) implemented in your region?

   ☐ The policy was fully implemented
   ☐ The policy was partially implemented
   ☐ The policy was not implemented

   (b) If the policy was partially or not implemented, please explain:

2. At what level of the RHA were decisions made about the implementation of preterm labour testing after the policy was issued by Alberta Health?

   ☐ Laboratory Services
   ☐ Senior RHA executive level
   ☐ Women's Health or Obstetrical Program area
   ☐ Other, please specify:

3. When was the policy for preterm labour testing implemented in your region? *If implementation spanned more than one year, select all years that apply.*

   ☐ 2006
   ☐ 2007
   ☐ 2008
   ☐ Not implemented

4. The 2006 policy directive stated that "*Given the potential for better and more appropriate care, potential savings, and the modest costs involved, RHAs are asked to fund the introduction and on-going operating costs of fFN testing from within existing budget allocations. The target date is intended to allow sufficient time for health regions to address implementation and budget issues.*"

   From which RHA program areas were funds allocated/reallocated?
   Please select all that apply:

   ☐ Each area of responsibility absorbed the cost within their existing budget.
   ☐ Additional dollars were found in the RHA and allocated to support implementation in Women's Health and/or Laboratory Services.
   ☐ Other, please specify:

5. In the table below, please indicate the availability of preterm labour testing kits and analyzers in your region at the time of implementation (by April 1, 2008):

Healthcare Facility:	Specimen Collection Kits on Site?	Analyzers on Site?
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know
	☐Yes ☐No ☐ Don't know	☐Yes ☐No ☐ Don't know

6. Comments about policy implementation in the region:

RHA Training for Preterm Labour Testing

7. Did your RHA provide training for preterm labour testing for any of the following groups?

8. For the groups who received training, please estimate the proportion of relevant* staff who were trained.

Group	Yes	No	Don't Know	All (90-100%)	Most (50-89%)	Some (25-49%)	Few (<25%)
a. Obstetricians	☐	☐	☐	☐	☐	☐	☐
b. Family physicians	☐	☐	☐	☐	☐	☐	☐
c. Obstetrical nurses	☐	☐	☐	☐	☐	☐	☐
d. Lab services staff	☐	☐	☐	☐	☐	☐	☐
e. Other, please specify:	☐	☐	☐	☐	☐	☐	☐

**'Relevant' refers to those who would have been involved in obstetrical deliveries, or in the collection and analysis of tests for preterm labour.*

9. What training resources were used for each group?

Group	RHA generated training package	Alberta Perinatal Health Program (APHP) materials	Vendor materials (i.e., Adeza)	Other, please specify:
a. Obstetricians	☐	☐	☐	
b. Family physicians	☐	☐	☐	
c. Obstetrical nurses	☐	☐	☐	
d. Lab services staff	☐	☐	☐	
e. Other, please specify:	☐	☐	☐	

10. Comments about training for preterm labour testing when introduced in the region:

RHA Protocols for Preterm Labour Testing

11. Were protocols established for preterm labour testing for any of the following groups?

12. If yes, at what level did the protocol(s) originate (i.e., were they developed and/or implemented at the RHA level or hospital level)?

Group	Yes	No	Don't Know	RHA level	Hospital level	Other (*Indicate who developed the protocol*):
a. Obstetricians	☐	☐	☐	☐	☐	
b. Family physicians	☐	☐	☐	☐	☐	
c. Obstetrical nurses	☐	☐	☐	☐	☐	
d. Lab service staff	☐	☐	☐	☐	☐	
e. Other, please specify:	☐	☐	☐	☐	☐	

13. Comments about protocols used for preterm labour testing in the region:

RHA Preterm Labour Testing Data

14. (a) Did your region conduct a study or formally monitor the implementation of the preterm labour testing (fetal fibronectin or Actim™ Partus) policy (e.g., for quality assurance purposes, evaluating implementation, evaluating outcomes, determining the impact on hospital utilization)?

    ☐ No
    ☐ Yes
    ☐ Don't know

    (b) If yes, briefly describe:

    (c) Would you be willing to share the findings with the research team?

    ☐ Yes
    ☐ No

SECTION B: Opinion Questions

RHA Policy Outcomes

15. Did policy implementation occur as planned or expected in your health region? Why or why not?

16. What factors facilitated the implementation of preterm labour testing policy in your region?

17. What barriers or challenges were encountered when implementing the policy in your region?

18. Do you believe obstetricians and family physicians changed their practice patterns as a result of availability of preterm labour testing; that is, did the test alter their management of women presenting with symptoms of preterm labour? On what do you base your answer?

19. Did you observe any unintended consequences as a result of preterm labour testing policy implementation? If yes, please describe.

20. Do you have any final comments about the implementation of preterm labour testing policy in your region?

Your involvement in the project is of value and your insights are important. Thank you for sharing your knowledge with us!

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

Post-Policy Implementation Review (PPIR) of Preterm Labour Testing in Alberta

DATA COLLECTION TEMPLATE – 2013 CURRENT STATUS OF TESTING IN ALBERTA HEALTH SERVICES (AHS) ZONES

AHS Zone:		
Key Informants:	Women's Health Administrative Lead or Representative:	Name: Current title:
	Women's Health Medical Lead or Representative:	Name: Current title:
	Laboratory Services Lead or Representative:	Name: Current title:

INTRODUCTION

In September 2006, Alberta Health and Wellness issued the following policy directive to all RHAs: *Fetal fibronectin testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008.* A subsequent letter was issued in March 2008, indicating that health regions could choose implementation of either fetal fibronectin or Actim™ Partus testing.

Alberta Health has now asked the Institute of Health Economics (IHE) to conduct a retrospective review of the implementation of this policy. This review is being guided by a provincial working group that includes representatives from Women's Health, Obstetrics and Laboratory Services. The goal of the review is to determine the impact of the preterm labour testing policy, as well as explore how the policy currently operates in health zones across the province.

The IHE has asked Charis Management Consulting Inc. (Charis) to support the retrospective review by contacting key informants in the current health zones. We will be interviewing representatives from Women's Health, Obstetrics and Laboratory Services to obtain a complete picture of how the policy to implement preterm labour testing is currently working. We would like to speak with you about the questions outlined in this data collection template. You may choose to complete as much of the template as possible before the telephone interview. If you are willing to participate in the project, we will ask you to take part in a 20-30 minute telephone interview or assist us in collecting information.

The template entails both descriptive and opinion-based questions. In the final project write-up, descriptive information (Section A) will be presented in template form, while all opinion-based information (Section B) will be compiled and presented in aggregate form. Responses to opinion-based questions will not be presented in individual form, nor will AHS zones be identified.

We may ask to audio-tape the interview conversation to better capture the discussion details. The audio file will be held confidential and will only be used for the purpose of this project. If you choose to not have the interviews recorded, we will only take notes during the conversation. Your completion of the form or willingness to participate in the telephone interview will constitute your consent for taking part in the project.

SECTION A: Descriptive Questions

Zone Policy Status

1. Since the preterm labour policy was implemented in about 2007, has there been any change in access to preterm labour testing among women who present with symptoms of preterm labour?

   Please estimate whether the percent of women with access to preterm labour testing has:

   ☐ Stayed the same
   ☐ Increased
   ☐ Decreased
   ☐ Cannot provide an estimate

2. In the table below, please indicate the availability of preterm labour testing kits and analyzers at each site in your zone:

Healthcare Facility:	Specimen Collection Kits on Site?	Analyzers on Site?
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know
	☐ Yes ☐ No ☐ Don't know	☐ Yes ☐ No ☐ Don't know

3. Provide any comments about policy implementation in your zone:

Zone Training for Preterm Labour Testing				
4. Please estimate the proportion of current staff who are currently trained in preterm labour testing. *Please provide information about the groups you can speak to:*				
Group	**All (90-100%)**	**Most (50-89%)**	**Some (25-49%)**	**Few (<25%)**
a. Obstetricians	☐	☐	☐	☐
b. Family physicians	☐	☐	☐	☐
c. Obstetrical nurses	☐	☐	☐	☐
d. Lab service staff	☐	☐	☐	☐
e. Other, please specify:	☐	☐	☐	☐
5. Comments about training for preterm labour testing for current staff in the zone:				

Zone Protocols for Preterm Labour Testing									
6. In which of the following healthcare facilities and staff groups in your zone are protocols for preterm labour testing in place? *Please provide information about the groups you can speak to:*									
Healthcare Facility	**Obstetrical/Family physicians**			**Obstetrical nurses**			**Lab services staff**		
	Yes	**No**	**Don't know**	**Yes**	**No**	**Don't know**	**Yes**	**No**	**Don't know**
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐	☐	☐	☐	☐
7. Comments about the protocols currently used for preterm labour testing in your zone:									

Zone Preterm Labour Testing Data

8. (a) Has anyone in your zone conducted a study or formally monitored any aspect of preterm labour testing (fetal fibronectin or Actim™ Partus) since 2006 (e.g., for quality assurance purposes, evaluating implementation, evaluation outcomes, determining the impact on hospital utilization)?

☐ No
☐ Yes

☐ Don't know

(b) If yes, briefly describe:

(c) Would you be willing to share the findings with the research team by providing any materials related to your study?

☐ Yes
☐ No

SECTION B: Opinion Questions

Zone Policy Outcomes

9. Are you experiencing any barriers or challenges with respect to preterm labour testing in your zone at the present time? If so, please describe.

10. To what extent have obstetricians and family physicians in your zone accepted preterm labour testing and are managing women with signs or preterm labour according to test results; that is, are obstetricians and family physicians managing and discharging patients appropriately based on the results of the test? Please elaborate.

11. Have you observed any unintended consequences as a result of the preterm labour testing policy in your zone? If yes, please describe.

12. Do you have any final comments about the current use of preterm labour testing in your zone?

Your involvement in the project is of value and your insights are important. Thank you for sharing your knowledge with us!

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

Post-Policy Implementation Review (PPIR) of Preterm Labour Testing in Alberta

ALBERTA HEALTH KEY INFORMANT INTERVIEW GUIDE

Name and Title:
Date and Time:

INTRODUCTION

In September 2006, Alberta Health and Wellness (AHW) issued a policy directive to all RHAs. The policy directive was as follows: *Fetal fibronectin testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008*. A subsequent letter was issued in March 2008, indicating that health regions could choose implementation of either Fetal Fibronectin or Actim™ Partus testing.

Alberta Health has now asked the Institute of Health Economics (IHE) to conduct a retrospective review of the implementation of this policy. This review is being guided by a provincial working group that includes representatives from Women's Health, Obstetrics and Laboratory Services. The goal of the review is to review the implementation process and determine the impact of the preterm labour testing policy.

The IHE has asked Charis Management Consulting Inc. (Charis) to support the retrospective review by contacting key informants who can speak to policy implementation in 2006-2008 and/or to the current status of policy implementation in 2013. We would like to arrange a time for a telephone interview with you. Our conversation would be guided by the questions listed below. If you agree to participate, you have the right to not answer any question, to conclude the interview at any point, or withdraw your information at any time during or after the interview.

Key informant interviews will provide insight about the implementation process as well as the desired and unexpected outcomes of the preterm labour testing policy. Only Charis researchers will have access to the raw data from individual interviews. Overall findings will be presented in summarized form in the main body of the report. If you have any questions or concerns please contact either of the following PPIR Research Team members:

Dr. Anderson Chuck
PPIR Project Lead
Director of Economic Evaluation and Analytics,
Institute of Health Economics
1200-10405 Jasper Ave., Edmonton, Alberta T5J 3H1
Phone: (780) 448-4881
E-mail: achuck@ihe.ca

Lindsay Wodinski
PPIR Project Coordinator
Research Analyst,
Charis Management Consulting Inc.
418, 10123 99 St., Edmonton, AB T5J 3H1
Phone: (780) 496-9067, Ext.231
E-mail: lindsay@charismc.com

SECTION 1: Policy Decision and Communication

1. What factors led to preterm labour policy implementation in Alberta? *Why was this policy decision implemented?*
2. What were the main goals of the policy? What was Alberta Health (then, Alberta Health and Wellness) hoping to achieve by implementing the policy?
3. Describe the background that led to the issuing of two policy statements – one in September 2006 regarding fetal fibronectin testing, and a second policy statement, issued in March 2008, that broadened the type of testing to include Actim™ Partus?
4. How was each of these policy decisions communicated to RHAs? *What written communications were issued, by whom and to whom? Did AH sponsor or participate in any meetings to communicate the policy decisions? If yes, please describe.*

SECTION 2: Policy Implementation

5. Was policy implementation monitored? If yes, describe. *Did AH/AHW follow-up with RHAs after the policy was issued? If so, how?*
6. Did AHW offer or provide any support to RHAs or others (e.g., Alberta Perinatal Health Program) to implement the policy? If yes, please describe.
7. What factors contributed to the policy implementation process? *What factors enabled RHAs to adopt the policy?*
8. What were the challenging aspects of policy implementation, if any? *Were you aware of any challenges or barriers that impeded RHA adoption of the policy?*

SECTION 3: Policy Outcomes

9. Do you perceive the policy achieved its intended outcomes? Explain.
10. Were there any the unexpected outcomes or consequences of policy implementation? If yes, describe.

SECTION 4: Conclusion

11. Do you have any final comments or questions? Would you to note any other observations about preterm labour policy implementation in Alberta?

Thank you for taking the time to assist with this policy review.

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

Post-Policy Implementation Review (PPIR) of Preterm Labour Testing in Alberta

AHS Provincial Laboratory Services Key Informant Interview Guide

Name and Title:
Date and Time:

INTRODUCTION

In September 2006, Alberta Health and Wellness (AHW) issued a policy directive to all RHAs. The policy directive was as follows: *Fetal fibronectin testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008*. A subsequent letter was issued in March 2008, indicating that health regions could choose implementation of either Fetal Fibronectin or Actim™ Partus testing.

Alberta Health has now asked the Institute of Health Economics (IHE) to conduct a retrospective review of the implementation of this policy. This review is being guided by a provincial working group that includes representatives from Women's Health, Obstetrics and Laboratory Services. The goal of the review is to review the implementation process and determine the impact of the preterm labour testing policy.

The IHE has asked Charis Management Consulting Inc. (Charis) to support the retrospective review by contacting key informants who can speak to policy implementation in 2006-2008 and/or to the current status of policy implementation in 2013. We would like to arrange a time for a telephone interview with you. Our conversation would be guided by the questions listed below. If you agree to participate, you have the right to not answer any question, to conclude the interview at any point, or withdraw your information at any time during or after the interview.

Key informant interviews will provide insight about the implementation process as well as the desired and unexpected outcomes of the preterm labour testing policy. Only Charis researchers will have access to the raw data from individual interviews. Overall findings will be presented in summarized form in the main body of the report. If you have any questions or concerns please contact either of the following PPIR Research Team members:

Dr. Anderson Chuck
PPIR Project Lead
Director of Economic Evaluation and Analytics,
Institute of Health Economics
1200-10405 Jasper Ave., Edmonton, Alberta T5J 3H1
Phone: (780) 448-4881
E-mail: achuck@ihe.ca

Lindsay Wodinski
PPIR Project Coordinator
Research Analyst,
Charis Management Consulting Inc.
418, 10123 99 St., Edmonton, AB T5J 3H1
Phone: (780) 496-9067, Ext.231
E-mail: lindsay@charismc.com

SECTION 1: Policy Decision and Communication

1. Can you please describe the extent to which lab data related to preterm labour testing is available?
2. To what extent are you aware of RHA decisions to implement the fFN testing policy and the difference approaches they may have used?

SECTION 2: Policy Implementation

3. Did laboratory staff receive any training about preterm labour testing and analysis? If yes, please describe.
4. Are you aware of any protocols established for lab staff regarding preterm labour testing and analysis?

SECTION 3: Policy Outcomes

5. What factors facilitated implementation of the preterm labour testing policy?
6. What implementation barriers or challenges were observed by lab staff?
7. Were there any the unexpected outcomes or consequences of policy implementation? If yes, describe.

SECTION 4: Conclusion

8. Do you have any final comments or questions?

Thank you for taking the time to assist with this policy review.

INSTITUTE OF HEALTH ECONOMICS
ALBERTA CANADA

Post-Policy Implementation Review (PPIR) of Preterm Labour Testing in Alberta

ALBERTA PERINATAL HEALTH PROGRAM KEY INFORMANT INTERVIEW GUIDE

Name and Title:
Date and Time:

INTRODUCTION

In September 2006, Alberta Health and Wellness (AHW) issued a policy directive to all RHAs. The policy directive was as follows: *Fetal fibronectin testing should be introduced as a publicly funded service available to all Alberta women and through all RHAs at the earliest possible date, but no later than April 1, 2008*. A subsequent letter was issued in March 2008, indicating that health regions could choose implementation of either Fetal Fibronectin or Actim™ Partus testing.

Alberta Health has now asked the Institute of Health Economics (IHE) to conduct a retrospective review of the implementation of this policy. This review is being guided by a provincial working group that includes representatives from Women's Health, Obstetric and Laboratory Services. The goal of the review is to review the implementation process and determine the impact of the preterm labour testing policy.

The IHE has asked Charis Management Consulting Inc. (Charis) to support the retrospective review by contacting key informants who can speak to policy implementation in 2006-2008 and/or to the current status of policy implementation in 2013. We would like to arrange a time for a telephone interview with you. Our conversation would be guided by the questions listed below. If you agree to participate, you have the right to not answer any question, to conclude the interview at any point, or withdraw your information at any time during or after the interview.

Key informant interviews will provide insight about the implementation process as well as the desired and unexpected outcomes of the preterm labour testing policy. Only Charis researchers will have access to the raw data from individual interviews. Overall findings will be presented in summarized form in the main body of the report. If you have any questions or concerns please contact either of the following PPIR Research Team members:

Dr. Anderson Chuck
PPIR Project Lead
Director of Economic Evaluation and Analytics,
Institute of Health Economics
1200-10405 Jasper Ave., Edmonton, Alberta T5J 3H1
Phone: (780) 448-4881
E-mail: achuck@ihe.ca

Lindsay Wodinski
PPIR Project Coordinator
Research Analyst,
Charis Management Consulting Inc.
418, 10123 99 St., Edmonton, AB T5J 3H1
Phone: (780) 496-9067, Ext.231
E-mail: lindsay@charismc.com

SECTION 1: Policy Decision and Communication

1. What kind of support did the APHP offer to the RHAs to support the implementation of the preterm labour testing policy?

2. Are you aware of any other supports provided by the APHP?

SECTION 2: Policy Implementation

3. Did the APHP follow up with RHAs after the policy was implemented?

SECTION 3: Policy Outcomes

4. What factors facilitated implementation of the preterm labour testing policy in the province?

5. What implementation barriers or challenges did you observe during the implementation of the preterm labour testing policy?

6. Are you aware of any unexpected outcomes or consequences of policy implementation? If yes, describe.

SECTION 4: Conclusion

7. Do you have any final comments or questions?

Thank you for taking the time to assist with this policy review.

Appendix 3.B: Data Collection Representation from RHA, Zone, and Covenant Health Areas

Table 3.B.1. Key informant interview data collection regarding 2006-2008 policy implementation

RHA	Women's Health Director or designate	Obstetrical Lead or designate	Director of Laboratory Services or designate	Data collected
Aspen	x	x	✓	*Partial*
Peace Country	x	x	x	*No data*
Northern Lights	✓	x	x	*Partial*
Capital Health	✓	✓	✓	*Complete*
Covenant Health Edmonton area	✓	x	x	*Partial*
Calgary	✓	✓	x	*Partial*
David Thompson	✓	x	x	*Partial*
East Central	✓	x	x	*Partial*
Chinook	x	x	x	*No data*
Palliser	✓	x	x	*Partial*

Table 3.B.2. Key informant interview data collection regarding 2013 policy status

Zone	Women's Health Director or designate	Obstetrical Lead or designate	Director of Laboratory Services or designate	Data collected
North	✓	x	✓	*Partial*
Edmonton	✓	✓	✓	*Complete*
Covenant Health Edmonton area	✓	x	✓ (2 KIIs)	*Partial*
Calgary	✓	✓	✓	*Complete*
Central	✓	x	✓	*Partial*
South	x	x	x	*No data*

Appendix 3.C: fFN Test Kit and Analyzer Availability in Zone Obstetrical Facilities

Table 3.C.1. Zone availability of fFN test kits and analyzers sites

Sites	fFN test kit availability	fFN test kit and analyzer availability
North (19 sites)		
Athabasca	✓	✓
Barrhead	✓	✓
Beaverlodge	✓	x
Bonneyville	✓	✓
Cold Lake	✓	✓
Edson	✓	✓
Fairview	✓	x
Fort McMurray	-	-
Fort Vermilion	-	-
Grande Prairie	✓	✓
High Level	-	-
Hinton	✓	✓
Lac La Biche	✓	✓
Peace River	✓	✓
St. Paul	✓	✓
Slave Lake	✓	✓
Valleyview	✓	x
Westlock	✓	✓
Whitecourt	✓	✓
Total	**16 (84%)**	**13 (68%)**
Edmonton (5 sites)		
Fort Saskatchewan	✓	x
Grey Nuns	✓	✓
Misericordia Hospital	✓	✓
Royal Alexandra Hospital	✓	✓
Total	**5 (100%)**	**4 (80%)**
Central (20 sites)		
Camrose	✓	✓
Daysland	✓	✓
Drayton Valley	x	x
Drumheller	x	x
Hanna	✓	x
Lacombe	✓	x

Lloydminster	-	-
Olds	✓	✓
Ponoka	✓	x
Provost	✓	✓
Red Deer	✓	✓
Rimbey	✓	x
Rocky Mountain House	✓	x
Stettler	✓	x
Sundre	✓	x
Three Hills	✓	x
Vermilion	✓	✓
Viking	✓	✓
Wainwright	✓	✓
Wetaskiwin	x	x
Total	**16 (80%)**	**8 (40%)**
Calgary (5 sites)		
Foothills Hospital	✓	✓
High River General Hospital	✓	✓
Peter Lougheed Hospital	✓	✓
Rockyview General Hospital	✓	✓
South Health Campus	✓	✓
Total	**5 (100%)**	**5 (100%)**
South (8 sites)		
Brooks	✓	✓
Cardston	x	x
Crowsnest Pass	x	x
Lethbridge	✓	✓
Medicine Hat	✓	✓
Raymond	x	x
Taber	x	x
Pincher Creek	✓	✓
Total	**4 (50%)**	**4 (50%)**
All zones	**46 of 57 (81%)**	**34 of 57 (60%)**

x = unavailable; - = unknown data

SECTION FOUR: Economic Analysis

Anderson Chuck, PhD, MPH; Thanh Nguyen, MD, MPH, PhD

Objectives

To determine the impact of fFN testing on ambulance transfers, hospital admissions, hospital length of stay, and health system costs.

Methods

The impact of fFN testing on ambulance transfers, hospital admissions, and hospital length of stay is predicated on how it influences clinical decision-making beyond what would have occurred in the absence of that information. However, health system impact is not based on clinical decision-making alone, but is a function of a variety of factors. These include: the epidemiology surrounding the distribution between true labour and false labour in the Alberta population, the fFN test result, physician decision-making, and the costs of the associated health services.

Accordingly, the economic analysis consists of two parts. The first is to measure the degree to which clinical decision-making is impacted by fFN testing information. The second is to incorporate this information into a framework that accounts for underlying distribution between true labour and false labour in the AB population and costs of health services.

We used three administrative health databases: the Discharge Abstract Database (DAD), which contains records for services provided for patients admitted into hospital; The Ambulatory Care Classification System (ACCS), which contains records for services provided for patients at a hospital facility but who were not admitted; and the Practitioners Claims Databases (PCD), which contains physician billing information for insured medical services. Fields contained in both the DAD and ACCS include information pertaining to the recipient and provider, service provided, date of service, International Classification of Disease (ICD) diagnosis, procedure interventions classified by CCI coding, and costs. The DAD also contains data on the patient length of stay in hospital. Fields contained in the PCD include information pertaining to the recipient and provider, billing code, date of service and ICD diagnosis, and costs.

These databases were supplemented with data from the provincial Laboratory Information System (LIS) and the Alberta Perinatal Health Program (APHP). The LIS database provided fFN testing data for the corresponding cohort of pregnancies that received fFN testing. Specifically, we extracted a dataset of all fFN tests done by Millennium (for the Calgary zone), Meditech (for the North, South, and Central zones), and Sunquest (for the Edmonton, Central, and North zones) from January 2008 to October 2013 (the most recent time the data were available). The fields contained in the dataset include the date the test was performed, test result, test site, and zone. Laboratory services also provided data that identified the capacity of the listed AHS facilities to perform fFN testing (that is, test kits alone versus test kits plus analyzers). The APHP provided further clinical information related to pregnancy.

All health services utilization records from April 2002 to March 2013 with at least one relevant ICD code within the first three diagnostic fields (Table 4.1) were extracted. Records with a case mix group code of preterm labour (599) or false labour (619) were also extracted.

Table 4.1: List of ICD and CMG codes for data extraction

		Description	Note
ICD 9 code	**ICD 10 code**		
644.2	O60.X	Preterm delivery	Early onset of delivery in ICD9
644.0, 644.1	O47.X	False labour	Threatened labour in ICD9
658.3, 659.X, 669.X	O75.X	Other complications of labour and delivery, not elsewhere classified	
V22.0, V72.4	Z32.X	Pregnancy examination & test	
V22.2	Z33.X	Pregnant state, incidental	
V22.0, V22.1	Z34.X	Supervision of normal pregnancy	
V23.X	Z35.X	Supervision of high-risk pregnancy	
658.1, 658.2	O42.X	Premature rupture of membranes	
CMG code			
599		Preterm labour	Same code over time
619		False labour	Same code over time

Impact on Clinical Decision-Making

Ambulance transfers

We compared ambulance transfers between the cohort of preterm pregnancies (<37 weeks gestation) presenting with signs of labour that received fFN testing to those that did not receive fFN testing. The analysis was conducted at the episode level opposed to the patient level, given that there could be multiple episodes of preterm labour within a single course of pregnancy. Accordingly, we assessed the clinical utility of fFN testing for each episode of preterm labour. The date of fFN testing did not directly correspond to the date of service for a particular episode. fFN testing data was linked to an episode if the test date occurred within 48 hours of the corresponding episode service date. This created the cohort that received fFN testing. Episodes that did not have a corresponding fFN test represented the cohort that did not receive fFN testing.

Multilevel logistic regression models were used to determine the likelihood of ambulance transfer from a lower acuity facility (that is, levels 0, A, B, or C) to a higher acuity facility (that is, level D) following a positive or negative test result between the two study cohorts.[j] This was conducted separately for the DAD and ACCS data. Further, within each dataset, the analysis was conducted separately for the subgroup of preterm pregnancies that were eventually found to be in true labour or false labour for our tested and untested study cohorts (identified using the codes listed in Table

[j] Alberta Perinatal Program Definitions:
Level 0 = Acute care hospital without an elective obstetrical service
Level A = Low Risk obstetrics without cesarean section capability
Level B = Rural facility with cesarean section capability
Level C = Regional hospital with specialist i.e. obstetricians, pediatricians etc.
Level D = Calgary and Edmonton hospitals where all women/babies would have access to tertiary care.

4.1). This was conducted because the utility of fFN testing is predicated on facilitating clinical judgement. For instance, a negative test result for a preterm pregnancy that is in false labour provides useful information, but is useless/potentially harmful if the preterm pregnancy were in true labour. Hence we examine how fFN testing influences the clinical decision to transfer or not transfer to a higher acuity facility, and whether during that episode there was or was not, in fact, a delivery.

In the regression model, each episode is nested within the individual patient to control for random effects or individual differences. Other covariates included patient characteristics including age and co-morbidity, and hospital episode characteristics including service episode admitting category (urgent versus others), service episode admitting entry code (emergency, clinic, and others), method of admission (arrived at hospital via ambulance or not), characteristics of hospital (levels of care and zone where hospital is located), and time (before/after 2008, to control for variation over time). These covariates were chosen based on their potential associations with the dependent variable (that is, ambulance transfer) and their availability in the datasets. All analyses were conducted with Stata 13.1 (www.stata.com).

Hospital admission and length of stay

The analysis for hospital admission and hospital length of stay is identical to the analysis of ambulance transfers with the following exceptions. For hospital admissions the outcome of interest is the likelihood of being admitted and the dataset used was the ACCS. For hospital length of stay (count data) we employed a multilevel negative binomial regression to compare the average length of stay in hospital between the study cohorts. The dataset used was the DAD.

Impact on Health System Costs

We employed a decision analytic approach to estimate health system cost impacts to account for the epidemiology surrounding the distribution between true labour and false labour in the Alberta population, the fFN test result, physician decision-making, and the costs of the associated health services. The decision analytic models were populated with our data and results calculated from the multivariate regressions.

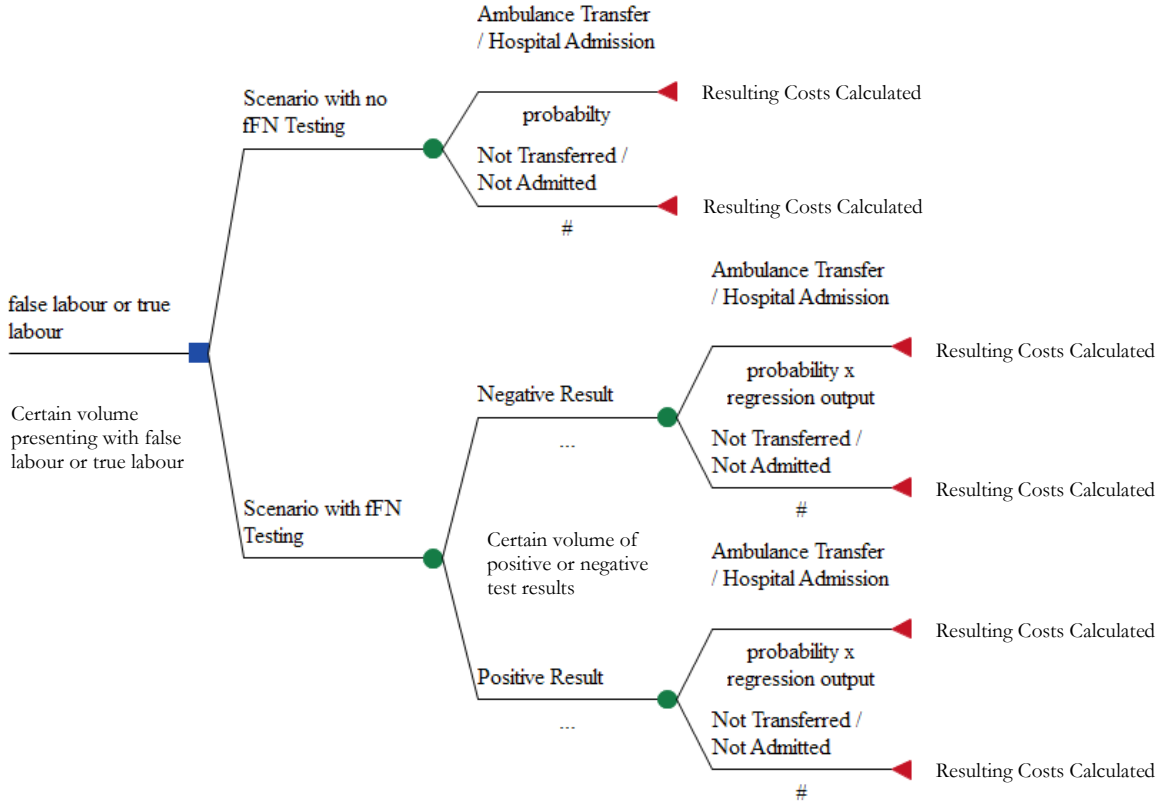
Figure 4.1 illustrates the general decision model. In the scenario with no fFN testing, there is baseline probability for physicians, in the absence of fFN testing information, to transfer the patient to a higher acuity facility or, alternatively, admit the patient into hospital. In the scenario with fFN testing, the baseline probability to transfer or admit is modified by the additional fFN testing information. The regression results provide the appropriate modifier. One-way sensitivity analysis was conducted on key model inputs to determine how ranges in model inputs impacted the cost calculation. Sensitivity analysis results are presenting in a tornado diagram. A tornado diagram ranks the variables that are associated with the greatest range (i.e. greatest variability) in results from highest to lowest. Hence the name due to the graph resembling a tornado.

Health system costs are calculated separately for ambulance transfer, and hospital admissions (including hospital length of stay) are conducted separately for episodes of true labour (considered appropriate utilization) and episodes of false labour (considered unnecessary utilization). Note that costs were not calculated for any impacts where testing reduces hospital admissions for episodes of true labour (that is, false negative result), because it is assumed that these episodes will be ultimately admitted in the immediate term (that is, admission costs are delayed but will be realized). The

specific decision trees and inputs used in the analysis (including sensitivity analysis) can be found in Appendix 4.A.

We estimated the cost impact of fFN testing from January 2008 (policy implementation) to March 2013 (most current data at time of analysis). All costs were converted to 2014 Canadian dollars using the Canadian Price Index. TreeAge Pro Suite 2014 (www.treeage.com) was used to conduct the analysis.

Figure 4.1: General decision analytic model



Results

Testing Volumes

In total, 15,042 tests were performed on 13,131 women between January 2008 and October 2013 (Table 4.2), with 65% of tests being conducted at a level D hospital. Testing volumes increased from the policy implementation date in 2008 to 2010, and then stabilized afterwards (note that volumes for 2013 did not reflect a full year). Similarly, the number of tests done at level D hospitals decreased from 80% in 2008 to about 60% between 2010 and 2013. Testing frequency was greatest in the Edmonton zone (43%), followed by the Calgary zone (25%), North zone (17%), Central zone (11%), and South zone (5%) (Table 4.3). Further descriptive results can be found in Appendices 4.C through G.

Table 4.2 Test results by year

Year		Test result					Total
		Positive	Negative	Invalid	Other	Missing	
2008	n	294	1,894	1	0	2	**2,191**
	%	13.42	86.44	0.05	0	0.09	**100**
2009	n	244	2,328	1	1	3	**2,577**
	%	9.47	90.34	0.04	0.04	0.12	**100**
2010	n	387	2,515	4	4	3	**2,913**
	%	13.29	86.34	0.14	0.14	0.1	**100**
2011	n	472	2,281	3	0	2	**2,758**
	%	17.11	82.7	0.11	0	0.07	**100**
2012	n	488	2,297	4	2	1	**2,792**
	%	17.48	82.27	0.14	0.07	0.04	**100**
2013*	n	317	1,492	1	1	0	**1,811**
	%	17.5	82.39	0.06	0.06	0	**100**
Total	**n**	**2,202**	**12,807**	**14**	**8**	**11**	**15,042**
	%	**14.64**	**85.14**	**0.09**	**0.05**	**0.07**	**100**

*Does not reflect a full year of data

Table 4.3: Number of tests by zone and year

Zone		2008	2009	2010	2011	2012	2013*	Total
Calgary	n	678	624	757	683	549	399	**3,690**
	%	30.94	24.21	25.99	24.76	19.66	22.03	**24.53**
Central	n	126	266	324	324	342	200	**1,582**
	%	5.75	10.32	11.12	11.75	12.25	11.04	**10.52**
Edmonton	n	1,094	1,143	1,237	1,045	1,173	777	**6,469**
	%	49.93	44.35	42.46	37.89	42.01	42.9	**43.01**
North	n	192	454	480	565	572	353	**2,616**
	%	8.76	17.62	16.48	20.49	20.49	19.49	**17.39**
South	n	101	90	115	141	156	82	**685**
	%	4.61	3.49	3.95	5.11	5.59	4.53	**4.55**
Total	**n**	**2,191**	**2,577**	**2,913**	**2,758**	**2,792**	**1,811**	**15,042**
	%	**100**	**100**	**100**	**100**	**100**	**100**	**100**

*Does not reflect a full year of data

Validity Checking

The analysis on ambulance transfers, hospital admissions, and hospital length of stay was conducted separately for the subgroup of preterm pregnancies that were eventually found to be in true labour

or false labour for our tested and untested study cohorts. We tested the validity of using the ICD coding to differentiate true labour (ICD code of preterm) versus false labour (ICD code of false labour) by cross tabulating these diagnostic codes with a separate code indicating "delivery" in the DAD. There were 39,369 total episodes of preterm and false labour in the DAD dataset between April 2002 and March 2013, which were accurate 85.9% and 99.9% of the time, respectively (Table 4.4). Of the episodes that received fFN testing in the DAD, 67% had positive results with a diagnosis of preterm, and 61% had a negative result with a diagnosis of false labour (Table 4.5). Of the episodes that received fFN testing in the ACCS, 32.1% had a positive result with a diagnosis of preterm, and 89.3% had a negative result with a diagnosis of false labour (Table 4.6).

Table 4.4: Main service code of hospital episodes by diagnosis

Diagnosis		Main service code		Total
		Other	Delivery	
Preterm	n	4,673	28,368	**33,041**
	%	14.14	85.86	**100**
False labour	n	6,326	2	**6,328**
	%	99.97	0.03	**100**
Total	**n**	**10,999**	**28,370**	**39,369**
	%	**27.94**	**72.06**	**100**

Table 4.5: Test results by diagnosis in the inpatient dataset (DAD)

Diagnosis		Test result					Total
		Positive	Negative	Invalid	Other	Missing	
Preterm	n	248	117	0	2	2	**369**
	%	67.21	31.71	0	0.54	0.54	**100**
False labour	n	888	1,423	2	1	2	**2,316**
	%	38.34	61.44	0.09	0.04	0.09	**100**
Total	**n**	**1,136**	**1,540**	**2**	**3**	**4**	**2,685**
	%	**42.31**	**57.36**	**0.07**	**0.11**	**0.15**	**100**

Table 4.6: Test results by diagnosis in the outpatient dataset (ACCS)

Diagnosis		Test result		Total
		Positive	Negative	
Preterm	n	373	788	**1,161**
	%	32.13	67.87	**100**
False labour	n	177	1,473	**1,650**
	%	10.73	89.27	**100**
Total	**n**	**550**	**2,261**	**2,811**
	%	**19.57**	**80.43**	**100**

Ambulance transfers

Episodes of true labour

There were 7,044 and 4,617 episodes presenting at facilities graded less than D that were determined to be true labour in the DAD and ACCS, respectively. Table 4.7a shows the proportion of these that were transferred to a higher acuity facility for those that were not tested and for those that received a positive or negative test result.

In the DAD and ACCS (only statistically significant results at $p<.05$ are reported), compared to the non-tested study cohort, the likelihood of an ambulance transfer to a higher acuity facility was more likely following a positive (7.45 and 3.68 times greater, respectively) than a negative (1.91 and 1.26 times greater, respectively) test result (Table 4.7b). When comparing the results among the zones in the DAD, ambulance transfers were more likely to occur in the Edmonton and North zones, compared to the Calgary zone. In the ACCS, ambulance transfers were less likely to occur in the North, South, and Central zones and more likely in the Edmonton zone, compared to the Calgary zone.

Table 4.7a: Number of ambulance transfers for tested/non-tested cohorts (true labour)

	Transferred	Not transferred	Total
DAD			
N	6,596	448	**7,044**
Test results:			
Negative (%)	2.43	9.82	**2.90**
Positive (%)	1.43	10.49	**2.00**
Non-tested (%)	96.15	79.69	**95.10**
ACCS			
N	3,541	1,076	**4,617**
Test results:			
Negative (%)	13.27	14.41	**13.54**
Positive (%)	4.97	10.13	**6.17**
Non-tested (%)	81.76	75.46	**80.29**

Table 4.7b: Likelihood of ambulance transfer (true labour)

Independent variables	Odds ratio	P-value	95% confidence interval	
			Lower	Upper
DAD[a]				
Test results: Non-tested as reference				
Negative	1.91	0.020	1.11	3.29
Positive	7.45	0.000	3.88	14.30
Zones: Calgary as reference				
North	2.25	0.019	1.15	4.40

South	1.33	0.429	0.66	2.67
Central	0.95	0.883	0.48	1.88
Edmonton	10.25	0.000	3.14	33.48
ACCS[b]				
Test results: Non-tested as reference				
Negative	1.26	0.089	0.96	1.66
Positive	3.68	0.000	2.55	5.32
Zones: Calgary as reference				
North	0.19	0.000	0.14	0.26
South	0.07	0.000	0.04	0.11
Central	0.19	0.000	0.14	0.26
Edmonton	2.36	0.000	1.59	3.50

a. Number of episodes = 7,044; Number of women = 6,347; Wald chi2 (12) = 100.55; P = 0.000. Controlled for age, year, admitting category, admitting entry code, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

b. Number of episodes = 6,205; Number of women = 5,014; Wald chi2 (11) = 238.61; P = 0.000. Controlled for age, year, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

Episodes of false labour

There were 3,523 and 2,671 episodes presenting at facilities graded less than D that were determined to be false labour in the DAD and ACCS respectively. Table 4.8a shows the proportion of these that were transferred to a higher acuity facility for those that were not tested and for those that received a positive or negative test result.

In the DAD and ACCS (only statistically significant results at p <.05 are reported), compared to the non-tested study cohort the likelihood of an ambulance transfer to a higher acuity facility was greater following a positive test result (2.22 and 10.81 times, respectively) (Table 4.8b). When comparing the results among the zones in the DAD, ambulance transfers were more likely to occur in the North, South, and Central zones, compared to the Calgary zone. In the ACCS, ambulance transfers were more likely to occur in the South and Edmonton zone, compared to the Calgary zone.

Table 4.8a: Number of ambulance transfers for tested/non-tested cohorts (false labour)

	Transferred	Not transferred	Total
Inpatients			
N	497	3,026	3,523
Test results:			
Negative	10.26%	9.62%	9.71%
Positive	10.66%	4.10%	5.02%
Non-tested	79.07%	86.29%	85.27%

Outpatients			
N	93	2,578	2,671
			Test results:
Negative	17.20%	13.38%	13.52%
Positive	9.68%	1.86%	2.13%
Non-tested	73.12%	84.76%	84.35%

Table 4.8b: Likelihood of ambulance transfers (false labour)

Independent variables	Odds ratio	P-value	95% confidence interval	
			Lower	Upper
DAD[a]				
Test results: Non-tested as reference				
Negative	0.78	0.247	0.51	1.19
Positive	2.22	0.001	1.38	3.57
Zones: Calgary as reference				
North	5.98	0.000	2.76	12.96
South	5.03	0.000	2.24	11.32
Central	3.67	0.001	1.69	7.98
Edmonton	4.70	0.062	0.92	23.95
ACCS[b]				
Test results: Non-tested as reference				
Negative	1.53	0.190	0.81	2.88
Positive	10.81	0.000	3.96	29.51
Zones: Calgary as reference				
North	0.70	0.062	0.48	1.02
South	0.03	0.000	0.01	0.11
Central	0.68	0.062	0.46	1.02
Edmonton	3.62	0.000	2.28	5.77

a. Number of episodes = 3,523; number of women = 2,961; Wald chi2 (14) = 91.09; P = 0.000. Controlled for age, year, admitting category, admitting entry code, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for full regression model.

b. Number of episodes = 6,173; number of women = 4,912; Wald chi2 (12) = 164.70; P = 0.000. Controlled for age, year, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

Hospital admissions

Episodes of true labour

There were 7,626 episodes that were determined to be true labour in the ACCS. Table 4.9a shows the proportion of these that were admitted into hospital for those that were not tested, and for those that received a positive or negative test result. Compared to the non-tested study cohort (only

statistically significant results at p<.05 are reported), the likelihood of hospital admission was greater following a positive test result (1.68 times), and lower following a negative test result (0.44 times) (Table 4.9b). When comparing the results among the zones hospital admission was more likely to occur in the North, Central and Edmonton zones, compared to the Calgary zone.

Table 4.9a: Hospital admissions for tested/non-tested cohorts (true labour)

	Not admitted	Admitted	Total
N	5,085	2,541	**7,626**
Test results:			
Positive (%)	4.58	5.47	**4.88**
Negative (%)	11.98	7.04	**10.33**
Non-tested (%)	83.44	87.49	**84.79**

Table 4.9b: Likelihood of hospital admissions (true labour)

Independent variables	Odds ratio	P-value	95% confidence interval	
			Lower	Upper
Test results: Non-tested as reference				
Positive	1.68	0.002	1.22	2.32
Negative	0.44	0.000	0.35	0.56
Zones: Calgary as reference				
North	4.59	0.000	3.29	6.38
South	1.41	0.066	0.98	2.02
Central	2.30	0.000	1.67	3.18
Edmonton	3.47	0.000	2.48	4.85

Number of episodes = 7,626; number of women = 6,218; Wald chi2 (8) = 265.17; P = 0.000. Controlled for age, year, arrived via ambulance, diagnosis, comorbidity, and transfer-up. Refer to Appendix 4.B for further details.

Episodes of false labour

There were 19,666 episodes that were determined to be false labour in the ACCS. Table 4.10a shows the proportion of these that were admitted into hospital for those that were not tested and for those that received a positive or negative test result. Compared to the non-tested study cohort (only statistically significant results at p<.05 are reported), the likelihood of hospital admission was greater following a positive test result (5.38 times), and lower following a negative test result (0.47 times) (Table 4.10b). When comparing the results among the zones, hospital admission was more likely to occur in the North, South, Central, and Edmonton zones, compared to the Calgary zone.

Table 4.10a: Hospital admissions for tested/non-tested cohorts (false labour)

Variables	Not admitted	Admitted	Total
N	16042	3624	**19,666**
Test results:			
Positive	0.71%	1.68%	**0.89%**
Negative	8.64%	2.4%	**7.49%**
Non-tested	90.65%	95.92%	**91.62%**

Table 4.10b: Likelihood of hospital admissions (false labour)

Independent variables	Odds ratio	P-value	95% confidence interval	
			Lower	Upper
Test results: Non-tested as reference				
Positive	5.38	0.000	3.65	7.95
Negative	0.47	0.000	0.37	0.60
Zones: Calgary as reference				
North	3.11	0.000	2.15	4.52
South	4.12	0.000	2.63	6.45
Central	2.15	0.000	1.45	3.18
Edmonton	4.69	0.000	3.28	6.71

Number of episodes = 19,630; Number of women = 13,902; Wald chi2 (11) = 606.28; P = 0.000. Controlled for age, year, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

Hospital length of stay

Episodes of true labour

There were 33,019 episodes that were determined to be true labour in the DAD. Table 4.11a shows the mean length of stay for those that were not tested, and for those that received a positive or negative test result. Compared to the non-tested study cohort (only statistically significant results at $p<.05$ are reported), the mean length of stay was longer following a positive (1.34 times) and negative test result (1.13 times) (Table 4.11b). When comparing the results among the zones, hospital length of stay was longer in the South and Edmonton zones and shorter in the Central zone, compared to the Calgary zone.

Table 4.11a: Hospital length of stay for tested/non-tested cohorts (true labour)

Variables	N (%)	Mean (days)	Median (days)
All samples	33,019 (100%)	3.84	2
Test results:			
Negative	925 (3%)	3.89	2
Positive	731 (2%)	4.85	2
Non-tested	31,363 (95%)	3.81	2

Table 4.11b: Hospital length of stay incidence rate ratio (true labour)

Independent variables	IRR	P-value	95% confidence interval	
			Lower	Upper
Test results: non-tested as reference				
Positive	1.34	0.000	1.25	1.43
Negative	1.13	0.000	1.06	1.20
Zones: Calgary as reference				
North	0.98	0.481	0.94	1.03
South	1.21	0.000	1.15	1.27
Central	0.91	0.000	0.87	0.96
Edmonton	1.23	0.000	1.20	1.26

IRR = incidence rate ratio; Number of episodes = 33,019; number of women = 28,548; Wald chi2 (15) = 2479.51; P = 0.000. Controlled for age, year, admitting category, admitting entry code, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

Episodes of false labour

There were 6,312 episodes that were determined to be false labour in the DAD. Table 4.12a shows the mean length of stay for those that were not tested, and for those that received a positive or negative test result. Compared to the non-tested study cohort (only statistically significant results at p<.05 are reported), the mean length of stay was longer following a positive (1.20 times) test result (Table 4.12b). When comparing the results among the zones, hospital length of stay was longer in the Edmonton zones and shorter in the North, South, and Central zone, compared to the Calgary zone.

Table 4.12a: Hospital length of stay for tested/non-tested cohorts (false labour)

Variables	N (%)	Mean (days)	Median (days)
All sample	6312 (100%)	2.14	1
Test results:			
Negative	480 (8%)	1.88	1
Positive	313 (5%)	2.37	2
Non-tested	5519 (87%)	2.15	1

Table 4.12b: Hospital length of stay incidence rate ratio (false labour)

Independent variables	IRR	P-value	95% confidence interval	
			Lower	Upper
Test results: non-tested as reference				
Negative	1.01	0.885	0.91	1.11
Positive	1.20	0.002	1.07	1.34
Zones: Calgary as reference				
North	0.89	0.020	0.81	0.98

South	0.81	0.001	0.72	0.92
Central	0.83	0.000	0.75	0.91
Edmonton	1.21	0.000	1.13	1.30

IRR = incidence rate ratio; Number of episodes = 6,312; number of women = 5,437; Wald chi2 (15) = 599.43; P = 0.000. Controlled for age, year, admitting category, admitting entry code, arrived via ambulance, and comorbidity. Refer to Appendix 4.B for further details.

Health System Costs

Episodes of false labour (inappropriate resource utilization)

Ambulance transfers

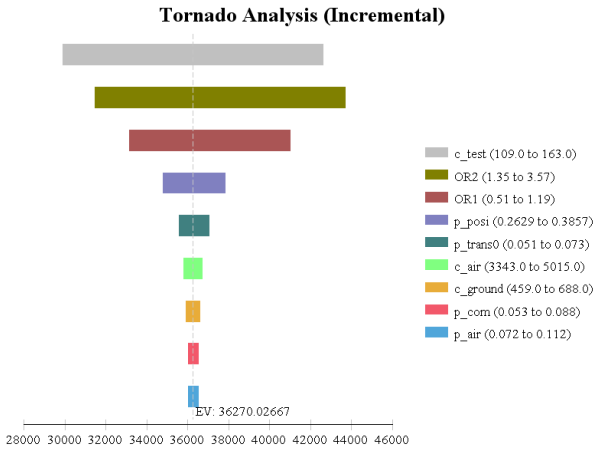
There were an estimated 54 unnecessary ambulance transfers associated with fFN testing between 2008 and 2013 (that is, false PTL that was transferred despite having a negative fFN test result). Had this group not been tested, the number would have decreased by 27, for a cost increase of $120,976 (Table 4.13). The variation in costs is shown in Figure 4.2. Note that a tornado diagram ranks the variables that are associated with the greatest range (that is, greatest variability) in results from highest to lowest; hence the name, due to the graph resembling a tornado.

Table 4.13: Impact of fFN testing on ambulance costs among episodes of false labour

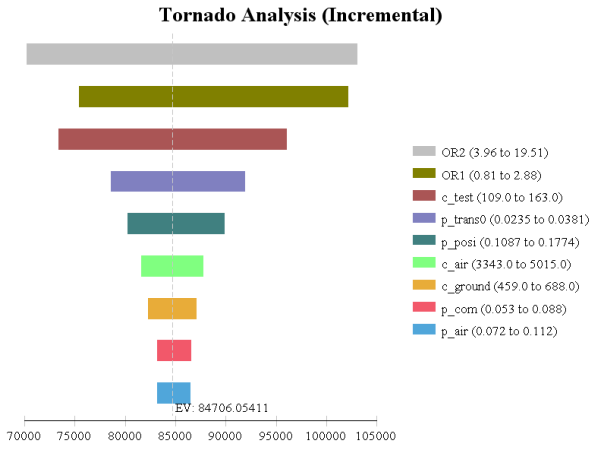
Variable	Had tested cohort not been tested	Tested cohort (includes test cost)	Difference
DAD			
Cost	$17,129	$53,399 (n=236)	$36,270 ($29,898-$47,027)
Number of transfer-up	14	18	4 (-1 to 10)
ACCS			
Cost	$15,043	$99,749 (n=420)	$84,706 ($70,229-$103,094)
Number of transfer-up	13	36	23 (11 to 39)

Figure 4.2: Tornado diagrams of ambulance costs among episodes of false labour

DAD



ACCS



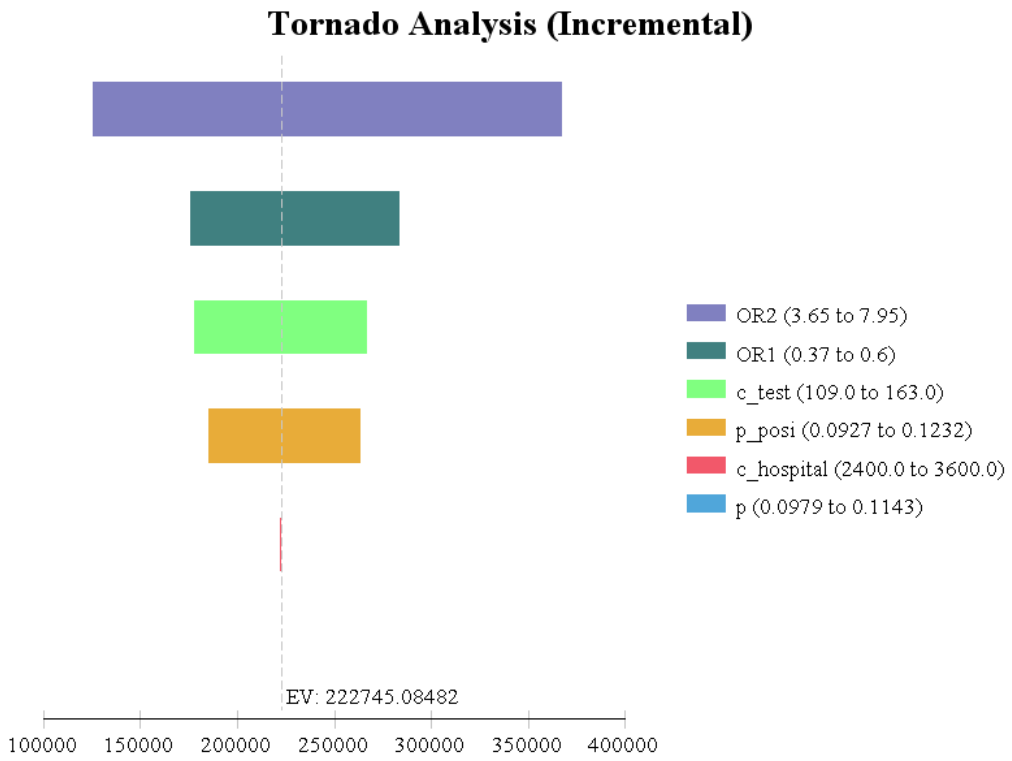
Hospital admissions

There were an estimated 174 unnecessary hospital admissions with fFN testing between 2008 and 2013. Had this group not been tested, there would have been one less hospital admission. However, there is still a cost increase of $222,745, due to the costs of testing (Table 4.14). The variation in costs is shown in Figure 4.3.

Table 4.14: Impact of fFN testing on hospital admission costs among episodes of false labour

Variable	Had tested cohort not been tested	Tested cohort (n=1,650) (includes test cost)	Difference (range)
Cost	$524,205	$746,950	$222,745 ($125,437-$367,300)
Number of admissions	175	174	-1 (-33 to 48)

Figure 4.3: Tornado diagram of hospital admission costs among episodes of false labour



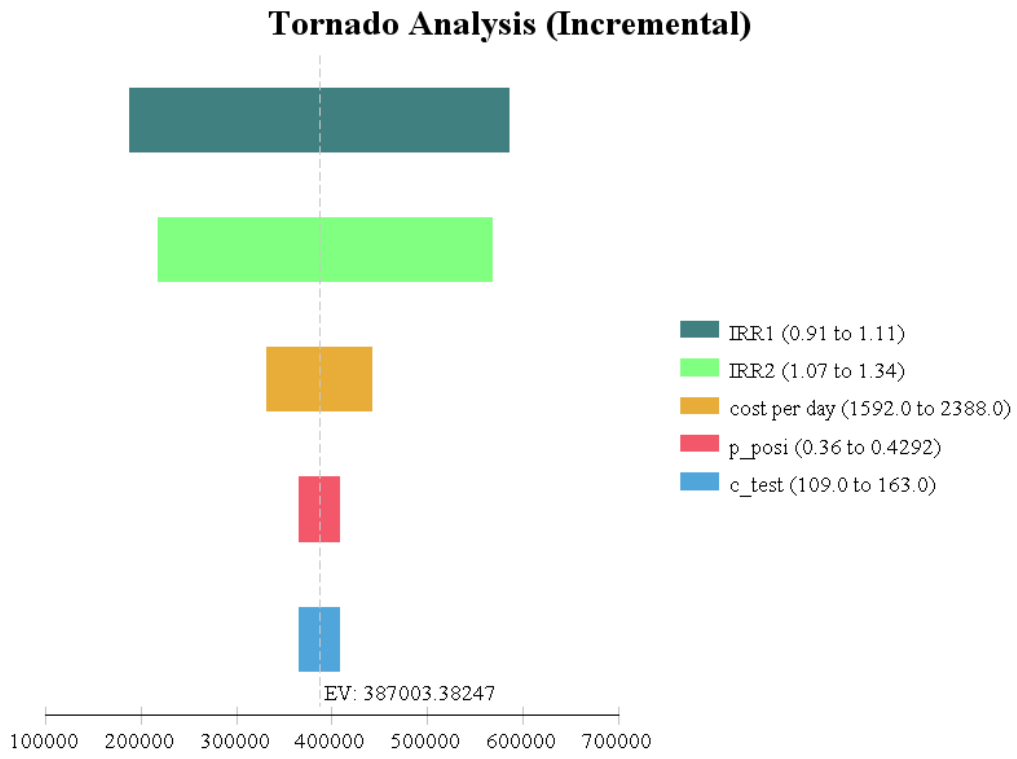
Hospital length of stay

There were an estimated 1,794 hospital days associated with fFN testing between 2008 and 2013. Had this group not been tested, the total number of hospital days would have been decreased by 143. This corresponds to a cost increase of $387,003 (Table 4.15). The variation in costs is shown in Figure 4.4.

Table 4.15: Impact of fFN testing on hospital length of stay costs among episodes of false labour

Variable	Had tested cohort not been tested	Tested cohort (n=794) (includes test cost)	Difference (range)
Cost	$3,286,525	$3,673,528	$387,003 ($187,906-$586,101)
Number of days	1,652	1,794	143

Figure 4.4: Tornado diagrams of hospital length of stay cost among episodes of false labour



Episodes of true labour (appropriate resource utilization)

Ambulance transfers

There were an estimated 494 necessary ambulance transfers associated with fFN testing between 2008 and 2013. Had this group not been tested, the number would have decreased by 272, for a cost increase of $496,640 (Table 4.16). The variation in costs is shown in Figure 4.5.

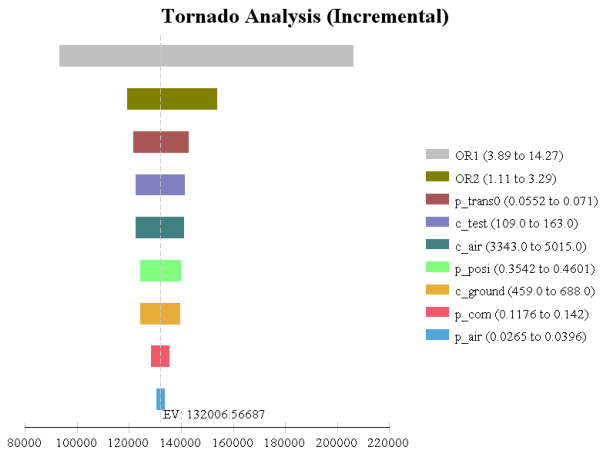
Table 4.16: Impact of fFN testing on ambulance transfer costs among episodes of true labour

Variable	Had tested cohort not been tested	Tested cohort (includes test cost)	Difference (range)
DAD			
Cost	$26,832	$158,839 (n=347)	$132,007 ($93,195-$206,358)
Number of transfer-up	22	91	69 (37 to 129)

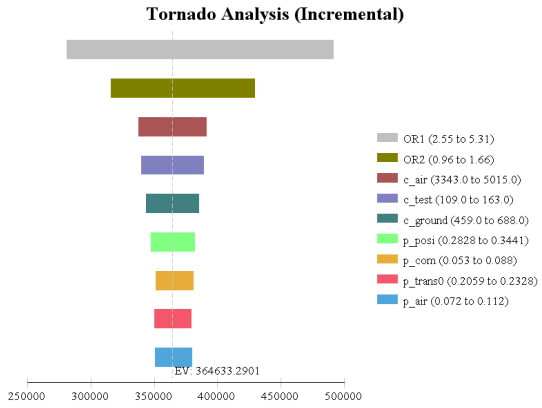
ACCS			
Cost	$236,719	$601,352 (n=911)	$364,633 ($280,962-$491,661)
Number of transfer-up	200	403	203 (132 to 305)

Figure 4.5: Tornado diagrams of ambulance transfer costs among episodes of true labour

DAD



ACCS



Hospital length of stay

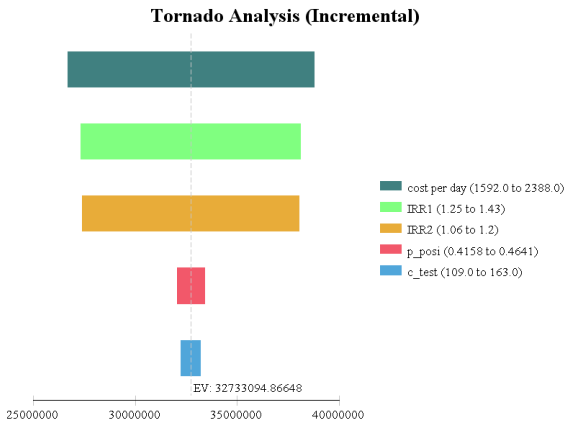
There were an estimated 7,562 hospital days associated with fFN testing between 2008 and 2013. Had this group not been tested, the total number of hospital days would have decreased by 1,379. This corresponds to a cost increase of $2,961,803 (Table 4.17). The variation in costs is shown in Figure 4.6.

Table 4.17: Impact of fFN testing on hospital length of stay costs among episodes of true labour

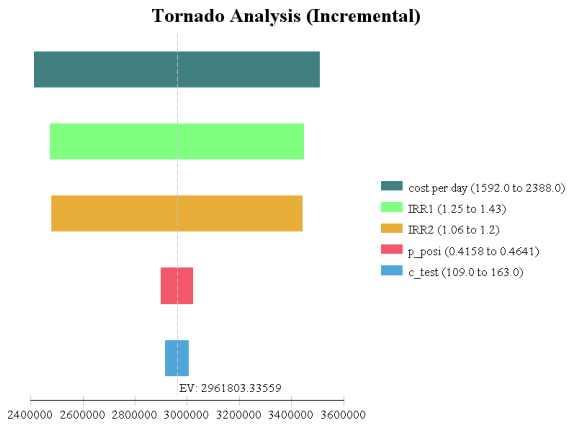
Variable	Had tested cohort not been tested	Tested cohort (n=1,662) (includes test cost)	Difference (range)
Cost	$12,303,454	$15,265,257	$2,961,803 ($2,414,649-$3,508,958)
Number of days	6,183	7,562	1,379

Figure 6: Tornado diagrams of hospital length of stay cost among episodes of true labour

If 100% tested



Current practice (9% tested)



Summary of estimated health system impact of fFN testing

Table 4.18: Summary of estimated health system impact of fFN testing

Outcomes	Rate if not tested	Impact of test*		Health system cost impact		
		Positive	Negative	Base	Low[a]	High[a]
False labour (unnecessary health service utilization including costs of testing)						
Ambulance transfer						
Inpatient dataset	6.12%	2.22	0.78**	$36,270	$29,898	$47,027
Outpatient dataset	3.02%	10.81	1.53**	$84,706	-$70,229	$103,094
Subtotal				**$120,976**	**-$40,331**	**$150,121**
Hospital admission	10.59%	5.38	0.47	$222,745	$125,437	$367,300
Length of hospital stay	2.08 days	1.2	1.01**	$387,003	$187,906	$586,101
Total				**$730,724**	**$273,012**	**$1,103,522**
True labour (appropriate health service utilization including costs of testing)						
Ambulance transfer						
Inpatient dataset	6.28%	7.45	1.91	$132,007	$93,195	$206,358
Outpatient dataset	21.91%	3.68	1.26**	$364,633	-$280,962	$491,661
Subtotal				**$496,640**	**-$187767**	**$698,019**
Hospital admission***	35.87%	1.68	0.44			
Length of hospital stay	3.72 days	1.34	1.13	$2,961,803	$2,414,649	$3,508,958
Total				**$2,961,803**	**$2,226,882**	**$3,508,958**

All labour (total health system impact)						
Ambulance transfer						
Inpatient dataset				$168,277	$123,093	$253,385
Outpatient dataset				$449,339	-$351,191	$594,755
Subtotal				**$617,616**	**-$228,098**	**$848,140**
Hospital admission				$222,745	$125,437	$367,300
Length of hospital stay				$3,348,806	$2,602,555	$4,095,059
Total				**$4,189,167**	**$2,499,894**	**$5,310,499**

[a]Calculated from sensitivity analysis

*Measured by odd ratios or incidence rate ratios in comparison with non-tested.

**Not statistically significance at p<.05. Non-statistically significant values calculated because they are assumed to represent the best unbiased estimator of impact.

***Not calculated because true labour will ultimate be admitted in the immediate term.

Discussion

The clinical and economic burden associated with spontaneous preterm labour is significant. Preterm birth has been associated with 60% to 80% of deaths in infants without congenital anomalies, it accounts for up to 75% of neonatal morbidity, and it contributes to neuro-developmental problems, respiratory/pulmonary dysfunction, hearing and visual impairment, and other long-term health problems (refer to the Literature Review section [Section Two] of this report).

The goal of clinical management is to differentiate between true (that is, will result in delivery) and false preterm labour so appropriate care can occur while minimizing the use of unnecessary services. fFN testing was adopted in Alberta between 2006 and 2008 to help clinicians rule out false preterm labour. Between January 2008 and October 2013, approximately 15,000 fFN tests were performed in Alberta.

Key Findings

The economic analysis evaluates how fFN testing influences clinical decision-making beyond what would have occurred in its absence, and quantifies the resulting health system cost impact. The results indicate that the additional information provided by fFN testing does influence clinical decision-making. However, physicians placed greater significance (that is, will decide to transfer to a higher acuity facility or admit into hospital) following a positive test result, compared to a negative test result (that is, confirmatory bias).

This is a noteworthy finding, because the utility of fFN testing is based on enabling physicians to better rule out false preterm labour (that is, high specificity) and not on the ability to identify cases of true preterm labour, as the sensitivity of fFN testing is low (60% in the literature, using a cut-off of delivery occurring within 1 week). The real world sensitivity of fFN testing is even lower at 16.2%, according to the administrative data (refer to Appendix 4.I), adding further credence that physicians should not be placing greater significance on a positive test result (**Key Finding #1 – confirmatory bias despite poor sensitivity**). We also determined that, although the real world specificity was very high at approximately 98% (using a cut-off of delivery occurring within 1 week),

the likelihood of an ambulance transfer to a higher acuity centre was greater following a negative test result, compared to episodes that were not tested at all.

There are several possible explanations for these results. The first is that there is significant risk to the mother and infant if a case of true spontaneous preterm labour is misdiagnosed, and thus there would be an inherent tendency to err on the side of caution (personal communication, fFN working group, October 29, 2014) and be influenced more greatly by a positive test result. Further to this point, clinicians do not consider the results of a positive fFN test alone, but consider other clinical information (for example, cervical length, bleeding, cramps); it is the combination of these factors that leads to the decision to act. The second explanation is that there are other factors in addition to the fFN test result that are being considered simultaneously. The third is that there was poor knowledge translation and an ill-defined implementation plan across RHAs.

One factor that the data suggests plays a significant role is geographic distance to a level D hospital (facility with full obstetrical services and access to tertiary care). When examining hospital admissions among episodes that were found to be in false preterm labour, the Calgary zone consistently outperformed the Edmonton zone by having a lower admission rate following a negative fFN test result. The data revealed that the Calgary zone had geographic catchment area that was predominantly localized within its own zone (92% within Calgary), whereas Edmonton had a larger geographic catchment area (75% within Edmonton) providing service to the North (15%) and Central (10%) zones (Appendix 4.D). Due to the greater difficulty in this subpopulation to re-access a level D hospital and the associated risks if they are mistakenly sent home, the admission rates overall would be expected to be higher in the Edmonton zone, compared to the Calgary zone (**Key Finding #2 – geographic distance to accessing level D facility may outweigh fFN negative findings**). It should be noted however that differences in obstetric criteria to approve a request for transfer will also impact differences between zones.

Altogether, the resulting impact of fFN testing to the health system was that there was a cost increase to the system. This increase, however, was not only attributable to the costs associated with testing, but also to the costs associated with the corresponding increases in both appropriate and unnecessary resource use following the test result (**Key Finding #3 – costs increase due to increases in both appropriate and unnecessary health care utilization**).

As discussed above, fFN testing did not prevent ambulance transfers of episodes that were found to be false labour, but in fact increased them (that is, more likely to transfer despite a negative test result). Although the overall rate of hospital admissions were lower for these episodes following a negative test (decreased by one admission), the decrease was not enough to offset the cost of testing. Consequently, the estimated health system cost impact attributable to unnecessary resource utilization for episodes of false labour was approximately $730,724 ($273,012-$1,103,522).

There was also a greater likelihood of ambulance transfer and a longer length of hospital stay following both a positive and negative test result for episodes that were found to be true labour. Although the overall rate of hospital admissions were lower for these episodes following a negative test (decreased by one admission), the decrease was not enough to offset the cost of testing. Consequently, the estimated health system cost impact attributable to appropriate care for episodes of true labour was approximately $2,961,803 ($2,226,882-$3,508,958). However, as stated above, the utility of fFN testing was based on a negative test result opposed to a positive result. Hence this cost increase was not an anticipated consequence (**Key Finding #4 – increase in appropriate resource utilization was an unintended outcome**). Note that the rate of preterm births being delivered in

facilities with < D level acuity decreased by approximately 2% post-implementation (21% versus 19%; refer to Appendix 4.K).

Adoption of fFN testing was to better allow physicians to rule in episodes of false preterm labour, thereby improving clinical management while leading to cost savings to the system from the reduction in unnecessary care. As it turns out, the magnitude of potential cost savings was not significant to begin with (**Key Finding #5 – potential maximum cost savings were small at the outset**). In the absence of fFN testing, the probability of ambulance transfers for episodes of false preterm labour was only 3% to 6%. Hence even at a specificity of 98%, and assuming that physicians acted accordingly, there still would be no net cost savings from the reduction of ambulance transfers after accounting for the cost of testing, but rather a cost increase of $76,100 (see Appendix 4.J). We also determined that there was no significant change in average hospital length of stay, and so, when accounting for the cost of testing, there is a cost increase of $153,338 (note that the evidence assessment conducted at the time of the policy decision anticipated a reduction in hospital length of stay).

With hospital admissions, the probability of hospital admission in the absence of testing was approximately 10%, and with the high costs associated with an episode of admission, there is a potential cost saving of $243,401 (at a specificity of 98% and physicians acting accordingly). Altogether, this means that the maximum cost saving that could have been observed between January 2008 and October 2013 was approximately $13,963. Instead, what was observed to happen was an increase in total health system costs of approximately $4,189,167 ($2,499,894-$5,310,499), of which increased appropriate (but unanticipated) resource utilization accounted for 71% of the costs.

Caveats

The key findings need to be examined in light of the following caveats:

1. The analysis is retrospective in design. Consequently, we cannot interpret the findings as being absolutely causal, as there is no means of accounting for all potential confounding factors.
2. The analysis relies on the accuracy of the ICD codes to differentiate between true and false preterm labour. We found the accuracy of the ICD codes of preterm and false labour to be 85.9% and 99.9%, respectively.
3. Although there were approximately 15,000 tests conducted, only 30% could not be linked to the health databases. Hence the clinical outcome and health system impact of 4,500 tests are unknown.
4. It is acknowledged that there may be other clinical reasons that justify an ambulance transfer or hospital admission that superseded a negative test result. Hence episodes that have been characterized as "unnecessary" may be misclassified.
5. The regression analyses were conducted at the episode and not patient level, because multiple episodes of preterm labour could occur during a course of pregnancy. The algorithm for linking these episodes to their corresponding subsequent health service utilization and outcomes may result in some degree of misclassification.
6. The system cost impact of fFN testing does not include:
    a) potential cost or benefits associated with differences in health outcomes resulting from the increase in preterm births being delivered in a level D facility, compared to

had they been delivered in a <D facility. Potential benefits are considered minimal, as the rate of preterm births being delivered in facilities with <D level acuity only decreased by approximately 2%.

b) the costs or benefits associated with false negative test results resulting in preterm pregnancies delivery being delivered in facilities with <D level acuity.

Appendix 4.A: Decision Trees and Model Inputs

Figure 4.A.1: Decision tree comparing ambulance transfer costs between tested and non-tested cohorts among episodes of false labour

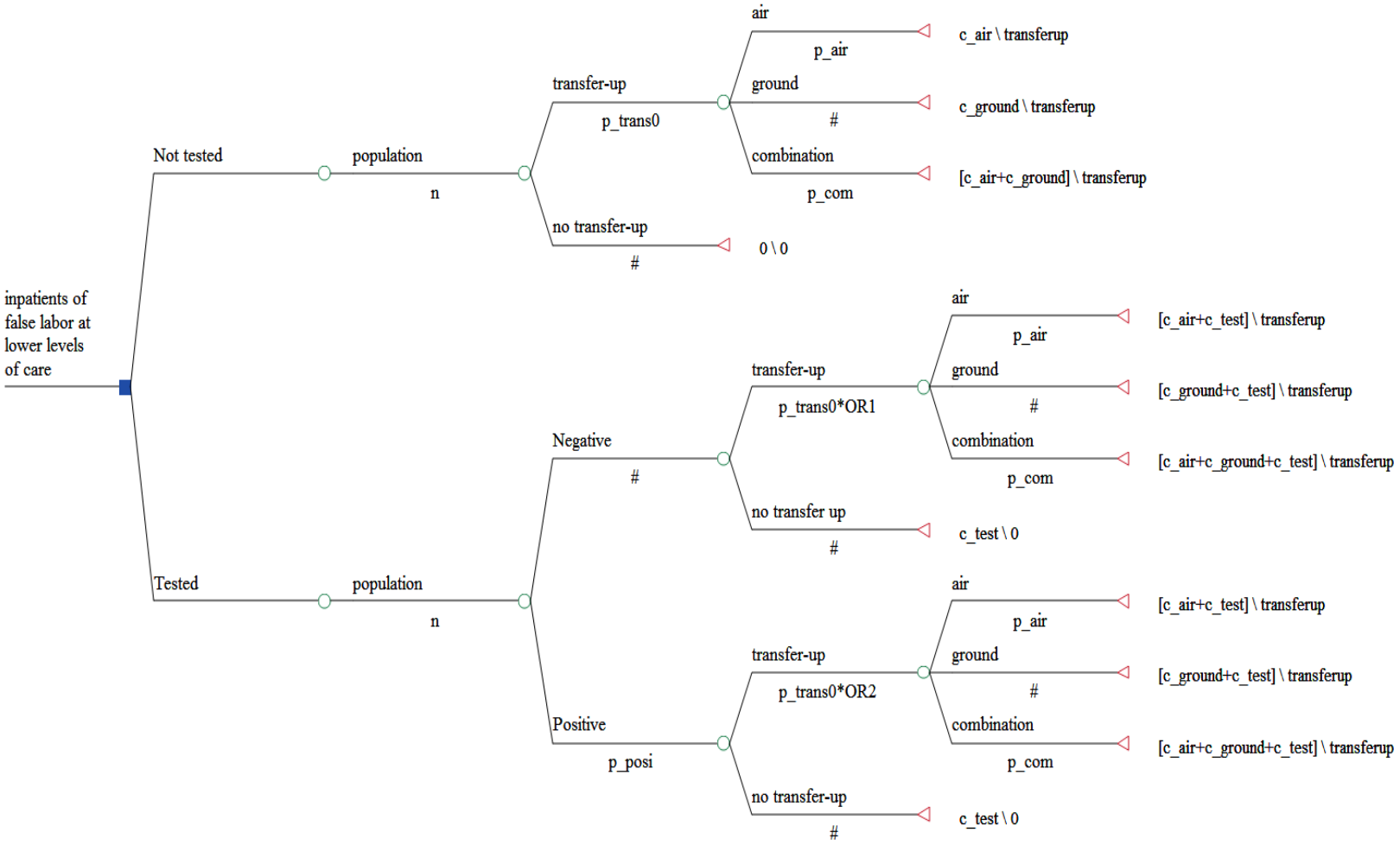


Table 4.A.1a: Inputs for decision model comparing ambulance transfer costs between tested and non-tested cohorts among episodes of false labour (DAD)

Variable	Description	Value (95%CI)	Sources
n	Number of hospital episodes with a diagnosis of false labour < 37 weeks of gestation at lower level of care hospitals in Alberta from January 2008 to March 2013	813	calculated from databases
n1	Number of n tested	236	calculated from databases
p_posi	Probability of positive results among tested	0.322 (0.2629-0.3857)	95% CI
p_trans0	Probability of transfer-up if not tested	0.0612 (0.051-0.073)	95% CI
OR1	Odd ratio between negative and not tested	0.78 (0.51-1.19)	95% CI
OR2	Odd ratio between positive and not tested	2.22 (1.38-3.57)	95% CI
p_air	Probability of air transfer	0.091 (0.072-0.112)	95% CI
p_com	Probability of combined (ground + air) transfer	0.068 (0.053-0.088)	95% CI
p_ground	Probability of ground transfer	0.841	= 1-(p_air + p_com)
c_air	Cost of air transfer-up in 2014 CA$	$4,179 ($3,343-$5,015)	Currie 2006 (±20%)
c_ground	Cost of ground transfer-up in 2014 CA$	$573 ($459-$688)	Currie 2006 (±20%)
c_test	Cost of fFN test	$136 ($109-$163)	ProvLab (±20%)

Table 4.A.1b: Inputs for decision model comparing ambulance transfer costs between tested and non-tested cohorts among episodes of false labour (ACCS)

Variable	Description	Value (range)	Sources
n	Number of outpatient visits with a diagnosis of false labour <37 weeks of gestation at lower levels of care in Alberta from January 2008 to March 2013	2,671	calculated from databases
n1	Number of n tested	420	calculated from databases
p_posi	Probability of positive results among tested	0.1405 (0.1087-0.1774)	95% CI
p_trans0	Probability of transfer-up if not tested	0.0302 (0.0235-0.0381)	95% CI
OR1	Odd ratio between negative and not tested	1.53 (0.81-2.88)	95% CI
OR2	Odd ratio between positive and not tested	10.81 (3.96-19.51)	95% CI
p_air	Probability of air transfer	0.0326 (0.0265-0.0396)	95% CI
p_com	Probability of combined (ground + air) transfer	0.1294 (0.1176-0.1420)	95% CI
p_ground	Probability of ground transfer	0.838	= 1-(p_air + p_com)
c_air	Cost of air transfer-up in 2014 CA$	$4,179 ($3,343-$5,015)	Currie 2006 (±20%)
c_ground	Cost of ground transfer-up in 2014 CA$	$573 ($459-$688)	Currie 2006 (±20%)
c_test	Cost of fFN test	$136	ProvLab (±20%)

Figure 4.A.2: Decision tree comparing hospital admission costs between tested and non-tested cohorts among episodes of false labour

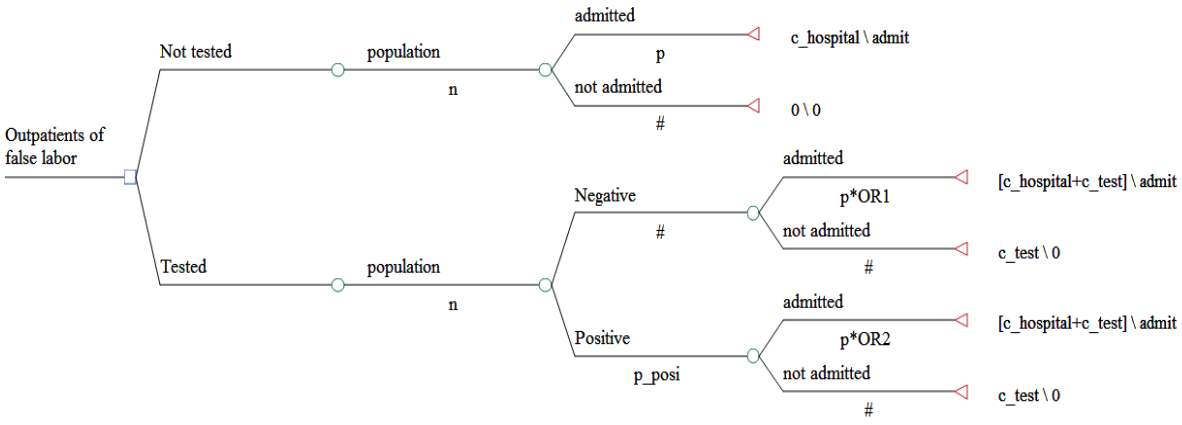


Table 4.A.2: Inputs for decision model comparing hospital admission costs between tested and non-tested cohorts among episodes of false labour

Variable	Description	Value (95%CI)	Sources
n	Number of outpatient visits with a diagnosis of false labour < 37 weeks of gestation in Alberta from January 2008 to March 2013	7,204	calculated from databases
n1	Number of n tested	1,650	calculated from databases
p_posi	Probability of positive results among tested	0.1073 (0.0927-0.1232)	95% CI
p_admit	Probability of hospital admission if not tested	0.1059 (0.0979-0.1143)	95% CI
OR1	Odd ratio between negative and not tested	0.47 (0.37-0.60)	95% CI
OR2	Odd ratio between positive and not tested	5.38 (3.65-7.95)	95% CI
c_test	Cost of fFN test	$136 ($109-$163)	ProvLab (±20%)
c_hospital	Cost of hospitalization for false labour	$3,000 ($2,400-$3,600)	calculated from databases (±20%)

Figure 4.A.3: Decision tree comparing hospital length of stay costs between tested and non-tested cohorts among episodes of false labour

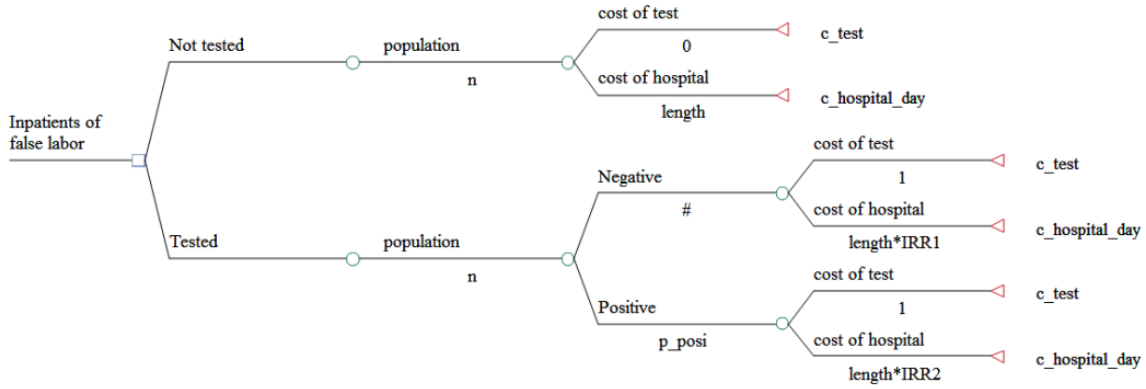


Table 4.A.3: Inputs for decision model comparing hospital length of stay costs between tested and non-tested cohorts among episodes of false labour

Variable	Description	Value (range)	Sources
n	Number of hospital episodes with a diagnosis of false labour in Alberta from January 2008 to March 2013	2,503	calculated from databases
n1	Number of n tested	794	calculated from databases
p_posi	Probability of positive results among the tested	0.3942 (0.36-0.4292)	95% CI
length	LOS if not tested	2.08 days	calculated from databases
IRR1	Incidence rate ratio between negative and not tested	1.01 (0.91-1.11)	95% CI
IRR2	Incidence rate ratio between positive and not tested	1.20 (1.07-1.34)	95% CI
c_test	Cost of an fFN test	$136 ($109-$163)	ProvLab (±20%)
C_hospital_day	Hospital cost per day for false labour	$1,990 ($1,592-$2,388)	calculated from databases (±20%)

Figure 4.A.4: Decision tree comparing ambulance transfer costs between tested and non-tested cohorts among episodes of true labour

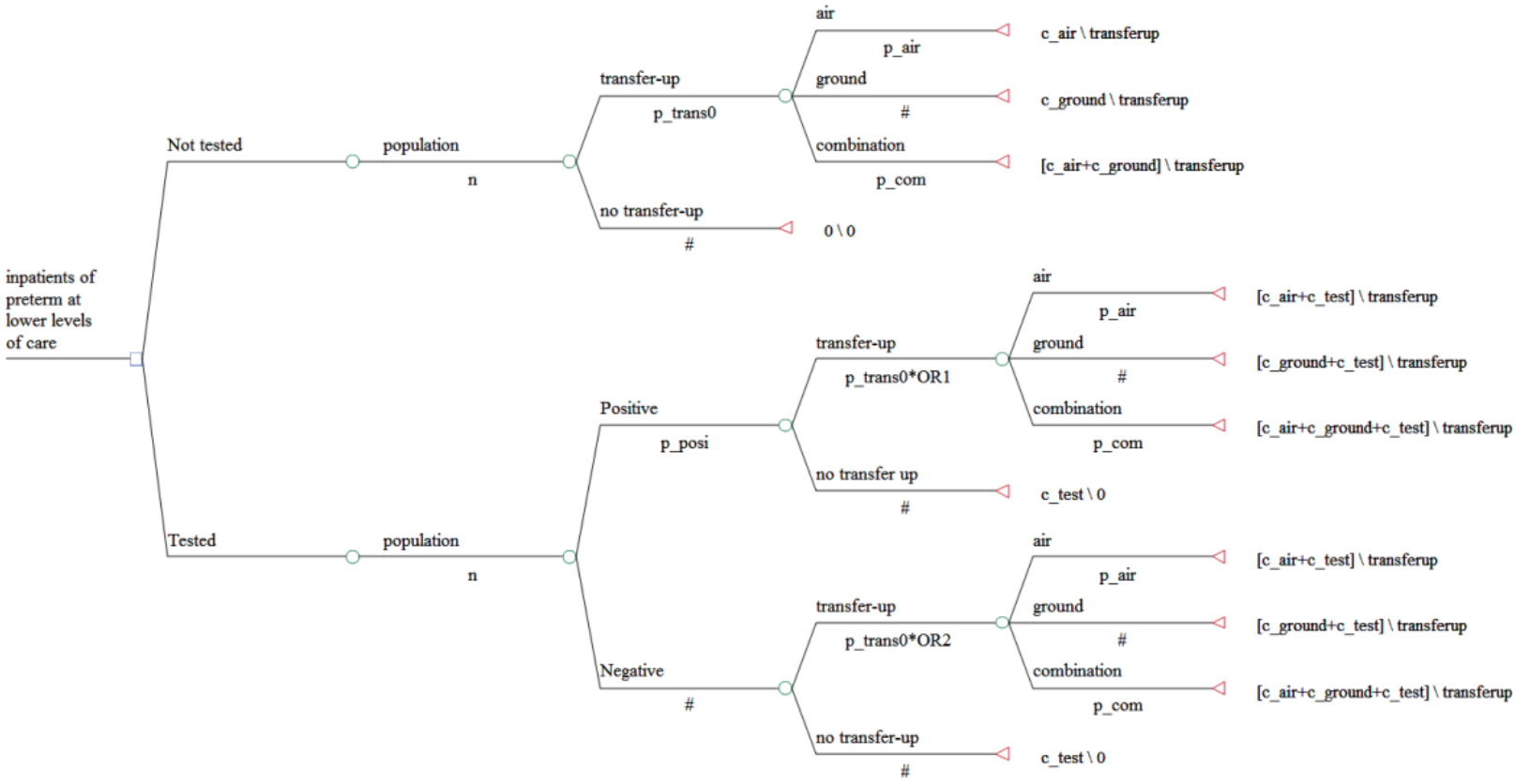


Table 4.A.4a: Inputs for decision model comparing ambulance transfer costs between tested and non-tested cohorts among episodes of true labour (DAD)

Variable	Description	Value (range)	Sources
n	Number of hospital episodes with a diagnosis of preterm at lower levels of care in Alberta from January 2008 to March 2013	4107	calculated from databases
n1	Number of n tested	347	calculated from databases
p_posi	Probability of positive results among tested	0.4063 (0.3542-0.4601)	95% CI
p_trans0	Probability of transfer-up if not tested	0.0628 (0.0552-0.071)	95% CI
OR1	Odd ratio between positive and not tested	7.45 (3.89-14.27)	95% CI
OR2	Odd ratio between negative and not tested	1.91 (1.11-3.29)	95% CI
p_air	Probability of air transfer	0.0326 (0.0265-0.0396)	95% CI
p_com	Probability of combined (ground + air) transfer	0.1294 (0.1176-0.1420)	95% CI
p_ground	Probability of ground transfer	0.838	= 1-(p_air + p_com)
c_air	Cost of air transfer-up in 2014 CA$	$4179 ($3343-$5015)	Currie 2006 (±20%)
c_ground	Cost of ground transfer-up in 2014 CA$	$573 ($459-$688)	Currie 2006 (±20%)
c_test	Cost of fFN test	$136 ($109-$163)	ProvLab (±20%)

Table 4.A.4b: Inputs for decision model comparing ambulance transfer costs between tested and non-tested cohorts among episodes of true labour (ACCS)

Variable	Description	Value (range)	Sources
n	Number of outpatient visits with a diagnosis of preterm at lower levels of care in Alberta from January 2008 to March 2013	4617	calculated from databases
n1	Number of n tested	911	calculated from databases
p_posi	Probability of positive results among tested	0.3128 (0.2828-0.3441)	95% CI
p_trans0	Probability of transfer-up if not tested	0.2191 (0.2059-0.2328)	95% CI

OR1	Odd ratio between positive and not tested	3.68 (2.55-5.31)	95% CI
OR2	Odd ratio between negative and not tested	1.26 (0.96-1.66)	95% CI
p_air	Probability of air transfer	0.0326 (0.0265-0.0396)	95% CI
p_com	Probability of combined (ground + air) transfer	0.1294 (0.1176-0.1420)	95% CI
p_ground	Probability of ground transfer	0.838	= 1-(p_air + p_com)
c_air	Cost of air transfer-up in 2014 CA$	$4,179 ($3343-$5015)	Currie 2006 (±20%)
c_ground	Cost of ground transfer-up in 2014 CA$	$573 ($459-$688)	Currie 2006 (±20%)
c_test	Cost of fFN test	$136 ($109-$163)	ProvLab (±20%)

Figure 4.A.5: Decision tree comparing hospital length of stay costs between tested and non-tested cohorts among episodes of true labour

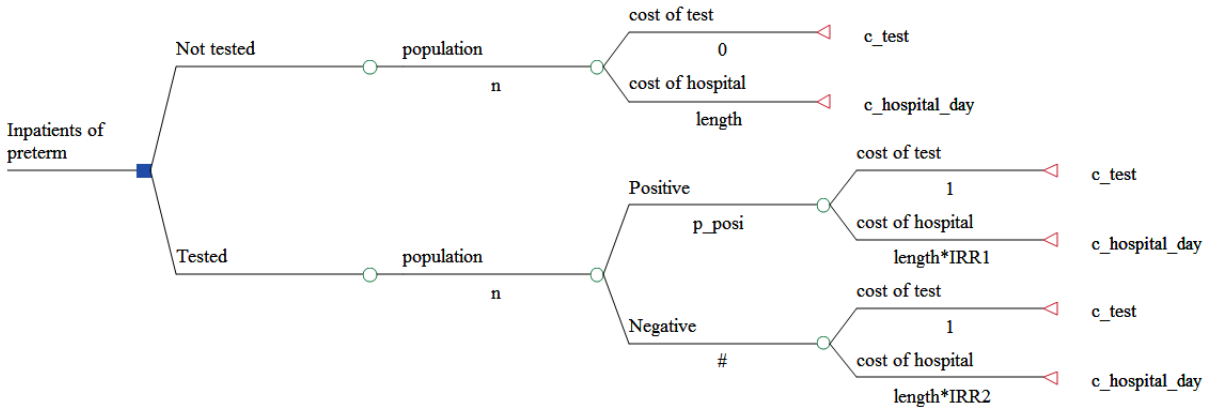


Table 4.A.5: Inputs for decision model comparing hospital length of stay costs between tested and non-tested cohorts among episodes of true labour

Variable	Description	Value (range)	Sources
n	Number of hospital episodes with a diagnosis of preterm in Alberta from January 2008 to March 2013	18368	calculated from databases
n1	Number of n tested	1662	calculated from databases
p_posi	Probability of positive results among the tested	0.4398 (0.4158-0.4641)	95% CI
length	LOS if not tested	3.72 days	calculated from databases

IRR1	Incidence rate ratio between positive and not tested	1.34 (1.25-1.43)	95% CI
IRR2	Incidence rate ratio between negative and not tested	1.13 (1.06-1.20)	95% CI
c_test	Cost of an fFN test	$136 ($109-$163)	ProvLab (±20%)
C_hospital_day	Hospital cost per day for preterm	$1990 ($1592-$2388)	calculated from databases (±20%)

Appendix 4.B: Variables Included in Regression Models

Table 4.B.1: Ambulance transfers among episodes of true labour

	Not transferred	Transferred	Total
DAD			
N	6,596	448	**7,044**
Test results:			
Negative (%)	2.43	9.82	**2.90**
Positive (%)	1.43	10.49	**2.00**
Non-tested (%)	96.15	79.69	**95.10**
Age (mean, sd)	27.11(5.69)	26.62(5.49)	**27.08(5.68)**
Years:			
2002 (%)	4.38	0.45	**4.13**
2003 (%)	6.05	0.89	**5.72**
2004 (%)	6.12	1.56	**5.83**
2005 (%)	6.47	0.22	**6.08**
2006 (%)	9.34	10.49	**9.41**
2007 (%)	10.32	13.39	**10.52**
2008 (%)	11.01	17.19	**11.4**
2009 (%)	10.96	9.6	**10.87**
2010 (%)	10.69	14.51	**10.93**
2011 (%)	10.63	14.06	**10.85**
2012 (%)	11.42	14.29	**11.6**
2013 (%)	2.61	3.35	**2.65**
Admitting categories:			
Urgent (%)	64.11	86.83	**65.56**
Others (%)	35.89	13.17	**34.44**
Comorbidity:			
Yes (%)	92.47	69.42	**91.00**
No (%)	7.53	30.58	**9.00**
Admitting entry code:			
Emergency room (%)	17.75	35.49	**18.88**
Clinic (%)	43.27	32.37	**42.58**
Others (%)	38.98	32.14	**38.54**
Arrived via ambulance			
Yes (%)	12.77	8.93	**12.52**
No (%)	87.23	91.07	**87.48**

ACCS			
N	4,800	1,405	**6,205**
Test results:			
Negative (%)	11.03	9.79	**10.07**
Positive (%)	3.67	7.76	**4.59**
Non-tested (%)	85.3	82.45	**85.34**
Age (mean, sd)	26.02(5.53)	26.22(5.77)	**26.06(5.59)**
Years:			
2002 (%)	0.67	0.21	**0.56**
2003 (%)	0.44	0.71	**0.5**
2004 (%)	0.44	0.57	**0.47**
2005 (%)	0.44	0.36	**0.42**
2006 (%)	9.33	8.61	**9.17**
2007 (%)	14.92	12.95	**14.47**
2008 (%)	13.08	14.09	**13.31**
2009 (%)	12.46	14.8	**12.99**
2010 (%)	13	12.6	**12.91**
2011 (%)	14.58	17.37	**15.21**
2012 (%)	16.23	14.66	**15.87**
2013 (%)	4.42	3.06	**4.11**
Comorbidity:			
Yes (%)	19.96	27.97	**21.77**
No (%)	80.04	72.03	**78.23**
Arrived via ambulance			
Yes (%)	6.77	5.48	**6.48**
No (%)	93.23	94.52	**93.52**
Zones:			
Calgary (%)	6.98	17.65	**9.4**
Edmonton (%)	2.71	13.45	**5.14**
North (%)	42.17	34.66	**40.47**
South (%)	13.17	4.91	**11.3**
Central (%)	34.94	28.75	**33.54**
Unknown (%)	0.03	0.58	**0.15**

Table 4.B.2: Ambulance transfers among episodes of false labour

	Not transferred	Transferred	Total
DAD			
N	3,026	497	**3,523**
Test results:			
Negative (%)	9.62	10.26	**9.71**
Positive (%)	4.10	10.66	**5.02**
Non-tested (%)	86.29	79.07	**85.27**
Age (mean, sd)	25.47 (5.28)	26.44 (5.63)	**25.61 (5.34)**
Years:			
2002 (%)	6.44	6.04	**6.39**
2003 (%)	8.39	4.43	**7.83**
2004 (%)	8.13	2.01	**7.27**
2005 (%)	8.39	4.43	**7.83**
2006 (%)	11.20	11.27	**11.21**
2007 (%)	10.84	10.46	**10.79**
2008 (%)	10.48	14.49	**11.04**
2009 (%)	8.39	9.46	**8.54**
2010 (%)	8.03	11.67	**8.54**
2011 (%)	8.26	10.87	**8.63**
2012 (%)	8.59	11.67	**9.03**
2013 (%)	2.84	3.22	**2.90**
Admitting categories:			
Urgent (%)	82.49	86.92	**83.11**
Others (%)	17.51	13.08	**16.89**
Comorbidity:			
Yes (%)	40.02	59.15	**42.72**
No (%)	59.98	40.85	**57.28**
Admitting entry code:			
Emergency room	36.09	36.22	**36.11**
Clinic (%)	32.32	31.59	**32.22**
Others (%)	31.59	32.19	**31.67**
Arrived via ambulance			
Yes (%)	13.88	10.06	**13.34**
No (%)	86.12	89.94	**86.66**

Zones:			
Calgary (%)	1.98	2.01	**1.99**
Edmonton (%)	0.83	0.60	**0.79**
North (%)	47.88	57.95	**49.30**
South (%)	14.87	15.49	**14.96**
Central (%)	29.02	23.94	**28.30**
Unknown (%)	5.42	0.01	**4.66**
ACCS			
N	5,600	573	**6,173**
Test results:			
Negative (%)	6.16	2.79	**5.85**
Positive (%)	0.86	1.57	**0.92**
Non-tested (%)	92.98	95.64	**93.23**
Age (mean, sd)	25.66(5.45)	25.82(5.39)	**25.68(5.44)**
Years:			
2002 (%)	8.84	16.75	**9.57**
2003 (%)	9.21	15.88	**9.83**
2004 (%)	8.98	20.07	**10.01**
2005 (%)	9.88	15.71	**10.42**
2006 (%)	7.57	10.82	**7.87**
2007 (%)	9.48	4.54	**9.02**
2008 (%)	8.71	4.01	**8.28**
2009 (%)	7.95	2.79	**7.47**
2010 (%)	7.54	2.62	**7.08**
2011 (%)	8.48	2.62	**7.94**
2012 (%)	10.64	3.14	**9.95**
2013 (%)	2.71	1.05	**2.56**
Comorbidity:			
Yes (%)	14.54	16.23	**14.69**
No (%)	85.46	83.77	**85.31**
Arrived via ambulance			
Yes (%)	4	6.81	**4.26**
No (%)	96	93.19	**95.74**
Zones:			
Calgary (%)	8.73	11.34	**8.97**

Edmonton (%)	5.91	20.42	**7.26**
North (%)	48.84	42.23	**48.23**
South (%)	7.84	0.52	**7.16**
Central (%)	28.39	24.61	**28.04**
Unknown (%)	0.29	0.88	**0.34**

Table 4.B.3: Hospital length of stay among episodes of true labour

Variables	N (%)	Mean (days)	Median (days)
All sample	33,019 (100%)	3.84	2
Test results:			
Negative	925 (3%)	3.89	2
Positive	731 (2%)	4.85	2
Non-tested	31,363 (95%)	3.81	2
Years:			
2002	1,659 (5%)	4.32	3
2003	2,083 (6%)	4.03	3
2004	2,122 (6%)	3.88	3
2005	2,320 (7%)	3.99	3
2006	3,131 (9%)	3.77	2
2007	3,336 (10%)	3.73	2
2008	3,608 (11%)	3.74	2
2009	3,482 (11%)	3.86	2
2010	3,487 (11%)	3.79	2
2011	3,455 (10%)	3.91	2
2012	3,510 (11%)	3.70	2
2013	826 (3%)	3.33	2
Admitting categories:			
Urgent	28,183 (85%)	3.90	2
Others	4,836 (15%)	3.46	2
Comorbidity:			
Yes	30,831 (93%)	3.96	3
No	2,188 (7%)	2.17	1
Admitting entry code:			
Emergency room	1,445 (4%)	2.49	1
Clinic	10,143 (31%)	3.30	2
Others	21,431 (65%)	4.18	3

Arrived via ambulance			
Yes	6,394 (19%)	5.13	3
No	26,625 (81%)	3.53	2
Zones:			
Calgary	12,051 (36%)	3.63	2
Edmonton	14,158 (43%)	4.48	3
North	2,473 (7%)	2.64	2
South	1,907 (6%)	3.41	3
Central	2,214 (7%)	2.68	2
Unknown	216 (1%)	2.65	2

Table 4.B.4: Hospital length of stay among episodes of false labour

Variables	N (%)	Mean (days)	Median (days)
All sample	6,312 (100%)	2.14	1
Test results:			
Negative	480 (8%)	1.88	1
Positive	313 (5%)	2.37	2
Non-tested	5,519 (87%)	2.15	1
Years:			
2002	707 (11%)	2.50	1
2003	697 (11%)	2.39	1
2004	643 (10%)	1.96	1
2005	617 (10%)	2.10	1
2006	543 (9%)	1.94	1
2007	602 (10%)	2.09	1
2008	474 (8%)	1.93	1
2009	422 (7%)	2.24	1
2010	444 (7%)	2.19	1
2011	460 (7%)	2.23	1
2012	557 (9%)	1.95	1
2013	146 (2%)	1.82	1
Admitting categories:			
Urgent	5,009 (79%)	2.09	1
Others	1,303 (21%)	2.34	1
Comorbidity:			
Yes	2,503 (40%)	2.90	1
No	3,809 (60%)	1.64	1

Admitting entry code:			
Emergency room	769 (12%)	1.49	1
Clinic	1,956 (31%)	1.95	1
Others	3,587 (57%)	2.38	1
Arrived via ambulance			
Yes	1,165 (18%)	2.69	1
No	5,147 (82%)	2.01	1
Zones:			
Calgary	2,011 (32%)	2.11	1
Edmonton	2,133 (34%)	2.73	1
North	1,023 (16%)	1.59	1
South	358 (6%)	1.53	1
Central	703 (11%)	1.58	1
Unknown	94 (1%)	1.84	1

Table 4.B.5: Hospital admissions among episodes of true labour

	No admit	Admit	Total
N	5,085	2,541	**7,626**
Test results:			
Positive (%)	4.58	5.47	**4.88**
Negative (%)	11.98	7.04	**10.33**
Non-tested (%)	83.44	87.49	**84.79**
Age (mean, sd)	26.29(5.59)	26.72(5.65)	**26.44(5.62)**
Years:			
2002 (%)	0.55	0.35	**0.49**
2003 (%)	0.45	0.31	**0.41**
2004 (%)	0.43	0.47	**0.45**
2005 (%)	0.28	0.51	**0.35**
2006 (%)	8.38	6.18	**7.64**
2007 (%)	12.39	11.37	**12.05**
2008 (%)	11.92	9.88	**11.24**
2009 (%)	13.53	13.97	**13.68**
2010 (%)	15.12	15.35	**15.2**
2011 (%)	16.83	16.25	**16.64**
2012 (%)	16.03	20.78	**17.61**
2013 (%)	4.09	4.57	**4.25**

Arrived via ambulance:			
Yes (%)	4.11	8.54	**5.59**
No (%)	95.89	91.46	**94.41**
Comorbidity:			
Yes (%)	18.54	20.03	**19.04**
No (%)	81.46	79.97	**80.96**
Zones:			
Calgary (%)	10.72	3.50	**8.31**
Edmonton (%)	20.33	25.78	**22.15**
North (%)	29.24	40.30	**32.93**
South (%)	10.32	6.93	**9.19**
Central (%)	29.18	23.49	**27.29**
Unknown (%)	0.20	0.00	**0.13**

Table 4.B.6: Hospital admissions of stay among episodes of false labour

Variables	No admit	Admit	Total
N	16,042	3,624	**19,666**
Test results:			
Positive (%)	0.71	1.68	**0.89**
Negative (%)	8.64	2.4	**7.49**
Non-tested (%)	90.65	95.92	**91.62**
Age (mean, sd)	26.82 (5.58)	27.02 (5.67)	**26.86 (5.60)**
Years:			
2002 (%)	7.51	7.53	**7.51**
2003 (%)	9.47	12.11	**9.96**
2004 (%)	9.91	12.44	**10.37**
2005 (%)	10.4	14.07	**11.08**
2006 (%)	10.21	18.68	**11.77**
2007 (%)	12.19	14.85	**12.68**
2008 (%)	10.7	10.38	**10.64**
2009 (%)	8.33	3.53	**7.45**
2010 (%)	7.18	1.88	**6.2**
2011 (%)	5.95	1.46	**5.13**
2012 (%)	6.55	2.4	**5.79**
2013 (%)	1.6	0.66	**1.42**

Arrived via ambulance:			
Yes (%)	1.3	1.79	**1.39**
No (%)	98.7	98.21	**98.61**
Comorbidity:			
Yes (%)	9.12	10.13	**9.31**
No (%)	90.88	89.87	**90.69**
Zones:			
Calgary (%)	3.44	1.13	**3.02**
Edmonton (%)	68.84	78.64	**70.65**
North (%)	15.78	12.44	**15.17**
South (%)	2.21	2.43	**2.25**
Central (%)	9.58	5.35	**8.8**
Unknown (%)	0.14	0	**0.11**

Appendix 4.C: Descriptive Statistics of Health Databases

1. Inpatients

Based on data extraction criteria shown in Table 4.1, 161,677 inpatient episodes from 1 April 2002 to 31 March 2013 were extracted, of which 2,685 (~2%) were tested (see Table 4.C.1a).

Table 4.C.1a: Distribution of tests among all inpatient episodes

Year		Non-tested cohort	Tested cohort	Total
2002	n	9,683	0	**9,683**
	%	100	0	**100**
2003	n	13,200	0	**13,200**
	%	100	0	**100**
2004	n	12,805	0	**12,805**
	%	100	0	**100**
2005	n	13,055	0	**13,055**
	%	100	0	**100**
2006	n	14,264	0	**14,264**
	%	100	0	**100**
2007	n	14,650	0	**14,650**
	%	100	0	**100**
2008	n	15,121	428	**15,549**
	%	97.25	2.75	**100**
2009	n	15,402	460	**15,862**
	%	97.1	2.9	**100**
2010	n	15,507	576	**16,083**
	%	96.42	3.58	**100**
2011	n	15,732	531	**16,263**
	%	96.73	3.27	**100**
2012	n	15,841	559	**16,400**
	%	96.59	3.41	**100**
2013	n	3,732	131	**3,863**
	%	96.61	3.39	**100**
Total	**n**	**158,992**	**2,685***	**161,677**
	%	**98.34**	**1.66**	**100**

*Note: of 15,042 tests, 2685 (18%) can be linked to the inpatient database (DAD)

Diagnoses of tested episodes are shown in Table 4.C.1b. There were a wide range of diagnoses for which the fFN tests were indicated. However, most of them were preterm (52%) and false labour (30%). Of the tested hospitalizations with a false labour diagnosis, 95% were false labour before 37 weeks of gestation.

Table 4.C.1b: ICD Diagnoses of tested cohort among inpatient episodes

Most responsible diagnoses	Number	%	Cum.%
Preterm labour and delivery	1,405	52.33	52.33
False labour:	797	29.68	82.01
- before 37 completed weeks of gestation	760		
- at or after 37 completed weeks of gestation	37		
Premature rupture of membranes	134	4.99	87
Labour and delivery complicated by fetal stress [distress]	35	1.3	88.31
Maternal care for known or suspected abnormality of pelvic organs	34	1.27	89.57
Supervision of normal pregnancy	26	0.97	90.54
Other maternal diseases classifiable elsewhere but complicating pregnancy, childbirth and the puerperium	25	0.93	91.47
Infections of genitourinary tract in pregnancy	24	0.89	92.36
Premature separation of placenta [abruptio placentae]	23	0.86	93.22
Antepartum haemorrhage, not elsewhere classified	18	0.67	93.89
Maternal care for known or suspected malpresentation of fetus	16	0.6	94.49
Obstructed labour due to malposition and malpresentation of fetus	15	0.56	95.05
Gestational [pregnancy-induced] hypertension with significant proteinuria	14	0.52	95.57
Abdominal and pelvic pain	11	0.41	95.98
Maternal care for other known or suspected fetal problems	10	0.37	96.35
Supervision of high-risk pregnancy	9	0.34	96.69
Other complications of labour and delivery, not elsewhere classified	8	0.3	96.98
Gestational [pregnancy-induced] hypertension without significant proteinuria	7	0.26	97.24
Multiple gestation	6	0.22	97.47
Other disorders of amniotic fluid and membranes	6	0.22	97.69
Abnormalities of forces of labour	5	0.19	97.88
Others	57	2.13	100
Total	**2,685**	**100**	

In total, there were 38,902 inpatient episodes with a diagnosis of preterm or false labour < 37 weeks of gestation from 1 April 2002 to 31 March 2013 (see Table 4.C.1c). Of these, 2,422 (~6%) had been tested for fFN since 2008. The percentage of tested increased over the years, from 9% in 2008 to 13% in 2013.

Table 4.C.1c: Inpatient episodes between tested and non-tested cohorts with a diagnosis of preterm or false labour by year

Year		Non-tested cohort	Tested cohort	Total
2002	n	2,323	0	**2,323**
	%	100	0	**100**
2003	n	2,731	0	**2,731**
	%	100	0	**100**
2004	n	2,713	0	**2,713**
	%	100	0	**100**
2005	n	2,886	0	**2,886**
	%	100	0	**100**
2006	n	3,653	0	**3,653**
	%	100	0	**100**
2007	n	3,902	0	**3,902**
	%	100	0	**100**
2008	n	3,662	382	**4,044**
	%	90.55	9.45	**100**
2009	n	3,456	411	**3,867**
	%	89.37	10.63	**100**
2010	n	3,391	516	**3,907**
	%	86.79	13.21	**100**
2011	n	3,410	474	**3,884**
	%	87.8	12.2	**100**
2012	n	3,510	518	**4,028**
	%	87.14	12.86	**100**
2013	n	843	121	**964**
	%	87.45	12.55	**100**
Total	**n**	**36,480**	**2,422**	**38,902**
	%	**93.77**	**6.23**	**100**

In total, there were 2,503 inpatient episodes with a diagnosis of false labour < 37 weeks from January 2008 to March 2013 (Table 4.C.1d). Of these, 794 (32%) were tested for fFN. Within false labour < 37 weeks, the percentage of tested episodes was higher in level D hospitals in comparison with the lower level of care hospitals (33% versus 29%).

Table 4.C.1d: Inpatient episodes between tested and non-tested cohorts with a diagnosis of false labour from January 2008 to March 2013 by level of acuity

Level of care		Non-tested cohort	Tested cohort	Total
< Level D	n	577	236	**813**
	%	70.97	29.03	**100**
Level D	n	1,132	558	**1,690**
	%	66.98	33.02	**100**
Total	**n**	**1,709**	**794**	**2,503**
	%	**68.28**	**31.72**	**100**

2. Outpatients

Based on the data extraction criteria, 929,627 outpatient visits were extracted, of which 6,169 (1%) had been tested for fFN since 2008 (see Table 4.C.2a).

Table 4.C.2a: Distribution of tests among all outpatient episodes

Year		Non-tested cohort	Tested cohort	Total
2002	n	26,229	0	**26,229**
	%	100	0	**100**
2003	n	55,341	0	**55,341**
	%	100	0	**100**
2004	n	70,812	0	**70,812**
	%	100	0	**100**
2005	n	72,442	0	**72,442**
	%	100	0	**100**
2006	n	80,257	0	**80,257**
	%	100	0	**100**
2007	n	93,267	0	**93,267**
	%	100	0	**100**
2008	n	98,197	994	**99,191**
	%	99	1	**100**
2009	n	102,247	1,204	**103,451**
	%	98.84	1.16	**100**
2010	n	98,140	1,243	**99,383**
	%	98.75	1.25	**100**
2011	n	98,719	1,234	**99,953**
	%	98.77	1.23	**100**
2012	n	102,213	1,213	**103,426**

	%	98.83	1.17	**100**
2013	n	25,594	281	**25,875**
	%	98.91	1.09	**100**
Total	**n**	**923,458**	**6,169***	**929,627**
	%	**99.34**	**0.66**	**100**

*Note: of 15,042 tests, 6,169 (41%) can be linked to the outpatient database (ACCS)

Among the tested outpatient visits, supervision of normal pregnancy diagnosis accounted for the most (32%), followed by false labour (29%) and preterm (19%). Of note, there was a long list of diagnoses for which the fFN tests were indicated (see Table 4.C.2b).

Table 4.C.2b: ICD Diagnoses of tested cohort among outpatient episodes

Most responsible diagnoses	Number	%	Cum.%
Supervision of normal pregnancy	1,976	32.03	32.03
False labour:	1,787	28.97	61
- before 37 completed weeks of gestation	1,650		
- at or after 37 completed weeks of gestation	137		
Preterm labour and delivery	1,161	18.82	79.82
Premature rupture of membranes	265	4.3	84.12
Supervision of high-risk pregnancy	158	2.56	86.68
Maternal care for other known or suspected fetal problems	151	2.45	89.13
Abdominal and pelvic pain	144	2.33	91.46
Other special examinations and investigations of persons without complaint or reported diagnosis	70	1.13	92.59
Other maternal diseases classifiable elsewhere but complicating pregnancy, childbirth and the puerperium	58	0.94	93.53
Other medical care	51	0.83	94.36
Infections of genitourinary tract in pregnancy	42	0.68	95.04
Dorsalgia	29	0.47	95.51
Antepartum haemorrhage, not elsewhere classified	19	0.31	95.82
Examination and observation for other reasons	18	0.29	96.11
Maternal care for known or suspected abnormality of pelvic organs	17	0.28	96.39
Persons encountering health services for other counselling and medical advice, not elsewhere classified	16	0.26	96.65
Need for other prophylactic measures	15	0.24	96.89
Multiple gestation	14	0.23	97.12
Other disorders of amniotic fluid and membranes	13	0.21	97.33
Maternal care for other conditions predominantly related to pregnancy	10	0.16	97.49

Other disorders of urinary system	8	0.13	97.62
Other complications of labour and delivery, not elsewhere classified	8	0.13	97.75
Other noninflammatory disorders of vagina	7	0.11	97.86
Pain, not elsewhere classified	7	0.11	97.97
Diabetes mellitus in pregnancy	4	0.06	98.03
Maternal care for known or suspected malpresentation of fetus	4	0.06	98.09
Nausea and vomiting	4	0.06	98.15
Excessive vomiting in pregnancy	3	0.05	98.2
Pain in throat and chest	3	0.05	98.25
Others	107	1.73	100
Total	**6,169**	**100**	

Table 4.C.2c shows that of total 27,292 outpatient visits with a diagnosis of preterm or false labour < 37 weeks of gestation, 2,811 (10%) were tested for fFN. The percentage of tested visits increased from 17% in 2008 to more than 20% in later years.

Table 4.C.2c: Outpatient episodes between tested and non-tested cohorts with a diagnosis of preterm or false labour by year

Year		Non-tested cohort	Tested cohort	Total
2002	n	1,514	0	**1,514**
	%	100	0	**100**
2003	n	1,989	0	**1,989**
	%	100	0	**100**
2004	n	2,074	0	**2,074**
	%	100	0	**100**
2005	n	2,206	0	**2,206**
	%	100	0	**100**
2006	n	2,898	0	**2,898**
	%	100	0	**100**
2007	n	3,412	0	**3,412**
	%	100	0	**100**
2008	n	2,442	508	**2,950**
	%	82.78	17.22	**100**
2009	n	1,947	561	**2,508**
	%	77.63	22.37	**100**
2010	n	1,809	570	**2,379**
	%	76.04	23.96	**100**

2011	n	1,744	533	**2,277**
	%	76.59	23.41	**100**
2012	n	1,974	507	**2,481**
	%	79.56	20.44	**100**
2013	n	472	132	**604**
	%	78.15	21.85	**100**
Total	**n**	**24,481**	**2,811**	**27,292**
	%	**89.7**	**10.3**	**100**

The number of outpatient visits with a diagnosis of false labour < 37 weeks of gestation from January 2008 to March 2013 by level of care is shown in Table 4.C.2d. Of the total 7204 visits, 1,650 (23%) were tested for fFN. The percentage of outpatient visits with a diagnosis of false labour < 37 weeks which were tested was higher in the highest level of care (level D) in comparison with lower level of care hospitals (27% versus 16%).

Table 4.C.2d: Outpatient episodes between tested and non-tested cohorts with a diagnosis of false labour from January 2008 to March 2013 by level of acuity

Level of care		Non-tested cohort	Tested cohort	Total
< Level D	n	2,251	420	**2,671**
	%	84.28	15.72	**100**
Level D	n	3,303	1,230	**4,533**
	%	72.87	27.13	**100**
Total	**n**	**5,554**	**1,650**	**7,204**
	%	**77.1**	**22.9**	**100**

3. Physician visits

Based on the data extraction criteria, totally 4,080,818 practitioner visits were extracted, of which 10,498 (0.3%) were tested for fFN (Table 4.C.3a).

Table 4.C.3a: Distribution of testing among all practitioner visits

Year		Non-tested cohort	Tested cohort	Total
2002	n	239,219	0	**239,219**
	%	100	0	**100**
2003	n	336,240	0	**336,240**
	%	100	0	**100**
2004	n	342,737	0	**342,737**
	%	100	0	**100**
2005	n	357,493	0	**357,493**
	%	100	0	**100**

2006	n	382,596	0	**382,596**
	%	100	0	**100**
2007	n	394,389	0	**394,389**
	%	100	0	**100**
2008	n	393,014	1,682	**394,696**
	%	99.57	0.43	**100**
2009	n	392,961	1,914	**394,875**
	%	99.52	0.48	**100**
2010	n	377,248	2,173	**379,421**
	%	99.43	0.57	**100**
2011	n	369,991	2,014	**372,005**
	%	99.46	0.54	**100**
2012	n	390,830	2,194	**393,024**
	%	99.44	0.56	**100**
2013	n	93,602	521	**94,123**
	%	99.45	0.55	**100**
Total	**n**	**4,070,320**	**10,498***	**4,080,818**
	%	**99.74**	**0.26**	**100**

*Note: Of 15,042 tests, 10,498 (70%) can be linked to the practitioner claims database.

Of those tested for fFN, false labour accounted for the most (54%), followed by supervision of normal pregnancy (19%), supervision of high risk pregnancy (14%), and other complications of labour and delivery (10%) (Table 4.C.3b).

Table 4.C.3b: ICD Diagnosis of tested cohort among practitioner visits

Diagnosis	Freq.	%	Cum. %
False labour	5,630	54	54
-Before 37 completed weeks of gestation	4,088		
-At or after 37 weeks	1,542		
Supervision of normal pregnancy	1,958	19	72
Supervision of high risk pregnancy	1,430	14	86
Other complications of labour & delivery	1,101	10	96
Multiple gestation	106	1	97
Other symptoms involving abdomen and pelvis	87	1	98
Others	190	2	100
Total	**10,498**	**100**	

Table 4.C.3c shows the number of fFN tested claims by functional center. Of the total, 40% were from inpatient service centers (IPSR) and another 40% were from outpatient service centers. In

theory, these are also found in DAD and/or ACCS databases. There were 2,074 tests (20%) performed in physician offices.

Table 4.C.3c: Number of tested claims by functional center

Centre type	Freq.	%	Cum. %
AMBU	4,098	39.69	39.69
DGTS	11	0.11	39.8
IPSR	4,142	40.12	79.91
POFF	2,074	20.09	100
Total	**10,325**	**100**	

Of the fFN tests done in physician offices, visits for supervision of normal pregnancy accounted for the most (68%), followed by supervision of high risk pregnancy (18%), and false labour (10%) (Table 4.C.3d).

Table 4.C.3d: Distribution of ICD diagnosis for tested cohort conducted in physician offices

Diagnosis	Freq.	%	Cum. %
Supervision of normal pregnancy	1,413	68.13	68.13
Supervision of high risk pregnancy	375	18.08	86.21
False labour	204	9.84	96.05
Other complications of labour & delivery	52	2.51	98.56
Multiple gestation	7	0.34	98.9
Others	23	1.1	100
Total	**2,074**	**100**	

Of the fFN tests performed at physician offices, general practitioners accounted for 43% and obstetricians and gynecologists for 57% (Table 4.C.3e). Of the tests, about 15% were positive and 85% were negative.

Table 4.C.3e: Test result by physician specialty (physician offices)

POFF		Positive	Negative	Invalid	Missing	Total
GP	n	128	762	1	1	**892**
	%	42.24	43.1	100	50	**43.01**
OBY	n	175	997	0	0	**1,172**
	%	57.76	56.39	0	0	**56.51**
Other	n	0	9	0	1	**10**
	%	0	0.51	0	50	**0.48**
Total	**n**	**303**	**1,768**	**1**	**2**	**2,074**
	%	**100**	**100**	**100**	**100**	**100**

Appendix 4.D: Preterm Deliveries by Patient Zone and Hospital Zone (APHP data 2002-2011)

Table 4.D.1: Distribution of patient zones by hospital zone

Patient zones		Hospital zones					Total
		South	Calgary	Central	Edmonton	North	
South	n	2,195	706	107	34	3	**3,045**
	%	90.14	4.25	3.59	0.19	0.12	**7.26**
Calgary	n	229	15,288	121	56	3	**15,697**
	%	9.4	92.09	4.06	0.32	0.12	**37.42**
Central	n	7	566	2,700	1,811	9	**5,093**
	%	0.29	3.41	90.7	10.34	0.37	**12.14**
Edmonton	n	1	17	17	13,053	157	**13,245**
	%	0.04	0.1	0.57	74.56	6.47	**31.58**
North	n	3	24	32	2,553	2,253	**4,865**
	%	0.12	0.14	1.07	14.58	92.91	**11.6**
Total	**n**	**2,435**	**16,601**	**2,977**	**17,507**	**2,425**	**41,945**
	%	**100**	**100**	**100**	**100**	**100**	**100**

Table 4.D.2: Distribution of hospital zones by patient zones

Patient zones		Hospital zones					Total
		South	Calgary	Central	Edmonton	North	
South	n	2,195	706	107	34	3	**3,045**
	%	72.09	23.19	3.51	1.12	0.1	**100**
Calgary	n	229	15,288	121	56	3	**15,697**
	%	1.46	97.39	0.77	0.36	0.02	**100**
Central	n	7	566	2,700	1,811	9	**5,093**
	%	0.14	11.11	53.01	35.56	0.18	**100**
Edmonton	n	1	17	17	13,053	157	**13,245**
	%	0.01	0.13	0.13	98.55	1.19	**100**
North	n	3	24	32	2,553	2,253	**4,865**
	%	0.06	0.49	0.66	52.48	46.31	**100**
Total	**n**	**2,435**	**16,601**	**2,977**	**17,507**	**2,425**	**41,945**
	%	**5.81**	**39.58**	**7.1**	**41.74**	**5.78**	**100**

Appendix 4.E: Number of Hospital Episodes by Level of Acuity between Tested and Non-tested Cohorts (2008-2013)

Table 4.E.1: Preterm diagnosis

Level of care		Non-tested cohort	Tested cohort	Total
< Level D	n	3,760	347	**4,107**
	%	22.51	20.88	**22.36**
Level D	n	12,946	1,315	**14,261**
	%	77.49	79.12	**77.64**
Total	**n**	**16,706**	**1,662**	**18,368**
	%	**100**	**100**	**100**

p = 0.129

Table 4.E.2: False labour diagnosis

Level of care		Non-tested cohort	Tested cohort	Total
< Level D	n	577	236	**813**
	%	33.76	29.72	**32.48**
Level D	n	1,132	558	**1,690**
	%	66.24	70.28	**67.52**
Total	**n**	**1,709**	**794**	**2,503**
	%	**100**	**100**	**100**

p = 0.045

Appendix 4.F: Number of Inpatient Episodes Before and After 2008 by Level of Hospital Acuity

Table 4.F.1: Preterm diagnosis

Level of care		After	Before	Total
< Level D	n	4,107	2,937	**7,044**
	%	22.36	20.02	**21.32**
Level D	n	14,261	11,736	**25,997**
	%	77.64	79.98	**78.68**
Total	**n**	**18,368**	**14,673**	**33,041**
	%	**100**	**100**	**100**

p = 0.000

Table 4.F.2: False labour diagnosis

Level of care		After	Before	Total
< Level D	n	813	1,449	**2,262**
	%	32.48	37.88	**35.75**
Level D	n	1,690	2,376	**4,066**
	%	67.52	62.12	**64.25**
Total	**n**	**2,503**	**3,825**	**6,328**
	%	**100**	**100**	**100**

p = 0.000

Appendix 4.G: Number of Ambulance Transfers to Level C Hospitals

Table 4.G.1: Preterm diagnosis

Diagnosis		Transferred		Total
		No	Yes	
Non-tested	n	3,112	76	**3,188**
	%	97.62	2.38	**100**
Tested	n	178	5	**183**
	%	97.27	2.73	**100**
Total	**n**	**3,290**	**81**	**3,371**
	%	**97.6**	**2.4**	**100**

p = 0.765

Table 4.G.2: False labour diagnosis

Diagnosis		Transferred		Total
		No	Yes	
Non-tested	n	1,043	24	**1,067**
	%	97.75	2.25	**100**
Tested	n	76	0	**76**
	%	100	0	**100**
Total	**n**	**1,119**	**24**	**1,143**
	%	**97.9**	**2.1**	**100**

p = 0.186

Appendix 4.H: Hospital Centres by Levels of Acuity Over Time

Zone	Facility	Community	2008	2009	2010	2011	2012
South	Fort MacLeod Health Centre	Fort MacLeod	0	0	0	0	0
	Magrath Health Centre	Magrath	0	0	0	0	0
	Milk River Health Centre	Milk River	0	0	0	0	0
	Cardston Health Centre	Cardston	A	A	A	A	A
	Raymond Health Centre	Raymond	A	A	A	A	A
	Crowsnest Pass Health Centre	Blairmore	B	B	B	B	B
	Pincher Creek Health Centre	Pincher Creek	B	B	B	B	B
	Taber Health Centre	Taber	B	B	B	B	B
	Chinook Regional Hospital	Lethbridge	C	C	C	C	C
	Bassano Health Centre	Bassano	0	0	0	0	0
	Bow Island Health Centre	Bow Island	0	0	0	0	0
	Big Country Health Centre	Oyen	A	A	A	A	0
	Brooks Health Centre	Brooks	B	0	B	B	B
	Medicine Hat Regional Hospital	Medicine Hat	C	C	C	C	C
Calgary	Oilfields General Hospital	Black Diamond	0	0	0	0	0
	Strathmore District Health Services	Strathmore	0	0	0	0	0
	Didsbury District Health Services	Didsbury	0	0	0	0	0
	Vulcan Community Health Centre	Vulcan	0	0	0	0	0
	Claresholm General Hospital	Claresholm	0	0	0	0	0
	Banff Mineral Springs Hospital	Banff	B	0	B	B	B
	Canmore General Hospital	Canmore	B	B	B	B	B
	FMC, PLC and RGH	Calgary	D	D	D	D	D
	High River General Hospital	High River	B	B	B	B	B
Central	Our Lady of the Rosary Hospital & Care Centre	Castor	0	0	0	0	0
	Coronation Hospital & Care Centre	Coronation	0	0	0	0	0
	Consort Hospital & Care Centre	Consort	0	0	0	0	0
	Hanna Health Centre	Hanna	A	A	A	A	0
	Innisfail Health Centre	Innisfail	A	0	0	0	0
	Rimbey Hospital & Care Centre	Rimbey	A	A	A	A	A
	Sundre Hospital & Care Centre	Sundre	A	A	A	A	A
	Drayton Valley Hospital & Care Centre	Drayton Valley	B	B	B	B	A
	Drumheller Health Centre	Drumheller	B	B	B	B	B
	Lacombe Hospital & Care Centre	Lacombe	B	B	B	B	B
	Olds Hospital & Care Centre	Olds	B	B	B	B	B
	Ponoka Hospital & Care Centre	Ponoka	B	B	B	B	B
	Rocky Mountain House Health Centre	Rocky Mountain House	B	B	B	B	B
	Stettler Hospital & Care Centre	Stettler	B	B	B	B	B

	Three Hills Health Centre	Three Hills	B	B	B	B	B
	Wetaskiwin Hospital & Care Centre	Wetaskiwin	B	B	B	B	B
	Red Deer Regional Hospital Centre	Red Deer	C	C	C	C	C
	Hardisty Health Centre	Hardisty	0	0	0	0	0
	Killam Health Centre	Killam	0	0	0	0	0
	Lamont Health Centre	Lamont	0	0	0	0	0
	Tofield Health Centre	Tofield	0	0	0	0	0
	Two Hills Health Centre	Two Hills	0	0	0	0	0
	St. Joseph's General Hospital	Vegreville	0	0	0	0	0
	St. Mary's Hospital	Camrose	B	B	B	B	B
	Daysland Health Centre	Daysland	B	B	B	B	B
	Lloydminster Hospital	Lloydminster	B	B	B	B	B
	Provost Health Centre	Provost	B	B	B	B	B
	Vermilion Health Centre	Vermilion	B	B	B	B	B
	Viking Health Centre	Viking	B	B	B	B	B
	Wainwright Health Centre	Wainwright	B	B	B	B	B
Edmonton	Devon General Hospital	Devon	0	0	0	0	0
	Leduc Community Hospital	Leduc	0	0	0	0	0
	Redwater Health Centre	Redwater	0	0	0	0	0
	Westview Health Centre	Stony Plain	A	A	A	0	0
	Fort Saskatchewan Health Centre	Fort Saskatchewan	B	B	B	B	B
	Sturgeon Community Hospital	Edmonton / St Albert (RAH, GNCH, MCH, SCH)	D	D	D	D	D
North	Boyle Healthcare Centre	Boyle	0	0	0	0	0
	Elk Point Healthcare Centre	Elk Point	0	0	0	0	0
	Seton Jasper Healthcare Centre	Jasper	0	0	0	0	0
	Mayerthorpe Healthcare Centre	Mayerthorpe	0	0	0	0	0
	George McDougall Healthcare Centre	Smoky Lake	0	0	0	0	0
	Swan Hills Healthcare Centre	Swan Hills	0	0	0	0	0
	Wabasca/Desmarais Healthcare Centre	Desmarais/ Wabasca	0	0	0	0	0
	Athabasca Healthcare Centre	Athabasca	A	A	A	A	A
	Barrhead Healthcare Centre	Barrhead	B	B	B	B	B
	Bonnyville Healthcare Centre	Bonnyville	B	B	B	B	B
	Cold Lake Healthcare Centre	Cold Lake	B	B	B	B	B
	Edson Healthcare Centre	Edson	B	B	B	B	B
	Hinton Healthcare Centre	Hinton	B	B	B	B	B
	William J. Cadzow Healthcare Centre	Lac La Biche	B	B	B	B	B
	Slave Lake Healthcare Centre	Slave Lake	B	B	B	B	A
	St. Therese St. Paul Healthcare Centre	St. Paul	B	B	B	B	B

	Westlock Healthcare Centre	Westlock	B	B	B	B	B
	Whitecourt Healthcare Centre	Whitecourt	B	B	B	B	B
	Manning Community Health Centre	Manning	0	0	0	0	0
	Central Peace Health Complex	Spirit River	0	0	0	0	0
	Grande Cache Community Health Complex	Grande Cache	0	0	0	0	0
	Grimshaw Berwyn Community Health Complex	Grimshaw	0	0	0	0	0
	Fox Creek Healthcare Centre	Fox Creek	0	0	0	0	0
	Beaverlodge Municipal Hospital	Beaverlodge	A	A	A	A	A
	Fairview Health Complex	Fairview	A	A	A	A	0
	Valleyview Health Centre	Valleyview	A	A	A	A	A
	High Prairie Health Complex	High Prairie	0	0	0	0	0
	Sacred Heart Community Health Centre	McLennan	B	B	A	A	A
	Peace River Community Health Centre	Peace River	B	B	B	B	B
	Queen Elizabeth II Hospital	Grande Prairie	C	C	C	C	C
	St. Theresa General Hospital	Fort Vermilion	A	A	A	A	A
	Northwest Health Centre	High Level	B	B	B	B	B
	Northern Lights Regional Health Centre	Fort McMurray	C	C	C	C	C

Appendix 4.I: Sensitivity and Specificity of fFN testing in Alberta

Sensitivity and specificity of fFN testing was calculated, using the available administrative data. Sensitivity and specificity were defined as follows:

- **True positive**: the test is positive and preterm delivery is within (≤) 7 days of the date of test
- **False positive**: the test is positive and preterm delivery is NOT within (>) 7 days of the date of test
- **True negative**: the test is negative and preterm delivery is NOT within (>) 7 days of the date of test
- **False negative**: the test is negative and preterm delivery is within (≤) 7 days of the date of test

Table 4.I.1: Sensitivity and specificity of fFN testing in Alberta

	Delivery ≤ 7 days*		Delivery ≤ 14 days*		Delivery ≤ 21 days*	
	n	%	n	%	n	%
True positive	312	**Sn=16.2**	405	**Sn=21.1**	487	**Sn=25.6**
False positive	1,615	83.8	1,517	78.9	1,415	74.4
Total**	**1,927**	**100**	**1,922**	**100**	**1,902**	**100**
True negative	10,916	**Sp=98.4**	10,785	**Sp=97.3**	10,553	**Sp=95.2**
False negative	174	1.6	305	2.8	537	4.8
Total	**11,090**	**100**	**11,090**	**100**	**11,090**	**100**
Of the false positive, number and % women that are still preterm delivery						
Yes	484	30.0	391	25.8	309	21.8
No	1,131	70.0	1,126	74.2	1,106	78.2
Total	**1,615**	**100**	**1,517**	**100**	**1,415**	**100**

Calculations were conducted by Alberta Health, due to data privacy. Calculations were conducted at the patient and not episode level. For pregnancies with multiple fFN tests, the date of the last fFN test was used in the calculation.

*time from the date of test to the date of delivery

**after excluded deliveries within the cut-off, but at >37 weeks gestation

Appendix 4.J: Conceptual Maximum Of System Cost Savings From Reducing False Positives

Specificity of fFN testing was estimated to by approximately 98%, according to the administrative data. The figures show the corresponding cost impact using a specificity of 98%, and assuming that physicians completely base course of management on a negative test result. All other model inputs are unchanged (refer to Appendix 4.A).

Figure 4.J.1: Cost impact of minimizing unnecessary ambulance transfers (inpatient episodes)

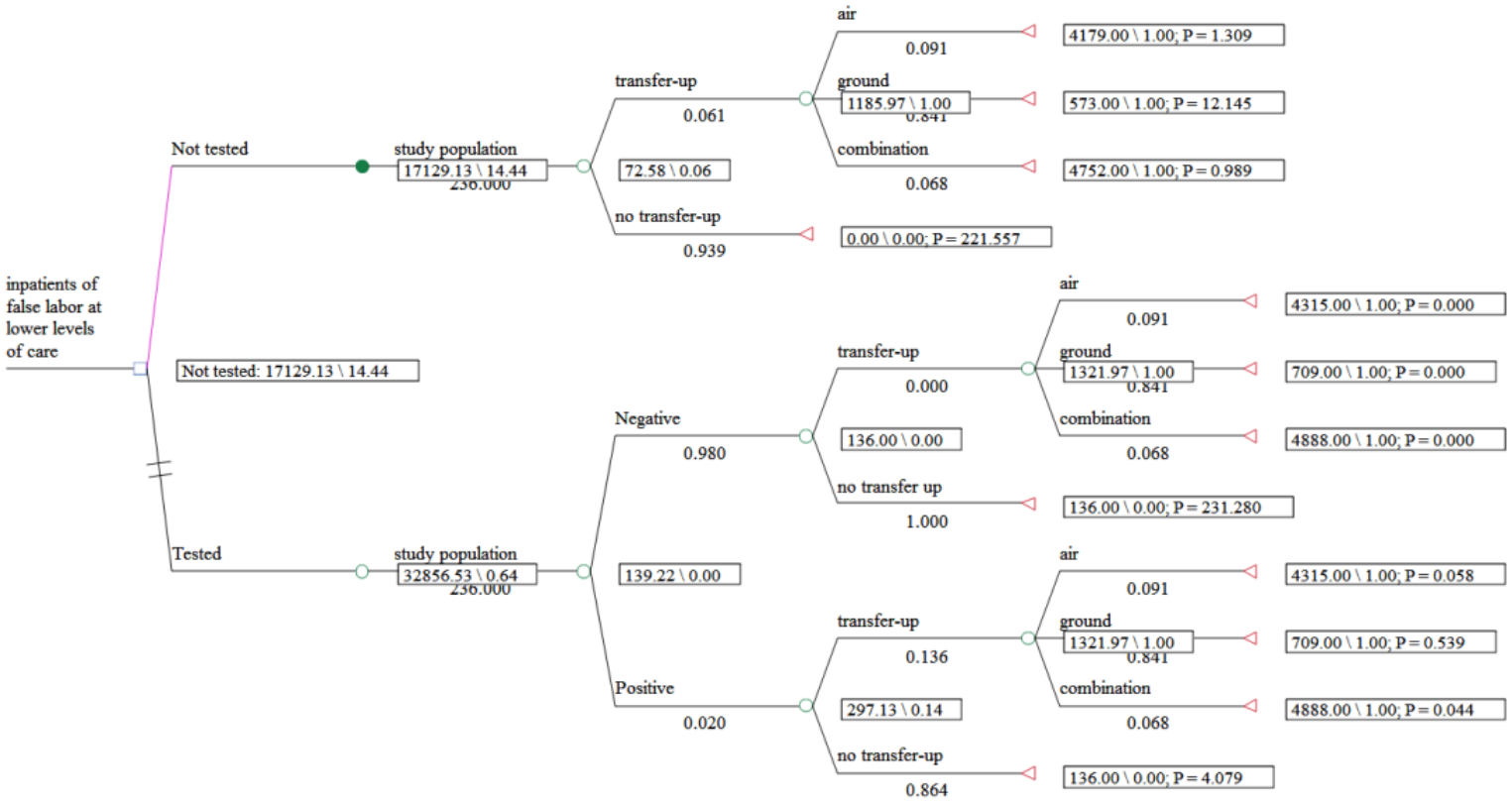


Figure 4.J.2: Cost impact of minimizing unnecessary ambulance transfers (outpatient episodes)

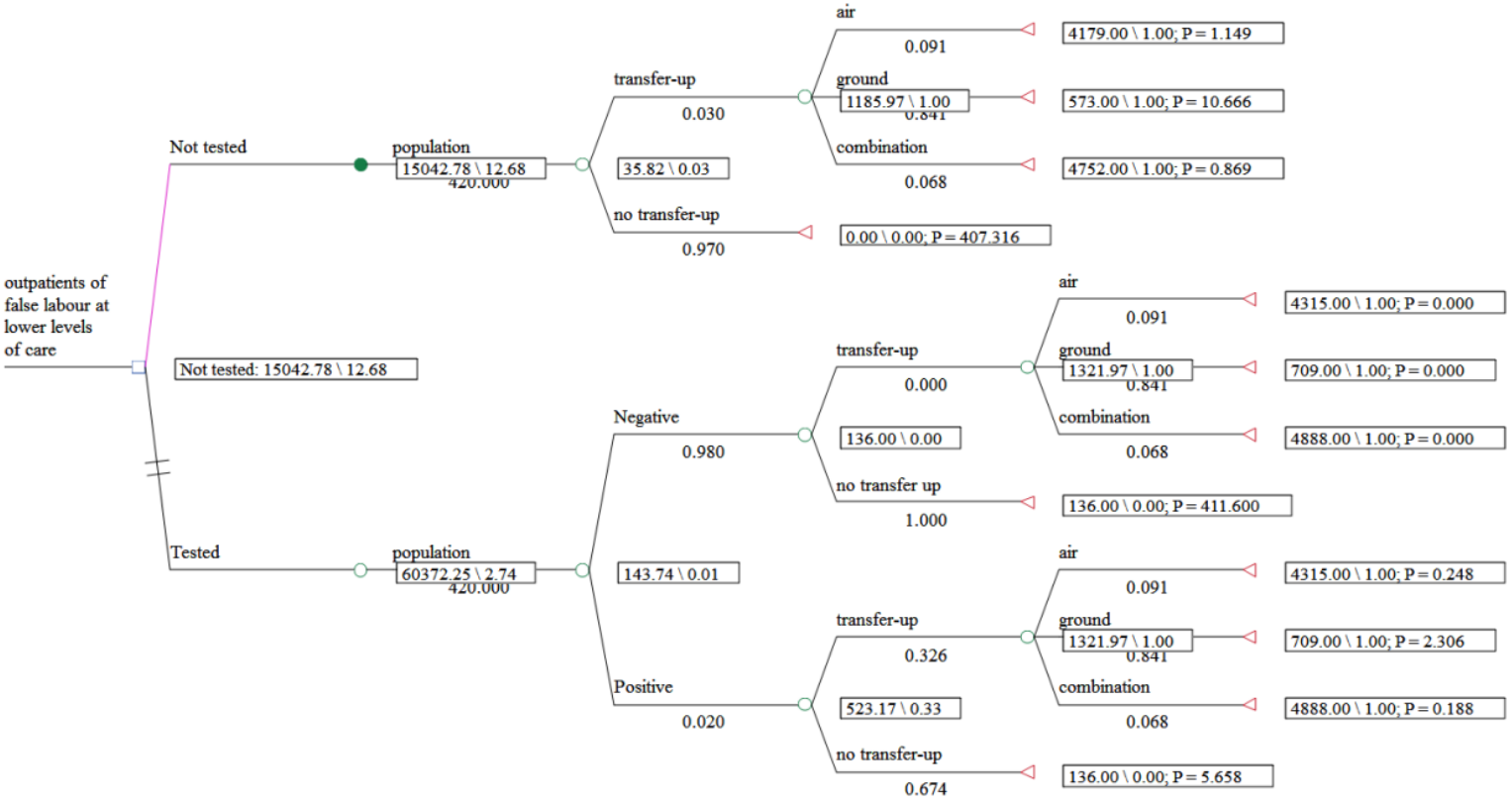


Figure 4.J.3: Cost impact of minimizing unnecessary hospital admissions

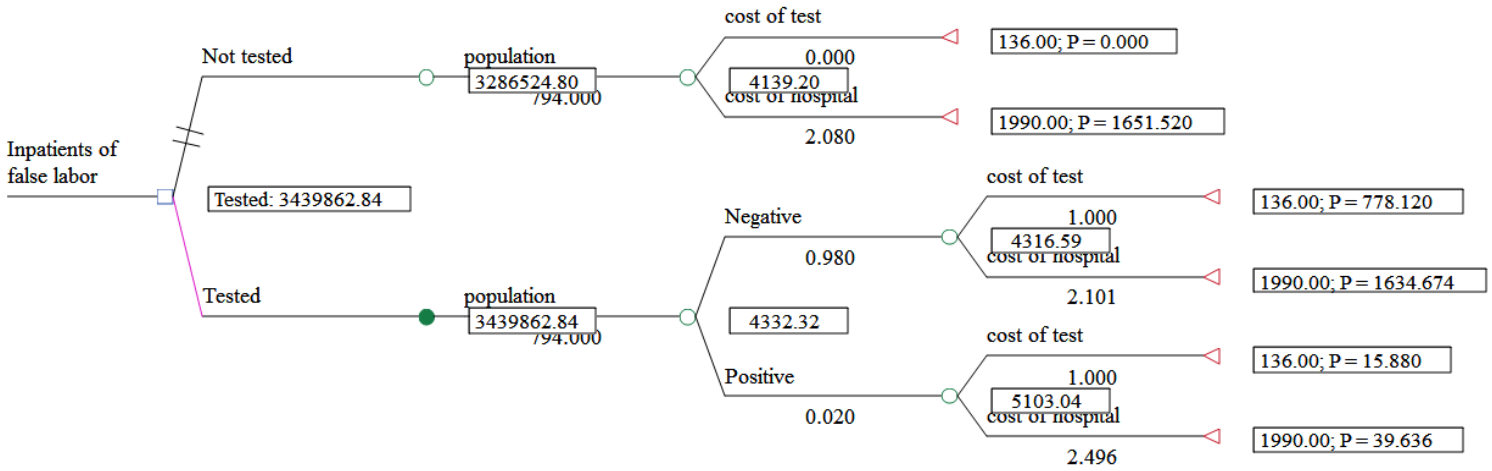
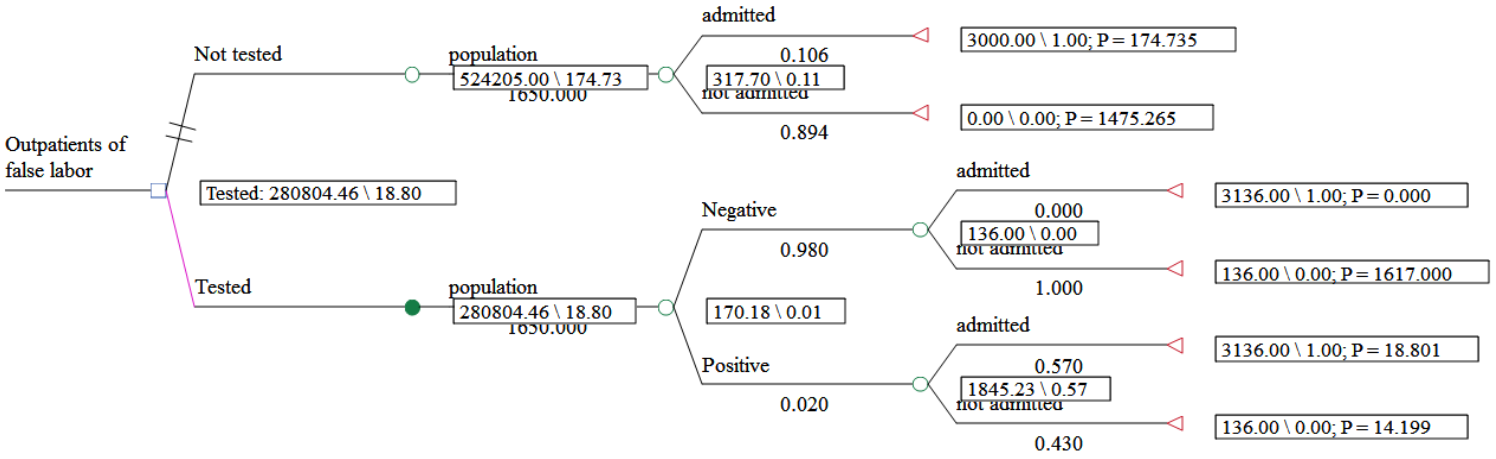


Figure 4.J.4: Cost impact of minimizing unnecessary hospital length of stay



Appendix 4.K: The Impact of fFN Testing On Rate Of Preterm Delivery Before And After Implementation By Level Of Acuity

Table 4.K.1: Preterm deliveries before and after fFN implementation by level of acuity

fFN Implementation	Level D Facility		Total
	No	Yes	
Pre	3,205	11,747	14,952
	21.44%	78.56%	100%
Post	2,578	10,838	13,416
	19.22%	80,78%	100%
Total	5,783	22,585	28,368
	20.39%	79.61%	100%

Pearson chi2(1) = 21.4609 Pr = 0.000

Appendix 4.L: The Impact of fFN Testing Capacity on Receiving an fFN Test

To investigate impact of availability of kit and analyzer on likelihood of getting the test, we used the data on hospital episodes with a diagnosis of preterm or false labour < 37 weeks of gestation at all levels of care. In total, there were 39,369 of such episodes, of which 2,456 (6%) were tested for fFN (Table 4.17). Of the tested, 0.77% were in hospitals with kit only, and 90.88% in hospitals with both kit and analyzer. Of note, 1.34% of the tests were in hospitals where both kit and analyzer were unavailable, and 7% were unknown.

Table 4.L.3 shows the impact of availability of kit and analyzer on likelihood of being tested among hospital episodes with a diagnosis of preterm and false labour < 37 weeks of gestation. The likelihood of being tested in hospitals with kit only was 0.5 times, and in hospitals with no kit nor analyzer was 0.23 times, lower than that of being tested in hospitals with both kit and analyzer. The difference was statistically significant.

Table 4.L.1: Distribution of tests conducted by fFN testing capacity

Variables	Tested cohort	Non-tested cohort	Total
N	2,456	36,913	**39,369**
Availability of test kit			
None (%)	1.34	5.38	**5.13**
Kit only (%)	0.77	1.15	**1.12**
Kit and analyzer (%)	90.88	87.23	**87.46**
Unknown (%)	7	6.25	**6.29**

Table 4.L.2: Variables included in regression model examining the impact of the availability to fFN testing on the likelihood of receiving a test

Variables	Tested cohort	Non-tested cohort	Total
N	2,456	36,913	**39,369**
Availability of test kit			
None (%)	1.34	5.38	**5.13**
Kit only (%)	0.77	1.15	**1.12**
Kit and analyzer (%)	90.88	87.23	**87.46**
Unknown (%)	7	6.25	**6.29**
Age (mean, sd)	27.31 (5.59)	28.39 (5.86)	**28.32 (5.85)**
Years:			
2002 (%)	0	6.45	**6.05**
2003 (%)	0	7.55	**7.08**
2004 (%)	0	7.49	**7.02**
2005 (%)	0	7.96	**7.46**
2006 (%)	0	10	**9.38**

2007 (%)	0	10.67	**10**
2008 (%)	15.84	10	**10.37**
2009 (%)	17.02	9.44	**9.92**
2010 (%)	21.21	9.24	**9.99**
2011 (%)	19.46	9.31	**9.94**
2012 (%)	21.42	9.59	**10.33**
2013 (%)	5.05	2.3	**2.47**
Admitting categories:			
Urgent (%)	91.21	83.91	**84.37**
Others (%)	8.79	16.09	**15.63**
Comorbidity:			
Yes (%)	54.15	86.78	**84.74**
No (%)	45.85	13.22	**15.26**
Admitting entry code:			
Emergency room (%)	7.94	5.50	**5.65**
Clinic (%)	36.56	30.37	**30.75**
Others (%)	55.50	64.13	**63.60**
Arrived via ambulance			
Yes (%)	28.79	18.57	**19.21**
No (%)	71.21	81.43	**80.79**
Diagnoses			
Preterm delivery (%)	15.02	75.85	**72.06**
Preterm diagnosis (%)	52.65	9.16	**11.87**
False labour < 37 wk (%)	32.33	14.99	**16.07**

Table 4.L.3: Likelihood of receiving an fFN test at varying capacity to conduct fFN testing

Being tested	Odds ratio	P-value	95% confidence interval	
			Lower	Upper
Availability: kit + analyzer as reference				
None	0.23	0.000	0.14	0.36
Kit only	0.50	0.040	0.25	0.97

Number of episodes = 39,369; number of women = 32,084; Wald chi2 (12) = 1341.86; P = 0.000. Controlled for age, year, admitting category, admitting entry code, arrived via ambulance, diagnosis, and comorbidity. Distributions of covariates are shown in Table 4.L.2.

References

Currie GR. *Fetal fibronectin testing for the diagnosis of preterm labour: A review of the economics literature and cost analysis for Alberta.* Edmonton AB: Institute of Health Economics, 2006.

Singer E, Pilpel S, Bsat F, Plevyak M, Healy A, Markenson G. Accuracy of fetal fibronectin to predict preterm birth in twin gestations with symptoms of labor. *Obstet Gynecol.* 2007;109(5):1083-7.

SECTION FIVE: Key Findings, Conclusion and Recommendations

Literature Review Update

- Specificity and negative predictive value estimates were high for both the $TLi_{IQ}^{®}$ System and the Actim™ Partus test at most clinical endpoints of interest and did not differ greatly between the two tests, meaning they performed well in predicting the majority of women who were not at risk of PTD.
- Sensitivity and positive predictive values were poor for both tests at all clinical endpoints of interest meaning they did not perform well in predicting the majority of women who were at risk of PTD.
- Compared to the $TLi_{IQ}^{®}$ System, the performance of the Actim™ Partus test was associated with a greater number of false positive results for PTD before 35 or 37 weeks of gestation, and for PTD within 7 days or within 14 days from sampling/testing.
- The LR+ values the $TLi_{IQ}^{®}$ System also had were greater than 6.0 for predicting risk of delivery before 35 or 37 weeks of gestation and within 7 days from testing, while the LR+ values for the Actim™ Partus test were lower than 3.0 for these clinical endpoints, meaning that the $TLi_{IQ}^{®}$ System was more accurate in predicting risk of PTD.
- According to results reported by three Canadian comparative studies, the overall accuracy of the $TLi_{IQ}^{®}$ System in predicting PTD in symptomatic women appears to be higher in comparison to the Actim™ Partus test. Hence, in terms of diagnostic performance, there is no evidence to suggest that the system adopted in Alberta should be changed.

Key Informant and Stakeholder Interviews

- All RHAs with representatives who participated in the study fully implemented testing for preterm labour policy within the timeframe indicated by AH (2006-2008). All RHAs chose fFN testing to manage patients with symptoms of preterm labour. Availability of fFN testing equipment (that is, specimen collection kits and analyzers) varied across RHAs.
- The majority of RHAs absorbed fFN equipment costs through the existing budgets of Women's Health, Obstetrical Programs, and Laboratory Services Program areas. One RHA allocated additional dollars to programs to purchase fFN test kits and analyzers.
- RHAs trained staff on fFN test collection and analysis using multiple lines of communication (for example, memos and rounds, vendor presentations, orientation session, clinical staff educators, video health teleconferencing, and e-learning opportunities). RHAs mainly drew from the Alberta Perinatal Health Program (APHP) and vendor fFN testing educational resources to train staff and develop fFN testing protocols.
- All RHAs provided training to obstetrical physicians and nurses and laboratory staff during policy implementation. Informants explained that family physicians may not have received training directly, but would have had access to APHP, vendor, and More^{OB} Program materials. After the policy implementation period, training for fFN testing mainly occurred through new staff orientations.

- Policy implementation was not formally monitored by the government or the APHP; however, some RHAs recorded information pertaining to fFN test usage for procurement purposes; there was no formal process, follow up, or evaluation plan of the service impacts.
- A few factors that facilitated policy implementation in the RHAs included:
    - existing research that demonstrated the efficacy of the fFN testing and encouraged perceptions of its legitimacy among health care providers;
    - local proponents of fFN testing who championed the establishment of this preterm labour testing option;
    - intensive staff education efforts;
    - existence of different groups within RHAs that helped organize and communicate policy implementation;
    - immediate availability of testing equipment after the policy directive was issued; and,
    - an organizational culture that aimed to increase the efficiency of health care service delivery.
- Policy implementation barriers identified by informants included:
    - the costs of fFN testing equipment;
    - lack of access to fFN analyzers in some of the facilities in RHAs;
    - determining how many tests to order and process at the site level;
    - training staff to use the test only when appropriate; and,
    - ongoing staff education, particularly in sites that manage fewer births on average.
- According to key informants, obstetricians and family physicians generally trusted the fFN test and consider fFN test results as part of their routine for managing patients with symptoms of suspected preterm labour.
- While most informants did not believe that policy implementation resulted in unintended consequences, a few individuals suspected that the policy has not prevented all unnecessary hospital transfers, as not all hospital sites are equipped with an fFN test analyzer.

Economic Analysis

- Physicians placed greater significance on positive test results (inappropriate use of test) compared to negative test results (appropriate use of test), despite the fact that the clinical utility of fFN testing is predicated on a high specificity (approximately 90% reported in the literature, and 98% calculated from administrative databases). Note literature review results that test has low sensitivity and positive predictive values.
- fFN testing did not reduce the number of unnecessary ambulance transfers or admissions for preterm pregnancies in false labour. Unnecessary ambulance transfers increased due to the significance placed on a positive fFN test result. Considerations of other factors such as geographic distance to a level D facility may attenuate the utility of fFN testings because of the time needed to return to a level D hospital if sent home. Level D facilities are those with full obstetrical services and access to tertiary care.

- fFN testing increased the number of appropriate ambulance transfer and admissions for preterm pregnancies in true labour.
- Total health system costs increased due to the purchasing of the fFN kits and analyzers, increased resource utilization associated with the unanticipated increase in appropriate clinical care (for example, appropriate transfers and admissions), and increased resource utilization associated with increasing unnecessary care. Increase in appropriate care resulting from positive fFN test results was an unintended consequence.
- The potential maximum cost savings resulting from fFN adoption were small at the outset.

Conclusion

The PPIR suggests that from a test performance characteristics perspective, there is no new evidence from the literature suggesting that the system adopted in Alberta should be changed. Nevertheless, the adoption of fFN testing in Alberta did not achieve its intended aims of reducing unnecessary utilization of health services to achieve health system savings due to physicians placing greater significance on positive test results compared to negative test results. This bias also contributed to the inadvertent increase in health care utilization overall. Thus, when factoring the costs of fFN testing as well, the total cost for the health system increased. Although there were efforts to provide adequate staff training during implementation, the lack of formally monitoring the implementation, utilization and associated outcomes of fFN testing was a barrier to achieving the intended objectives.

Recommendations

1. If the current way that fFN testing is used in the province continues, then fFN testing should be eliminated from the system as it has not produced any of the intended benefits nor would it in the future.
2. There is potential value in fFN testing given its high specificity. We recommend that education and training to clinicians be conducted to mitigate the inherent confirmatory bias particularly in light of fFN having poor sensitivity. Any potential benefits of fFN testing will be completely dependent on clinicians using and interpreting the test appropriately. Ideally, education and training would be conducted periodically.
    - Note that if clinicians ultimately do not trust the test (negative test result), trust it wrongly (positive test result), or if there are other factors that supersede the fFN test result, then fFN testing provides no value that would justify its cost and investment, and education and training may not produce the desired effect.
3. Assuming Recommendation #2 is followed, all centres that perform fFN testing should have both kits and analyzers available. Equipping centres with kits but no analyzers is a sunk cost (testing is not ordered in these centres – refer to Appendix 4.K), and may actually create confusion in subsequent care (for example, women presenting at level D hospitals holding their fFN sample to be analyzed, personal communication, fFN working group, October 29, 2014).
4. Assuming Recommendation #2 is followed, ongoing monitoring and assessment of its performance should be conducted. This prospective approach would remedy some of the limitations of our retrospective approach while providing information closer to real time that would allow for early, iterative, and ongoing adjustment of fFN testing services. There must

be a mechanism for ongoing management and assessment of fFN testing that can feed back to ordering physicians as well as health system managers.

5. If fFN testing were to be removed from the system, an assessment should also be conducted to evaluate the impact of its disinvestment.

Author Contribution Statements

Paula Corabian contributed to study conception and design, data analysis, and interpretation of the literature review update.

Bing Guo contributed to study conception and design, data analysis, and interpretation of the literature review update.

Dagmara Chojecki conducted the search for the literature review update.

Lindsay Wodinski contributed to the study conception and design, data analysis, and interpretation of the key informant and stakeholder interviews.

Margaret Wanke contributed to study conception and design, data analysis, and interpretation of the key informant and stakeholder interviews.

Thanh Nguyen contributed to study conception and design, statistical analysis, and interpretation of the economic analysis.

Anderson Chuck contributed to the conception and design of the overall evaluation. He also contributed to the statistical analysis and interpretation of the economic analysis.

This report provides an assessment of how the adoption of fFN testing in Alberta impacted the clinical management of preterm labour and health system resources

Institute of Health Economics
1200 – 10405 Jasper Avenue
Edmonton AB Canada T5J 3N4
Tel. 780.448.4881 Fax. 780.448.0018
info@ihe.ca

www.ihe.ca

ISBN 978-1-926929-52-1 (online)