

INNKOMST-CTG

**En vurdering av testens prediktive verdier,
reliabilitet og effekt**

Betydning for jordmødre i deres daglige arbeid



Ellen Blix

**DrPH-avhandling ved Nordiska högskolan för folkhälsovetenskap,
Göteborg, 2006**



NHV
Nordiska högskolan
för folkhälsovetenskap



INNKOMST-CTG

**En vurdering av testens prediktive verdier,
reliabilitet og effekt**

Betydning for jordmødre i deres daglige arbeid

Ellen Blix

DrPH-avhandling ved Nordiska Högskolan för folkhälsovetenskap,
Göteborg, 2006



INNKOMST-CTG

En vurdering av testens prediktive verdier, reliabilitet og effekt. Betydning for jordmødre i deres daglige arbeid
DrPH-avhandling

©Ellen Blix

Nordiska högskolan för folkhälsovetenskap
Box 12133
SE-402 42 Göteborg
Sverige
www.nhv.se

Digitaltrykk: Lundblad Media AS, Tromsø, Norge
Tlf: (+47) 77 75 32 50

ISBN 91-7997-149-0
ISSN 0283-1961

SAMMENDRAG

Formål: Innkomst-kardiotokografi (CTG) innebærer at kvinnen får tatt en CTG-registrering av 20-30 minutters varighet ved innleggelsen i fødeavdelingen. Formålet med denne avhandlingen var å vurdere i hvilken grad rutinemessig innkomst-CTG kan forutsi uønska fødselsutfall og om testen kan forebygge disse utfallene (effektivitet). Avhandlingen undersøkte også inter-observatør enighet (reliabilitet) for testen, og hvilken betydning den kan ha for jordmødre i deres daglige arbeide.

Materiale og metoder: Journalene til alle kvinnene som fødte ved Hammerfest sykehus i årene 1996-1998 ble gjennomgått retrospektivt. Det ble registrert hvordan innkomst-CTG ble tolket, inngrep i fødselsforløpet og utfall for mor og barn (studie I). To erfarne leger tolket et utvalg CTG-utskrifter uavhengig av hverandre (studie II). Seks jordmødre og leger som hadde gjennomgått et opplæringsprogram i tolking av CTG, tolket et annet utvalg utskrifter (studie III). Det ble kalkulert inter-observatør enighet mellom observatørene, og i hvilken grad hver observatør kunne forutsi uønska fødselsutfall. Det ble gjort en systematisk oversikt over forskningsresultater om innkomst-CTG (studie IV). Det ble også gjort kvalitative intervjuer med 12 jordmødre, intervjuene ble tolket i følge "grounded theory" metoden (studie V).

Resultater: Med noen få unntak, var sensitivitet, positiv prediktiv verdi og likelihood ratios for innkomst-CTG dårlig (studie I-IV). Testens evne til å forutsi uheldige utfall var altså lav. Enighet i tolking av innkomst-CTG var dårlig i studie II, mens den var rimelig god når observatørene hadde gjennomgått et opplæringsprogram i tolking av CTG (studie III). Kvinner som ble randomisert til innkomst-CTG ved innleggelsen i fødeavdelingen fikk oftere epiduralbedøvelse, kontinuerlig CTG-overvåking og tatt fosterblodprøve enn kvinner som ble randomisert til auskultasjon ved innleggelsen. Det var ingen forskjeller i utfallene hos de nyfødte (studie IV). Jordmødre kan oppleve motsetninger i seg selv eller mellom seg selv og andre når de forholder seg til rutinemessig bruk av innkomst-CTG. Motsetningene opplevdes mellom yrkesidentitet og den økende bruken av teknologi, mellom trygghet og utrygghet og mellom makt og avmakt (studie V).

Konklusjoner: Det finnes ikke vitenskapelig kunnskap som tyder på at innkomst-CTG er nyttig for lavrisiko fødende. Testen forebygger ikke uønska utfall og kan heller ikke forutsi disse utfallene. Det finnes for lite forskning til at man kan konkludere om testen er nyttig hos risikofødende. Når testene ble tolket av jordmødre og leger som hadde gjennomgått et standardisert opplæringsprogram i tolking av CTG, var testens reliabilitet (eller inter-observatør enighet) god. Fødselshjelpere må være klar over at det vil være en viss uenighet i tolking av innkomst-CTG, spesielt når det gjelder unormale tester. Intervjuene med de 12 jordmødrene avdekket at noen jordmødre syntes det var vanskelig å tolke CTG.

Makthierarkiet i fødeavdelingen påvirket bruk og tolking av innkomst-CTG, og noen jordmødre opplevde at jordmødre generelt ble dårligere i tradisjonelle kliniske ferdigheter på grunn av økende bruk av obstetrisk teknologi.

Nøkkelord: Kardiotokografi; innkomst-CTG; screening; prediktive verdier; inter-observatør enighet; kappa; jordmorpraksis.

DrPH-avhandling ved Nordiska högskolan för folkhälsovetenskap, Göteborg, 2006.

Ellen Blix, Hammerfest sykehus, Helse Finnmark HF.

E-post: eblix@barentsnett.no

ABSTRACT

Aims: The labour admission test comprises a cardiotocography (CTG) of 20-30 minutes' duration carried out upon admission to the maternity ward. The aim of the present dissertation was to assess to what degree the labour admission test can predict adverse outcomes, the effectiveness of the test in preventing adverse outcomes, the inter-observer agreement (or the test's reliability) and to explore what information and knowledge the labour admission test is perceived to provide, and what meaning it carries in the daily work of practicing midwives.

Material and methods: The patient records of all women who gave birth at Hammerfest Hospital in 1996, 1997 and 1998 were retrospectively reviewed. Variables like how the labour admission test was classified, obstetric interventions, maternal and neonatal outcomes were registered (Paper I). Two obstetricians assessed a sample of CTG-tracings (Paper II), and six midwives and obstetricians who had completed a standardised training program assessed another sample of tracings (Paper III). The observers did not have any contact with each other while doing the assessments. Inter-observer agreement, and to what degree each observer could predict adverse outcomes, was calculated. A systematic review was performed to assess the effectiveness of the labour admission test in preventing adverse outcomes compared with auscultation only and to assess the test's prognostic values (Paper IV). Twelve practising midwives were interviewed in order to explore how the labour admission test is perceived. The interviews were analysed using the grounded theory technique.

Findings: Except for a few single study outcomes, sensitivity, positive predictive values and likelihood ratios were low (Papers I-IV). Paper II found poor inter-observer agreement between the two observers, while inter-observer agreement were fair to good when the tracings were assessed by six observers who had completed a standardised training program. Women randomised to the labour admission test were more likely to have minor obstetric interventions like epidural analgesia, continuous electronic fetal monitoring and fetal blood sampling while there were no significant differences in neonatal outcomes (Paper IV). The core category "experiencing contradictions" was identified, indicating that the midwives found conflicting interests within themselves, or between themselves and others when using the labour admission test. They experienced contradictions between professional identity and the increasing use of obstetric technology, between feeling safe and feeling unsafe and between having power and being powerless (Paper V).

Conclusions: There is no scientific evidence of benefit from the labour admission test in low risk women. The test performs poorly in preventing adverse outcomes and is a poor predictor of these adverse outcomes. There is not enough research available to draw any conclusions about its use in high risk women. Medical staff should be aware that there is a certain disagreement in interpreting labour admission tests, especially in tests assessed not normal. Interviews with 12 midwives found that interpreting CTG tracings could be difficult, especially for newly qualified midwives. The hierarchy of power in the labour ward influenced the use and interpretation of the test, and some of the midwives felt their professional identity threatened and that midwives in general are losing their traditional skills because of the increasing use of obstetric technology.

Key words: Cardiotocography; labour admission test; screening; inter-observer agreement; systematic review; midwifery practice.

Doctoral thesis at the Nordic School of Public Health, Gothenburg, Sweden, 2006.

Ellen Blix, Hammerfest Hospital, Finnmark Health Trust, Norway.

E-mail: eblix@barentsnett.no

AVHANDLINGEN ER BASERT PÅ FØLGENDE ORIGINALE PUBLIKASJONER

- I Blix E, Øian P. Labor admission test: An assessment of the test's value as screening for fetal distress in labor. Acta Obstet Gynecol Scand 2001;80:738-43.
- II Blix E, Sviggum O, Koss KS, Øian P. Inter-observer agreement in assessment of 845 labour admission tests: comparison between midwives/obstetricians in the clinical setting and two experts. BJOG 2003;110:1-5.
- III Blix E, Øian P. Interobserver agreements in assessing 549 labor admission tests after a standardized training program. Acta Obstet Gynecol Scand 2005;84:1087-92.
- IV Blix E, Reinart LM, Klovning A, Øian P. Prognostic value of the labour admission test and its effectiveness compared with auscultation only: a systematic review. BJOG 2005;112:1595-1604.
- V Blix E, Öhlund LS. Norwegian midwives' perception of the labour admission test. Midwifery 2006 (Epub ahead of print).

Artiklene er gjengitt med tillatelse fra forleggerne.

INNHold

	ORDLISTE	1
1	INTRODUKSJON	2
1.1	Innledning	2
1.2	Undersøkelser av fosterets tilstand	3
1.3	Kardiotokografi	4
1.4	Innkomst-CTG	8
1.5	Formål med fosterovervåking under fødselen	9
1.5.1	Forebygge sykdom og skade	9
1.5.2	Forebygge død	10
1.6	Screening/diagnostiske tester	11
1.6.1	Definisjoner	11
1.6.2	Testegenskaper, prediktive verdier og likelihood ratios	13
1.7	Oppfatning av risiko	14
1.8	Fødselsparadigmet	17
1.9	Oppfatning av begrepet ”normal fødsel”	19
1.10	Medikalisering	20
1.11	Syn på teknologi	21
1.12	Fødselsomsorgen i Norge	23
1.13	Jordmor	24
1.14	Formålet med denne avhandlingen	26
1.14.1	Noen bemerkninger	26
1.14.2	Hovedmål	27
1.14.3	Delmål	27
2	DESIGN, STUDIEPOPULASJONER OG METODE	28
2.1	Oversikt over enkeltstudiene	28
2.2	Studie I	28
2.2.1	Design	28
2.2.2	Studiepopulasjon	28
2.2.3	Metoder	30
2.3	Studie II og III	31
2.3.1	Design	31
2.3.2	Studiepopulasjoner	31

2.3.3	Metoder	32
2.4	Studie IV	34
2.4.1	Design	34
2.4.2	Data	35
2.4.3	Metoder	35
2.5	Studie V	36
2.5.1	Design	36
2.5.2	Metode og utvalg	36
2.5.3	Intervju	37
2.5.4	Dataanalyse	37
2.6	Etiske overveielser	38
3	RESULTATER	39
3.1	Testens evne til å forutsi uønska utfall	39
3.1.1	Når testene ble tolket av fødselshjelperne i den kliniske situasjonen	39
3.1.2	Når testene ble tolket blindt av uavhengige ”eksperter” i en ikke-klinisk situasjon	40
3.1.3	I en systematisk oversikt	40
3.2	Inter-observatør variabilitet	41
3.2.1	Når testene ble tolket av to uavhengige eksperter og av jordmødre/leger	41
3.2.2	Når testene ble tolket av seks observatører som hadde gjennomgått et opplæringsprogram i tolking av CTG	41
3.3	Effekt av innkomst-CTG sammenliknet med auskultasjon ved innleggelsen	45
3.4	Hvordan innkomst-CTG oppleves av jordmødre i deres daglige arbeid	48
3.4.1	Utvalg	48
3.4.2	Kjernekategori og underkategorier	48
4	DISKUSJON	53
4.1	Oppsummering	53
4.2	Diskusjon av resultatene	53
4.2.1	I hvilken grad testen kan forutsi uønska utfall	53
4.2.2	Inter-observatør variabilitet	56
4.2.3	Innkomst-CTG sammenliknet med auskultasjon ved innleggelsen	58

4.2.4	Betydning for jordmødre i deres daglige arbeide	59
4.3	Metodekritikk	62
4.3.1	Styrker og svakheter i eget materiale	62
4.3.2	Studier som vurderer testegenskaper, prediktive verdier og likelihood ratios	63
4.3.3	Studier som vurderer inter-observatør enighet	65
4.3.4	Systematisk oversikt som vurderer effekten av innkomst-CTG sammenliknet med auskultasjon ved innleggelsen	66
4.3.5	Studie som undersøkte hvordan jordmødre opplevde innkomst-CTG i deres daglige arbeid	67
4.4	Generell diskusjon	68
5	KONKLUSJONER	71
5.1	Følger for praksis	71
5.2	Forslag til framtidig forskning	71
5.3	Anbefalinger	72
	TAKK	73
	LITTERATUR	74
	ENGELSK SAMMENDRAG (SUMMARY IN ENGLISH)	86
	ERRATA	93
	 VEDLEGG	
	Originale publikasjoner I-V	
	NHV-rapporter	

ORDLISTE	
Apgar skåre	En systematisk, klinisk vurdering av barnet ett, fem og ti minutter etter fødselen. Hjerterefreknens, respirasjon, muskeltonus, refleksirritabilitet og farge vurderes og gis fra null til to poeng for hver av observasjonene. Fra sju til ti poeng etter fem minutter regnes som normalt
Asfyksi	Den alvorligste formen for oksygenmangel hos fosteret, og som påvirker de sentrale organene
Cerebral parese	En affeksjon av motoriske funksjoner av en ikke-progredierende defekt/lesjon i den umodne hjerne i henhold til etablerte kriterier
CTG	Cardiotokografi: elektronisk, kontinuerlig registrering av fosterhjerter- og rifreknens
EKG	Elektrokardiogram, registrering av den elektriske aktiviteten i hjertemuskelen
Hypoksemi	Oksygeninnholdet i fosterblod faller og påvirker utelukkende arterieblodet
Hypoksi	Oksygenmangel hos fosteret som påvirker de perifere vevne
Intermitterende auskultasjon	Fosterhjerterefreknens blir undersøkt intermitterende, vanligvis hver halve time eller hvert kvarter i åpningsfasen og etter hver ri i pressefasen. Hjerterefreknens undersøkes med jordmorstetoskop eller med dopplerutstyr
Intrapartum dødelighet	Antall fostre som dør under fødselsforløpet per 1000 fødte (både levende og døde)
Kontinuerlig elektronisk fosterovervåking	At fosterhjerterefreknens, rienes hyppighet og eventuelt fosterets EKG blir registrert kontinuerlig og over lengre tid. Dette blir gjort ved hjelp av CTG-apparat eller STAN-apparat
Metabolsk acidose	Opphopning av melkesyre i fosteret som følge av anaerobt (ikke-oksygenavhengig) stoffskifte, som igjen er følge av nedsatt oksygentilførsel til fosteret
Neonatal encefalopati	Forstyrret nevrologisk funksjon hos den nyfødte i de første leveukene og kan gi respirasjonsproblemer, nedsatt tonus og refleks, endret bevissthetstilstand og kramper. Tilstanden graderes i tre grupper
Perinatal dødelighet	Dødsfall hos fosteret/barnet mellom 28. svangerskapsuke og første leveuke. Noen land har endret grensen fra 28. til 24. svangerskapsuke. I den vestlige verden er perinatal mortality rate (PMR) beregnet som: $(\text{Intrauterine dødsfall fra 28. svangerskapsuke} + \text{dødsfall i første leveuke} / \text{dødfødte} + \text{levende fødte}) \times 1000$
STAN [®]	I tillegg til fosterhjerterefreknens og rienes hyppighet registreres foster-EKG kontinuerlig. Dette brukes for beregning av T/QRS-kvoten samt for analyse av ST-segmentet

1 INTRODUKSJON

1.1 Innledning

Acheson¹ har definert folkehelsevitenskap slik:

“Public health is the science and art of preventing disease, prolonging life and promoting health through organised efforts of society” (s 1).

Grunnideen i folkehelsearbeidet er at folkets helse kan påvirkes av organiserte aktiviteter for å forebygge sykdom og for å fremme helse². Å føde barn er en viktig begivenhet i menneskers liv. Vi må anta at både foreldre, fødselshjelpere og samfunnet ønsker at fødselsforløpet skal gå godt for mor og barn. Samfunnet har tatt ansvar gjennom å organisere helsekontroller i svangerskapet, fødselshjelp og helsestasjonstjenester med vaksinasjonsprogram og barnekontroller. Formålet er at svangerskapet, fødselen og barseltiden ikke skal føre til sykdommer eller skader som kunne vært unngått, og at barnet fødes friskt og starter livet med så gode forutsetninger som mulig. I Norge er all helsehjelp i forbindelse med svangerskap, fødsel og helsestasjonstjeneste gratis. Sosial- og helsedirektoratet³ har definert formålet med svangerskapsomsorgen som:

“...å sikre at svangerskap og fødsel forløper på en naturlig måte, slik at morens somatiske og psykiske helse, og hennes sosiale velvære blir best mulig, sikre fosterets helse, slik at det kan fødes levedyktig og uten sykdom eller skade som kunne vært forhindre, oppdage og behandle sykdom og andre helsetruende forhold hos moren, slik at svangerskapet medfører minst mulig risiko for henne og barnet” (s 8).

Fødselsforløpet med livmorsammentrekninger, barnet som skal dreie seg ned fødselskanalen og inn i verden, morkaka med fosterhinner som skal løsne og blødningen fra livmorslimhinnene idet morkaka løsner, regnes som den mest risikofulle perioden både for mor og barn, og den perioden som trenger mest intensiv oppfølging og overvåking⁴. Fødselen deles vanligvis inn i tre faser: Åpningsfasen, pressefasen og etterbyrdsfasen⁴. Åpningsfasen er når livmormunnen åpner seg, livmormunnen regnes som utslettet når den ikke kan palperes ved siden av ledende fosterdel (vanligvis bakhodet). Åpningsfasen deles i latent og aktiv fase. Fødselens start defineres som når kvinnen har regelmessige rier med ti minutters intervall eller mindre⁴, mens start av aktiv fase regnes fra kvinnen har regelmessige rier og mormunnen er minst fire cm åpen⁵. Pressefasen er fra mormunnen er helt åpen og ledende fosterdel er sluttrotert og på bekkenbunnen. Etterbyrdsfasen er tiden fra barnet er født og til etterbyrden, det vil si morkake, navlesnor og fosterhinner er født og blødningen er stoppet⁴.

I åpningsfasen og i pressefasen blir hjerterefrekvensen overvåket intermitterende eller kontinuerlig. Å overvåke fosteret i fødselsforløpet må regnes som forebyggende helsearbeide. Formålet med fosterovervåking er å sikre fødselen av et friskt barn med et minimum av intervensjoner⁶. En av måtene å overvåke tilstanden til fosteret på, er å ta en innkomst-CTG tidlig i fødselsforløpet og i forbindelse med at kvinnen legges inn på fødeavdelingen⁷.

Denne avhandlingen vurderer i hvilken grad innkomst-CTG kan predikere uønska fødselsutfall, om den kan forhindre uønska utfall, hvilken reliabilitet testen har, og hvilken betydning den kan ha for jordmødre.

1.2 Undersøkelser av fosterets tilstand

Fosterhertet er ferdigdannet 46 dager etter konsepsjonen og fosterets puls kalles fosterlyden eller fosterhertelyden. Fosterhjerterefrekvensen forandrer seg hele tiden og reagerer på stimuli som nedsatt tilførsel av oksygen og også på rievirksomhet. Den franske legen Lannec beskrev i 1819 at man kunne få kunnskap om fosterets tilstand ved å lytte på fosterhertelyden⁸. Etter hvert ble det vanlig å undersøke fosterlyden ved hjelp av et kort stetoskop (føtoskop) i svangerskapet og under fødselen, og i begynnelsen av det 20. århundret var auskultasjon etablert praksis i Europa⁹. Det antas at dette er nyttig selv om det aldri har vært vitenskapelig prøvd. I Norden brukes vanligvis et 20-50 cm langt stetoskop dreid i tre når fosterlyden skal telles. I Norge kalles det trestetoskop eller jordmorstetoskop (figur 1). De siste tiårene er auskultasjon ved hjelp av dopplerutstyr blitt stadig mer utbredt.



Figur 1. Undersøkelse av fosterhertelyden med jordmorstetoskop

I svangerskapet kan man også vurdere tilstanden til fosteret ved å undersøke hvor ofte det sparker, måle om livmora vokser som den skal og gjøre ultralydundersøkelse for å vurdere foster, fostervann og morkake. Dette blir gjort rutinemessig, ved tegn til avvik henvises kvinnen til spesialundersøkelser som for eksempel ultralyd eller dopplerundersøkelse av blodgjennomstrømming av blodkar i det føto-maternelle kretsløpet. I

fødselsforløpet er undersøkelse av fosterhertelyden den vanligste måten å vurdere tilstanden til fosteret. Man kan også rive hull på fosterhinnene for å vurdere mengde og farge på foster-

vannet, ta blodprøver fra fosterskalpen for å vurdere om fosteret får tilført nok oksygen, ta elektrokardiogram (EKG) av fosterhjertet, gjøre ultralydundersøkelser og dopplerundersøkelser.

Man skiller mellom kontinuerlig fosterovervåking og intermitterende auskultasjon. Ved kontinuerlig overvåking blir fosterhertelyden (og eventuelt foster-EKG) registrert hele eller store deler av fødselsforløpet ved hjelp av CTG-apparat eventuelt STAN[®]-apparat¹⁰. Intermitterende auskultasjon innebærer at fosterhertelyden blir undersøkt med jevne mellomrom ved hjelp av jordmorstetoskop eller dopplerutstyr. Man kan også bruke CTG-apparatet til intermitterende auskultasjon, ved å holde transduceren på magen til mor. Vanligvis blir fosterhertefrekvensen undersøkt i 15-60 sekunder med 15 til 30 minutters intervall i åpningsfasen, og etter hver ri i pressefasen.

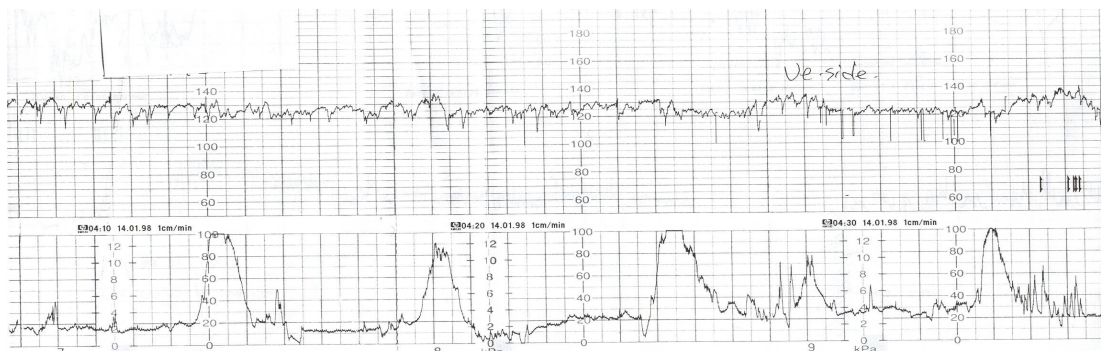
1.3 Kardiotokografi



Figur 2. CTG-registrering med utvendige transducere

Kardiotokografi (kardio=hjerte, toco=ri og grafi=utskrift) er teknologi utviklet for overvåking av fosteret under svangerskap og fødsel. Fosterhertelyden og rienes hyppighet overvåkes kontinuerlig. Det er to måter å overvåke fosterhertelyden med kardiotokografi (CTG): Utvendig registrering, hvor to utvendige transducere blir plassert på magen til kvinnen og festet ved hjelp av elastiske belter (figur 2). Fosterhertelyden

blir registrert ved hjelp av ultralydsignaler. Ved innvendig registrering blir en spiralelektrode plassert direkte på fosterhodet etter at fosterhinnene er punktert, og fosterhertelyden blir registrert ved hjelp av elektriske impulser. Riene registreres enten ved hjelp av en utvendig trykkmåler eller et innvendig kateter som legges opp i livmoren¹¹. Transducerne eller elektrodene blir kobla til CTG-apparater. På ett display kan man lese fosterhertelyden i antall slag per minutt rienes hyppighet. Begge deler skrives grafisk ut på en papirstrimmel (figur 3).



Figur 3. Utskrift fra CTG-apparat. Den øverste linja viser fosterhjerterefrekvensen, den nederste riens hyppighet

CTG-strimla skal tolkes, og dette skal gjøres i den kliniske konteksten¹². Som regel trengs minst 15-20 minutters registrering for å få et bilde av hva som skjer. Man vurderer frekvens, variabilitet, akselerasjoner og deselerasjoner sammen med rifrekvensen og ristryken. Dette vurderes i sammenheng med det kliniske bildet (for eksempel farge og mengde på fostervann, om fosteret er vekstretardert eller ikke, ved termin osv). Basalfrekvensen er gjennomsnittlig hjerterefrekvens i minuttet, denne skal være registrert i minst ti minutter og mellom riene. Normal basalfrekvens er mellom 110 og 150 slag per minutt. Bradykardi er langsom hjerterefrekvens (under 110 slag per minutt) mens takykardi er hurtig hjerterefrekvens (over 150 slag per minutt). Variabilitet er slag-til-slag variasjoner, og båndbredden for variasjonene er et mål på variabiliteten. Normal variabilitet under fødselen er en båndbredde på mellom fem og 25 slag per minutt. Kortidsvariabilitet, eller slag-til-slag variabilitet, er tidsforskjellen mellom hvert hjerteslag. Langtidsvariabilitet er svingninger i basalfrekvens i løpet av lengre tid. En akselerasjon er en forbigående økning av hjerterefrekvensen på mer enn 15 slag i mer enn 15 sekunder. Akselerasjoner er et viktig tegn på at fosteret får tilført tilstrekkelig med oksygen^{10; 13}. En deselerasjon er en forbigående senking av hjerterefrekvensen med mer enn 15 slag i mer enn 15 sekunder. Deselerasjoner kan ha en sammenheng med utvikling av hypoksi. Uniforme deselerasjoner har et avrundet utseende og medfører sjelden et uttalt tap av slag. Utseendet er likt fra gang til gang. Uniforme deselerasjoner deles inn i tidlige og seine deselerasjoner. Tidlige deselerasjoner kommer samtidig med ri og innebærer en refleksutløst senking av hjerterefrekvensen. Tidlige deselerasjoner har vanligvis ingen sammenheng med hypoksi. Seine deselerasjoner starter etter at rien har nådd sitt maksimum, og kan være relatert til hypoksi på grunn av nedsatt blodstrøm gjennom morkaka. Ved variable deselerasjoner reduseres hjerterefrekvensen raskt, og deselerasjonen har ofte et varierende utseende fra gang til gang. Variable deselerasjoner deles

inn i ukompliserte og kompliserte deselerasjoner. En ukomplisert variabel deselerasjon har et slagtap på mindre enn 60 slag og varighet under 60 sekunder. Dette regnes som ufarlig for fosteret. En komplisert variabel deselerasjon varer over 60 sekunder og gir økt risiko for at fosteret blir utsatt for hypoksi. Riregistreringen gir et bilde av rienes hyppighet eller frekvens. Den gir ikke noe sikkert bilde av varighet og intensitet¹⁰.

CTG kan klassifiseres på flere måter. I Norge er det vanlig å klassifisere strimlene som normal, avvikende, patologisk eller preterminal¹⁰ (tabell 1). Innkomst-CTG klassifiseres vanligvis i tre kategorier: Normal, avvikende (eller suspekt) eller patologisk. Kategorien preterminal inngår som patologisk CTG. Tolkning av innkomst-CTG er nærmere beskrevet i studie I, II og III.

Tabell 1. Klassifisering av CTG i følge Sundström og medarb.¹⁰

	Basalfrekvens	Variabilitet Reaktivitet	Deselerasjoner
Normalt CTG	110-150 slag/min	5-25 slag/min Akselerasjoner	Tidlige deselerasjoner, ukompliserte variable deselerasjoner med varighet <60 sek og tap av slag <60 slag
Avvikende CTG	100-110 slag/min 150-170 slag/min Kort episode med bradykardi	>25 slag/min uten akselerasjoner <5 slag/min >40 min	Ukompliserte variable deselerasjoner med varighet <60 sek og tap av slag <60 slag
	En kombinasjon av flere avvikende faktorer gir et patologisk CTG		
Patologisk CTG	150-170 slag/min og nedsatt variabilitet >170 slag/min Bestående bradykardi	<5 slag/min >60 min Sinusoidalt mønster	Kompliserte variable deselerasjoner med varighet >60 sek. Tilbakevendende seine deselerasjoner
Preterminalt CTG	Fullstendig fravær av variabilitet og reaktivitet med eller uten bradykardi		

CTG regnes for å gi god informasjon ved normalt mønster og ved svært avvikende mønster. Ved oksygenmangel hos fosteret er ikke CTG spesifikt nok til å gradere belastningen fosteret utsettes for¹⁰. I slike tilfeller kan måling av pH eller laktat av fosterblod eller måling av foster-EKG gi viktig tilleggsinformasjon^{14; 15}.

Flere studier har vist at det kan være vanskelig å oppnå enighet mellom forskjellige personer i tolking av CTG¹⁶⁻¹⁹. Dette vil bli drøftet nærmere i studie II og III. Tolkningen av CTG blir ofte diskutert og kan være et problem i erstatnings- og rettssaker^{20; 21}.

Kardiotokografen (CTG-apparatet) ble utviklet sist i 1950-årene og kom i kommersiell produksjon fra 1968²². Den ble innført som alternativ til tradisjonell auskultasjon med jordmorstetoskop. Teknologien spredte seg raskt i den vestlige verden uten at det ble gjort noen vitenskapelige vurdering av nytteverdien på forhånd, men med støtte fra obstetrikere som mente at den var nyttig²³. Formålet med denne nye teknologien var å redusere perinatal dødelighet og sykkelighet, spesielt cerebral parese, ved å oppdage oksygenmangel tidlig hos fosteret⁹. Tidlig på 1970-tallet ble det publisert en rekke retrospektive observasjonsstudier som rapporterte færre perinatale dødsfall når kvinnene ble overvåket med kontinuerlig CTG⁹. Allerede i 1979 gjorde de to amerikanske legene Banta og Thacker en systematisk oversikt hvor de vurderte effekten av kontinuerlig CTG-overvåking i fødselsforløpet²⁴. De konkluderte at kontinuerlig CTG-overvåking ikke hadde noen særlig nytteverdi sammenliknet med auskultasjon, og de viste til en økende keisersnittfrekvens som følge av CTG. Banta og Thacker har seinere dokumentert hvordan både rapporten og de som personer ble angrepet av fødselsleger^{23; 25}. Seinere randomiserte studier har heller ikke entydig vist noen nytteeffekt, derimot har en meta-analyse av ni randomiserte, kontrollerte studier som inkluderte 18 561 kvinner som fødte 18 659 barn vist at kontinuerlig CTG-overvåking fører til flere obstetriske inngrep som keisersnitt, vakuumporøsning (sugekoppforløsning) og tang¹⁴. Den økte inngrepsfrekvensen var mest uttalt hos lavrisikokvinner og der foster-pH ikke ble målt i tillegg til kontinuerlig overvåking. Den eneste fordelen ved kontinuerlig CTG-overvåking var at det var færre nyfødtkramper blant de nyfødte som ble overvåket med CTG sammenliknet med de som ble overvåket med intermitterende auskultasjon (23/9461 vs 47/9466, relativ risiko (RR) 0,51, 95 % konfidensintervall (KI) 0,32-0,82). Denne virkningen ble bare sett i studier av høy kvalitet og hvor det var muligheter til å måle pH-verdier i fosterblod. Den største studien inkludert i meta-analysen var den såkalte Dublin-studien som inkluderte 12 964 kvinner og 13 084 nyfødte²⁶. De fant at risiko for nyfødtkramper bare var forhøyet ved langvarige fødsler (over fem timer) og ved oksytocinstimulerte fødsler som ble overvåket ved hjelp av intermitterende auskultasjon. Meta-analysen¹⁴ fant ikke at kontinuerlig CTG-overvåking gav beskyttelse mot cerebral parese (28/6649 av barna som ble CTG-overvåket vs 17/6676 som ble intermitterende auskultert, RR 1,66, 95 % KI 0,92-3,00). En studie²⁷ har gjort en fire-års oppfølging av barna fra Dublin-studien²⁶. De fant at av de ni barna i CTG-gruppen som hadde nyfødtkramper, var det tre som utviklet cerebral parese. Av de 21 barna i auskultasjonsgruppen med nyfødtkramper var det også tre som fikk cerebral parese. I tillegg var det 16 barn (åtte fra hver gruppe) som utviklet cerebral parese uten at de hadde hatt noen tegn til oksygenmangel under fødselen. Forfatterne konkluderte at

kontinuerlig CTG-overvåking ikke hadde noen beskyttende effekt mot cerebral parese sammenliknet med intermitterende auskultasjon.

Det blir ikke systematisk registrert hvor stor andel av norske kvinner som blir CTG-overvåket under fødselen, men det dreier seg muligens om over 60 % (Ulla Vinding Huseby, personlig meddelelse og Ellen Blix, upubliserte data). Nå er det utviklet teknologi som i tillegg til å registrere fosterlyd og rier også registrerer og tolker foster-EKG. Denne teknologien ser ut til å være noe bedre sammenliknet med CTG alene og har vist å redusere andelen barn født med alvorlig metabolsk acidose¹⁵, men resultatene bør bekreftes i flere studier.

Denne avhandlingen dreier seg om innkomst-CTG, og ikke om CTG-overvåking ellers i fødselsforløpet.

1.4 Innkomst-CTG

Innkomst-CTG er en del av undersøkelsen av kvinnen ved innleggelsen i fødeavdelingen. Ved innkomsten undersøkes vanligvis kvinnens blodtrykk, puls og urin. Fosterets leie, presentasjon og størrelse blir vurdert ved hjelp av Leopolds håndgrep. Rienes hyppighet og intensitet vurderes, og det blir gjort vaginal undersøkelse for å vurdere hvor langt fødselen er kommet⁴. Man vurderer hvordan kvinnen har det og hvordan hun mestrer fødselsarbeidet.

Innkomst-CTG innebærer at kvinnen får gjort en CTG-registrering av 20-30 minutters varighet (noen ganger lengre) ved innleggelsen i fødeavdelingen. Denne undersøkelsen ble introdusert i Norge i slutten av 1980-årene. Det vitenskapelige grunnlaget for å ta testen i bruk var hovedsakelig en observasjonsstudie fra Singapore⁷, samt antakelig en sterk tro på at den var nyttig for de fødende og barna deres. I Sverige er innkomst-CTG standard praksis på alle fødeavdelinger²⁸, det er vanlig ved de fleste avdelinger i Norge og i det øvrige Norden, og også svært utbredt i resten av den vestlige verden²⁹.

Den internasjonale sammenslutningen av gynekologer og obstetrikere publiserte retningslinjer for bruk av elektronisk fosterovervåking i 1987¹². De skriver at en kort, ikke-invasiv CTG-registrering ved innkomst for å identifisere uventa kronisk føtal hypoksi hos lavrisiko kvinner blir av mange ansett for å være nyttig.

I 1997 publiserte Statens helsetilsyn utredningen ”Faglige krav til fødeinstitusjoner”³⁰ Her anbefales det at det ved kvinneklinikker og fødeavdelinger tas innkomst-CTG på alle kvinner uavhengig av risiko, mens det på fødestuer ikke er nødvendig. Anbefalingene blir altså gitt ut fra nivå på fødeinstitusjonen, og ikke ut fra hvilken risikogruppe kvinnen er i.

Veilederen fra Den norske lægeforening fra 1998 skriver at en innkomst-CTG av 10-20 minutters varighet kan være nyttig³¹. Nye retningslinjer er underveis og vil bli publisert i løpet av 2006³². Her vil ikke innkomst-CTG bli anbefalt til lavrisiko kvinner.

I mai 2001 publiserte den britiske foreningen for obstetrikere og gynekologer forskningsbaserte retningslinjer for bruk av elektronisk fosterovervåking⁹. Retningslinjene slår fast at innkomst-CTG dårlig kan forutsi hvilke fostre som blir utsatt for hypoksi under fødselen, og at det ikke finnes vitenskapelig grunnlag for å anbefale rutinemessig innkomst-CTG for lavrisiko fødende. Anbefalingen var basert på tre studier^{7; 33; 34}.

En måned seinere, i juni 2001, publiserte den svenske Socialstyrelsen retningslinjer hvor innkomst-CTG ble anbefalt for alle fødende kvinner³⁵. De svenske anbefalingene var basert på sju studier^{7; 36-41}.

Det er gjort en rekke observasjonsstudier som vurderer prediktive verdier og noen få randomiserte studier som vurderer effekten av innkomst-CTG; disse blir gjort rede for i studie IV. Gjennom grundige litteratursøk (studie IV) ble det også funnet en studie⁴² som vurderte enighet i tolking av innkomst-CTG mellom jordmødre i den kliniske situasjonen og en uavhengig ekspert. Studien inkluderte 231 lavrisikofødende og hadde data slik at kappa-koeffisienten kunne regnes ut. Kappa var 0,47 som uttrykker rimelig til god enighet mellom observatørene. Kappa-koeffisienten forklares nærmere i kapittel 2.3.3. Det ble også funnet en studie som blant annet undersøkte jordmødres holdninger til innkomst-CTG⁴³. Den fant at hvorvidt jordmødrene brukte innkomst-CTG var påvirket av retningslinjer og ledernes holdninger. Det ble ikke funnet studier om hvordan kvinner opplever innkomst-CTG.

1.5 Formål med fosterovervåking under fødselen

1.5.1 Forebygge sykdom og skade

Formålet med fosterovervåking under fødselen er å oppdage hypoksi hos fosteret slik at fødselshjelperne kan gripe inn før dette utvikler seg til en varig hjerneskade⁶.

Cerebral parese er definert som en affeksjon av motoriske funksjoner av en ikke-progredierende defekt/lesjon i den umodne hjerne i henhold til etablerte kriterier⁴⁴. Tidligere antok man at den viktigste årsaken til cerebral parese var skader under fødselen⁹. En studie⁴⁵ fra Vestfold fylke fant en ikke-signifikant endring av cerebral parese hos nyfødte med fødselsvekt ≥ 2500 g, fra 1,8 per 1000 i 1970-74 til 1,4 per 1000 i 1995-99. Den samme studien viser at av til sammen 70 824 fødte, var årsaken til cerebral parese hypoksisk-iskemisk encefalopati hos 33 barn, altså muligens direkte forbundet til forhold under fødselen. Dette gir en forekomst på 0,5 per 1000. Andre studier har vist lavere forekomst av tilfeller

som kan relateres til fødselen, 0,1 per 1000⁹. Det er ofte vanskelig å finne ut når skaden som forårsaket cerebral parese oppsto, og det er strenge kriterier som skal oppfylles før man med sikkerhet kan si at skaden oppsto under fødselen⁴⁶. Det antas at i 90 % av tilfellene er det ikke intrapartum årsaker til cerebral parese. I de resterende 10 % oppsto skaden enten før fødselen eller under fødselen, ofte er det vanskelig å finne tidspunktet i ettertid⁴⁶.

Apgar skåre er en systematisk, klinisk vurdering av barnet ett, fem og ti minutter etter fødselen⁴⁷. Hjerterefrekvens, respirasjon, muskeltonus, refleksirritabilitet og farge vurderes og gis fra null til to poeng for hver av observasjonene. Apgar skåre ved fem minutter regnes som mer prediktiv for barnets prognose enn etter ett minutt. Fra sju til ti poeng etter fem minutter regnes som normalt. Undersøkelsen er subjektiv, og forskjellige fødselshjelpere vil nok vurdere litt forskjellig. Forekomsten av nyfødte med Apgar skåre < 7 etter fem minutter har vært svært stabil over tid. I 1980 hadde 1,0 % av alle fødte under sju poeng etter fem minutter, i 1990 1,0 % og i 2002 1,3 % (Steinar Nilssen, Medisinsk fødselsregister, 24.11.05).

Når det blir oppdaget indirekte tegn til oksygenmangel hos barnet under fødselen, vil fødselshjelperne vurdere hvilke tiltak som skal gjøres for å forhindre at dette utvikler seg til varig skade eller død. Det kan være aktuelt å forsøke stillingsendring hos den fødende, redusere eller stoppe pågående oksytocininfusjon, men ofte vil tiltaket være å forløse operativt, enten ved hjelp av keisersnitt eller operativ vaginal forløsning dersom forholdene ligger til rette for det. Andelen kvinner forløst med keisersnitt har vært sterkt stigende de siste tiårene. I 1970 ble 2,2 % forløst ved keisersnitt, 8,3 % i 1980, 12,6 % i 1990⁴⁸ og 15,8 % i 2002⁴⁹. Andelen operative vaginale forløsninger (tang- og vakuumforløsning) var 3,8 % i 1970, 7,1 % i 1980, 7,7 % i 1990⁴⁸ og 7,8 % i 2002⁴⁹.

1.5.2 Forebygge død

Dersom en hypoksi ikke oppdages, eller at hensiktsmessige tiltak ikke blir satt i verk, kan fosteret utvikle asfyksi og i verste fall dø under eller etter fødselen. Et av formålene med fosterovervåking er å forebygge perinatale dødsfall⁹. Perinatal dødelighet er antall dødfødte og døde i løpet av første leveuke per 1000 fødte (både levendefødte og dødfødte)⁵⁰. Perinatal dødelighet er et av verdens største helseproblemer og rammer først og fremst fattige land. I 1999 var det omkring 126,6 millioner fødsler i verden, av disse barna døde 6,9 millioner perinatalt⁵¹. Fire og en halv million av dødsfallene skjedde i Afrika og i Sørøst-Asia. Perinatal dødelighet var i 1999 vurdert til å være 52 per 1000 fødsler i hele verden, 79 i Afrika, 67 i Sørøst-Asia og 15 i Europa⁵¹. Man anslår at omkring en tredjedel av dødsfallene skyldes komplikasjoner under fødselen som fører til oksygenmangel hos fosteret⁵². Andre viktige

årsaker er for tidlig fødsel, sykdommer hos mor, infeksjoner og misdannelser. Hypotermi (lav kroppstemperatur) er en viktig medvirkende årsak til perinatale dødsfall⁵². Forhold som tilgjengelighet av forebyggende og behandlende helsehjelp, fattigdom, analfabetisme, krig og kvinnenenes stilling i samfunnet påvirker også perinatal dødelighet⁵³.

I 1900 var perinatal dødelighet i Norge 40 per 1000 fødsler, i 1950 30 per 1000. Fra 1950 og fram til i dag har reduksjonen i perinatal dødelighet vært tre ganger større enn i de foregående 50 år^{48; 49; 54; 55}. I 1970 var perinatal dødelighet 16,7 per 1000, 9,0 i 1980, 5,7 i 1990⁴⁸ og 4,8 i 2002⁴⁹. Årsakene til redusert dødelighet i Norge kan ikke automatisk kun tilskrives bedre obstetrisk og pediatrik innsats⁵⁰. Årsakene er mange, som bedre økonomi for folk flest, høyere utdanningsnivå, bedre boliger, andre sosiale forhold, primær- og sekundær helsetjeneste av høy kvalitet, bedre standarder og organisering av helsetilbudet i svangerskap og fødsel samt bedre standarder for nyfødtnedisin. Medisinsk fødselsregister hevder at den perinatale dødeligheten nå er så lav at man ikke kan forvente noe særlig videre fall⁵⁶.

Intrapartum dødelighet er dødfødte som dør under fødselsforløpet (pr 1000 fødte), og er altså en del av den perinatale dødeligheten. Andel fostre som dør under fødselen i Norge, var 1,6 i 1970, 0,7 i 1980, 0,2 i 1990 og 0,1 i 2002 (Steinar Nilssen, Medisinsk fødselsregister, 24.11.05). Andre har rapportert forekomsten på intrapartum dødelighet til 0,8 per 1000⁹. Disse estimatene vil være usikre ettersom det kan være vanskelig å vite om fosteret døde før eller etter at fødselen startet. Tiltak som har vist seg å være nyttige for å forebygge perinatal dødsfall på grunn av oksygenmangel under fødselen er at kvinnen har en kyndig fødselshjelper ("skilled birth attendant") og at hensiktsmessige tiltak blir satt i verk når det oppstår komplikasjoner⁵⁷.

1.6 Screening/diagnostiske tester

1.6.1 Definisjoner

Screening kan defineres som undersøkelse av mennesker uten symptomer for å finne ut om det er sannsynlig om de har en bestemt sykdom/tilstand. Mennesker som tester positivt, undersøkes nærmere for å bekrefte eller avkrefte diagnosen, og de som har sykdommen behandles. Målet med et screening program er å redusere dødelighet og sykkelighet ved å oppdage og behandle sykdommen tidlig, og det stilles vanligvis krav om at testen skal være ufarlig og at sykdommen det screenes etter skal kunne behandles⁵⁸. Når store befolkningsgrupper eller hele populasjoner screenes, kalles dette for massescreening eller populasjonsscreening. Multippel eller multifasisk screening er når man bruker flere forskjellige screening tester ved samme anledning. "Prescriptive" screening er et program

som har som mål å oppdage en definert sykdom hos tilsynelatende friske individer. Sykdommen skal være av en slik natur at tidlig oppdaging er fordelaktig for behandling og prognose⁵⁹.

En screening test er ikke ment å være diagnostisk. Diagnostisk testing er å finne mer informasjon for å klargjøre tilstand og prognose. Målene ved diagnostisk testing kan defineres som å oppdage eller utelukke sykdom, medvirke til videre diagnostisk og terapeutisk håndtering og vurdere prognose⁶⁰.

Resultatet av en screening test eller diagnostisk test vil være positivt eller negativt, og det er fire muligheter: Det kan være sann-positivt, falsk-positivt, sann-negativt eller falsk-negativt. Sann-positive og sann-negative tester vil sjelden være problematiske. En falsk-positiv urinprøve på eggehvite vil sjelden føre til store problemer, men en falsk-positiv HIV test eller ultralydfunn vil kunne føre til store påkjenninger for kvinnen og familien hennes. Falsk-negative tester kan føre til at tegn til sykdom eller avvik ikke blir tatt alvorlig og at tilstanden man screener for blir oppdaget seint.

I svangerskaps- og fødselsomsorgen utføres en rekke screeninger som for eksempel blodtrykksmåling, undersøkelse av urin for å oppdage infeksjoner eller tegn til preeklampsi (svangerskapsforgiftning), symfyse-fundusmål for å finne vekstretarderte fostre, veiing av kvinnen, blodtype, HIV, rubella, syfilis, ultralyd omkring 18. svangerskapsuke for å bestemme fødselstermin og avdekke store utviklingsavvik, Leopolds håndgrep for å oppdage avvikende fosterleie, helseundersøkelse av de nyfødte og screening for Føllings sykdom og hypothyreose. Det ble ikke nødvendigvis gjort en vurdering av nytteverdien av disse undersøkelsene før de ble innført og gjort til rutine, eller nytteverdien blir ikke tatt i betraktning og undersøkelsene blir videreført på grunn av tradisjon. For eksempel har den svenske jordmor og professor Ulla Waldenström nylig påpekt at rutinemessig ultralyd av gravide ikke har ført til nedgang i sykелighet og dødelighet blant de nyfødte selv om metoden fører til en mer nøyaktig bestemmelse av fødselsterminen⁶¹.

Ingemarsson og medarbeidere⁷ har definert innkomst-CTG som en screening test for føtal distress i fødselsforløpet. Tilstanden "føtal distress" ble definert som operativ forløsning på indikasjon forandret/dårlig fosterhertelyd, eller Apgar skåre < 7 etter fem minutter. De skriver også i samme artikkel at testen kan finne fostre med oksygenmangel allerede ved innkomsten, og derved unngå unødige forsinkelser i håndteringen av disse. Innkomst-CTG kan altså defineres både som screening og som diagnostisk test: Screening fordi den omfatter hele populasjonen av fødende kvinner og fostrene deres, og skal predikere hvorvidt det er sannsynlig at det oppstår sykdom/uønska tilstander i fødselsforløpet (føtal distress).

Diagnostisk test fordi den skal bekrefte eller utelukke sykdom/oksygenmangel hos fosteret allerede ved innkomst, og også medvirke til avgjørelsen om hvilken overvåkingsmetode som er best gjennom fødselsforløpet. Innkomst-CTG defineres som populasjonsscreening ettersom målgruppa er hele populasjonen av fostre. Testen kan også defineres som "prescriptive" screening i hvert fall når truende eller manifest oksygenmangel er sykdommen eller tilstanden det screenes for. Det er rimelig å anta at det er fordelaktig for barnet at oksygenmangel blir oppdaget så tidlig som mulig. Behandlinga vil vanligvis være operativ forløsning.

En screening test eller diagnostisk test kan ha tre mulige følger for populasjonen der den blir gjort:

1. En bedring av sykkelighet og dødelighet. Testen vil da være til fordel for populasjonen, og vil antakelig være verd økonomiske kostnader og eventuelt ubehag og engstelse
2. Testen har ingen effekt på sykkelighet og dødelighet.
3. Testen viser seg å ha skadelige effekter.

1.6.1 Testegenskaper, prediktive verdier og likelihood ratios

Testegenskaper (sensitivitet og spesifisitet), positiv prediktiv verdi (PPV), negativ prediktiv verdi (NPV), likelihood ratio for positiv test (LR+) og likelihood ratio for negativ test (LR-) er nærmere forklart i tabell 2.

Sensitivitet, eller sann positiv rate, er andelen individer med sykdommen/tilstanden som er test-positive. Spesifisitet, eller sann negativ rate, er andelen individer som ikke har sykdommen, og som er test-negative. PPV er andelen individer med positivt testresultat som virkelig har sykdommen, og NPV er andelen individer med negativt testresultat som ikke har sykdommen. Likelihood ratio for positiv test (LR+) er sannsynligheten for positiv test hos et individ som har sykdommen/tilstanden dividert med sannsynligheten for positiv test hos et individ som ikke har sykdommen/tilstanden. Likelihood ratio for negativ test (LR-) er sannsynligheten for negativ test hos et individ som har sykdommen/tilstanden dividert med sannsynligheten for negativ test hos et individ som ikke har sykdommen/ tilstanden⁶².

Testegenskaper og likelihood ratios påvirkes ikke av prevalensen av sykdommen/tilstanden man screener for. De prediktive verdiene påvirkes av prevalensen, derfor regnes likelihood ratios som mer robuste og informative enn de prediktive verdier⁶³. For fødselshjelperne vil det være nyttig å vite hvordan et testresultat predikerer risiko for uønska utfall. Sensitivitet og spesifisitet gjør ikke det, de beskriver hvordan abnormalitet (eller normalitet) predikerer spesielle testresultater⁶⁴. PPV og NPV beskriver sannsynligheten for uønska utfall etter et positivt eller negativt testresultat, men påvirkes av prevalensen i studiepopulasjonen. Jo

høyere prevalens, jo høyere PPV og lavere NPV, og jo lavere prevalens, jo høyere NPV og lavere PPV⁶².

Det er vanskelig å si hva som er ”gode” verdier. Noen av verdiene vil avhenge av prevalensen i populasjonen, og man kan til en viss grad bestemme hvor mange falsk positive og falsk negative tester man tillater. I noen tilfeller kan det viktigste være å finne absolutt alle med sykdommen, og da må man akseptere at andelen falsk-positive kan bli høy. Noen hevder at en generell regel for likelihood ratio er at dersom LR+ er over 10 og LR- under 0,1, er testen overbevisende. Dersom verdiene er henholdsvis over 5 og under 0,2 må testene betegnes som gode. Dette avhenger imidlertid av pre-test sannsynlighet og i hvilken sammenheng testen ble utført⁶⁵.

Tabell 2. Nærmere forklaring av sensitivitet, spesifisitet, positiv prediktiv verdi (PPV), negativ prediktiv verdi (NPV), likelihood ratio for positiv test (LR+) og for negativ test (LR-)

		SYK		
		+	-	
TEST	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	

a=individer som har sykdommen og er test-positive (sann positive)

b=individer som ikke har sykdommen, men er test-positive (falsk positive)

c=individer som har sykdommen, men er test-negative (falsk negative)

d=individer som ikke har sykdommen og er test-negative (sann negative)

a+b=alle test-positive

c+d=alle test-negative

a+c=alle som har sykdommen

b+d=alle som ikke har sykdommen

Sensitivitet= $a/a+c$

Spesifisitet= $d/b+d$

Positiv prediktiv verdi (PPV)= $a/a+b$

Negativ prediktiv verdi (NPV)= $d/c+d$

Likelihood ratio for positiv test (LR+)=sensitivitet/1-spesifisitet, eller $(a/a+c)/(b/b+d)$

Likelihood ratio for negativ test (LR-)=1-sensitivitet/spesifisitet, eller $(c/a+c)/(d/b+d)$

1.7 Oppfatning av risiko

Risikobegrepet brukes når vi vurderer utfall av hendelser som skal skje, eller kommer til å skje i framtiden. Det brukes vanligvis om dårlige eller uønska utfall. Ingen kan vurdere nøyaktig alle farer man kommer til å møte, så hvordan avgjør vi hvilke forholdsregler vi skal ta?

Oppfatningen av risiko i forbindelse med barnefødsler er hele tiden i forandring. I Norge var det fram til 1970 vanlig å føde på fødestue. Ei fødestue har jordmor til stede og

allmennlege i vakt, og er uten obstetriker, anestesilege, kirurg eller pediater. I 1970 foregikk 13,5 % av fødslene på fødestue, i 1980 3,5 %, i 1990 1,2 % og i 1998 1,0 %⁴⁸. I begynnelsen av 1970-tallet startet sentraliseringen av fødeinstitusjonene i Norge. Fra 1971 til 2000 ble antallet redusert fra omkring 150 til 66⁶⁶. Fra 1980 til 1994 ble antallet fødestuer redusert fra 30 til ti⁶⁷. Hovedargumentet var at det er vanskelig å forutsi fødselskomplikasjoner og at det er tryggest å føde på steder hvor det er tilgang til operative forløsninger og anestesitjeneste. Det finnes svært lite litteratur som underbygger denne påstanden⁶⁷. Sosiologen Lise Kjølørød har lansert begrepet ”verstefallstenkning” i fødselsomsorgen og hevder at dette var grunnen til at reduksjonen i antall fødestuer startet i 1970-årene⁶⁸. Verstefallstenkningen baserer seg på at fødsel innebærer risikomomenter, og det eneste forsvarlige er å sikre mor og barn mot det som hele tiden i verste fall kan skje.

I 2001 vedtok Stortinget at fødselsomsorgen i Norge skal være desentralisert og differensiert⁶⁹. En del små fødeavdelinger med færre enn 400 fødsler er blitt omorganisert til jordmorstyrte fødestuer med fødetilbud til lavrisikofødende. Mange støtter reformen, men den har også ført til protestaksjoner fra leger, jordmødre og legfolk. Et av hovedargumentene er at reformen fører til risikofylte fødsler på fødestuene. Norske forskere konkluderer forskjellig med hensyn til hvor trygt det er å føde på fødestuer.

En stor retrospektiv studie⁷⁰ som inkluderte 1,25 millioner fødsler i årene 1972 til 1995 fant at det var lavest neonatal dødelighet etter fødsler ved institusjoner som hadde mellom 2000 og 3000 fødsler årlig. Det var høyest dødelighet ved institusjoner som hadde mindre enn 100 fødsler årlig, og nest høyest ved institusjoner som hadde over 3000 eller mellom 100 og 500 fødsler årlig. Denne studien⁷⁰ opplyste ikke hvorvidt kvinnene faktisk fødte der de var selektert til å føde. En prospektiv studie inkluderte 1275 kvinner som startet fødselen ved norske fødestuer⁷¹. Denne fant gode utfall hos lavrisikofødende som selekteres til å føde på fødestue, og konkluderte at dette var forsvarlig praksis. Det er mulig at andel fødende som regnes som lavrisiko er i ferd med å endre seg slik at stadig færre kvinner regnes som lavrisikofødende. En nylig publisert studie fra USA viser at andelen kvinner vurdert som lavrisiko gjennom svangerskap og fødsel har sunket fra 46 % i 1991 til 42 % i 2001⁷².

Antropologen Mary Douglas beskriver risikobegrepet ut fra en kulturell teori⁷³. Det sosiale miljøet, seleksjonsprinsippene og det oppfattende individ blir sett på som ett system. Oppfatning av risiko er en kulturell konstruksjon, hva som oppfattes som farlig og risikofullt blir til gjennom en sosial prosess. Som eksempel nevner Douglas Lele-folket i Zaire, som var utsatt for mange sykdommer, feber, tuberkulose, lungebetennelse, sår, spedalskhet med mer. Det de fryktet mest, var bronkitt, infertilitet og å bli rammet av lynet. Årsaken var at disse

tilstandene var forbundet med spesielle moralske forestillinger hvor medlemmer av samfunnet ville bli klandret.

Karl Halvor Teigen beskriver hvordan risiko kan bedømmes i hverdagen⁷⁴. En måte å bedømme risiko er subjektiv oppfatning av sannsynligheter. De fleste mennesker har ikke oversikt over hvor store disse sannsynlighetene faktisk er, og gjør subjektive vurderinger av sjanser. På denne måten kan det lett bli over- og undervurderinger av hva som er risikabelt. En annen måte å vurdere risiko på, er tanker på konsekvensene. Teigen skriver at det i blant kan se ut som om folk legger liten vekt på sannsynlighetene og legger vekt på hva som "kan" skje. Ei jordmor som tidligere har opplevd å ta i mot et uventet dårlig barn, eller er svært engstelig for å oppleve dette, vil antakelig legge stor vekt på denne risikoen. Hun vil kanskje ta flere forholdsregler for å unngå en slik situasjon, som for eksempel lytte på hjertelyden ofte eller kjøre CTG-registreringer.

Denne måten å oppfatte risiko på, står i motsetning til den epidemiologiske forståelse av risiko som et fenomen som kan måles og forklares i tallestimater. Viktige epidemiologiske begreper for å beskrive risiko er absolutt risiko, relativ risiko, absolutt risikoreduksjon og relativ risikoreduksjon. Absolutt risiko er sannsynligheten for at en hendelse skal skje, og sannsynligheten for en hendelse eller et utfall er prevalens dividert med populasjonen⁵⁹. For eksempel var sannsynlighet for perinatal død i Norge i 2002: $267 \text{ (perinatale dødsfall)} / 56\,158 \text{ (fødte)} = 4,2$ perinatalt døde per 1000 fødsler. Relativ risiko er ratio for et utfall hos en gruppe sammenliknet med en annen gruppe⁵⁹. Sannsynlighet for perinatal død var i følge WHO i 2000 80/1000 i Somalia og 6/1000 i Norge⁵¹. Relativ risiko for perinatal død i Somalia sammenliknet med Norge er $80/6 = 13,5$. Risiko for perinatal død var altså 13 ganger høyere i Somalia enn i Norge.

Absolutt risikoreduksjon er den gevinsten man får ved å gjøre en behandling eller en forebygging⁷⁵. Det er for eksempel to måter å forebygge etterbyrdsblødning i forbindelse med fødsel, enten å gi oksytocinprofylakse til alle fødende (generell profylakse), eller å gi oksytocinprofylakse til de kvinnene som har forhøya risiko for blødning (selektiv profylakse). En norsk studie⁷⁶ som omfattet 11 493 kvinner har blant annet sammenliknet disse to regimene. Ved generell profylakse var det 8,2 % av kvinnene som blødde mer enn 500 ml og 0,4 % som blødde mer enn 1500 ml og ved selektiv profylakse var tallene henholdsvis 9,4 og 0,7 %. Dersom man regner selektiv profylakse som en eksponering (for høyere sjanse for å få etterbyrdsblødning), kan man regne ut hvor stor andel etterbyrdsblødninger selektiv profylakse forårsaker: $(9,4-8,2)/9,4 = 13 \%$ av blødningene over 500 ml og $(0,7-0,4)/0,7 = 43 \%$ av blødningene over 1500 ml. Generell profylakse gir altså en absolutt risikoreduksjon på

(9,4 % - 8,2 % =) 1,2 % for blødning over 500 ml og (0,7 % - 0,4 % =) 0,3 % for blødning over 1500 ml. Relativ risikoreduksjon vurderer behandlingsgevinsten i forhold til bakgrunnsrisikoen⁷⁵. Relativ risikoreduksjon ved generell profylakse sammenliknet med selektiv profylakse var (1,2 %/9,4 % =) 13 % for etterbyrdsblødning over 500 ml og (0,3 %/0,7 % =) 43 % for blødning over 1500 ml.

For fødselshjelperne kan det være viktig å vite om kvinnen er risikoføddende eller ikke. Risikovurdering av den enkelte kvinne har imidlertid vist seg å være upresis, bare mellom 10 og 30 % av risikoføddende opplever utfallet de har forhøyet risiko for⁷⁷. For den føddende er opplysningen bare nyttig dersom noe kan gjøres for å redusere den aktuelle risikoen. I verste fall kan risikovurderingen være mer til skade enn til nytte i det den kan føre til unødvendige inngrep og engstelse hos kvinnen⁷⁷.

1.8 Fødselsparadigmet

Begrepet "fødselsparadigmet" brukes for å beskrive at fødselsprosessen kan oppfattes på kvalitativt forskjellige måter. Vanligvis beskrives to ytterpunkter i oppfatningen av fødselsprosessen. Bente Haug⁷⁸ beskriver to ulike oppfatninger hun fant blant norske jordmødre. Den ene oppfatningen var at fødselen er en naturlig, biologisk hendelse som enhver normalt sunn og frisk kvinne vil kunne beherske. Fødselshjelperne skal bare gripe inn i fødselsforløpet dersom det er nødvendig. Den andre oppfatningen var at fødselen er en medisinsk tilstand forbundet med risiko. Denne risikoen reduseres ved at fødselen kontrolleres av fødselshjelpere som gir råd og veiledning underveis.

Gunnhild Blåka drøfter det hun kaller den jordmorfaglige diskursen og den fødselsvitenskapelige diskursen⁷⁹. Ifølge den fødselsvitenskapelige diskursen kan kroppen til den føddende kvinnen feile på ethvert tidspunkt. Fødselshjelperens oppmerksomhet rettes mot de eventuelle uheldige utfall fødselen kan tenkes å få. Den jordmorfaglige diskursen innebærer at jordmor anser at et sett av faktorer er avgjørende for et heldig fødselsutfall. Ifølge Blåka er det viktig å sette en ramme omkring det ventende forløpet, at jordmor tar imot kunnskap fra kvinnen, er avventende og oppmerksom, er i forkant og har et godt overblikk. Jordmor er forberedt på å møte de situasjoner som kan komme og vet at man aldri kan ha full kontroll.

Marsden Wagner⁸⁰ beskriver de to ytterpunktene i oppfatningen av fødselsprosessen som den medisinske og den sosiale modellen. Tilhengere av den medisinske modellen ser fødselen som et medisinsk anliggende. Kvinnekroppen er et komplekst maskineri og kan lett feile, derfor må fødselen foregå under tett medisinsk overvåking. Informasjon fra kvinnen

regnes som kvalitativ og upresis, mens fødselsteknologien gir kvantitativ og objektiv informasjon direkte fra kvinnens kropp og er mer troverdig. Tilhengere av den sosiale modellen hevder at fødsel ikke nødvendigvis er en sykdom eller et medisinsk anliggende og at fødekvinnen ikke nødvendigvis er en pasient. Fødselen er en biososial prosess, det betyr en biologisk, anatomisk, fysiologisk og biokjemisk hendelse som også har mentale og spirituelle komponenter. Fødselsteknologien fører til etiske og praktiske dilemma: Teknologien innebærer også usikkerhet, fordeler og ulemper og er ikke alltid utredet og det kan være vanskelig for kvinnen å ta informerte valg. Den sosiale modellen fokuserer på hensiktsmessig teknologi, helst skal teknologien være ikke-invasiv og akseptabel for kvinnene.

Denis Walsh⁸¹ kaller de to motsatte synene i fødselsparadigmet for den biomedisinske og den kvinnesentrerte modellen. Den biomedisinske modellen innebærer at kroppen blir sett på som en maskin, den er reduksjonistisk i det den legger vekt på de rent fysiske prosessene. Den legger vekt på sikkerhet, fokuserer på patologi og teknologien er en partner. Den kvinnesentrerte modellen blir beskrevet som holistisk idet den integrerer de fysiologiske, psykososiale og spirituelle aspektene, legger vekt på selvrealisering, fokuserer på normalitet og teknologien er en tjener.

Cecil G. Helman⁸² kaller det for den obstetriske fødselskultur og en legkvinnekultur. Motsetningene beskrives som at den obstetriske fødselskulturen ser fødselen som en isolert medisinsk hendelse, mens for kvinnene er den integrert i resten av livet hennes.

Disse to oppfatningene, eller måtene å risikogrudere fødselen på er ikke nødvendigvis to strengt atskilte skoler. Noen mener litt av det ene og litt av det andre avhengig av situasjonen, og er ikke bevisst de forskjellige synene på fødselsprosessen. Et eksempel hvor begge syn støttes er Statens helsetilsyns ”Faglige krav til fødeinstitusjoner”³⁰. Her anbefales at det ved fødeavdelinger og kvinneklinikker bør gjøres innkomst-CTG på alle fødende uavhengig av risiko. Dette kan tolkes som et uttrykk for at fødselen innebærer faremomenter, i dette tilfellet at fosteret av en eller annen grunn har oksygenmangel ved innkomsten i fødeavdelingen. Samtidig anbefales det at det ved fødestuer bør være muligheter til CTG-registrering uten at dette vil være obligatorisk ved alle fødsler. Dette kan tolkes som om at fødselen kan forløpe helt normalt, og at fødende kan selekteres med hensyn til risiko.

Hvilken oppfatning eller diskurs som råder, vil være bestemmende for hvordan fødselshjelpen organiseres og for hvilke observasjoner, tiltak og behandlinger som anbefales for de fødende. I 2001 vedtok Stortinget at Norge skulle ha en desentralisert og differensiert fødselshjelp⁶⁹. Nasjonalt råd for fødselsomsorg hevder at Stortinget herved har tatt et oppgjør med tradisjonell verstefallstenking, og at vi sannsynligvis står overfor et paradigmeskifte i

synet på den normale fødsel⁸³. Det er imidlertid ikke enighet om dette i det norske fagmiljøet, og debatten om risiko forbundet med fødsler og sentralisering pågår^{84; 85}.

1.9 Oppfatning av begrepet ”normal fødsel”

Normal fødsel er ikke et begrep som det hersker enighet om. Verdens helseorganisasjon har laget det de kaller en arbeidsdefinisjon av den normale fødselen⁸⁶:

“We define normal birth as: spontaneous in onset, low-risk at the start of labour and remaining so throughout labour and delivery. The infant is born spontaneously in the vertex position between 37 and 42 completed weeks of pregnancy. After birth mother and infant are in good condition”. (s 4)

I den norske læreboken ”Obstetikk og gynekologi”⁸⁷ defineres den normale fødsel slik:

”Normal fødsel starter og avsluttes spontant, etter fullgått svangerskap, uten instrumenter, med ett levende barn i hode- eller issepresentasjon, etter en varighet på under 20 timer for førstegangsfødende og under 15 timer for annen- og flergangsfødende. Etterbyrden skal fødes hel, uten inngrep, og blødningsmengden hos mor etter fødselen skal ikke overstige 500 ml”. (s 200)

I følge denne læreboken kunne vel 60 % av fødslene i Norge i 2003 kalles normale. Ingen av disse definisjonene tar hensyn til intervensjoner som CTG-overvåking, smertelindring eller stimulering av rier.

Selv om det finnes autoriserte definisjoner på hva som er normalt, vil nok oppfatningen om hva som er normal fødsel variere fra land til land og blant annet avhenge av hvilke ressurser og hjelpemidler fødselshjelpen har. I år 2000 døde 529 000 kvinner som følge av svangerskap eller fødsel, 95 % av dødsfallene var i fattige land, eller det vi kaller for utviklingsland⁸⁸. Dette er selvfølgelig med på å forme oppfatninga av hva som er normalt, både hos fødselshjelperne og hos befolkninga ellers. Noen steder i verden er keisersnitt så utbredt at det antakelig blir regna som en normal måte å få barn på. En tredel av alle fødslene i Brasil foregår ved private sykehus hvor keisersnittfrekvensen er over 75 %⁸⁹.

Keisersnittfrekvensen var 40 % i 2000 for hele Brasil⁹⁰. I enkelte områder i Afghanistan er maternell dødelighet (dødsfall hos en kvinne mens hun er gravid eller inntil 42 dager etter fødselen/aborten) helt oppe i 6500 per 100 000 levende fødte, og kvinnene har verken tilgang til keisersnitt eller særlig mange andre former for inngrep⁹¹. Her regnes antakelig et mødredødsfall som sørgelig, men normalt. I Norge var mødredødeligheten 16 per 100 000 levende fødte i 2000⁸⁸.

1.10 Medikalisering

En definisjon av medikalisering er at moderne medisin har ekspandert og nå tar seg av forhold som tidligere ikke ble regnet som medisinske problemer⁸². ”Medikalisering” har i dag en negativ klang, men det er verd å huske at for noen få generasjoner siden i Norge var fødselen et sosialt, og ikke et medisinsk anliggende. Kvinnene fødte med hjelpekoner og jordmødre som ikke hadde så mye å hjelpe seg med når det oppsto komplikasjoner. Dette gjelder fremdeles for det store flertallet av fattige kvinner i tredje verden. At fødselskomplikasjoner ikke lenger blir regnet som skjebne, men som noe som kan forebygges og behandles, må regnes som et gode.

Den moderne medisinen ekspanderte og tok seg av kompliserte svangerskap og fødsler, men den stanset ikke der. I de siste tiårene har det i den vestlige verden vært en økende bruk av teknikker for å starte, stimulere, regulere, overvåke og gripe inn i den fysiologiske fødselsprosessen. En annen, og kanskje mer tidsriktig definisjon er at medikalisering kan betegnes som uhensiktsmessig bruk av inngrep, teknologi og/eller helsepersonell i fødselsomsorgen⁹². Eksempelvis har andel barn født ved hjelp av keisersnitt økt fra 4,0 % i 1975⁴⁸ til 15,8 % i 2002⁴⁹. Stadig flere kvinner får gjort keisersnitt på eget ønske, uten at det er noen medisinsk indikasjon for inngrepet⁹³. At kvinnen er engstelig for fødselen, eller at hun av andre grunner ikke ønsker å føde, blir altså kirurgisk behandlet. Et annet eksempel er fødselssmerten, som tidligere ble regnet som fysiologisk og tiltakene var trøst, oppmuntring og lindring. Nå kan den ”behandles” helt vekk ved hjelp av epiduralbedøvelse. Denne bedøvelsen gis i stort omfang på norske fødeavdelinger. Keisersnitt kan være livreddende og epiduralbedøvelse kan i noen tilfeller hjelpe kvinner gjennom en vaginal fødsel med fysisk og psykisk helse i behold. Det er når disse inngrepene tilbys friske kvinner uten for eksempel alvorlig fødselsangst, og at de regnes som en normal måte å føde på, at de inngår i det som kalles medikalisering av den fysiologiske prosessen.

For 20 år siden lagde en perinatal studiegruppe i Verdens helseorganisasjons Europaavdeling en rapport om fødselsomsorgen i Europa⁹⁴. De skrev at de fleste helsearbeidere ikke lenger vet hva en ikke-medikalisert fødsel er, og at den moderne obstetrikken er basert på observasjoner av den medikaliserte fødsel.

Overmedikalisering er i dag en av utfordringene i den perinatale omsorgen i Europa⁵². I 1998 utarbeidet Verdens helseorganisasjons Europaavdeling ti prinsipper for perinatal omsorg (tabell 3).

Disse ti prinsippene ble seinere ratifisert av medlemslandene i regionen⁹⁵. Her anføres at omsorgen i det normale svangerskap og fødsel skal være demedikalisert. I dette ligger at den

grunnleggende fødselsomsorgen skal gis med kun de helt nødvendige inngrep og med minst mulig bruk av teknologi. Et annet av de ti prinsippene er hensiktsmessig bruk av teknologi. Det beskrives som at avansert og komplisert teknologi ikke bør brukes der enklere prosedyrer vil fungere like bra eller bedre.

Tabell 3. Verdens helseorganisasjons Europaavdeling sine ti prinsipper for perinatal omsorg⁹⁵

1 Demedikalisering av normale fødsler	I dette ligger at den grunnleggende fødselsomsorgen skal gis med kun de helt nødvendige inngrep, og med minst mulig bruk av teknologi
2 Hensiktsmessig bruk av teknologi	Bruk av avansert og komplisert teknologi bør ikke brukes der enklere prosedyrer vil fungere like bra eller bedre
3 Kunnskapsbasert	Omsorgen bør være basert på beste tilgjengelig kunnskap, fortrinnsvis kunnskap fra randomiserte studier der det er hensiktsmessig.
4 Desentralisert	Omsorgen bør være desentralisert, og organisert etter prinsippet om lavest effektive omsorgsnivå, med utgangspunkt i primærhelsetjenesten, med henvisningsmuligheter til spesialisttjenesten.
5 Tverrfaglig	Omsorgen bør være tverrfaglig, og involvere helsearbeidere som jordmødre, obstetrikere, neonatologer, sykepleiere, samt fødselspedagoger og samfunnsvitere.
6 Helhetlig	Omsorgen bør omfatte hele mennesket; de intellektuelle, følelsesmessige, sosiale og kulturelle behov kvinnene, barna og familien har, ikke bare deres fysiske behov.
7 Familiefokusert	Omsorgen bør rettes inn for å møte behovene ikke bare for kvinnen og barnet, men også for kvinnens partner og for deres familie/ venner.
8 Kulturtilpasset	Omsorgen bør vurderes og tilpasses ut fra hva som passer best i den kulturelle sammenhengen den gis.
9 Medbestemmelse	Kvinner bør involveres i beslutningsprosessene.
10 Respekt	Omsorgen bør avspeile respekt for kvinners privatliv og hennes integritet og retten til konfidensiell behandling.

Oversatt fra engelsk ved Den Norske Jordmorforening

1.11 Syn på teknologi

Antropologen Birgitte Jordan⁹⁶ beskriver ”autoritativ kunnskap” som et kunnskapssystem som i et samfunn regnes som overlegent, som det ”riktige”, og som blir lagt til grunn for avgjørelser. Autoritativ kunnskap blir til gjennom en sosial prosess, denne prosessen gjenspeiler maktstrukturen i samfunnet. Fødselen og omstendighetene rundt den – fødesystemet – defineres som et kulturelt skapt system eller samfunn. I fødesystemet i USA er autoritativ kunnskap forvaltet av leger. Jordan hevder at i USA og i resten av den vestlige verden spiller teknologi en viktig rolle, og at maskinene (for eksempel CTG-apparater og infusjonspumper) spiller en viktig symbolsk rolle som kulturgjenstander. Antropologen Cecil G Helman⁸² skriver at i tillegg til å ha en medisinsk eller teknisk hensikt, kan mange av de

moderne obstetriske prosedyrene også beskrives som overgangsritualer (rites de passage). I alle samfunn er svangerskap og fødsel mer enn en biologisk hendelse, det er også en viktig overgang fra en sosial status til en annen, fra ”kvinne” til ”mor”. På den farefulle veien fra en status til den neste må individet beskyttes ved hjelp av ritualer⁸². Ritualer kan redusere engstelse og frykt for det ukjente hos kvinnen – og også hos behandlerne. Ritualer kan i tillegg ha andre formål, som å bekrefte og rebekrefte maktstrukturen i et samfunn – for eksempel fødesystemet⁹⁶. Professorstipendiat i jordmorfag Gunnhild Blåka⁹⁷ hevder at da teknologien ble innført i fødselsomsorgen, førte det til en forskyvning av arbeidsoppgavene og en holdningsendring som gikk i retning av en mer inngripende innstilling. Blåka skriver også:

”Vi kan videre snakke om at apparatene omsetter naturens sansekvaliteter, som farge, klang, lukt og smak til kvantitative målbare størrelser. Det som kan registreres objektivt, blir mest troverdig” (s 163).

Bjørn Hofmann har skrevet en doktorgradsavhandling om blant annet hvordan teknologi påvirker vår oppfatning av sykdom⁹⁸. Han hevder at teknologien ikke er verdinøytral, men at den er verdiladet og fører til handling⁹⁹. Han mener også at når skolemedisinen bruker teknologiske innretninger uten dokumentert effekt, skiller den seg ikke fra healing eller andre former for alternativ medisin. Medisinsk teknologi er et middel for å hjelpe mennesker, men i mange sammenhenger handler vi som om teknologien i seg selv har blitt målet¹⁰⁰.

Ulla Waldenström skriver at mulige årsaker til overbruk av medisinsk teknologi kan være redsel for feilbehandling, for liten tid, for liten kunnskap om alternative metoder, uvilje til å framstå som umoderne og bakstreversk, at det som er medisinsk mulig styrer våre handlinger og at innstillingen ”kunden har alltid rett” har fått større gjennomslag enn evidensbegrepet⁶¹.

Lærebøker i obstetrikk problematiserer ikke alltid bruk av teknologi innen fødselshjelpen. Den norske ”Obstetrikk og gynekologi”⁴ beskriver bare bruk og tolking av CTG. Det samme gjør den britiske ”Myles textbook for midwives”¹⁰¹. Den svenske ”Lärobok för barnmorskor”¹⁰² problematiserer at kontinuerlig overvåking med CTG uten at det foreligger medisinsk indikasjon er blitt alminnelig. De skriver at dette har forandret fødestua til ei intensivavdeling hvor lege og jordmor spiller en sentral rolle og at CTG-apparatet kan forandre relasjonene mellom kvinnen og hennes partner og mellom kvinnen og fødselshjelperne. Britiske ”Turnbull’s Obstetrics” presiserer at kontinuerlig CTG-overvåking fører til flere operative forløsninger⁵ og at jordmødre skal være selektive og kritiske i bruken av teknologi, spesielt i bruk og tolking av CTG¹⁰³.

De som produserer og selger CTG-apparater (og annen medisinsk teknologi) vil ha som mål å selge flest mulig av dem, og vil antakelig påvirke potensielle kjøpere til en mest mulig positiv holdning til apparatet. Wagner⁸⁰ hevder at et møte hvor Den internasjonale organisasjonen for gynekologer og obstetrikere (FIGO) skulle utvikle retningslinjer for bruk av CTG hovedsakelig var bekostet av produsentene av teknologien som skulle evalueres. Dette vil antakelig påvirke fødselshjelpenes oppfatning av CTG. Edward Hon var en amerikansk lege og ingeniør som oppfant skalpelektroden. Denne elektroden gjorde det mulig å koble fosteret direkte til CTG-apparatet. Banta og Thacker²⁵ hevder at Hon fikk offentlige midler for å evaluere bruken av kontinuerlig elektronisk fosterovervåking. Banta og Thacker stiller spørsmålsteget ved om en oppfinner av teknisk utstyr kan gjøre en upartisk evaluering av sin egen oppfinnelse.

1.12 Fødselsomsorgen i Norge

All helsehjelp i forbindelse med svangerskap, fødsel og barseltid er gratis. Det finnes 58 fødeinstitusjoner som er organisert på tre nivå³⁰:

1. Fødestuer med minimum 40 fødsler årlig. Fødestuene har vaktberedskap av jordmødre. Knappt 2 % av alle fødende føder ved fødestue⁴⁹.
2. Fødeavdelinger med minimum 400-500 fødsler årlig. De har vaktberedskap av fødsels- og anestesileger, nødvendig jordmor- og operasjonsstuebemannning og barnelege tilknyttet avdelinga. Mellom 30 og 40 % føder ved fødeavdelinger⁴⁹.
3. Kvinneklinner med minimum 1500 fødsler årlig, tilstedevakt av fødsels- og anestesileger og vaktberedskap av barnelege. Det er nødvendig jordmor- og operasjonsstuebemannning og barneavdeling med intensivbehandling av nyfødte. Mellom 60 og 70 % av alle fødende i Norge føder ved kvinneklinikker⁴⁹.

Dette er en inndeling som Stortinget har bestemt at skal være gjeldene⁶⁹ slik det ble forslått av Statens helsetilsyn³⁰, men den er ikke helt gjennomført. Hjemmefødsler organiseres ikke av det offentlige helsevesenet, men det finnes noen jordmødre som tilbyr hjemmefødsel. I 2002 var det 118 (0,2 %)⁴⁹ planlagte hjemmefødsler i Norge. Utgiftene blir dekket gjennom rikstrygdeverket.

Som tidligere nevnt, vedtok Stortinget i 2001 at Norge skulle ha en desentralisert og differensiert fødselshjelp⁶⁹. Dette innebærer en viss omorganisering av fødselsomsorgen, blant annet er enkelte fødeavdelinger i lokalsykehus omgjort til fødestuer. Ved flere fødeavdelinger

og kvinneklinikker gir man differensiert tilbud til de fødende avhengig om de blir vurdert til å ha forhøyet sannsynlighet for uønska utfall eller ikke. Flere større fødeavdelinger (Ullevål, Haukeland, Fredrikstad, Tønsberg og Stavanger) har opprettet egne jordmorstyrte fødeavdelinger ved siden av de vanlige, eller tradisjonelle, avdelingene. Dette er et tilbud til kvinner uten risikofaktorer i svangerskapet eller fødselen, som ønsker å føde uten medikamentell smertelindring og som kanskje legger vekt på forhold som kontinuitet mellom svangerskap, fødsel og barseltid¹⁰⁴. Slike avdelinger kalles ABC-avdeling (alternative birth care) eller fødestue i sykehus. I forbindelse med omorganiseringen av fødselsomsorgen har regjeringen opprettet et Nasjonalt råd for fødselsomsorgen, deres mandat er å bistå i omstillingsprosessen.

1.13 Jordmor

Verdens helseorganisasjon har slått fast at den beste måten å forebygge unødvendig død og sykdom er at svangerskap, fødsel og barseltid følges opp av utdanna jordmødre (health professionals with midwifery skills)⁵². I Norge har hver fødende kvinne ei jordmor sammen med seg, eller i umiddelbar nærhet. Jordmor skal på eget ansvar bistå kvinnen under fødsel, fremme den normale fødsel, oppdage komplikasjoner, skaffe medisinsk assistanse og utføre nødtiltak¹⁰⁵. Jordmødre er definert som helsepersonell etter helsepersonelloven § 3¹⁰⁶. Jordmødre omfattes av EØS avtalens sektordirektiv¹⁰⁷; i artikkel 4 heter det blant annet:

”Medlemslandene skal påse at jordmødrene minst har rett til å starte og utøve følgende virksomhet: (...) å pleie og bistå kvinnen under fødselen og overvåke fosterets tilstand i livmoren ved hjelp av egnede kliniske og tekniske midler”.

I Norge kreves det autorisasjon som sykepleier samt ett års arbeidserfaring for å komme inn ved landets fire jordmorhøyskoler. Sykepleierutdanninga tar tre år og jordmorutdanninga to år. Det internasjonale jordmorforbundet har definert jordmor som en person som har skaffet seg nødvendige kvalifikasjoner og er blitt autorisert for å utøve jordmorvirksomhet. Jordmor har et selvstendig arbeidsansvar så lenge det dreier seg om normale svangerskap og fødsler uten komplikasjoner. Her kan jordmor arbeide selvstendig og uten innblanding fra lege, og lege har ikke instruksjonsmyndighet i spørsmål der jordmødre har like høy eller høyere kompetanse¹⁰⁸. Det kan imidlertid være gråsoner hvor det ikke er enighet om hva som er normalt eller avvikende forløp. Det har ingen dokumentert nytteverdi å involvere fødselsleger i alle fødselsforløp⁷⁷. Det er vist at kvinner som har en støtteperson sammen med seg under hele fødselsforløpet, sjeldnere får smertelindring, keisersnitt og operative vaginale

forløsninger sammenliknet med kvinner som ikke hadde støtteperson sammen med seg¹⁰⁹.

Støttepersonen behøver ikke å være jordmor eller annet utdannet helsepersonell.

I Norge er det vanligvis et tett samarbeid mellom jordmødre og fødselsleger.

Formålet med dette samarbeidet er å sikre kvinnen og barnet best mulig helsehjelp. Verdens helseorganisasjon har vurdert at omkring 15 % av alle kvinner utvikler komplikasjoner som er så alvorlige at behandling av medisinsk ekspertise er nødvendig for å unngå død eller uførhet¹¹⁰. Ved komplikasjoner i fødselsforløpet blir lege vanligvis tilkalt og overtar ansvaret for kvinnen. Det varierer fra institusjon til institusjon hvor mye legene er involvert i kvinner med normale fødselsforløp. Ved fødestuer og ABC-enheter arbeider jordmødrene selvstendig og overflytter kvinnene dersom det oppstår avvik eller komplikasjoner. Andre steder er legene aktivt med i omsorg og behandling av de normale fødekvinnene, går visitt på fødestuene, undersøker vaginalt og bestemmer inngrep og tiltak som jordmødrene skal utføre eller følge opp. Under slike forhold kan det være vanskelig for jordmødrene å ivareta sitt selvstendige arbeidsansvar.

Ikke alle jordmødre ønsker nødvendigvis å arbeide autonomt og selvstendig. Bente Haug⁷⁸ har beskrevet at norske jordmødre representerer begge ytterpunktene i fødselsparadigmet (se kapittel 1.8). Mander & Fleming¹¹¹ beskriver to typer jordmødre: Den obstetriske sykepleieren som styres av frykt og forholder seg til rutiner og sykehusets behov, og den ”sanne jordmor” som setter kvinnen og hennes behov i sentrum for sin praksis. Heptinstall og Lee¹¹² skriver at det finnes to forskjellige ”jordmortyper”. Den ene har utviklet sosiale og mellommenneskelige ferdigheter og arbeider best utenfor sykehus. Den andre typen arbeider i sykehus og har utviklet gode tekniske ferdigheter og kan gi omsorg til kvinner med kompliserte svangerskap og fødsler.

Virkeligheten er vel neppe så dikotom som disse framstillingene. Jordmødre formes og påvirkes antakelig av systemene de jobber i. De som arbeider ved store avdelinger hvor de ofte møter kvinner med kompliserte svangerskap og fødsler, vil bli flinke til å ta seg av disse kvinnene og til å håndtere avansert teknologi. Dette er helt sikkert til fordel for kvinner med kompliserte svangerskap og fødsler. Jordmødre ved fødestuer og ABC-enheter, hvor et av formålene er å ikke gripe inn i fødselsforløpet dersom det ikke er god grunn til det, vil bli flinke til å fremme den normale fødselen, oppmuntre kvinnene og til å unngå inngrep og komplikasjoner.

1.14 Formålet med denne avhandlingen

1.14.1 Noen bemerkninger

Arbeidet med studie I begynte allerede i 1998, mens selve avhandlingen ble planlagt i 2000-2001. På det tidspunktet fantes det få studier om innkomst-CTG, og det datasettet som hyppigst ble referert til i lærebøker, var Ingemarsson og medarbeideres observasjonsstudie fra Singapore som omfattet 1041 lavrisikofødende⁷. Siden da har det kommet flere viktige publikasjoner om innkomst-CTG.

I mai 2001 utgav Royal College of Obstetricians and Gynaecologists retningslinjer om bruk av elektronisk fosterovervåking⁹. Retningslinjene var kunnskapsbasert (evidence-based), det vil si at det ble søkt etter tilgjengelig litteratur, litteraturen ble kvalitetsvurdert og rådene ble gitt ut fra hvilken kunnskap som finnes. Retningslinjene anbefalte ikke innkomst-CTG til lavrisikofødende ettersom testens evne til å forutsi føtal distress var dårlig og fordi det ikke fantes forskning som viste at testen var nyttig. Retningslinjene viste hovedsakelig til Ingemarsson og medarbeideres studie⁷, men nevnte også to andre studier^{33; 34}. Seinere i 2001 kom den første randomiserte kontrollerte studien som vurderte effekten av innkomst-CTG hos lavrisikofødende¹¹³ og i 2003 kom to til^{114; 115}. Ingen av disse studiene tydet på at innkomst-CTG var nyttig.

I januar 2003 ble studie II publisert i British Journal of Obstetrics and Gynaecology. Studien vurderte reliabiliteten til innkomst-CTG gjennom å undersøke inter-observatør variabilitet. (Inter-observatør variabilitet vil bli beskrevet i kapittel 2.2.3.) I en leder i samme nummer av tidsskriftet, ble studien kritisert for at det kunne være systematiske bias mellom de to ekspertene som vurderte testene, og at observatørene burde fått lik opplæring før vurdering av testene¹¹⁶.

Dette gjorde at den opprinnelige planen for avhandlingen ble forandret noe. Det var planlagt en kvalitativ studie for å undersøke hvordan kvinnene opplevde innleggelsen i fødeavdelingen og spesielt å få kjørt innkomst-CTG. Ettersom det begynte å bli ganske klart at innkomst-CTG ikke var nyttig for lavrisikofødende, var det ikke så aktuelt å undersøke hvordan kvinnene opplevde testen. Denne studien ble derfor fjernet fra den opprinnelige planen. En ny studie som vurderte inter-observatør enighet i tolking av CTG etter at observatørene hadde gjennomgått et standardisert opplæringsprogram, ble planlagt og gjennomført istedenfor (studie III). Selv om det ikke finnes forskning som vurderer effekten av kontinuerlig CTG sammenliknet med intermitterende auskultasjon i risikosvangerskap, anbefales kontinuerlig overvåking av visse risikofødende⁹. Anbefalingene er gjort på grunnlag av at det finnes assosiasjoner mellom en rekke faktorer i svangerskapet og cerebral parese,

neonatal encefalopati og perinatal død⁹. CTG-overvåking av risikofødende vil altså fortsatt være en viktig del av fødselshjelpen, og hvordan CTG-strimlene tolkes og i hvilken grad de kan predikere uønska utfall er fremdeles svært relevant.

I dette sammendraget blir de fem studiene detaljert beskrevet. Årsaken er at jeg ønsker at denne forskningen skal være tilgjengelig for norske jordmødre, og det er fremdeles få som leser litteratur på engelsk.

1.14.2 Hovedmål

Målet for avhandlingen er å finne ut i hvilken grad rutinemessig innkomst-CTG kan forutsi uønska fødselsutfall, undersøke inter-observatør enighet (testens reliabilitet) om innkomst-CTG kan forebygge uønska fødselsutfall (testens effektivitet), og hvilken betydning den kan ha for jordmødre i deres daglige arbeide.

1.14.3 Delmål

Studie I: Å vurdere i hvilken grad innkomst-CTG kan forutsi føtal distress hos kvinner i fødsel, hos lavrisikofødende og risikofødende samt hos kvinner som føder innen seks timer etter at testen er tatt samt hos kvinner som føder mellom seks og 24 timer etter at testen er tatt.

Studie II: Å finne grad av enighet i vurdering av innkomst-CTG mellom jordmødre og leger i den kliniske situasjonen og to eksperter i en ikke-klinisk situasjon; samt å finne grad av enighet mellom de to ekspertene og i hvilken grad de kunne forutsi føtal distress i fødselsforløpet.

Studie III: Å undersøke grad av enighet i vurdering av innkomst-CTG hos jordmødre og leger som hadde gjennomgått et opplæringsprogram i tolking av CTG.

Studie IV: Å gjøre en systematisk litteraturgjennomgang for å finne ut om innkomst-CTG forbedrer fødselsutfallene for mor og barn sammenliknet med auskultasjon ved innkomst i fødeavdelingen og derved få en samlet oversikt over testens nytteverdi.

Studie V: Å undersøke hvordan innkomst-CTG oppleves av, og hvilken betydning testen kan ha for jordmødre i deres daglige arbeide.

2 DESIGN, STUDIEPOPULASJONER OG METODER

2.1 Oversikt over enkeltstudiene

Tabell 4 viser en oversikt over design, studiepopulasjoner, datainnsamling og måleenheter. Studie I, II og III har overlappende studiepopulasjoner idet alle er utvalg fra de 1639 kvinnene som fødte ved Hammerfest sykehus i årene 1996, 1997 og 1998. Figur 4, 5 og 6 gjør rede for inklusjonen til de tre studiene.

2.2 Studie I

2.2.1 Design

Studie I vurderer i hvilken grad innkomst-CTG kan forutsi føtal distress i fødselsforløpet. Føtal distress ble definert i følge Ingemarsson og medarbeidere⁷, som operativ forløsning på indikasjon forandret/dårlig fosterhertelyd, eller Apgar skåre < 7 etter fem minutter. Studien ble gjort ved Hammerfest sykehus. Sykehuset har en kombinert avdeling med tre fødestuer, 12 barselsenger og fire senger for innlagte gravide. I studieperioden var det mellom fem og seks hundre fødsler pr år. Omkring 15 % av fødslene i opptaksområdet foregår på fødestua i Alta. Dette gjør at fødeavdelingen har en litt større andel førstegangs fødende og risikofødende enn en normalpopulasjon.

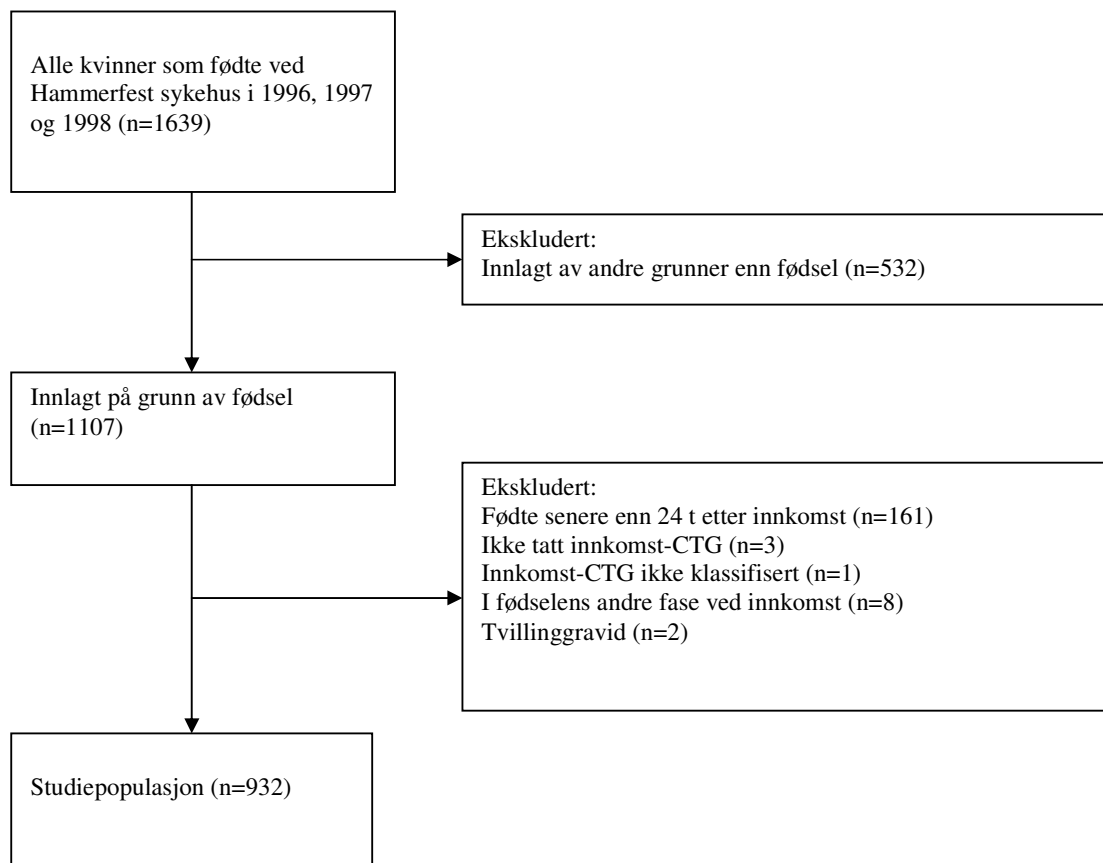
2.2.2 Studiepopulasjon

Alle kvinner som fødte ved Hammerfest sykehus i årene 1996, 1997 og 1998 ble i utgangspunktet inkludert. Eksklusjonskriterier var at kvinnen ble innlagt av andre grunner enn fødsel, at det gikk mer enn 24 timer fra innkomst-CTG til fødsel, ikke fikk tatt innkomst-CTG, innkomst-CTG var ikke klassifisert, i fødselens andre fase (mormunn utslettet), tvillinggravid eller svangerskapsvarighet mindre enn 196 dager (figur 4). Til sammen 932 av 1639 kvinner ble inkludert i studien. Ut fra opplysninger fram til innleggelsen i fødeavdelingen, ble kvinnene vurdert som lavrisikofødende eller risikofødende. Kvinnene ble vurdert til å være risikogravide dersom det var forhold til stede som kunne true fosterets helse: Hypertensiv sykdom hos mor, diabetes, for tidlig fødselsstart (< 259 dager), overtidig svangerskap (> 294 dager), vaginal blødning i siste trimester, veksthemming hos fosteret, setepresentasjon eller andre forhold som kan true fosterets helse. Til sammen 13 % av de inkluderte var risikogravide mot 28 % i totalpopulasjonen. Årsaken til denne forskjellen var at risikogravide oftere ble innlagt av andre årsaker enn fødsel. Studiepopulasjonen ble også delt inn i to tidsgrupper, de kvinnene som fødte innen seks timer etter at innkomst-CTG var tatt og de som fødte mellom seks og 24 timer etter testen.

Tabell 4. Oversikt over design, studiepopulasjoner, datasamling og måleenheter

	Design og formål	Populasjon/utvalg	Datasamling	Måleenheter
Studie I	Observasjonsstudie som vurderer i hvilken grad inntakst-CTG kan forutsi føtal distress i fødselsforløpet	932 (av 1639) kvinner som fødte ved Hammerfest sykehus 1996, 1997 og 1998. Jordmødre og leger i den kliniske situasjonen	Data fra pasientjournaler	Sensitivitet, spesifisitet, PPV, NPV, LR+ og LR-
Studie II	Observasjonsstudie som vurderer inter-observatør enighet, og i hvilken grad observatørene kunne forutsi føtal distress i fødselsforløpet	845 (av 1639) kvinner som fødte ved Hammerfest sykehus 1996, 1997 og 1998. Jordmødre og leger i den kliniske situasjonen samt to eksperter fra andre institusjoner	Data fra pasientjournaler og fra de to ekspertene	Vekta kappa-koeffisient, proportion of agreement, sensitivitet, spesifisitet, PPV, NPV, LR+ og LR-
Studie III	Observasjonsstudie som vurderer inter-observatør enighet hos forskjellige bedømmere som hadde gjennomgått et opplæringsprogram i tolking av CTG. Studien vurderer også i hvilken grad observatørene kunne forutsi føtal distress i fødselsforløpet	549 (av 1639) kvinner som fødte ved Hammerfest sykehus 1996, 1997 og 1998. Tre jordmødre og tre leger som hadde gjennomgått et opplæringsprogram i tolking av CTG	Data fra pasientjournaler og fra de seks observatørene	Vekta kappa-koeffisient, proportion of agreement, sensitivitet, spesifisitet, PPV, NPV, LR+ og LR-
Studie IV	Systematisk oversikt som vurderer effekten av inntakst-CTG sammenliknet med auskultasjon samt i hvilken grad testen kan forutsi uønska fødselsutfall	Tre randomiserte kontrollerte studier som inkluderte 11259 kvinner og 11 observasjonsstudier som inkluderte 5831 kvinner	Litteratursøk i elektroniske databaser, referanselister og ved kontakt med eksperter	RR, sensitivitet, spesifisitet, PPV, NPV, LR+ og LR-
Studie V	Kvalitativ studie som undersøker hvilken kunnskap jordmødre fikk fra inntakst-CTG, og hvilken betydning testen hadde i det daglige arbeidet	Tolv jordmødre. Utvalget ble først gjort strategisk, deretter teoretisk (strategically/theoretical sampling)	Kvalitative dybdeintervju	Kategorier som beskriver funnene

PPV=positiv prediktiv verdi, NPV=negativ prediktiv verdi, LR+=likelihood ratio for positivt testresultat, LR-=likelihood ratio for negativt testresultat, RR=relativ risiko



Figur 4. Flytskjema for utvelgelse av studiepopulasjon til studie I

2.2.3 Metoder

Journalene ble gjennomgått retrospektivt og følgende variabler ble registrert: Alder, paritet, risikogruppe, årsak til innleggelse i fødeavdelingen, tidsrom fra innkomst-CTG til fødsel, klassifikasjon av innkomst-CTG, forløsningsmåte, indikasjon for operative forløsninger og Apgar skåre. Det lyktes å finne journalene til alle kvinnene.

Innkost-CTG ble tatt og tolket av de jordmødre og leger som tok hånd om kvinnene. Testene ble tolket som enten normal, suspekt eller patologisk ifølge Ingemarsson og medarbeidere¹¹⁷. Det ble gjort en beskrivende analyse med utregning av sensitivitet, spesifisitet, PPV, NPV, LR+ og LR- i hele studiepopulasjonen, i de to risikogruppene, og i de to tidsgruppene.

Ettersom innkost-CTG ble tolket i tre kategorier (normal, suspekt og patologisk), ble analysene gjort på to forskjellige cut-off nivå. Ved cut-off suspekt test ble normale innkost-CTGer regnet som test-negativ, mens suspekter og patologiske ble regnet som test-positive. Ved cut-off patologisk test ble normale og suspekter innkost-CTGer regnet som test-negativ mens patologiske ble regnet som test-positive.

2.3 Studie II og III

2.3.1 Design

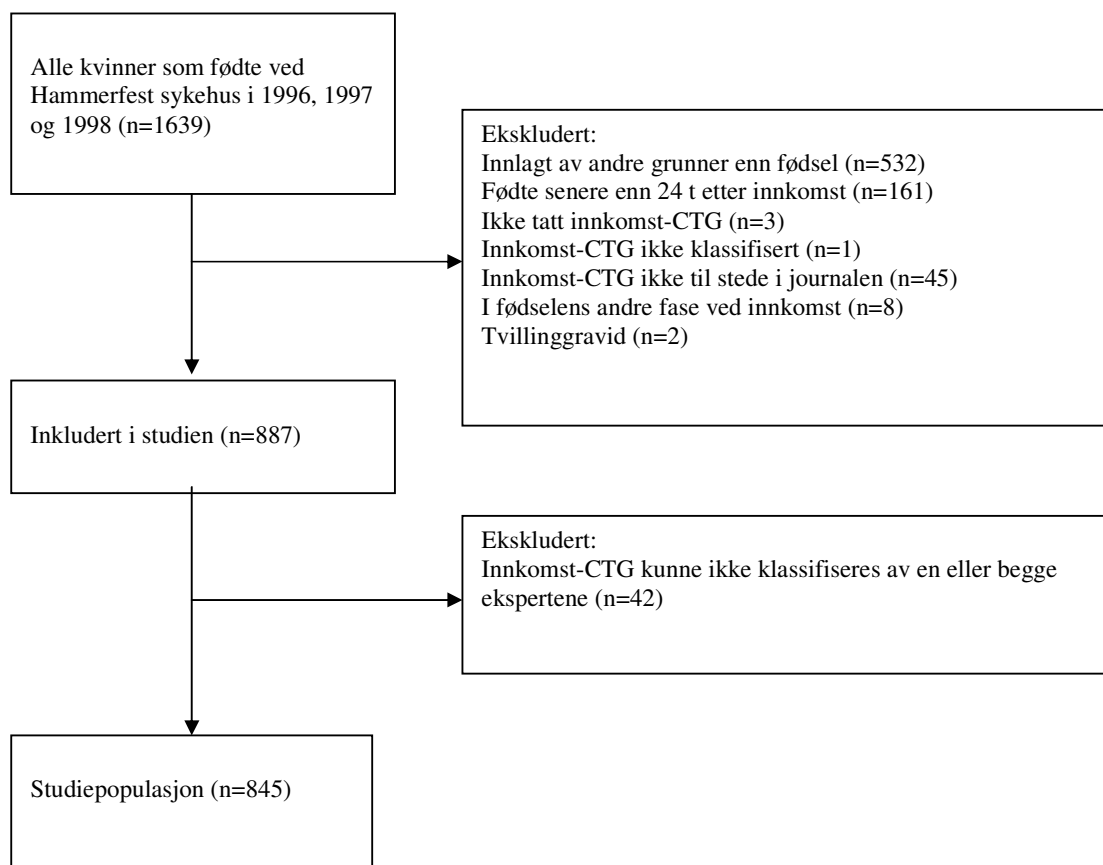
Begge studiene vurderte inter-observatør enighet i tolking av innkomst-CTG mellom forskjellige observatører. Inter-observatør enighet er et mål på en tests reliabilitet, altså evne til å gi samme resultat ved gjentatte målinger. For å vurdere testens evne til å forutsi føtal distress i fødselsforløpet, ble sensitivitet, spesifisitet, PPV, NPV, LR+ og LR- ble kalkulert for hver av observatørene.

2.3.2 Studiepopulasjoner

I utgangspunktet ble de 1639 kvinnene som fødte ved Hammerfest sykehus i årene 1996, 1997 og 1998 inkludert.

I studie II var eksklusjonskriteriene at kvinnen ble innlagt av andre grunner enn fødsel, at det gikk mer enn 24 timer fra innkomst-CTG til fødsel, ikke fikk tatt innkomst-CTG, innkomst-CTG var ikke klassifisert av jordmor eller lege, innkomst-CTG ikke til stede i journalen, i fødselens andre fase (mormunn utslettet), tvillinggravid eller svangerskapsvarighet mindre enn 196 dager (figur 5). Til sammen 887 kvinner oppfylte inklusjonskriteriene. Av disse ble 42 (5 %) ekskludert fordi en eller begge ekspertene fant at CTG-strimlene var umulige å klassifisere. Til sammen 845 kvinner ble inkludert i studien, av disse var 13 % risikogravide. Innkomst-CTG ble først klassifisert i den kliniske konteksten av de jordmødre og leger som tok hånd om kvinnene. Deretter ble de vurdert av to leger fra andre institusjoner. De ble valgt som eksperter fordi de hadde lang klinisk erfaring og fordi de var spesielt interessert i CTG.

I studie III var eksklusjonskriteriene at det var mer enn 24 timer fra innkomst-CTG til fødsel, ikke fikk tatt innkomst-CTG, innkomst-CTG var ikke klassifisert av jordmor eller lege, innkomst-CTG ikke til stede i journalen, innkomst-CTG av kortere varighet enn 15 minutter, i fødselens andre fase (mormunn utslettet), tvillinggravid eller svangerskapsvarighet mindre enn 196 dager (figur 6). Av de 1669 kvinnene oppfylte 1199 inklusjonskriteriene. Et utvalg på 572 CTG-strimler ble valgt ut fra et ønske om å inkludere så mange unormale strimler som mulig, og strimlene skulle være av rimelig god teknisk kvalitet. Av de 572 CTG-strimlene ble 23 (4 %) ekskludert fra analysen fordi en eller flere av observatørene vurderte at de ikke kunne klassifiseres. Av de 549 kvinnene som ble inkludert, var 18 % risikogravide.



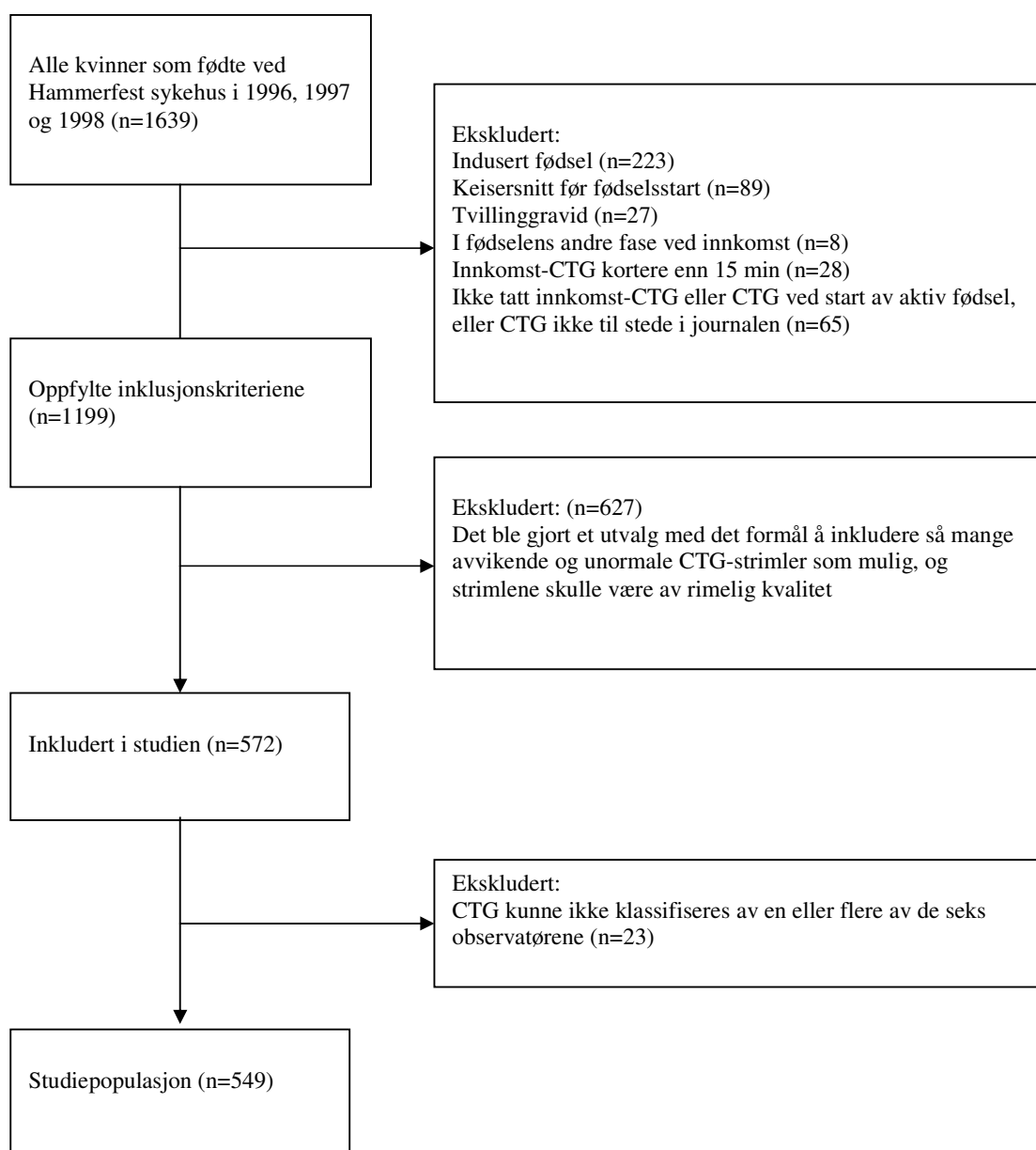
Figur 5. Flytskjema for utvelgelse av studiepopulasjon til studie II

Lederne av fødeavdelingene ved to universitetssykehus og et lokalsykehus ble bedt om å rekruttere seks fødselshjelpere som vurderte CTGer i sitt daglige arbeide og som ble vurdert til å være dyktige. Tre jordmødre og tre leger fra tre forskjellige sykehus ble bedt om å vurdere innkomst-CTGene. Alle hadde gjennomgått opplæringsprogrammet ”Fosterovervåking”¹⁰. To av dem var superbrukere, det vil si at de hadde mer opplæring og at de underviste og sertifiserte leger og jordmødre i blant annet tolking av CTG. Målet med opplæringsprogrammet er å sertifisere jordmødre og leger som STAN[®]-brukere. Programmet består av seks timer teoretisk og praktisk undervisning og en viktig del av undervisninga er om CTG-tolking og fysiologi, et hefte hvor ti av 30 sider er om CTG og et CD-rom som blant annet inneholder 50 CTG-strimler. Eksamen består av 22 multiple-choice spørsmål hvor ni er om tolking av CTG.

2.3.3 Metoder

CTGene ble først tolket i den kliniske konteksten av de jordmødre og leger som tok hånd om kvinnene. Observatørene fikk deretter kopier av CTG-strimlene hvor navn og kommentarer

var fjernet. De fikk vite at kvinnene var i fødselens første fase (åpningsfasen) og fikk ekstra informasjon dersom kvinnene hadde en svangerskapsvarighet mindre enn 36 uker eller hadde fått terbutalin (medikament for å stoppe riene) før eller under transporten til sykehus. Ellers fikk de ingen annen informasjon om kvinnene, fødselsutfallene eller studiepopulasjonen. De hadde ikke kontakt med hverandre mens de gjorde tolkningene av CTG-strimlene.



Figur 6. Flytskjema for utvelgelse av studiepopulasjon til studie III

De to observatørene i studie II ble bedt om å klassifisere testene som normale, suspekte eller patologiske ifølge Ingemarsson og Ingemarsson¹¹⁷. For studie III ble det utarbeidet en ny mal for tolking av inntakst-CTG basert på Ingemarsson og Ingemarsson¹¹⁷ og Sundström og medarbeidere¹⁰. Denne malen er gjort rede for i studie III. Her ble testene vurdert som normale, avvikende eller patologiske.

Inter-observatør enighet i en klinisk test er det samme som testens reliabilitet¹¹⁸. En tests reliabilitet er dens evne til å gi samme resultat gjentatte ganger. En tests reliabilitet avhenger av stabiliteten til det som blir målt (i dette tilfellet fosterhjerterefrekvensen), variabiliteten i målemetodene (varigheten på testen, mors leie) og dyktigheten hos de som tar testen (fødselshjelperne som vurderer testen)¹¹⁹.

Ettersom det ikke er enighet om hvilken metode som er best for å beregne inter-observatør enighet, ble både kappa-koeffisienten (κ)¹²⁰ og ”proportion of agreement” (Pa)¹²¹ beregnet. Grant¹²¹ mener at Pa er en bedre test for inter-observatør variasjon for kategoriske data, men at testen krever et utvalg på minst 200. Pa for normal CTG mellom to observatører uttrykkes som: Antall CTG-strimler begge definerte som normale/antall CTG-strimler en eller begge vurderte som normale. Kappa er definert som forskjellen mellom observert og forventet enighet, eller andelen enighet ut over forventet enighet⁶². Forventet enighet er det samme som tilfeldig enighet. Kappa uttrykkes som: $\kappa = \text{observert enighet} - \text{forventet enighet}$. Fleiss¹²⁰ beskriver kappa over 0,75 som utmerket enighet, kappa mellom 0,75 og 0,40 som rimelig til god enighet og under 0,40 som dårlig enighet. Fordi studiene hadde tre mulige klassifikasjoner (normal, suspekt eller avvikende og patologisk), ble vektet kappa (κ_w) med Ciccetti-Allisons vektorer (1, 0,5 og 0) beregnet.¹²⁰ Kappa-verdien påvirkes blant annet av fordelingen av kategorier (normal/suspekt eller avvikende/patologisk testresultat) i studiepopulasjonen og bør presenteres sammen med opplysninger om hvordan testene ble tolket¹²². For å vurdere i hvilken grad de forskjellige observatørene kunne forutsi føtal distress⁷ i fødselsforløpet, ble sensitivitet, spesifisitet, PPV, NPV, LR+ og LR- beregnet.

2.4 Studie IV

2.4.1 Design

Studie IV er en systematisk oversikt som vurderer hvilken nytteverdi inntakst-CTG har sammenliknet med auskultasjon ved inntakst samt i hvilken grad testen kan forutsi uønska utfall.

2.4.2 Data

Det ble gjort elektroniske søk i Medline, Embase, CINAHL, SweMed, the Cochrane central register, manuelle søk i referanselister og det ble tatt kontakt med eksperter på feltet.

2.4.3 Metoder

Både randomiserte kontrollerte studier og observasjonsstudier ble kvalitetsvurdert og inkludert. Kriterier for å inkludere randomiserte kontrollerte studier var: Populasjon: Fødende kvinner. Intervensjon: CTG av 20-40 minutters varighet ved innleggelse i fødeavdeling. Kontroll: Auskultasjon av fosterhjerterefrekvens ved innleggelse. De skulle inkludere et eller flere av følgende utfallsmål: "Perinatal mortalitet", "Apgar skåre < 7 etter fem min", pH <7,05 i navlesnorsarterie", "gjenoppliving av den nyfødte" "overflytting til nyfødtavdeling", "tykt, misfarget fostervann", "nyfødtkramper", "skalpblodpøve av fosteret", "operativ forløsning (keisersnitt, vakuum og tang)", "keisersnitt", "operativ forløsning (keisersnitt, vakuum og tang) på indikasjon føtal distress", "keisersnitt på indikasjon føtal distress", "stimulering av fødsel", "kontinuerlig CTG-overvåking", "epiduralanalgesi" og "føtal distress som definert av Ingemarsson et al.⁷. Studiene skulle også inneholde data som gjorde det mulig å lage firefeltstabeller over utfall og intervensjon/kontroll.

Kriterier for å inkludere observasjonsstudier var: Populasjon: Kvinner i fødsel. Test: Innkomst-CTG av 20-40 minutters varighet ved innleggelsen. De skulle inkludere et av følgende utfallsmål: "Apgar skåre < 7 etter 5 min", "pH <7,05 i navlesnorsarterie", "gjenoppliving av den nyfødte" "overflytting til nyfødtavdeling", "tykt, misfarget fostervann", "nyfødtkramper", "operativ forløsning (keisersnitt, vakuum og tang) på indikasjon føtal distress", "keisersnitt på indikasjon føtal distress", og "føtal distress som definert av Ingemarsson et al.⁷. De skulle også inneholde data så det var mulig å lage firefeltstabeller over testresultater og utfallsmål.

CONSORT kvalitetskriterier¹²³ ble brukt for å vurdere randomiserte studier og QUADAS kvalitetskriterier¹²⁴ for å vurdere observasjonsstudier. CONSORT¹²³ (Consolidated Standards of Reporting Trials) er anbefalinger for å forbedre rapporteringer av randomiserte studier, og inneholder en sjekkliste med 22 punkter for å kvalitetsvurdere studiene. Anbefalingene er utarbeidet av en internasjonal arbeidsgruppe. QUADAS^{124; 125} (Quality Assessment of Diagnostic Accuracy Studies) er et hjelpemiddel for å vurdere diagnostiske studier (studies of diagnostic accuracy) og inneholder en sjekkliste med 14 punkter for å kvalitetsvurdere studiene. QUADAS er utviklet av en gruppe britiske og nederlandske forskere.

Studier av dårlig metodologisk kvalitet ble ekskludert. Studiene ble kvalitetsvurdert uavhengig av to personer (EB og LMR). Eventuell uenighet ble avgjort av en tredjeperson (PØ). Uthenting av data ble også gjort uavhengig av to personer (EB og LMR), og forfatterne av originalstudiene ble kontaktet for å innhente ekstraopplysninger der det var nødvendig.

2.5 Studie V

2.5.1 Design

Studie V er en kvalitativ intervjustudie som undersøker hvilken informasjon og kunnskap innkomst-CTG kan gi, og hvilken betydning testen kan ha for jordmødre i deres daglige arbeide. De siste årene har det vært diskutert blant norske fødselshjelpere om innkomst-CTG fortsatt bør være rutine hos lavrisikofødende. I løpet av våren 2006 vil vi få retningslinjer som ikke anbefaler innkomst-CTG til lavrisikokvinner³². Mitt inntrykk er at selv om jordmødre vanligvis gir uttrykk for at de er tilbakeholdne med å bruke teknologi, er det få som ønsker å slutte med rutinemessig CTG. Jeg har ofte hørt argumenter som at selv om det ikke er vitenskapelig bevist at innkomst-CTG er nyttig, så gir den en viktig informasjon om hvordan fosteret har det, og også at det er bra med ”objektiv” informasjon. På min arbeidsplass ble det for noen år siden bestemt at vi skulle slutte med innkomst-CTG hos lavrisikofødende. For noen jordmødre var dette greit, men det var noen som fortsatte å ta rutinemessig CTG på alle ved innkomst. Det er mulig at innkomst-CTG bidrar med kunnskap og har betydninger som ikke er belyst i de kvantitative studier som er gjort om emnet. Dersom man skal slutte å ta innkomst-CTG på grupper av kvinner, vil det være nyttig å vite om mekanismer og forhold som gjør det vanskelig å avvike denne rutinen.

2.5.2 Metode og utvalg

Dersom vi vil vite mer om menneskelige erfaringer, oppfatninger, forestillinger eller forventninger, er kvalitative metoder egnet. Målet er å utforske meningsinnholdet i fenomener slik det oppleves for aktørene selv¹²⁶. Med kvalitative metoder kan man forstå og beskrive, ikke forklare eller predikere. I denne studien ble data samlet inn ved hjelp av kvalitative forskningsintervjuer. Intervjuer ble valgt fordi det er en egnet måte å finne ut hvordan jordmødre, med egne ord, opplever innkomst-CTG i deres daglige arbeide¹²⁷. Steinar Kvale¹²⁷ definerer forskningsintervju som ”et intervju som har som mål å innhente beskrivelser av den intervjuedes livsverden, med henblikk på fortolkning av de beskrevne fenomenene”. Han beskriver også et intervju som en konversasjon som har en viss struktur og hensikt.

Intervjuene ble analysert ved hjelp av grounded theory-metoden¹²⁸. Metoden er utviklet for å oppdage teorier, begreper og hypoteser direkte fra data¹²⁹. Grounded theory har røtter i symbolsk interaksjon som fokuserer på hvilken mening mennesker gir omverdenen¹³⁰. Datainnsamling og analyse foregår parallelt. Antallet informanter ble ikke bestemt på forhånd. I begynnelsen ble det gjort strategisk utvalg ("purposive sampling"), det vil si at man velger informanter med ulik erfaring og bakgrunn. På denne måten kan man få fram mange oppfatninger av og erfaringer omkring fenomenet som studeres. Etter hvert ble utvalget gjort etter prinsippet om "theoretical sampling", det vil si at informantene ble valgt med det formålet å finne variasjoner i begrepene og for å gi fylldige beskrivelser av kategoriene¹³⁰. Datainnsamlinga avsluttes når "metning" (theoretical saturation) oppnås, det er når nye data ikke tilfører ny kunnskap¹²⁸.

2.5.3 Intervju

Intervjuene ble gjort på arbeidsplassen eller hjemme hos informantene og varte i 30-90 minutter. De fokuserte på hvilken kunnskap og informasjon jordmor opplevde at hun fikk fra innkomst-CTG, hvordan hun informerte kvinnene om testen og hvilke konsekvenser resultatet av testen hadde for jordmor i den videre oppfølginga av kvinnen. Emner som kom opp og som var relevante for problemstillinga, ble fulgt opp. Intervjuene ble tatt opp på lydbånd, skrevet ordrett ut og analysert fortløpende. Alle intervjuene ble gjort og skrevet ut av EB. Informantene fikk tilbud om å få utskrift av intervjuet og ble bedt om å kontakte EB om de ønsket forandringer i teksten.

2.5.4 Dataanalyse

Datainnsamling og analyse ble gjort parallelt. Først ble hvert intervju lest nøye for å "få tak i helheten" i intervjuet, deretter ble analysen gjort i tre steg: 1) Åpen koding hvor hver setning ble lest nøye for å identifisere begreper og beskrive dem. Et begrep er en abstrakt benevnelse av en hendelse, objekt eller interaksjon som vurderes til å være viktig. Begrepene ble samlet i foreløpige kategorier, disse ble beskrevet i henhold til egenskaper og variasjoner. En kategori står for et fenomen, det kan være et problem, et mål eller en hendelse som vurderes å være viktig for de som blir intervjuet. 2) Neste steg var aksial koding, hvor det ble lett etter relasjoner mellom kategoriene. Kategoriene ble satt i sammenheng med underkategoriene og beskrevet. 3) Siste steg var selektiv koding, hvor en eventuell kjernekategori kommer fram. Kjernekategorien skal være sentral og kunne relateres til de andre kategoriene. Til slutt ble alle intervjuene lest grundig gjennom for å verifisere resultatene i data. De forskjellige

nivåene i analysen ble gjort delvis parallelt^{128; 130}. Tekstbehandlingsprogrammet NVivo ble brukt som støtte i analysen¹³¹. Analysene ble gjort av EB. LÖ leste alle intervjuene, diskuterte og bekreftet funnene i samarbeid med EB. I tillegg ble funnene presentert for ei gruppe av fire jordmødre; disse ble valgt ut fra klinisk erfaring og evne til kritisk refleksjon. Å la noen med liknende erfaring og arbeide som informantene validere resultatene regnes som nyttig¹³².

Intervjuene ble gjort mellom mai 2000 og april 2002, analysen og artikkelskriving var ferdig til innsending våren 2005. Årsaken til at det tok så lang tid, var at studien krevde konsentrasjon over lengre tid. I en oppstykket hverdag var det ikke så lett å få til, og studien ble først gjort ferdig da jeg fikk mulighet til å ta sammenhengende fri over en lengre periode.

2.6 Etiske overveielser

Studie I og III ble vurdert hos Norsk samfunnsvitenskapelige datatjeneste, studie II og V ble vurdert i Regional komité for medisinsk forskningsetikk. Det var ingen innvendinger mot noen av studiene. Studie IV er en systematisk oversikt og data besto av allerede publiserte artikler. Det var derfor ikke nødvendig med etisk vurdering av denne studien.

3 RESULTATER

Med noen få unntak, var sensitivitet, PPV og likelihood ratios for innkomst-CTG dårlig (studie I, II, III og IV). Testens evne til å forutsi uheldige utfall var altså lav. Enighet i tolking av innkomst-CTG var dårlig i studie II, mens den var rimelig god når observatørene hadde gjennomgått et opplæringsprogram i tolking av CTG (studie III). Kvinner som ble randomisert til innkomst-CTG ved innleggelsen i fødeavdelingen fikk oftere epiduralbedøvelse, kontinuerlig CTG-overvåking og tatt fosterblodprøve enn kvinner som ble randomisert til auskultasjon ved innleggelsen. Det var ingen forskjeller i utfallene hos de nyfødte (studie IV). Jordmødre kan oppleve motsetninger i seg selv eller mellom seg selv og andre når de forholder seg til rutinemessig bruk av innkomst-CTG. Motsetningene opplevdes mellom yrkesidentitet og den økende bruken av teknologi, mellom trygghet og utrygghet og mellom makt og avmakt (studie V).

3.1 Testens evne til å forutsi uønska utfall

3.1.1 Når testene ble tolket av fødselshjelperne i den kliniske situasjonen (studie I)

Utfallsmålet var føtal distress definert som operativ forløsning på indikasjon forandret/dårlig fosterhertelyd, eller Apgar skåre < 7 etter fem minutter⁷.

Resultatene er vist i tabell 5. Sensitivitet var gjennomgående lav. Høyest var den hos kvinner som fødte innen seks timer etter innkomst-CTG, her var den 36. Det vil si at testen identifiserte 36 % av kvinnene som utviklet føtal distress. Hele 64 % av de som utviklet føtal distress, testet altså negativt. Positiv prediktiv verdi var også gjennomgående lav, noe som viser at de fleste som testet positivt ikke utviklet føtal distress. Ved cut-off patologisk test var positiv prediktiv verdi høy i et par tilfeller, men andel falsk-negative var over 80 % i disse tilfellene. Likelihood ratio for positiv innkomst-CTG var 7,3 hos kvinnene som fødte innen seks timer etter testen. Det betyr at det var 7,3 ganger høyere sannsynlighet for at testen kom fra en kvinne som utviklet føtal distress som fra en som ikke gjorde det. Ellers var LR+ mellom 4,1 og 0,3 ved cut-off suspekt test (tabell 5). Dette betyr at det ikke er så mye høyere sannsynlighet for at en kvinne med føtal distress hadde en suspekt eller patologisk innkomst-CTG enn at en kvinne som ikke utviklet føtal distress hadde positiv test. Undergruppen av høyrisikokvinner hadde sensitivitet 8, PPV 10 og LR+ 1,0. Ved cut-off patologisk test var LR+ høy i noen tilfeller, men også i disse tilfellene var sensitiviteten svært lav og LR- tett på

1. Dette indikerer at det er omtrent like sannsynlig at en kvinne som utvikler føtal distress hadde en normal innkomst-CTG som at den var suspekt eller patologisk (tabell 5).

Spesifisitet og negativ prediktiv verdi var gjennomgående høye, noe som indikerer at innkomst-CTG i høy grad kan identifisere de kvinnene som ikke utvikler føtal distress. LR-ligger imidlertid tett på 1 i alle tilfellene (tabell 5).

3.1.2 Når testene ble tolket blindt av uavhengige ”eksperter” i en ikke-klinisk situasjon (studie II og III)

Utfallsmålet var også her føtal distress definert som operativ forløsning på indikasjon forandret/dårlig fosterhertelyd, eller Apgar skåre < 7 etter fem minutter⁷.

Resultatene er vist i Tabell 6. Også her er de prediktive verdiene lave. Ved cut-off suspekt/avvikende test er riktignok sensitivitet oppe mot 50 i et par tilfeller, men positiv prediktiv verdi er ikke høyere enn 15 og viser at hele 85 % var falsk-positive. Ved cut-off patologisk test var riktignok LR+ høy i et par tilfeller, mens sensitivitet var svært lav (mellom 21 og 5).

3.1.3 I en systematisk oversikt (studie IV)

Det ble funnet 22 observasjonsstudier om innkomst-CTG. Av disse ble 11 ekskludert på grunn av dårlig kvalitet, ikke relevante utfallsmål, eller av andre årsaker^{33; 34; 133-141}. Det er gjort grundig rede for inklusjonen i studie IV. Av de 11 studiene som ble inkludert var tre fra Singapore^{7; 142; 143}, tre fra Storbritannia^{42; 144; 145}, fire fra USA^{40; 41; 146; 147} og en fra Norge¹⁴⁸. De 11 studiene omfattet til sammen 5831 kvinner. De på forhånd definerte utfallsmålene var ”operativ forløsning (keisersnitt, tang eller vakuum) på indikasjon føtal distress”, ”keisersnitt på indikasjon føtal distress”, ”Apgar skåre < 7 etter fem minutter”, ”føtal distress”⁷, ”gjenoppliving av den nyfødte” og ”overflytting til nyfødtavdeling”. Det ble ikke funnet resultater for utfallsmålene ”pH < 7,05 i navlesnorsarterie”, og ”tykt, misfarget fostervann”. Den inkluderte norske studien¹⁴⁸ er den samme som studie II, i den systematiske oversikten er resultatene presentert på en litt annen måte og med flere utfallsmål. De inkluderte studiene er beskrevet i studie IV.

Resultatene er presentert i tabell 7. Sensitivitet varierte mellom 5 og 83 i de 28 enkeltutfallene og var over 50 fire enkeltutfall. PPV var mellom 1 og 28 % bortsett fra i ett tilfelle hvor det var 75 %. Spesifisitet varierte mellom 78 og 98 og NPV mellom 89 og 99. Likelihood ratio for positiv test var over 10 i to tilfeller, mellom 5 og 10 i seks tilfeller og

ellers lavere. I ett tilfelle var likelihood ratio for negativ test 0,2, i alle de andre enkeltutfallene var den mellom 0,4 og 1,0.

3.2 Inter-observatør variabilitet

3.2.1 Når testene ble tolket av to uavhengige eksperter og av jordmødre/leger i den kliniske situasjonen (studie II)

I studie II ble til sammen 845 innkomst-CTGer vurdert, først av jordmødrene og legene i den kliniske situasjonen og deretter to uavhengige eksperter. Av disse ble 660 (78 %) vurdert som normale, sju (1 %) som suspekter og to (<1 %) som patologiske av alle tre observatørene. I 158 av tilfellene (19 %) var det uenighet over en kategori, det vil si at de samme CTG-strimlene ble vurdert som enten normale eller suspekter (146 strimler) eller som suspekter eller patologiske (12 strimler). I 18 av tilfellene (2 %) var det uenighet over to kategorier, det vil si at de samme strimlene ble vurdert både som normale og patologiske av de tre observatørene.

I de tre observatørparene var vekta kappa mellom 0,25 og 0,38, Pa for normal test mellom 0,85 og 0,89 og Pa for suspekt/patologisk test mellom 0,18 og 0,33 (tabell 8).

3.2.2 Når testene ble tolket av seks observatører som hadde gjennomgått et opplæringsprogram i tolking av CTG (studie III)

I studie III ble til sammen 549 innkomst-CTGer vurdert av tre jordmødre og tre leger som hadde gjennomgått et standardisert opplæringsprogram i tolking av CTG. Av de 549 CTG-strimlene ble 308 (56 %) vurdert som normale, 36 (7 %) som avvikende og sju (1 %) som patologiske av alle seks observatørene. Det var uenighet over en kategori i 187 (34 %) ettersom 147 CTG-strimler ble vurdert som enten normale eller avvikende mens 40 ble vurdert som enten avvikende eller patologiske. I 11 tilfeller (2 %) var det uenighet over to kategorier, det vil si at de samme CTG-strimlene ble vurdert både som normale og patologiske av de forskjellige observatørene.

De seks observatørene utgjorde til sammen 15 observatørpar. Vekta kappa varierte mellom 0,57 og 0,75, Pa for normal test mellom 0,78 og 0,88 og Pa for avvikende/patologisk test mellom 0,56 og 0,66 (tabell 8).

Tabell 5. Sensitivitet, spesifisitet, prediktive verdier og likelihood ratios i hele studiepopulasjonen samt i undergrupper når inkomst-CTG ble tolket av fødselsjelperne i den kliniske situasjonen (Kilde: Studie I). Utfallsmålet er føtal distress definert som operativ forløsning på indikasjon forandret/dårlig fosterhertelyd, eller Apgar skåre < 7 etter 5 min⁷

	Sensitivitet	Spesifisitet	Positiv prediktiv verdi	Negativ prediktiv verdi	Likelihood ratio	
					For positiv test (LR+)	For negativ test (LR-)
Cut-off suspekt test						
Hele studiepopulasjonen (n=932)	15 (5-24)	95 (94-97)	16 (6-27)	95 (93-96)	3,2 (1,6-6,4)	0,9 (0,8-1,0)
Lav-risiko (n=806)*	17 (6-29)	96 (94-97)	18 (6-30)	96 (94-97)	4,1 (1,9-8,7)	0,9 (0,8-1,0)
Høy-risiko (n=124)*	8 (0-22)	92 (87-97)	10 (0-29)	90 (84-95)	1,0 (0,1-6,9)	1,0 (0,9-1,2)
Fødte innen 6 timer (n=461)	36 (16-57)	95 (93-97)	27 (11-43)	97 (95-98)	7,3 (3,7-14,4)	0,7 (0,5-0,9)
Fødte mellom 6-24 timer (n=471)	2 (0-6)	96 (94-98)	3 (0-9)	93 (91-95)	0,3 (0,1-5,5)	1,0 (1,0-1,1)
Cut-off patologisk test						
Hele studiepopulasjonen (n=932)	9 (2-17)	100 (99-100)	63 (29-96)	95 (93-96)	27,1 (6,7-110,4)	0,9 (0,8-1,0)
Lav-risiko (n=806)	12 (2-22)	100 (100-100)	83 (54-113)	96 (94-97)	93,3 (11,2-780,3)	0,9 (0,8-1,0)
Høy-risiko (n=124)	4 (-6-13)	98 (95-100)	17 (0-59)	89 (84-95)	1,6 (0,1-31,7)	1,0 (0,9-1,1)
Fødte innen 6 timer (n=461)	23 (5-40)	100 (99-100)	71 (38-105)	96 (95-98)	49,9 (10,3-242,9)	0,8 (0,6-1,0)
Fødte mellom 6-24 timer (n=471)	2 (0-6)	100 (99-100)	25 (0-85)	93 (91-95)	4,4 (0,2-107,0)	1,0 (1,0-1,0)

95 % konfidensintervaller i parentesene. *For to av kvinnene var det ikke nok opplysninger i pasientjournalen til at de kunne inndeles i risikogruppe

Tabell 6. Sensitivitet, spesifisitet, prediktive verdier og likelihood ratios når innkomst-CTG ble tolket av ”eksperter” i en ikke-klinisk situasjon (Kilder: Studie II (n=845) og III (n=549)). Utfallsmålet er føtal distress definert som operativ forløsning på indikasjon forandret/dårlig fosterhjerterlyd, eller Apgar skåre < 7 etter 5 min⁷

Cut-off suspekt/avvikende test	Sensitivitet	Spesifisitet	Positiv prediktiv verdi	Negativ prediktiv verdi	Likelihood ratio	
					For positiv test (LR+)	For negativ test (LR-)
Cut-off suspekt/avvikende test						
Ekspert I (studie II)	22 (11-34)	90 (88-93)	13 (6-20)	95 (93-96)	2,3 (1,3-4,1)	0,9 (0,7-1,0)
Ekspert II (studie II)	30 (17-43)	84 (82-87)	11 (6-16)	95 (93-97)	1,9 (1,2-3,0)	0,8 (0,7-1,0)
Super-bruker I (studie III)	48 (34-62)	69 (65-73)	12 (8-17)	94 (91-96)	1,5 (1,1-2,1)	0,8 (0,6-1,0)
Super-bruker II (studie III)	43 (30-58)	74 (70-77)	13 (8-19)	94 (91-96)	1,6 (1,1-2,4)	0,8 (0,6-1,0)
Jordmor I (studie III)	46 (32-60)	72 (68-76)	12 (8-18)	94 (91-96)	1,6 (1,1-2,3)	0,8 (0,6-1,0)
Lege I (studie III)	39 (26-53)	81 (77-84)	15 (9-22)	94 (91-96)	2,0 (1,3-3,0)	0,8 (0,6-1,0)
Jordmor II (studie III)	43 (30-58)	76 (72-79)	13 (9-20)	94 (91-96)	1,8 (1,2-2,6)	0,8 (0,6-1,0)
Lege II (studie III)	41 (28-56)	78 (74-81)	14 (9-21)	94 (91-96)	1,8 (1,2-2,7)	0,8 (0,6-1,0)
Cut-off patologisk test						
Ekspert I (studie II)	12 (3-21)	97 (96-98)	21 (6-35)	95 (93-96)	4,2 (1,8-9,7)	0,9 (0,8-1,8)
Ekspert II (studie II)	4 (1-11)	100 (100-100)	100 (100-100)	94 (93-96)	78,0 (3,8-1604,2)	1,0 (0,9-1,0)
Super-bruker I (studie III)	14 (6-27)	95 (92-96)	18 (8-34)	93 (90-95)	2,5 (1,1-5,6)	0,9 (0,8-1,0)
Super-bruker II (studie III)	21 (11-35)	95 (92-96)	24 (13-40)	93 (91-95)	3,7 (1,9-7,3)	0,8 (0,7-1,0)
Jordmor I (studie III)	21 (11-34)	96 (94-97)	29 (16-47)	93 (91-95)	4,7 (2,3-9,6)	0,8 (0,7-1,0)
Lege I (studie III)	11 (5-24)	99 (97-100)	46 (21-72)	93 (90-95)	9,6 (3,0-30,0)	0,9 (0,8-1,0)
Jordmor II (studie III)	16 (8-29)	98 (96-99)	41 (22-64)	93 (91-95)	8,0 (3,2-20,1)	0,9 (0,8-1,0)
Lege II (studie III)	16 (8-29)	98 (96-99)	37 (19-59)	93 (91-95)	6,7 (2,8-16,1)	0,9 (0,8-1,0)

95 % konfidensintervaller i parentesene. I studie II var 13 % av kvinnene risikofødende, i studie III var 18 % risikofødende. Studiepopulasjonene er delvis overlappende.

Tabell 7. Sensitivitet, spesifisitet, prediktive verdier og likelihood ratios for inkomst-CTG i 11 observasjonsstudier som omfattet til sammen 5831 kvinner. (Kilde: Studie IV)

	Sensitivitet	Spesifisitet	Positiv prediktiv verdi	Negativ prediktiv verdi	For positiv test (LR+)	Likelihood ratio	For negativ test (LR-)					
^aOperativ forløsning på indikasjon føtal distress												
^b Blix og medarb. 2003 ¹⁴⁸	31	(18-45)	84	(82-87)	10	(5-15)	96	(94-97)	2,0	(1,3-3,2)	0,8	(0,7-1,0)
^b Chua og medarb. 1996 ¹⁴²	31	(20-43)	95	(94-96)	28	(18-39)	96	(94-97)	6,3	(4,0-9,9)	0,7	(0,6-0,9)
Farrell og medarb. 1998 ¹⁴⁴	19	(0-38)	95	(91-98)	25	(1-50)	92	(88-96)	3,5	(1,0-11,5)	0,9	(0,7-1,1)
^b Ingemarsson og medarb. 1986 ⁷	15	(3-27)	95	(93-96)	9	(1-16)	97	(96-98)	2,7	(1,2-6,4)	0,9	(0,8-1,0)
Keisersnitt på indikasjon føtal distress												
^b Blix og medarb. 2003 ¹⁴⁸	44	(23-64)	84	(82-87)	7	(3-12)	98	(97-99)	2,8	(1,7-4,5)	0,7	(0,5-1,0)
Ducey og medarb. 1990 ⁴⁰	56	(39-73)	98	(97-100)	75	(58-92)	96	(94-98)	35,0	(14,9-81,9)	0,5	(0,3-0,7)
Elimitan og medarb. 2003 ¹⁴⁷	30	(16-49)	96	(93-97)	32	(17-52)	95	(93-97)	7,0	(3,3-14,6)	0,7	(0,6-0,9)
^b Ingemarsson og medarb. 1986 ⁷	28	(7-49)	95	(93-96)	9	(1-16)	99	(98-99)	5,3	(2,4-11,6)	0,8	(0,6-1,0)
Sarno og medarb. 1989 ¹⁴⁶	83	(54-100)	85	(78-92)	24	(6-42)	99	(97-100)	5,4	(3,0-9,5)	0,2	(0,0-1,2)
Sarno og medarb. 1990 ⁴¹	67	(48-86)	80	(76-84)	18	(10-26)	97	(96-99)	3,4	(2,4-4,8)	0,4	(0,2-0,7)
^b Somerset of medarb. 1993 ¹⁴⁵	38	(14-61)	89	(86-92)	15	(4-26)	97	(95-99)	3,4	(1,7-6,9)	0,7	(0,5-1,0)
Apgar skåre < 7 etter 5 min												
^b Blix og medarb. 2003 ¹⁴⁸	29	(0-62)	84	(81-86)	1	(0-3)	99	(99-100)	1,8	(0,5-5,7)	0,9	(0,5-1,4)
^b Chua og medarb. 1996 ¹⁴²	63	(39-86)	94	(93-96)	14	(6-22)	99	(99-100)	11,0	(7,0-17,3)	0,4	(0,2-0,8)
Elimitan og medarb. 2003 ¹⁴⁷	13	(0-45)	94	(92-96)	2	(0-7)	99	(98-100)	2,1	(0,2-28,6)	0,9	(0,6-1,4)
Farrell og medarb. 1995 ⁴²	17	(0-60)	89	(84-93)	2	(0-7)	99	(97-100)	1,5	(0,1-18,6)	0,9	(0,6-1,6)
Farrell og medarb. 1998 ¹⁴⁴	25	(0-85)	93	(90-97)	4	(0-14)	99	(98-100)	3,6	(0,3-42,6)	0,8	(0,4-1,8)
Sarno og medarb. 1989 ¹⁴⁶	25	(0-85)	80	(73-88)	2	(0-9)	98	(96-100)	1,3	(0,1-14,4)	0,9	(0,4-2,1)
Sarno og medarb. 1990 ⁴¹	30	(2-58)	78	(74-82)	3	(0-7)	98	(96-99)	1,4	(0,5-3,5)	0,9	(0,6-1,4)
^cFøtal distress												
^b Blix og medarb. 2003 ¹⁴⁸	30	(17-43)	84	(82-87)	11	(6-16)	95	(93-97)	1,9	(1,2-3,0)	0,8	(0,7-1,0)
^b Ingemarsson og medarb. 1986 ⁷	41	(20-62)	95	(94-96)	15	(6-24)	99	(98-99)	8,3	(4,7-14,8)	0,6	(0,4-0,9)
^b Ingemarsson og medarb. 1988 ¹⁴³	36	(16-57)	93	(92-95)	14	(5-23)	98	(97-99)	5,4	(2,9-10,0)	0,7	(0,5-0,9)
Gjenoppliving av den nyfødte												
^{b, d} Chua og medarb. 1996 ¹⁴²	27	(14-40)	94	(93-96)	17	(8-26)	97	(96-98)	4,8	(2,8-8,3)	0,8	(0,6-0,9)
Farrell og medarb. 1998 ¹⁴⁴	5	(0-15)	93	(89-97)	8	(0-24)	89	(85-94)	0,8	(0,1-5,7)	1,0	(0,9-1,1)
Overflytting til nyfødtavdeling												
^b Blix og medarb. 2003 ¹⁴⁸	39	(21-57)	84	(82-87)	8	(3-12)	98	(97-99)	2,5	(1,5-4,1)	0,7	(0,5-1,0)
^b Chua og medarb. 1996 ¹⁴²	31	(20-43)	95	(94-96)	27	(17-37)	96	(95-97)	6,2	(3,9-9,8)	0,7	(0,6-0,9)
Elimitan og medarb. 2003 ¹⁴⁷	9	(3-24)	94	(92-96)	12	(4-30)	93	(90-95)	1,6	(0,5-5,1)	1,0	(0,9-1,1)
Farrell og medarb. 1995 ⁴²	43	(6-80)	90	(86-94)	12	(0-24)	98	(96-100)	4,2	(1,6-10,8)	0,6	(0,3-1,2)
Farrell og medarb. 1998 ¹⁴⁴	17	(0-60)	93	(89-97)	4	(0-14)	99	(97-100)	2,4	(0,2-32,1)	0,9	(0,5-1,5)

^aVakuum, tang og keisersnitt, ^bCut-off susppekt, ^cFøtal distress som definert av Ingemarsson og medarb. 1986⁷, ^dassistert ventilasjon. Testene ble tolket blindt i seks av studiene^{7,41;42;143;147;148}, i de øvrige er det ikke opplyst hvem som tolket. 95 % konfidensintervall i parentesene.

Tabell 8. Vekta kappa (κ) and andel enighet (proportion of agreement (Pa)) i de forskjellige observatørparene. (Kilder: Studie II (n=845) og III (n=549))

Observatørpar	κ	Pa for normal innkomst-CTG	Pa for suspekt (avvikende)/patologisk innkomst-CTG*
Ekspert 1/Ekspert 2 (studie II)	0,38 (0,31-0,46)	0,86 (0,83-0,88)	0,33 (0,26-0,40)
Ekspert 1 /jordmødre og leger (studie II)	0,25 (0,15-0,36)	0,89 (0,87-0,91)	0,18 (0,11-0,25)
Ekspert 2/jordmødre og leger (studie II)	0,28 (0,20-0,37)	0,85 (0,82-0,88)	0,20 (0,14-0,26)
Superbruker 1/Superbruker 2 (studie III)	0,60 (0,54-0,67)	0,78 (0,74-0,82)	0,56 (0,49-0,63)
Superbruker 1/Jordmor 1 (studie III)	0,66 (0,60-0,72)	0,82 (0,78-0,85)	0,62 (0,55-0,68)
Superbruker 1/Lege 1 (studie III)	0,57 (0,51-0,63)	0,81 (0,77-0,85)	0,56 (0,48-0,63)
Superbruker 1/Jordmor 2 (studie III)	0,63 (0,57-0,69)	0,83 (0,79-0,86)	0,63 (0,56-0,69)
Superbruker 1/Lege 2 (studie III)	0,60 (0,53-0,66)	0,81 (0,77-0,85)	0,58 (0,51-0,65)
Superbruker 2/Jordmor 1 (studie III)	0,67 (0,60-0,73)	0,82 (0,78-0,85)	0,60 (0,53-0,67)
Superbruker 2/Lege 1 (studie III)	0,64 (0,57-0,70)	0,86 (0,82-0,89)	0,62 (0,54-0,69)
Superbruker 2/Jordmor 2 (studie III)	0,63 (0,57-0,70)	0,84 (0,80-0,87)	0,61 (0,53-0,67)
Superbruker 2/Lege 2 (studie III)	0,60 (0,53-0,66)	0,83 (0,79-0,86)	0,57 (0,50-0,64)
Jordmor 1/Jordmor 2 (studie III)	0,68 (0,62-0,74)	0,86 (0,82-0,89)	0,66 (0,59-0,73)
Jordmor 1/Lege 1 (studie III)	0,60 (0,54-0,67)	0,83 (0,79-0,86)	0,57 (0,49-0,64)
Jordmor 1/Lege 2 (studie III)	0,62 (0,56-0,69)	0,84 (0,81-0,87)	0,62 (0,55-0,69)
Jordmor 2/Lege 1 (studie III)	0,68 (0,61-0,74)	0,86 (0,83-0,89)	0,61 (0,53-0,68)
Jordmor 2/Lege 2 (studie III)	0,75 (0,70-0,81)	0,88 (0,85-0,91)	0,69 (0,61-0,75)
Lege 1/Lege 2 (studie III)	0,64 (0,57-0,71)	0,85 (0,82-0,88)	0,57 (0,49-0,64)

95 % konfidensintervall i parentesene. *I studie II er CTGene klassifisert som normal/suspekt/patologisk; i studie III som normal/avvikende/patologisk. Tolkingsmalene er ikke helt identiske i de to studiene. I studie II er 13 % av kvinnene risikofødende, i studie III er 18 % risikofødende. Studiepopulasjonene er delvis overlappende.

3.3 Effekt av innkomst-CTG sammenliknet med auskultasjon ved innleggelsen (studie IV)

Det ble inkludert tre randomiserte, kontrollerte studier og ingen ble ekskludert.

(Seleksjonsprosessen er vist i Figure 1, studie IV.) En studie var fra Irland¹¹⁴ og to fra Skottland^{113; 115}. De tre studiene omfattet til sammen 11 259 kvinner, alle var lavrisikofødende. En av studiene ble vurdert til å være av god kvalitet¹¹⁴ og de to andre av middels kvalitet^{113; 115} (Table 1 og Table 2, studie IV).

Den ene studien var fra Dundee i Skottland¹¹³ og inkluderte 3751 lavrisikofødende. Studien ble gjort ved fødeavdelingen i Ninewells Medical Hospital som er et universitetssykehus. Kvinnene ble enten randomisert til 20 minutter CTG eller til auskultasjon med doppler under og like etter minst ei ri ved innleggelsen i fødeavdelingen. Av de 1186 kvinnene som ble randomisert til CTG ved innleggelsen, hadde 255 (21,5 %) unormal innkomst-CTG. Av de som ble randomisert til auskultasjon, hadde 42 (3,6 %) unormal hjertelyd ved innleggelsen. Kvinnene som ble randomisert til CTG ved innkomst fikk oftere kontinuerlig elektronisk forsterovervåking (OR 1,49, 95 % KI 1,26-1,76), oksytocinstimulering av riene (OR 1,26, 95 % KI 1,02-1,56), epiduralanalgesi (OR 1,33, 95 %

KI 1,10-1,61) og operativ forløsning (vakuum, tang eller keisersnitt) (OR 1,36, 95 % KI 1,2-1,65) sammenliknet med kvinnene som ble randomisert til auskultasjon ved innkomst. Det var ingen signifikante forskjeller i utfallene hos de nyfødte.

Den andre studien var fra Irland¹¹⁴ og inkluderte 8580 lavisikofødende. Studien ble gjort på National Maternity Hospital i Dublin hvor førstegangsfødende blir behandlet etter prinsippet "active management of labour". Dette innebærer blant annet at det blir gjort amniotomi (punktering av fosterhinner) ved innleggelsen og at den fødende har ei jordmor sammen med seg under hele fødselsforløpet. Bare kvinnene med klart fostervann ble inkludert i studien. Kvinnene ble enten randomisert til 20 minutter CTG eller til auskultasjon like etter amniotomi. Av de 4254 kvinnene som ble randomisert til CTG, hadde 1360 (32 %) unormal innkomst-CTG. Det ble ikke oppgitt hvor mange i auskultasjonsgruppa som hadde unormale hjertelyd ved innleggelsen. Kvinnene som ble randomisert til innkomst-CTG fikk oftere kontinuerlig elektronisk overvåking (RR 1,39, 95 % KI 1,33-1,45) og skalpblodprøve av fosteret (RR 1,30, 95 % KI 1,14-1,47). Det var ingen signifikante forskjeller i utfallene hos de nyfødte.

Den tredje studien var fra Skottland¹¹⁵ og inkluderte 312 lavrisikofødende som fødte ved en jordmorstyrt fødeenhet ved Glasgow Royal Maternity Hospital. Kvinnene ble randomisert enten til 20 minutter CTG eller auskultasjon under og etter ei ri i minst 60 sekunder ved innleggelsen. Det ble ikke beskrevet hvor mange unormale funn det var i hver av gruppene. Kvinner randomisert til innkomst-CTG fikk oftere kjørt en ekstra CTG-registrering i fødselsforløpet enn kvinnene randomisert til auskultasjon ved innleggelsen (125/148 sammenliknet med 61/164, khikvadrattest = 73,8, $p = 0,001$). Ellers var det ingen signifikante forskjeller mellom de to gruppene.

Rådata fra de tre studiene ble slått sammen og det ble gjort meta-analyse med kalkulering av relativ risiko for hvert utfallsmål¹⁴⁹.

Kvinner randomisert til innkomst-CTG hadde oftere intervensjoner som epiduralbedøvelse, kontinuerlig CTG-overvåking og skalpblodprøve av fosteret enn kvinner randomisert til auskultasjon ved innkomst. Det var også en ikke-signifikant trend til flere keisersnitt blant kvinnene randomisert til innkomst-CTG. Ellers var det ingen forskjeller i noen av de neonatale utfallene mellom de to gruppene. Alle resultatene er vist i tabell 9.

Tabell 9. Meta-analyse av tre randomiserte, kontrollerte studier som til sammen inkluderte 11 259 kvinner. (Kilde: Studie IV)

	Antall tilfeller i gruppene (n/N)		Relativ risiko (95 % CI)
	Innkost-CTG	Auskultasjon ved innleggelse	
^aOperativ forløsning			
Cheyne og medarb. 2003 ¹¹⁵	23/148	30/164	0,9 (0,5 – 1,4)
Impey og medarb. 2003 ¹¹⁴	673/4298	634/4282	1,1 (1,0 – 1,2)
Mires og medarb. 2001 ¹¹³	313/1186	247/1181	1,3 (1,1 – 1,5)
Til sammen	1009/5632	911/5627	1,1 (1,0 – 1,3)
^aOperativ forløsning på indikasjon føtal distress			
Impey og medarb. 2003 ¹¹⁴	308/4298	273/4282	1,1 (1,0 – 1,3)
Keisersnitt			
Cheyne og medarb. 2003 ¹¹⁵	11/148	9/164	1,4 (0,6 – 3,2)
Impey og medarb. 2003 ¹¹⁴	180/4298	158/4282	1,1 (0,9 - 1,4)
Mires og medarb. 2001 ¹¹³	61/1186	43/1181	1,4 (1,0 - 2,1)
Til sammen	252/5632	210/5627	1,2 (1,0 - 1,4)
Keisersnitt på indikasjon føtal distress			
Impey og medarb. 2003 ¹¹⁴	57/4298	50/4282	1,1 (0,8 - 1,7)
Stimulering av fødsel			
Cheyne og medarb. 2003 ¹¹⁵	28/148	29/164	1,1 (0,7 - 1,7)
Impey og medarb. 2003 ¹¹⁴	1637/4298	1629/4282	1,0 (1,0 - 1,1)
Mires og medarb. 2001 ¹¹³	246/1183	202/1175	1,2 (1,0 - 1,4)
Til sammen	1911/5629	1860/5621	1,1 (0,9 - 1,2)
Epiduralanalgesi			
Cheyne og medarb. 2003 ¹¹⁵	26/148	22/164	1,3 (0,8 - 2,2)
Mires og medarb. 2001 ¹¹³	325/1186	261/1181	1,2 (1,1 – 1,4)
Til sammen	351/1334	283/1345	1,2 (1,1 – 1,4)
Kontinuerlig CTG-overvåking			
^b Cheyne og medarb. 2003 ¹¹⁵	7/148	5/164	1,6 (0,5 - 4,8)
Impey og medarb. 2003 ¹¹⁴	2511/4298	1802/4282	1,4 (1,3 - 1,5)
Mires og medarb. 2001 ¹¹³	672/1186	551/1178	1,2 (1,1 – 1,3)
Til sammen	3190/5632	2358/5624	1,3 (1,2 – 1,5)
Skalpblodprøve av fosteret			
Cheyne og medarb. 2003 ¹¹⁵	6/148	8/164	0,8 (0,3 - 2,3)
Impey og medarb. 2003 ¹¹⁴	457/4298	349/4282	1,3 (1,1 – 1,5)
Mires og medarb. 2001 ¹¹³	96/1186	76/1181	1,3 (0,9 - 1,7)
Til sammen	559/5632	433/5627	1,3 (1,1 – 1,5)
Perinatal dødelighet			
Cheyne og medarb. 2003 ¹¹⁵	0/148	0/164	Ikke estimerbar
^c Impey og medarb. 2003 ¹¹⁴	0/4298	1/4282	0,3 (0,0-8,2)
Mires og medarb. 2001 ¹¹³	2/1186	1/1181	2,0 (0,2-21,9)
Til sammen	2/5632	2/5627	1,1 (0,2-7,1)
Gjenoppliving av den nyfødte			
^d Impey og medarb. 2003 ¹¹⁴	8/4298	13/4282	0,6 (0,3 - 1,5)
^e Mires og medarb. 2001 ¹¹³	5/1185	4/1178	1,2 (0,3 - 4,6)
Til sammen	13/5483	17/5460	0,8 (0,4 - 1,6)
Nyfødtkramper			
Impey og medarb. 2003 ¹¹⁴	11/4298	14/4282	0,8 (0,4 - 1,7)
Apgar skåre <7 etter 5 min			
Cheyne og medarb. 2003 ¹¹⁵	0/148	2/164	0,2 (0,0 - 4,6)
Impey og medarb. 2003 ¹¹⁴	17/4298	11/4282	1,5 (0,7 - 3,3)
Mires og medarb. 2001 ¹¹³	25/1181	18/1171	1,4 (0,8 - 2,5)
Til sammen	42/5627	31/5617	1,4 (0,9 - 2,2)
Overflytting til nyfødtavdeling			
Cheyne og medarb. 2003 ¹¹⁵	1/148	3/164	0,4 (0,0 - 3,5)
Impey og medarb. 2003 ¹¹⁴	203/4298	197/4282	1,0 (0,8 - 1,2)
Mires og medarb. 2001 ¹¹³	46/1185	45/1175	1,0 (0,7 - 1,5)
Til sammen	250/5631	245/5621	1,0 (0,9 - 1,2)

^aVakuüm, tang og keisersnitt, ^bmer enn 75 % av tiden i fødeavdelingen, ^cdødfødsel eller tidlig dødsfall (innen 7 dager) hos et barn uten store medfødte misdannelser, ^dmekanisk ventilasjon i mer enn 15 minutter, ^e”Need for intermittent positive pressure ventilation at resuscitation”

3.4 Hvordan innkost-CTG oppleves av jordmødre i deres daglige arbeid (studie V)

3.4.1 Utvalg

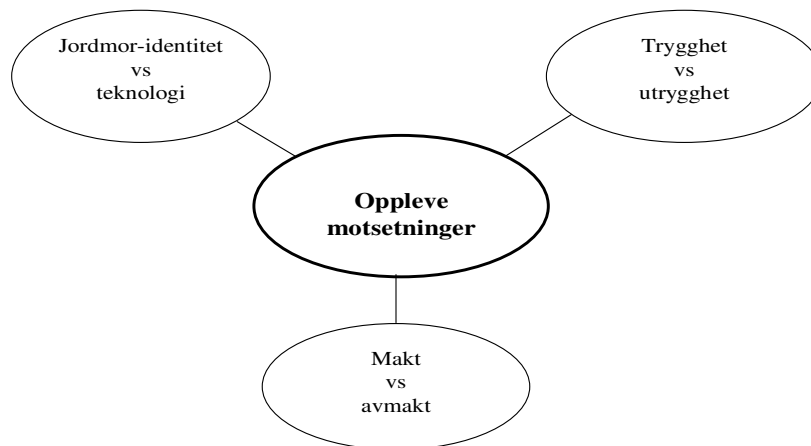
De 12 informantene var alle kvinner og arbeidet i fire forskjellige fødeinstitusjoner: Ei fødestue, ei fødeavdeling, en kvinneklinikk og en jordmorstyrt fødeenhet i universitetssykehus. Tre av fødeinstitusjonene var i Helse Nord mens den fjerde var i Helse Øst. Ei av jordmødrene arbeidet også som hjemmefødselsjordmor. Informantene hadde mellom to uker og 38 år arbeidserfaring (tabell 10). Samlet hadde de arbeidserfaring som omfattet 11 forskjellige fødeinstitusjoner i Norge, to i Sverige, en i Finland og en i Danmark. Ni var utdanna ved norske jordmorhøyskoler, mens to var utdanna i Sverige og en i Finland.

Tabell 10. De 12 jordmødrene i studiepopulasjonen

	Type fødeinstitusjon	Arbeidserfaring
Jordmor 1	Fødeavdeling	13 år
Jordmor 2	Fødeavdeling	38 år
Jordmor 3	Fødestue	3 år
Jordmor 4	Kvinneklinikk	11 år
Jordmor 5	Hjemmefødejordmor/kvinneklinikk	12 år
Jordmor 6	ABC-avdeling i kvinneklinikk	15 år
Jordmor 7	Fødeavdeling	2 uker
Jordmor 8	Fødeavdeling	1 år
Jordmor 9	Kvinneklinikk	11 år
Jordmor 10	Kvinneklinikk	10 år
Jordmor 11	Kvinneklinikk	1 år
Jordmor 12	Kvinneklinikk	31 år

3.4.2 Kjernekategori og underkategorier

Kjerne kategorien ”oppleve motsetninger” ble identifisert, den forklarer hvilken betydning innkost-CTG kan ha for jordmødre i deres daglige arbeide. Kjerne kategorien forklares av underkategoriene ”jordmoridentitet vs teknologi”, trygghet vs utrygghet” og ”makt vs avmakt” (figur 7). ”Å oppleve motsetninger” innebar at jordmødrene opplevde motsetninger i seg selv eller mellom seg selv og andre når de forholdt seg til rutinemessig bruk av innkost-CTG. Dette forklares nærmere i beskrivelsene av underkategoriene.



Figur 7. Betydning inntøms-CTG kan ha for jordmødre i deres daglige arbeid

”Jordmoridentitet vs teknologi”: Jordmoridentitet ble beskrevet som at jordmor var ansvarlig for normale fødselsforløp. Dersom det var avvik fra det normale, ble jordmors rolle oppfattet som å støtte kvinnen og derved fremme fødselsforløpet mens legen tok det medisinske ansvaret. Begrepet ”jordmorkunnskap” ble ofte nevnt og ble beskrevet som de kunnskapene jordmor har, personligheten hennes, at hun kan bruke hender, øyne, ører, nese og intuisjon for å hente informasjon fra kvinnen og at hun kan gjøre egne vurderinger ut fra den informasjonen hun får.

Jordmødrene mente at deres profesjonelle identitet ble truet på grunn av økende bruk av CTG og annen teknologi. Det ble uttrykt en bekymring over at inntøms-CTG ble gitt større betydning enn andre observasjoner som ble gjort i forbindelse med innleggelsen. Det var vanlig praksis å holde opp inntøms-CTG på rapportene slik at alle kunne se strimla. Andre dokumenter, som for eksempel helsekort for gravide eller partogrammet, ble aldri vist fram på samme måte. Det ble også uttrykt en bekymring over at jordmødre generelt blir mindre flinke til å stole på egne jordmorkunnskaper på grunn av økende bruk av teknologi. Ei av jordmødrene sa at hun selv hadde blitt dårligere til å lytte med trestetoskop fordi ”teknikken har tatt så mye over”. De andre informantene uttrykte bekymring på vegne av andre jordmødre, de nevnte for eksempel at unge jordmødre og jordmødre i store og travle avdelinger ikke lenger var så flinke til å lytte med stetoskop.

Innkomst-CTG kunne bli brukt fornuftig og ufornuftig. Fornuftig bruk innebar at den ble sett på som en del av et hele og ikke ble ”stolt for mye på”. Ufornuftig bruk ble beskrevet som at jordmor ikke kjente til begrensningene til CTG, at hun kunne bli fiksert på CTG og ikke sette den i sammenheng med resten av observasjonene.

”Trygghet vs utrygghet”: En normal innkomst-CTG kan gjøre at jordmora føler seg trygg på at barnet har det bra der og da. Dette ble også beskrevet som en falsk trygghetsfølelse ettersom innkomst-CTG ikke nødvendigvis sier noe om hvordan barnet kommer til å få det seinere i forløpet. Det ble beskrevet flere eksempler på at friske barn ble født etter unormale CTGer og ett eksempel på at et alvorlig sykt barn ble født etter at både innkomst-CTG og påfølgende CTGer hadde vært normale. Jordmødrene med lang arbeidserfaring sa at de syntes det var enkelt å gjenkjenne normale innkomst-CTGer, og også de som var klart patologiske. Det kunne imidlertid være vanskelig å tolke de som lå imellom, slike resultater kunne gjøre at de følte seg utrygge under fødselsforløpet. Jordmødrene med kortest arbeidserfaring (mindre enn tre år) uttrykte at de syntes det var vanskelig å tolke CTG i det hele tatt og at de ”manglet erfaring”. Ei av dem sa at hun helst ville at en lege skulle signere strimla så snart som mulig og gå god for at den var normal. Dette i tilfelle det skulle oppstå noe i fødselsforløpet og det i ettertid kunne vise at det var noe i CTGen hun ikke hadde fått med seg.

Ikke bare tolkingen av CTG kunne være en kilde til utrygghet, men også apparatet i seg selv. Det kunne ta inn mors hjerterefrekvens i stedet for barnets og en av jordmødrene fortalte at hun hadde opplevd at transduceren lå oppå apparatet og tok inn en lyd med normal frekvens.

Det var tre forskjellige oppfatninger av hvilken verdi CTG-strimla hadde som dokumentasjon. Den ene var at bare ei CTG-strimle signert av en lege gjaldt og at et journalnotat med beskrivelse av hjerterefrekvens etter auskultasjon ikke hadde juridisk gyldighet. Den andre oppfatninga var at auskultasjon dokumentert i journalen var akkurat like gyldig juridisk dokumentasjon som ei CTG-strimle. Det var også en oppfatning om at riktignok var et journalnotat like gyldig dokumentasjon som ei CTG-strimle, men jordmors evner og troverdighet kunne bli stilt i tvil. Informasjon fra CTG er dokumentert på papir og tilgjengelig for alle, ved auskultasjon har hjerterefrekvens bare vært i jordmors øre i ei kort stund og det kan være vanskelig for andre å vite hva jordmor virkelig hørte og hva hun sier at hun hørte.

”Makt vs avmakt”: På fødeavdelingen er det et makthierarki hvor jordmødrene står under legene, men over fødekvinne. Dette makthierarkiet påvirker bruk og tolking av innkomst-CTG. Informantene beskrev at rutinemessig bruk av innkomst-CTG var noe som var bestemt av legene. Noen syntes at dette var uproblematisk og beskrev testen som nyttig

fordi den gav informasjon som var umulig å få fra auskultasjon og kliniske undersøkelser, som korttidsvariabilitet og deselerasjoner. Andre følte seg tvunget til å gjøre en rutine som de ikke selv syntes var nyttig. Noen av jordmødrene mente at de fikk samme informasjon ved å auskultere, palpere magen for å vurdere rienes hyppighet og styrke og ved å spørre kvinnen om fosterbevegelser.

Informantene beskrev tilfeller hvor de vurderte at det ville være best for kvinnen at fosterhertelyden bare ble auskultert og at det ikke ble tatt innkomst-CTG. Årsaken til ikke å ta innkomst-CTG kunne være at kvinnen var kommet langt i fødselsforløpet eller hadde smertefulle rier og syntes det var ubehagelig å være i ro. Noen av jordmødrene var redde for å få kritikk dersom de vurderte å ikke ta innkomst-CTG. Denne redselen påvirket deres individuelle vurdering. Andre jordmødre hadde antakelig mer selvtillit og en høyere posisjon i hierarkiet og kunne derved gjøre mer individuelle vurderinger. De kunne også få kritikk fra legene, men de var ikke redd for det og forventet at det ble stolt på de vurderingene de hadde gjort.

De fleste informantene jobbet i avdelinger hvor innkomst-CTG alltid skulle signeres av lege. Jordmor vurderte CTG-strimla først, dersom den var unormal ville hun vanligvis tilkalle lege for vurdering og videre behandling. Dersom innkomst-CTG var normal, ville den bli signert når det passet legen og gjerne etter at barnet var født. Jordmødrene hadde forskjellige oppfatninger av denne praksisen. Noen mente at legen var den som hadde det medisinske ansvaret og hadde best kompetanse til å vurdere CTG. En annen oppfatning var at signering ved lege ikke hadde noe med jordmødrene å gjøre, men med opplæring av assistentleger. En tredje oppfatning var at legeunderskrift på normale innkomst-CTGer innebar en undervurdering av jordmors kompetanse til å ha ansvar for normale fødsler. Legen ble ikke nødvendigvis ansett som dyktigere enn jordmor til å tolke CTG, noen av informantene regnet seg selv som svært dyktige. Det ble sagt at når det gjelder tolking av CTG, er det ikke kunnskapen som teller, men hvilken plass du har i systemet.

Jordmødrene beskrev at de hadde makt over kvinnene på den måten at kvinnene stolte på dem. De fikk makt og tillit gjennom omsorgsrollen. Noen av informantene sa at de kunne få kvinnen til å gjøre akkurat som de ville. De kunne la kvinnen "være i kontroll", eller de kunne "ta kontrollen" om de fant det nødvendig. Et par av informantene beskrev at jordmødre kunne misbruke denne makten ved å ta den hele tiden og ved ikke å la kvinnen "være i kontroll". Det ble også nevnt at dette kunne være problematisk for kvinnene, men at de underordna seg og aksepterte at det var jordmor som hadde makta.

Jordmødrene ga kvinnen informasjon om innkomst-CTG på en positiv måte, at det var for å se om barnet hadde det bra og som noe man gjorde ved innleggelsen. Det ble spurt på en måte som gjorde at kvinnene ikke protesterte eller sa at de ikke ønsket å ta testen. Den ble for eksempel ikke presentert med fordeler og ulemper slik at kvinnen kunne gjøre et informert valg. Når jordmor påvirker kvinnen til å gjøre det ”systemet” forventer av henne, kan det hindre at jordmor får kritikk eller andre problemer.

4 DISKUSJON

4.1 Oppsummering

Denne avhandlingen har vurdert i hvilken grad inntakst-CTG kan forutsi uønska fødselsutfall, testens reliabilitet, hvilken effekt inntakst-CTG har sammenliknet med auskultasjon av fosterhertelyden ved inntakst, og hvilken betydning den kan ha for jordmødre i deres daglige arbeide.

Det er gjort en observasjonsstudie om i hvilken grad inntakst-CTG kan forutsi uønska fødselsutfall (studie I). Dette er den første studien om inntakst-CTG som er gjort i Norge (og i Norden). Det er gjort to studier som vurderer testens reliabilitet (studie II og III). Tidligere er det gjort en studie som vurderte inter-observatør enighet⁴². Det ble gjort en systematisk oversikt over i hvilken grad inntakst-CTG kan forutsi uønska utfall, og hvilken effekt testen har sammenliknet med auskultasjon ved inntakst (studie IV). Vi er ikke kjent med at det er gjort en så grundig kunnskapsoppsummering tidligere. The Royal College of Obstetricians and Gynaecologists gjorde en kunnskapsoppsummering om nytteverdien av inntakst-CTG i forbindelse med retningslinjene for bruk av elektronisk fosterovervåking som ble publisert i 2001⁹. De refererer kun til tre studier, og har nok ikke gjort så grundige litteratursøk som vi gjorde. Studie V i denne avhandlingen undersøker hvilken betydning inntakst-CTG kan ha for jordmødre i deres daglige arbeide. Vi er ikke kjent med at det er gjort studier om dette emnet tidligere.

Avhandlingen bidrar med ny kunnskap om nytteverdien og ny forståelse om den komplekse betydningen bruken av testen kan ha. I tillegg er det gjort en kunnskapsoppsummering av all tilgjengelig forskning om inntakst-CTG.

4.2 Diskusjon av resultatene

4.2.1 I hvilken grad testen kan forutsi uønska utfall

I studie I ble inntakst-CTG tolket av de jordmødre og leger som tok hånd om kvinnene, altså i den kliniske konteksten. For hele studiepopulasjon var sensitivitet 15, det vil si at 85 % av de som utviklet føtal distress i fødselsforløpet ikke ble funnet ved hjelp av testen. Av de som testet positivt, var 84 % falsk positive. I undergruppene var det de som fødte innen seks timer etter at inntakst-CTG var tatt som hadde best sensitivitet og PPV, likevel ble 64 % av de som utviklet føtal distress ikke funnet ved hjelp av testen, og tre fjerdedeler av de som testet positivt var falsk positive (tabell 5).

Når innkomst-CTG ble tolket av uavhengige eksperter (studie II og III), ble det funnet høyere verdier på sensitivitet ved cut-off suspekt/avvikende test. Årsaken er at en forholdsvis stor andel tester ble bedømt som unormale, og dette har gått på bekostning av positive prediktive verdier som er gjennomgående lave; dette betyr at langt de fleste som testet positivt var falsk positive. I studie II vurderte ekspert I 10 % av innkomst-CTGene til å være test positive (suspekt eller patologisk) og ekspert II vurderte 16 % til å være test positive. Da de samme CTGene ble vurdert etter samme kriterier av jordmødre og leger i den kliniske situasjonen, ble bare 5 % vurdert til å være suspekter/patologiske. I studie III vurderte de seks uavhengige observatørene at mellom 21 og 34 % av innkomst-CTGene var test positive (avvikende eller patologiske). Da de samme CTGene ble vurdert av jordmødre og leger i den kliniske situasjonen, ble 12 % vurdert til å være test positive. Det er vanskelig å vite årsaken til denne forskjellen. Vi vet imidlertid at i de studiene vi har sett, varierer andel innkomst-CTGer vurdert som unormale fra 5 %¹³⁸ til 32 %¹¹⁴.

I observasjonsstudiene inkludert i studie IV, er også sensitivitet og de prediktive verdiene for positiv test gjennomgående lave. Her er det imidlertid to studier som skiller seg ut ved å ha rimelig gode verdier på utfallsmålet "keisersnitt på indikasjon føtal distress": Ducey og medarbeidere⁴⁰ og Sarno og medarbeidere 1989¹⁴⁶ (tabell 7). Vi kunne ikke finne metodologiske årsaker til at akkurat de to studiene skilte seg ut.

Kulkarni & Shrotri³⁴ vurderte i hvilken grad innkomst-CTG kunne forutsi uheldige utfall blant 100 risikofødende i India. Utfallsmålet var perinatal morbiditet. Sensitivitet var 80, PPV 41, LR+ 2,6 og LR- 0,3. Kvinnene hadde komplikasjoner som blant annet diabetes, kronisk hypertensjon, preeklampsi, vekstretardasjon og oligohydramnion. Kvinnene fikk tatt CTG ved innkomst, resten av forløpet ble de auskultert. Forfatterne skriver at det ikke er praktisk mulig å overvåke alle risikofødsler i India med kontinuerlig overvåking. Dette gjør at resultatene ikke lar seg overføre til en vestlig kontekst, her ville alle kvinnene blitt nøye overvåket gjennom hele fødselsforløpet. Blant kvinnene som hadde reaktiv innkomst-CTG, var det ingen tilfeller av føtal distress i løpet av de neste fem timene (definert som fosterhertelyd som indikerte operativ forløsning og/eller tykt misfarget fostervann og/eller Apgar skåre < 7 etter fem minutter). Ingemarsson og medarbeidere⁷ konkluderte også at en reaktiv test var prediktiv for de neste timene, forutsatt at det ikke oppsto en akutt situasjon.

Tannirandorn og medarbeidere¹³⁷ vurderte innkomst-CTG i en studie som inkluderte 140 kvinner (45 av dem risikofødende) i Bangkok. Sensitivitet var 43, PPV 50, LR+ 19,0 og LR- 0,6. Utfallsmålet var perinatal død og/eller Apgar skåre <7 etter fem minutter og/eller keisersnitt på indikasjon føtal distress og/eller tykt misfarget fostervann og/eller overflytting

til nyfødtavdeling. Kushtagi & Naragoni¹³⁴ vurderte testen i en indisk populasjon bestående av 500 kvinner hvorav 166 var risikofødende. Utfallsmålet var føtal distress definert som misfarget fostervann og/eller unormal hjertelyd under fødselen og/eller Apgar skåre <7 etter ett og fem minutter og/eller overflytting til nyfødtavdeling. Sensitivitet var 53, PPV 61, LR+ 8,7 og LR- 0,5. Det er vanskelig å si hva som er årsaken til at disse to studiene har forholdsvis gode verdier. Begge er prospektive og fødselshjelperne fikk ikke vite resultatene av tolkningen. Det er imidlertid ikke opplyst hvordan kvinnene ble fulgt opp i fødselsforløpet. Begge studiene hadde brede definisjoner på føtal distress; i den ene studien var forekomsten 5 %¹³⁷, i den andre 15 %¹³⁴.

Samlet sett må man likevel kunne si at verken sensitivitet, positiv prediktiv verdi eller likelihood ratios støtter at innkomst-CTG er en nyttig screening for uønska utfall i fødselsforløpet. På den andre siden er spesifisitet og negativ prediktiv verdi gjennomgående høye, noe som betyr at innkomst-CTG i høy grad kan finne de fostrene som ikke vil utvikle føtal distress i fødselsforløpet. Ei lærebok skriver at dersom innkomst-CTG er normal og det ikke er andre risikofaktorer, regnes det som trygt å auskultere hvert 15. minutt og å kjøre 20 minutter CTG med to til tre timers intervaller²². En svensk studie¹⁵⁰ har sammenliknet innkomst-CTG samt intermitterende CTG-registrering i 10-30 minutter med to til to og en halv timer intervaller og kontinuerlig CTG-registrering i åpningstiden hos lavrisikofødende. De fant ingen forskjeller mellom gruppene og konkluderte at begge deler var like trygt. Praksisen med å kjøre innkomst-CTG samt intermitterende CTG-overvåking i fødselsforløpet hos lavrisikofødende har aldri blitt sammenliknet med auskultasjon alene.

Jeg har svært ofte hørt fødselshjelpere argumentere for rutinemessig innkomst-CTG ettersom testen har svært høy spesifisitet og NPV. De som har normal innkomst-CTG har en sannsynlighet på 95 % for å få et barn uten asfyksi eller for inngrep for å forhindre asfyksi. Dette er i og for seg riktig, men har neppe noen kliniske konsekvenser. Det er vel heller ikke alminnelig å screene befolkninger for å finne ut hvem som er friske.

Om vi ser nøyer på resultatene fra studie I, ser vi at det var en pre-test sannsynlighet på 5,8 % for å utvikle føtal distress. Etter en normal innkomst-CTG, var post-test sannsynlighet 5,2 %. Absolutt risikoreduksjon var altså svært beskjeden ($5,8\% - 5,2\% = 0,6\%$), mens relativ risikoreduksjon var ($0,6/5,8 = 10\%$). Man kan beregne hvor mange kvinner som må screenes⁷⁵ med innkomst-CTG for å finne ett barn som vil utvikle føtal distress: ($100/0,6 = 167$ kvinner. Dette er gitt at tegn til føtal distress ikke vil bli oppdaget på andre måter i fødselsforløpet.

Ingemarsson og medarbeidere⁷, hvis artikkel har hatt stor påvirkningskraft når det gjelder innføringen av inntomst-CTG som screening, hadde en pre-test sannsynlighet på 2,1 % og post-test sannsynlighet etter normal test var 1,3 %. Det gir en absolutt risikoreduksjon på 0,8 % og en relativ risikoreduksjon på 38 %. Når utgangsriskoen er så lav som 2,1 %, har en relativ risikoreduksjon på 38 % ubetydelig klinisk verdi.

En fødende kvinne skal ha jordmor sammen med seg eller i umiddelbar nærhet. En av jordmors oppgaver vil være å observere fosterets tilstand, og å sette i gang videre observasjon og tiltak dersom det er tegn til avvik som for eksempel unormal hjertelyd, misfarget fostervann eller unormalt hyppige rier. Dette vil antakelig føre til at de fleste fostre som er truet av asfyksi oppdages i fødselsforløpet. I studie I var det bare sju av de 54 nyfødte med føtal distress (som definert i studien) som hadde Apgar skåre <7 etter fem minutter. Av disse sju barna kom fire seg raskt og ble på barsel, to ble overflytta nyfødtavdeling og kom tilbake til barsel etter henholdsvis en og tre dager og en døde pga mekoniumaspirasjon. Alle sju hadde normal inntomst-CTG.

Sensitiviteten er gjennomgående lav i studiene, og det er antakelig to hovedårsaker til dette. Den ene er at utfallssmålene man screener etter ikke nødvendigvis er til stede når inntomst-CTG blir tatt. Fosteret kan ha det helt fint ved inntomsten og deretter kan det oppstå forhold som truer fosterets helse. Den andre årsaken til lav sensitivitet er antakelig at fosterhjerterefrekvensen ofte er uspesifikk for de tilstandene man screener etter.

Resultatene fra studie I, II og III samt fra de ti observasjonsstudiene inkludert i studie IV støtter ikke bruk av inntomst-CTG som screening for føtal distress (eller andre utfall som tyder på tegn til hypoksi/asfyksi i fødselsforløpet). Kulkarni & Shrotris studie³⁴ kan tyde på at inntomst-CTG kan være nyttig i fødeinstitusjoner i fattige land hvor det er begrensa muligheter for kontinuerlig fosterovervåking i fødselsforløpet. Dette må vurderes i forbindelse med en helhetlig vurdering av ressursene og behovene i de enkelte land. Det vil for eksempel være begrenset nytte av å bruke mange ressurser på CTG-apparater til noen få, mens hovedproblemet er at kvinner flest ikke har tilgang på kyndig fødselshjelper.

4.2.2 Inter-observatør variabilitet

I studie II var vekta kappa mellom 0,25 og 0,38, i studie III mellom 0,57 og 0,75. Andel enighet (Pa) for normal inntomst-CTG var gjennomgående høy i begge studiene, mellom 78 og 89 %. Andel enighet for suspekt (avvikende) og patologisk test var mellom 18 og 33 % i studie II og mellom 56 og 69 % i studie III.

Sammenlikninger av kappa-verdier mellom forskjellige studier er vanskelige å tolke¹²⁰, men i studie II var andelen positive tester lavere enn i studie III og dette kan føre til lavere verdier¹⁵¹. Det var langt større enighet om normale innkomst-CTGer enn om de unormale. I studie III var enigheten om de unormale testene langt høyere enn i studie II. Dette kan skyldes opplæringsprogrammet alle observatørene i studie III hadde gjennomgått. Studiepopulasjonen i studie III inkluderte flere kvinner med unormal innkomst-CTG, dette påvirker også resultatet. I tillegg ble svært mange av testene vurdert som unormale av observatørene, dette vil også påvirke resultatet.

En tidligere studie som vurderte enighet i tolking av 231 innkomst-CTGer mellom jordmødrene i den kliniske konteksten og en uavhengig ekspert, viste kappa 0,47⁴². Studier på inter-observatør enighet i tolking av antepartum og intrapartum CTGer tyder på at dersom observatørene blir bedt om å gjøre ”enkle” vurderinger, som for eksempel å finne basalfrekvensen, kan man oppnå høy grad av enighet. En studie fant kappa 1,00 og Pa 80 % da tre eksperter ble bedt om å vurdere basalfrekvensen i 17 CTG-strimler¹⁵². Andre studier har funnet lav grad av enighet også for vurdering av normale CTG-strimler og basalfrekvens. En studie fant Pa 62 % ved vurdering av normale CTG-strimler¹⁸. To andre studier fant kappa-verdier på 0,16 og 0,51 og Pa 63 % ved vurdering av basalfrekvens^{19; 153}. Ved vurdering av ikke-reaktive CTG-strimler fant en studie Pa 42 % for suspekt CTG og 25 % for patologisk CTG¹⁸. Ved vurderinger av deselerasjoner, varierte Pa mellom 51 og 30 %^{18; 153; 154}. En helt ny studie¹⁵⁵ har vurdert inter-observatør enighet mellom 28 jordmødre i vurdering av tre intrapartum CTGer. De fant kappaverdier på mellom 0,65 og 0,74. De fant høyest enighet i vurderingene av deselerasjoner (kappa = 0,79). Enighet i vurdering av normale CTGer var 0,54, 0,77 for avvikende og 0,75 for patologiske CTGer. Disse resultatene skiller seg fra våre studier. Årsaken kan være en reell forskjell i inter-observatør enighet, men resultatene er vanskelig å sammenlikne ettersom de bare vurderte tre CTG-strimler.

Jordmødre og leger som hadde gjennomgått opplæringsprogrammet ”Fosterovervåking”¹⁰ oppnådde høy grad av enighet i vurdering av 549 innkomst-CTGer (studie III). Det var høyere enighet i vurdering av normale tester enn avvikende/patologiske tester. Det er grunn til å være bekymret over den høye andelen av innkomst-CTGer som ble vurdert som unormale av de seks observatørene i studie III sammenliknet med når de samme testene ble vurdert av jordmødre og leger i den kliniske konteksten. Det er ikke klart hvorfor så mange tester ble vurdert som unormale av observatørene. CTG-strimlene var tatt (og vurdert av jordmødre og leger i den kliniske konteksten) seks til åtte år før de ble vurdert av de seks observatørene. Det er mulig at fødselshjelpere generelt er blitt mer engstelig når de

vurderer CTG, eller det kan være at det er noe ved opplæringsprogrammet som gjør at en stor andel blir vurdert som unormale.

Vi har funnet tre andre studier¹⁵⁶⁻¹⁵⁸ som har vurdert hvilken virkning opplæring har på å vurdere CTG. Fødselshjelpere som fikk opplæring opplevde også at de fikk økt kunnskap. En av studiene¹⁵⁶ fant bedre enighet i vurderinger av basalfrekvens etter opplæring. Ingen av de tre studiene vurderte om opplæring hadde noen virkning på fødselsutfallene. En annen studie¹⁵⁹ fant at månedlige audit-møter og kurser i CTG-tolking hver sjette måned reduserte antall tilfeller med dårlig oppfølging og behandling (sub-optimal care). De neonatale utfallene ble likevel ikke bedre.

Fødselshjelpere bør være klar over at det er en viss uenighet eller variasjon i vurdering av innkomst-CTG, spesielt gjelder dette tester som tolkes som unormale. Det er mulig at opplæringsprogrammer og framtidig bedre teknologi kan endre dette, likevel vil man neppe noen gang oppnå full enighet i tolking av CTGer. Uenighet eller variasjon i tolking kan føre til ulik oppfølging og behandling, og ikke minst kan det ha betydning ved eventuelle rettssaker hvor CTG-strimler kan brukes som viktige dokumenter. Selv om realibiliteten blir bedre, vil dette ikke nødvendigvis føre til forbedring av fødselsutfallene.

4.2.3 Innkomst-CTG sammenliknet med auskultasjon ved innleggelsen

Innkomst-CTG førte til flere obstetriske inngrep uten at det ble vist noen nytteverdi for de nyfødte. Det er vanskelig å forstå at en CTG-registrering på 20-40 minutter kan gjøre at flere kvinner ble liggende med kontinuerlig CTG og at flere fikk epidural, skalpblodprøver og keisersnitt. I studien til Mires og medarbeidere¹¹³ ble 21,5 % av innkomst-CTGene vurdert som unormale, mens det ble funnet 3,6 % unormale hjertelyder ved auskultasjon i kontrollgruppen. Hos Impey og medarbeidere¹¹⁴ var 32 % av innkomst-CTGene vurdert til å være unormale. Det var ikke oppgitt hvor mange som hadde unormal hjertelyd ved auskultasjon. Cheyne og medarbeidere¹¹⁵ oppga ikke hvor stor andel som ble vurdert som unormal. Andelen innkomst-CTGer som ble vurdert som unormal, varierte fra 5,7 til 32 % i de forskjellige studiene (studie IV).

Mires og medarbeidere¹¹³ beskrev ikke kriteriene for tolking av innkomst-CTG. Metodeavsnittet sier ikke om det var standard oppfølging for kvinnene med unormal innkomst-CTG, men i diskusjonen skriver de at ved unormal test vil man vanligvis fortsette med kontinuerlig CTG-overvåking. Impey og medarbeidere¹¹⁴ definerte innkomst-CTG som normal dersom basalfrekvens var mellom 110-160, variabilitet mer enn 5 slag/min, fravær av deselerasjoner og mer enn en akselerasjon i løpet av 20 minutter. Dersom innkomst-CTG ikke

ble vurdert som normal, ble kvinnen CTG-registrert fram til fødselen. Impey og medarbeideres kriterier for normal innkomst-CTG er noe forskjellig fra Ingemarsson og medarbeidere⁷ som vurderte testene som normale dersom det var to akselerasjoner (mer enn 15 slag, mer enn 15 sek) i løpet av 10-20 min eller dersom det ikke var akselerasjoner, men normal basalfrekvens og variabilitet. Cheyne og medarbeidere¹¹⁵ beskrev ikke kriterier for tolking av innkomst-CTG. Årsaken til den økte andelen obstetriske intervensjoner er sannsynligvis at så mange flere kvinner i intervensjonsgruppene fikk kontinuerlig CTG-overvåking i fødselsforløpet. En meta-analyse av ni randomiserte kontrollerte studier som omfattet mer enn 18 500 kvinner sammenliknet kontinuerlig CTG-overvåking med intermitterende auskultasjon under fødselsarbeidet¹⁴. De fant signifikant flere keisersnitt og operative vaginale forløsninger, og færre barn med nyfødtkramper hos kvinnene som hadde kontinuerlig CTG-overvåking under fødselen.

Impey og medarbeidere¹¹⁴ gjorde sin studie ved Dublin National Maternity Hospital hvor det gjøres "active management of labour"¹⁶⁰. Dette innebærer blant annet amniotomi ved innleggelsen, oksytocinstimulering dersom cervixdilatasjonen er <1 cm i timen og at det er ei jordmor eller jordmorstudent til stede hos kvinnen under hele fødselsforløpet. Bare kvinner med klart fostervann ble inkludert i studien. En systematisk oversikt fant at tidlig amniotomi var forbundet med en ikke-signifikant trend i øking av keisersnitt¹⁶¹. En annen systematisk oversikt fant at kvinner som hadde en omsorgsperson til stede under hele fødselsforløpet hadde færre operative forløsninger enn kvinner i kontrollgruppen¹⁰⁹. Det er ikke rutine ved alle fødeavdelinger å gjøre amniotomi ved innkomst, heller ikke at den fødende har ei jordmor eller en annen profesjonell omsorgsperson sammen med seg hele tiden. Resultatene fra Impey og medarbeideres studie¹¹⁴ kan derfor ikke automatisk overføres til andre avdelinger.

Den systematiske oversikten er for liten til å vurdere noen av de neonatale utfallene som perinatal død, gjenoppliving og Apgar skåre < 7 etter fem minutter. For små studier kan føre til type II feil, dvs at man ikke oppdager effekter som er til stede.

4.2.4 Betydning for jordmødre i deres daglige arbeid

Hovedfunnet var at jordmødre opplevde motsetninger mellom seg selv og andre eller i seg selv når de forholdt seg til rutinemessig bruk av innkomst-CTG. Det er mulig at jordmødre opplever liknende motsetninger i andre områder av yrkespraksisen. Jeg har opplevd at jordmødre har hatt delte meninger og heftige debatter om blant annet strenge rutiner for å henge opp oksytocindrypp for å framskynde fødselen, strenge regler for når kvinnen skulle begynne å presse aktivt, rutinemessig oksytocinprofylakse i fødselens tredje fase og

rutinemessig tidlig avklemming av navlesnor. Det var ikke mulig å finne publiserte studier som undersøkte dette.

Det ble uttrykt en bekymring om at jordmødres profesjonelle identitet ble truet på grunn av økende bruk av CTG og annen teknologi. Sinclair¹⁶² har undersøkt jordmødres holdning til bruk av CTG og fant at noen jordmødre mente at bruk av teknologi gjorde at jordmødre ble mindre dyktige. Sinclair har også undersøkt jordmorlederes synspunkter på bruk av teknologi i fødselsomsorgen¹⁶³. Også her fant hun en frykt for at bruk av CTG kunne svekke tradisjonelle ferdigheter som å lytte med trestetoskop og å palpere abdomen. Det fantes også en oppfatning om at teknologi kunne erstatte kliniske ferdigheter hos jordmødre som ikke var så flinke. Woodward¹⁶⁴ skriver at kliniske ferdigheter hos sykepleiere og jordmødre er langt mindre verdsatt enn kunnskaper og ferdigheter om teknologi. Det ble ikke funnet studier som vurderer virkningen av teknologi på jordmødres kliniske ferdigheter. Det er uten tvil en økende bruk av teknologi innen fødselshjelpen, men vi har ingen vitenskapelig kunnskap om hvordan det påvirker jordmødres kliniske ferdigheter. Det er mulig at bruk av teknologi svekker kliniske ferdigheter, men det er også mulig at jordmødre integrerer teknologi på en måte som bevarer og utvikler klinisk dyktighet.

Noen av jordmødrene syntes det var vanskelig å tolke inntakst-CTG som ikke var klart normale eller patologiske. Jordmødrene med kort arbeidserfaring syntes det generelt var vanskelig å tolke CTG. Studie II og III, samt en skotsk studie⁴², fant høy inter-observatør enighet i tolking av normale inntakst-CTG og lav enighet i tolking av unormale tester. Disse studiene sier ingenting om hvorvidt observatørene følte seg trygge eller ikke, men viser at tolking av inntakst-CTG innebærer usikkerhet. Sinclair¹⁶² fant at unge jordmødre (20-29 år) oftere syntes at CTG kunne føre til problemer enn hva deres eldre kolleger mente. Det kom ikke fram av studien hvorfor de mente dette. En annen studie¹⁶⁵ fant at en fjerdedel av en studiepopulasjon av 446 nordiske jordmødre mente at de trengte opplæring i tolking av CTG. To britiske studier viste at manglende opplæring er et viktig punkt ved rettstvister^{20; 21}.

Enkelte av informantene sa at de kunne føle seg utrygge over hvordan utfallet av fødselen ble, spesielt i tilfeller hvor de syntes at inntakst-CTG var vanskelig å tolke. Smythe hevder at innen fødselshjelpen vil usikkerhet alltid være til stede¹⁶⁶. Selv om jordmor arbeider på en måte som ansees som forsvarlig, kan fødselsutfallet være dårlig. Smythe gir også et eksempel på hvordan uforsvarlig fødselshjelp kan gi gode utfall: På ei travel aftenvakt ble en kvinne ”glemt” og var uten tilsyn i mange timer. Da jordmor kom inn på rommet presset kvinnen og fødte et friskt barn like etterpå. Fødselsutfallet trengte ikke forsvarlig fødselshjelp ettersom barnet og fødselsarbeidet var trygge (eller forsvarlige) i seg selv. Omvendt kan heller

ikke selv den beste fødselshjelp garantere gode utfall. Smythe hevder at vi må akseptere at jordmødre noen ganger føler seg utrygge med hensyn til egne avgjørelser og fødselsutfallet.

Det hevdes at det er et økende antall rettssaker hvor fødselshjelpere blir anmeldt, og at CTG-strimler er svært viktige dokumenter i slike tvister²⁰. Noen av jordmødrene i vår studie mente at CTG-strimler var viktigere og mer objektive enn annen dokumentasjon. Det er vanskelig å forstå årsaken til at jordmødrene hadde så forskjellig oppfatning av CTG-strimlenes juridiske gyldighet.

Jordmødrene beskrev rutinemessig bruk av innkomst-CTG som noe som var bestemt av legene. Antropologen Jordan⁹⁶ beskriver omstendighetene omkring fødselen som et kulturelt system. Hun hevder at i den vestlige verden er autoritativ kunnskap omkring fødselen vanligvis forvaltet av leger. Makten i autoritativ kunnskap er ikke at det er denne kunnskapen som er riktig eller sann, men at det er den som regnes for riktig. Autoritativ kunnskap blir til i en prosess som konstruerer og gjenspeiler maktrelasjoner, eller hvordan makta er fordelt. Jordan hevder at obstetrisk teknologi (for eksempel CTG-maskiner) har en symbolsk rolle som kultgjenstander. Eierskapet til kultgjenstandene definerer og viser hvem det er som forvalter autoritativ kunnskap. At alle innkomst-CTGer skal signeres av leger, kan sees i denne sammenhengen. Det er påfallende at når en normal hjerterefrekvens kommer på et stykke papir og gjennom en maskin, skal den vurderes og signeres av en lege. En lege vil neppe spørre ei jordmor nøyte ut om en normal fosterlyd hun har hørt gjennom et trestetoskop, og det forekommer svært sjelden at leger selv undersøker hjertelyder med trestetoskop. Dette kan tyde på at CTG-apparatet som kultgjenstand har en verdi i seg selv, ut over at den viser hjertelyd og rienes hyppighet. Sinclair¹⁶³ fant at jordmorledere mente at når jordmødre hadde kompetanse på og mestret obstetrisk teknologi, førte det til at de fikk et mer likeverdig forhold (power-sharing relationship) med legene. Vi fant ikke noe liknende i vår studie, men vi undersøkte heller ikke jordmødres ferdigheter i å bruke CTG. Studien vår begrensa seg til hvordan innkomst-CTG ble oppfattet.

Jordmødrene uttrykte at de hadde makt eller innflytelse på fødekvinnene, men bare et par av dem reflekterte over at dette kunne ha uheldige virkninger overfor kvinnen. Dette kan kanskje indikere at jordmødrene ønsker å beholde makta i forhold til kvinnene. Henderson¹⁶⁷ har gjort en studie som undersøkte sykepleieres og pasienters syn på at pasienter skulle få medansvar og innflytelse (partnership) i sykehus. Hun fant at de fleste sykepleierne ikke ønsket å dele sin makt med pasientene. I vår studie uttrykte jordmødrene at de hadde makt og innflytelse over kvinnene gjennom sin omsorgsrolle og fordi kvinnene stolte på dem. En

islandsk studie¹⁶⁸ fant at fødekvinne oppfattet at møtet med ei omsorgsfull jordmor ga dem en følelse av å bli styrket (empowered).

4.3 Metodekritikk

4.3.1 Styrker og svakheter i eget materiale (studie I, II og III)

Journalene til alle kvinnene som fødte ved Hammerfest sykehus i årene 1996, 1997 og 1998 ble gjennomgått retrospektivt. Alle papirjournaler (til sammen 1639) og datajournaler ble lest, data ble overført til et registreringsskjema og deretter lagt inn i en database. Dette ble gjort av en person (EB). Det ble tatt kopier av alle innkomst-CTGer, eller CTG ved fødselsstart der det gikk mer enn 24 timer fra innleggelse til fødsel. Det ble funnet journaler på alle kvinnene. I noen tilfeller manglet det dokumenter i papirjournalene, som helsekort for gravide og CTG-strimler. Operative inngrep, indikasjon for inngrep og Apgar skåre var alltid registrert i datajournalen.

Kvinnene ble inndelt i risikogrupper på grunnlag av opplysninger på helsekort for gravide og eventuelle journalnotater fram til innleggelsestidspunktet. Inndeling i risikogrupper ble gjort samtidig med datainnsamlingen. Risikoskåring ble gjort først og uten at resultatet av innkomst-CTG var kjent. Inndeling i risikogrupper ble altså gjort i ettertid, basert på journalopplysninger og ikke på en virkelig klinisk situasjon.

Styrken i datamaterialet er at det er en stor studiepopulasjon og at det ble funnet data for alle kvinnene. Svakheten er at materialet er samlet inn retrospektivt.

I treårsperioden var det en viss utskifting i staben av leger og jordmødre. Andel operative forløsninger og andre intervensjoner i fødselsforløpet (igangsettelse av fødsel, stimulering av rier, kontinuerlig CTG-overvåking) og andel barn med Apgar skåre <7 etter fem minutter var stabil gjennom perioden, mens det var en økning i andel kvinner som fikk epiduralbedøvelse (fra 14 % i 1996 til 20 % i 1998). Materialet er altfor lite til å vurdere endringer i dødelighet. Ved gjennomlesing av journalene fikk jeg inntrykk av at praksis varierte noe mellom forskjellige fødselshjelpere. Det virket som om noen var snarere enn andre til å intervenere og at det var forskjellige terskler for hva som ble definert som ”truende føtal distress” eller ”truende asfyksi”. Det virket også som om noen jordmødre brukte kontinuerlig CTG, epiduralbedøvelse og oksytocinstimulering oftere enn andre. Denne forskjellen i praksis har ikke gitt utslag i materialet, og er antakelig jevnt fordelt gjennom tidsperioden. Økning i andel kvinner som fikk epiduralbedøvelse er nok mer et uttrykk for en trend i tiden enn endringer i personalet.

4.3.2 Studier som vurderer testegenskaper, prediktive verdier og likelihood ratios

Det ideelle er om en klinisk test kan sammenliknes med en helt pålitelig gullstandard. Et eksempel på en pålitelig gullstandard er at testen er utvendig palpasjon av abdomen ved hjelp av Leopolds håndgrep for å finne de fostrene som ligger i seteleie. Etterpå kontrolleres resultatet opp mot gullstandarden ultralydundersøkelse som vil ha en sensitivitet og spesifisitet nær 100 % (forutsatt at den gjøres av en person som kan dette). En test bør altså ideelt sett vurderes opp mot en uavhengig standard test eller en diagnose med 100 % sensitivitet og spesifisitet. I praksis er dette oftest umulig, og det anbefales at begrepet referansestandard ("reference standard") brukes istedenfor gullstandard⁶⁰. I studiene som vurderer inntakst-CTG, er ulike utfallsmål brukt: Føtal distress i følge Ingemarsson og medarbeidere⁷, Apgar skåre <7 etter fem minutter, pH i navlesnor under en viss verdi, operativ forløsning på grunn av føtal distress, gjenoppliving av den nyfødte, overflytting til nyfødtavdeling og perinatal død. Perinatal død, lav navlesnors pH og Apgar skåre er forholdsvis standardiserte mål, men forekommer sjelden (i hvert fall i den vestlige verden) ettersom fødselshjelperne griper inn i forløpet for å forhindre død og skade hos fosteret. Operativ forløsning på indikasjon føtal distress er et mål med stor usikkerhet idet forskjellige fødselshjelpere vil vurdere dette forskjellig. Vi vet at det er store forskjeller i andel operative forløsninger mellom ulike institusjoner og mellom ulike land. Gjenoppliving av den nyfødte og overflytting til nyfødtavdeling er heller ikke standardiserte mål. Etter min oppfatning er Ingemarsson og medarbeideres definisjon av føtal distress⁷ det mest realistiske utfallsmålet ettersom det fanger opp både de som utvikler reell føtal distress (Apgar skåre <7 etter fem minutter) og der man forløser operativt for å forhindre død og skade. Dette utfallsmålet er likevel ingen gullstandard og gir betydelig rom for subjektive vurderinger.

Dersom tolkinga av inntakst-CTG skjer i den kliniske konteksten, påvirker dette antakelig prevalensen for utfallet som blir målt. For eksempel kan tolkinga føre til operativ forløsning for føtal distress. QUADAS-kriteriene for vurdering av diagnostiske studier¹²⁴ har som et av sine punkter at samme kliniske data skal være tilgjengelig ved en evaluering som når testen brukes i praksis. Et annet av QUADAS-kriteriene for evaluering av tester, er at den som vurderer utfallet (referansestandard) ikke skal vite testresultatet. I forskningssammenheng er dette ikke vanskelig å få til, men resultatene kan neppe overføres direkte til det virkelige liv. På fødeavdelingene vil resultatet fra inntakst-CTG alltid være tilgjengelig for fødselshjelperne. Her er det et dilemma: En studie som gjenspeiler de

virkelige forhold er tvilsom metodologisk, mens en god metodologisk studie har dårligere overføringsverdi.

I denne avhandlingen er inkomst-CTGene tolket blindt av eksperter i studie II og III samt i fem av de inkluderte observasjonsstudiene i studie IV^{7; 41; 42; 143; 147}. I studie I er de tolket av fødselshjelperne i den kliniske konteksten. I de øvrige inkluderte observasjonsstudiene i studie IV^{40; 142; 144-146} var det ikke beskrevet hvem som hadde tolket inkomst-CTGene. Blind tolking er selvfølgelig å foretrekke i forskningssammenheng, men det anbefales at CTGer alltid skal tolkes i den kliniske konteksten¹². Resultater fra studier med blind tolking vil derfor ha en noe begrenset overføringsverdi til klinisk praksis.

Når prediktive verdier blir beregnet for mange observatører, blir det mange estimat. Når man beregner så mange estimat, kan det bli betydelige tilfeldige variasjoner mellom observatørene.

Bias er forholdsvis vanlig i studier om diagnostiske tester, og de kan innvirke på resultatene. Lijmer og medarbeidere¹⁶⁹ vurderte hvilken innvirkning design-relaterte bias hadde på diagnostisk nøyaktighet (diagnostic accuracy), det vil si i hvilken grad et resultat basert på målinger representerer den sanne verdien på det som blir målt⁵⁹. Til sammen 184 studier som vurderte 218 diagnostiske tester ble vurdert, og resultater fra studier av dårlig metodologisk kvalitet ble sammenliknet med studier som oppfylte kvalitetskravene. De fant at dersom referanstandard ble tolket av en som visste resultatet av den diagnostiske testen, kunne diagnostisk nøyaktighet (accuracy) overvurderes¹⁶⁹.

De vanligste bias i studier av screening/diagnostiske tester er seleksjonsbias, spektrumbias og lead-timebias. Seleksjonsbias defineres som feil som følge av systematiske forskjeller mellom de som er inkludert i studien og de som ikke er det. Seleksjonsbias kan oppstå dersom det ikke er fortløpende (consecutive) rekruttering til studien, eller om det er en sammenheng mellom testresultatet og rekruttering til studien⁵⁹.

Spektrumbias kan oppstå når sykdomsspekteret blir vurdert i en studiepopulasjon med annerledes klinisk spektrum (andel personer med en bestemt sykdom, og alvorlighetsgraden av denne) enn i den populasjonen testen skal brukes¹⁶⁹. Det vanlige er at tester blir utprøvd og kalibrert i en sykehuskontekst hvor det vil være flere med sykdommen (utfallsmålet) og mer framskredne tilfeller enn i den populasjonen testen er ment å brukes. Prevalens påvirker ikke sensitivitet og LR, men spektrum kan gjøre det. Spektrum er fordelinga av alvorlighetsgrad av tilstanden. Et asfyktisk foster vil for eksempel oftere klassifiseres som test-positivt enn et foster med hypoksemi. Dette kan føre til at LR blir forskjellig for den samme testen i populasjoner med ulikt sykdomsspektrum⁶³.

Seleksjonsbias kan føre til spektrum bias dersom for eksempel flere med positivt testresultat, eller fortrinnsvis mennesker med høyere sannsynlighet for å ha sykdommen inkluderes i studien. Med hensyn til innkomst-CTG så kan spektrumbias være et problem, men kun som følge av seleksjonsbias. I studie I og II er det fortløpende inkludering, i studie III er utvalget gjort med et ønske om å få med så mange positive tester som mulig. Her er det sannsynlig at det er spektrumbias, og det er mulig at de prediktive verdiene er bedre enn de ville vært i en normalpopulasjon. Av observasjonsstudiene inkludert i studie IV, er det seks som har fortløpende (consecutive) inklusjon. Tre har bekvemmelighetsinkludering (convenient inclusion, i dagtiden eller når en bestemt lege har vakt)^{7; 143}, for de andre er det ikke opplyst^{142; 146}. Også her er det muligheter for spektrumbias.

En tredje vanlig bias er lead-timebias¹⁶⁹. Det er når man for eksempel får ett års lengre overlevelse fordi sykdommen ble oppdaget ett år tidligere enn den ville blitt uten testen. Individet får ikke ett års lengre levetid, men ett år lengre med diagnosen. Lead-timebias vil ikke være så veldig aktuelt når det gjelder innkomst-CTG. Likevel er det viktig å forholde seg til denne problematikken. Ved de fleste sykdommer har man et pre-klinisk stadium innen sykdommen gir symptomer. Slik kan det også være med føtal distress, og man kan oppdage en avvikende fosterlyd uten andre tegn på at fosteret har det dårlig. Under fødselen kan lead time være noen få timer, i sjeldne tilfeller bare få minutter. Føtal distress eller noen av de andre utfallsmålene egner seg egentlig ikke spesielt godt å bli screenet for ettersom lead time kan være så kort.

Man kan sette spørsmålstegn ved om det er riktig å bruke begrepet ”diagnostic accuracy” som er i hvilken grad testresultatet representerer den sanne verdien på det som blir målt⁵⁹. Forutsetningene for de utfallsmålene man screener etter (tabell 7), er ikke nødvendigvis til stede ved innleggelsen i fødeavdelingen. Et negativt testresultat kan altså være ”sant” selv om barnet fødes med utfallsmålet det screenes etter. Det er nok riktigere å si at resultatene er en refleksjon over testens prognostiske verdier.

4.3.3 Studier som vurderer inter-observatør enighet

Inter-observatør enighet svarer til testens reliabilitet¹¹⁸. En tests reliabilitet er dens evne til å gi samme resultatet ved gjentatte målinger, dette avhenger av stabiliteten på det som blir målt, på variabiliteten i målemetodene og på dyktigheten til de som tar testen¹¹⁹. I dette tilfellet vil stabiliteten til det som måles være stabiliteten til fosterhjerterefrekvensen. Det er kjent at tolking av CTG-mønstre kan være tvetydig. Nedsatt variabilitet og fravær av akselerasjoner kan skyldes nedsatt surstofftilførsel, medikamentpåvirkning, anomalier i sentralnervesystemet

eller fosterhjertet. Det kan også være et friskt foster som sover¹³. Ved en initial periode med nedsatt variabilitet og fravær av akselerasjoner vil jordmor vanligvis forsøke å vekke fosteret ved hjelp av utvendig manipulering eller ved å gjøre vaginalundersøkelse, men det er ikke kontrollert for dette i studie II og III. Variabiliteten i målemetodene er i dette tilfellet varigheten på testen og mors leie. I studie III ble det kun inkludert CTG-strimler som var 15 minutter eller mer. Vanligvis vil jordmor la CTG-registreringen fortsette til hun ser at testen er normal, men dette er ikke systematisk kontrollert for. Det anbefales at kvinnen skal ligge i sideleie ved CTG-registrering, men heller ikke dette er kontrollert for. Studie II og III vurderer grad av enighet mellom observatører, men man bør være klar over at stabilitet av hjertelyd og variabilitet i målemetode kan påvirke resultatene.

Kappa-verdier påvirkes av flere faktorer og skal tolkes med forsiktighet. Grant¹²¹ hevder at kappa-koeffisienten ikke måler enighet, men tester nullhypotesen om at sammenfall i vurderinger gjort av to eller flere observatører er tilfeldig. Prevalenseffekt oppstår når andelen positive resultater avviker betydelig fra 50 %; biaseffekt er når observatørene har forskjell i andel positive resultater. Begge deler fører til lavere kappa-verdier¹⁵¹. Både i studie II og III er det mulig at både prevalenseffekten og biaseffekten førte til lavere kappa-verdier.

4.3.4 Systematisk oversikt som vurderer effekten av inntakst-CTG sammenliknet med auskultasjon ved innleggelsen

Når man skal vurdere om inntakst-CTG er nyttig for kvinnene og de nyfødte, vil randomiserte kontrollerte studier være det beste. De kan finne ut om det er forskjeller i utfall mellom kvinner som blir screenet med inntakst-CTG og kvinner som ikke blir det. Man skal imidlertid huske at randomiserte kontrollerte studier måler effekten av noe, den gjenspeiler ikke nødvendigvis den kliniske situasjonen. Det er ikke sikkert at resultatene blir de samme som i en studie når intervensjonen settes ut i det virkelige liv (eller fjernes). I Impey og medarbeideres studie¹¹⁴ skulle alle som ikke hadde normal inntakst-CTG etter 20 minutter ha kontinuerlig CTG-overvåking til barnet var født. I virkeligheten ville jordmor kanskje ventet 10-20 minutter til, dersom CTG ble normal, ville hun ha tatt den av. Protokollen i seg selv førte kanskje til en praksis som påvirker resultatet. Når man vurderer nytteverdien av intervensjoner vurdert i randomiserte kontrollerte forsøk, drøfter man vanligvis intervensjonens ”efficacy”¹⁷⁰. Begrepet viser til at de gode resultatene fra studier ofte gjenspeiler resultatene under optimale forhold. Når testen gjøres ”i det virkelige liv” blir ikke alltid effekten av intervensjonen like bra som resultatene fra de randomiserte studiene. Det

kan tenkes at i tilfellet innkomst-CTG er det en omvendt ”efficacy”, at de dårlige resultatene (økt andel obstetriske intervensjoner) delvis skyldtes forhold bestemt av studieprotokollen.

Når man gjør en meta-analyse av randomiserte kontrollerte studier, er det muligheter for flere biaser som kan påvirke resultatene vi kommer fram til⁶⁵. Først er det kvaliteten på de enkeltstudiene man finner: Dersom randomiseringen ikke har vært gjort ordentlig, kan de to gruppene som sammenliknes være forskjellige i utgangspunktet (selection bias). Gruppene som sammenliknes kan få forskjellig oppfølging og behandling bortsett fra intervensjonen (performance bias), vurdering av utfall kan være påvirket av hvilken gruppe individet var randomisert til (detection bias), og resultatene kan påvirkes av bortfall etter randomisering (attrition bias). Dersom man slår sammen flere studier med bias, kan resultatene selvfølgelig bli misvisende. For å unngå disse biasene, gjorde vi kvalitetsvurderinger¹²³ av de tre studiene vi fant. Studier av dårlig metodologisk kvalitet blir vanligvis ekskludert for å unngå bias i resultatene. Andre bias som kan påvirke resultatene er ”reporting bias”. Noen studier er publiserte, andre er det ikke. Studier som finner en forskjell mellom to behandlinger blir oftere publisert enn dersom resultatet er negativt (publication bias). Studier som finner forskjeller blir også hurtigere publiserte enn studier med negative funn (time lag bias). Begge disse biasene kan gjøre at effekten av behandlingen eller intervensjonen overvurderes. Studier i svært kjente tidsskrift siteres ofte og blir mer kjent enn studier fra mindre kjente tidsskrifter (citation bias). Engelskspråklige studier er langt mer tilgjengelig enn for eksempel ungarske (language bias) og selektiv rapportering av noen utfall (outcome reporting bias). For å unngå dette, kontaktet vi eksperter på området for å høre om de visste av pågående og/eller upubliserte studier. Vi søkte aktivt etter studier på andre språk enn engelsk, leste referanselister grundig og kontaktet forfattere for å få rede på utfall de ikke rapporterte, blant annet perinatale dødsfall i to av studiene^{113; 115}. Likevel er det mulig at vi ikke har fått med oss alle kontrollerte studier som er gjort om innkomst-CTG.

4.3.5 Studie som undersøkte hvordan jordmødre opplevde innkomst-CTG i deres daglige arbeid

Informantene i denne studien var 12 jordmødre som arbeidet ved fire forskjellige fødeinstitusjoner i Norge. Selv om de 12 jordmødrene var utvalgt slik at flest mulige erfaringer og oppfatninger skulle representeres, er det mulig at andre jordmødre har andre oppfatninger om innkomst-CTG. Resultatene kan derfor ikke generaliseres til å gjelde alle jordmødre i Norge.

Grounded theory er en induktiv metode som tar utgangspunkt i empiriske data. Likevel vil det være nærmest umulig å hente teori bare fra data. Kvalitativ forskning vil aldri være objektiv. Forskerens perspektiv er alltid begrenset og bestemmer hva som kan sees¹⁷¹. Data i denne studien ble til i møtet mellom de 12 jordmødrene og meg. Mine spørsmål under intervjuene har selvsagt påvirket hvilke emner som ble berørt. Jeg har alltid hatt en kritisk holdning til medikalisering av graviditet og fødsel. Jeg har vært jordmor i snart 20 år, og den nærmest grenseløse tilliten jeg har møtt fra foreldrene har ofte undret meg. Jeg har også lurt på hvordan det kan ha seg at fødselslegene har en så stor makt og innflytelse i fødselsomsorgen som de har, og også på den tro og tillitt gravide og fødselshjelpere har overfor teknologiske hjelpemidler som for eksempel CTG, ultralyd og medikamenter. Det er kanskje ikke så merkelig at intervjuene blant annet har vært innom områdene makt og forhold til teknologi. Forskerens interesser påvirker både spørsmålsstillinga og hvilke data som blir produsert i intervjuene. Dette er ikke nødvendigvis galt, men man skal være klar over at det kan finnes mange ulike versjoner av virkeligheten og at alle kan være gyldige¹²⁶. Resultatene vi kom fram til er ikke nødvendigvis de eneste, de representerer *måter* innkomst-CTG kan oppfattes, ikke *måten* testen oppfattes. For å sikre gyldighet, ble det underveis i analyseprosessen hele tiden ”stilt spørsmål til data”, lett etter data som eventuelt motsa teorier og kategorier som vokste fram¹²⁸ og intervjuene ble grundig gjennomgått etter analysen for å undersøke om kategorier og modeller virkelig kom fra data. Alle intervjuer ble lest av to personer (EB og LÖ). Analysene ble gjort av EB, resultatene ble diskutert og bekreftet av LÖ. Den eksterne valideringa hvor resultatene ble drøftet med fire praktiserende jordmødre, var også nyttig. De fire jordmødrene bekreftet at resultatene kunne være gyldige.

4.4 Generell diskusjon

Jeg har ofte hørt argumenter for at det bør gjøres en randomisert kontrollert studie om effekten av innkomst-CTG i Norge og at de to skotske^{113; 115} og den irske¹¹⁴ studien ikke er representative for oss. Perinatal dødelighet i Norge er 4,8/1000⁴⁹ Dersom vi ønsker å påvise at rutinemessig innkomst-CTG reduserer dødeligheten til det halve, og går ut fra at alle dødsfallene kan forebygges, må vi ha en studie som inkluderer 9772 kvinner. Alle dødsfall kan ikke forebygges, og innkomst-CTG vil bare kunne forebygge dødsfall etter at fødselen er startet. Intrapartum dødelighet var 0,1/1000 i 2002 (Steinar Nilssen, Medisinsk fødselsregister, 24.11.05). Om vi vil gjøre en studie og undersøke om innkomst-CTG reduserer intrapartum dødelighet til det halve, må vi inkludere nærmere en halv million kvinner. Det er mulig at dersom vi hadde gjort en randomisert studie i en fødeavdeling hvor

ikke så mange innkomst-CTGer ble vurdert som unormale, så ville vi ikke funnet den samme økning i intervensjoner hos kvinnene randomisert til innkomst-CTG. Det er lite sannsynlig at vi ville funnet bedre neonatale utfall, og det vil være uetisk med så store og kostbare studier som kun vil kunne påvise marginale nytteverdier.

Det er ikke kjent hva fødekvinnene synes om innkomst-CTG og hvordan de vil reagere dersom rutinen blir avskaffet. Ved noen få fødesteder har man sluttet å ta rutinemessig innkomst-CTG i løpet av de siste årene. Meg bekjent har det ikke vært reaksjoner blant fødekvinnene på grunn av dette.

Det er trygt å føde i Norge, både perinatal og maternell dødelighet er blant de laveste i verden. Om vi skal peke på problemer i dagens fødselshjelp, er det ikke at altfor mange barn får surstoffmangel under fødselen uten at det blir oppdaget. I 2002 var det sju intrapartum dødsfall i Norge. Problemet er snarere at det gjøres for mange intervensjoner og inngrep på kvinner som ikke trenger det^{95; 172}. For mange kvinner blir operativt forløste, spesielt ved hjelp av keisersnitt, og uten at det nødvendigvis gjør at utfallene blir bedre⁹³. En femtedel av kvinner som blir forløste ved hjelp av keisersnitt, opplever komplikasjoner som infeksjoner og større blødninger i forbindelse med keisersnitt, og faren for alvorlige komplikasjoner i forbindelse med narkose, tromboemboli og livmorrupstur ved seinere fødsler er økt¹⁷³⁻¹⁷⁵. Årsakene til den økende keisersnittfrekvensen er mange. Denne avhandlingen viser at rutinemessig bruk av innkomst-CTG antakelig er en av årsakene. I tabell 9 ser vi at det var en økt risiko for keisersnitt blant kvinnene som ble randomisert til innkomst-CTG (RR 1,2, 95 % KI 1,0-1,4). Denne forskjellen var ikke statistisk signifikant på 5 %-nivået. En annen årsak er at flere kvinner blir forløst med keisersnitt uten at det foreligger medisinsk indikasjon, oftest fordi de ikke våger eller ønsker å føde vaginalt. Den aktive og intervensjonerende fødselshjelpen som praktiseres ved de fleste fødeavdelinger med høyt bruk av epiduralbedøvelse og oksytocinstimulering^{176; 177}, som igjen krever kontinuerlig fosterovervåking er antakelig også en årsak. Det har også vært nevnt at engstelse hos fødselshjelperne kan være en medvirkende årsak til rådene kvinnen får og vurderingene som blir gjort¹⁷⁸.

Fødselsomsorgen i Norge er i endring⁶⁹. Noen fødeavdelinger i lokalsykehus er gjort om til fødestuer og større fødeavdelinger har skilt ut jordmorstyrte enheter for lavrisikokvinner. Andre avdelinger skiller kvinnene i lavrisiko som jordmødrene har ansvaret for og høyrisiko som også skal følges opp av legene. Meninga er at kvinnene skal få det tilbudet som er mest hensiktsmessig, og hindre unødvendige intervensjoner¹⁷⁹. Det er mulig at dette vil føre til en framtidig demedikalisering av fødselsomsorgen, noe som er i tråd med

Verdens helseorganisasjons anbefalinger⁹⁵. Det er ikke enighet om fødestuer og lavrisikoenheter er like trygt som vanlige fødeavdelinger for alle lavrisikofødende^{70; 180}.

Det er vist at dersom kvinnen hadde en støtteperson sammen med seg gjennom hele fødselsforløpet, ble behovet for medikamentell smertelindring redusert, det ble færre operative forløsninger og kvinnene var mer fornøyd med fødselsopplevelsen¹⁰⁹. Det ble imidlertid ikke funnet forskjeller i utfallene hos de nyfødte. Virkningene var best dersom støttepersonen ikke var en av personalet på fødeavdelingen, når støttepersonen var til stede fra tidlig i fødselsforløpet og i avdelinger hvor epiduralbedøvelse ikke var et tilbud til alle. Så vidt jeg kjenner til, er det ingen norske fødeinstitusjoner som tilbyr støtteperson til kvinnen under fødselen. Mange avdelinger arbeider mot at kvinnen skal få ha jordmor sammen med seg så mye som mulig i fødselsforløpet og at jordmor ikke skal ha andre oppgaver når hun tar seg av en kvinne som er i aktiv fødsel. Dette kalles en-til-en omsorg.

En måte å forbedre fødselsutfallet (friskt barn med minst mulig intervensjoner), er å utvikle enda bedre teknologi for fosterovervåking. En stor svensk randomisert kontrollert studie¹⁸¹ fant at CTG sammen med ST-analyse av foster-EKG (STAN[®]) var bedre enn CTG alene. Blant kvinner randomisert til STAN[®]-overvåking var det færre nyfødte med metabolsk acidose og neonatal encefalopati, og færre ble operativt forløst på indikasjon føtal distress sammenliknet med kvinnene randomisert til CTG alene. STAN[®]-overvåking er nå i utbredt bruk i Norden. Overvåkingen krever spesialkompetanse, og det er ikke sikkert at metoden fungerer like bra overalt. Resultatene bør bekreftes ved flere studier. Resultatene fra studien¹⁸¹ er basert på at 33 % av alle fødende har behov for CTG-overvåking med skalpelektrode og bygger muligens på et overforbruk, eller en praksis som nødvendiggjør det. Det er en kjensgjerning at det som er nyttig for noen få, som kunstig igangsettelse av fødselen, hinneriving, epiduralbedøvelse, stimulering med oksytocin, keisersnitt og kontinuerlig fosterovervåking, har en tendens til å bre seg videre til friske fødekvinne^{172; 182}. Ved fødeavdelingen i Haugesund og ved Kvinneklubben, Haukeland universitetssykehus blir 20 % av kvinnene overvåket med STAN[®] (Torunn Eikeland og Gunn Gabrielsen, personlig meddelelse). Dersom fødselshjelperne ikke har en kritisk og reflektert holdning, er det mulig at tidligere overforbruk av CTG erstattes av et overforbruk av STAN[®]. Dette kanskje vil være litt bedre, men antakelig ikke den beste løsningen.

5 KONKLUSJONER

5.1 Følger for praksis

Rutinen med å ta innkomst-CTG på lavrisikofødende bør avvikles. Rutinemessig innkomst-CTG må betegnes som bruk av uhensiktsmessig teknologi, og enklere prosedyrer som auskultasjon med trestetoskop eller doppler bør brukes. Dette vil være et skritt i retning av demedikalisering av den normale fødsel, og vil antakelig føre til færre unødvendige intervensjoner. Det finnes ikke tilstrekkelig forskning til å konkludere noe om bruk av innkomst-CTG til risikofødende. Det er gjort svært lite forskning om effekten av kontinuerlig CTG-overvåking til risikofødende. Det er kjent at det er sammenhenger mellom faktorer i svangerskapet og under fødselen og cerebral parese, perinatal død og neonatal encefalopati⁹. Derfor anbefales kontinuerlig fosterovervåking gjennom fødselsforløpet når det er økt risiko for noen av disse tilstandene (for eksempel ved vaginal blødning under fødselen, intrauterine infeksjoner, overtidig svangerskap, mekoniumtilblandet fostervann, oksytocinstimulering)⁹. Disse kvinnene skal overvåkes nøye gjennom forløpet, og innkomst-CTG som screening vil neppe ha noen betydning.

Det finnes jordmødre som mener de får mer informasjon fra en CTG enn fra auskultasjon, og det finnes jordmødre som opplever innkomst-CTG som en trygghet. Derfor kan det være hensiktsmessig å drøfte dette i avdelinga, finne ut hvordan jordmødre og leger forholder seg til dette og eventuelt tilby opplæring i auskultasjon. Det bør være skriftlige retningslinjer for hvilke undersøkelser som bør gjøres ved innleggelsen og hvordan disse skal dokumenteres i journalen. Fødekvinnene bør informeres om hvorfor testen ikke lenger anbefales.

5.2 Forslag til framtidig forskning

Framtidig forskning bør fokusere på hvordan man best kan overvåke høyrisikofødende samtidig som man unngår unødvendige inngrep. Det bør også forskes mer på hvilke kvinner som har risiko for å få et barn med surstoffmangel og derved har nytte av kontinuerlig overvåking under fødselen. Det vil også være nyttig å undersøke hvilken innvirkning den økende bruken av obstetrisk teknologi har på legers og jordmødres tradisjonelle kunnskaper og ferdigheter, som lytting med jordmorstetoskop, vurdere fosterets leie og størrelse, vurdere rier og evne til å beskrive og bruke disse vurderingene. Det bør også forskes på hvordan enighet i tolking av CTG kan bli bedre, spesielt CTGer som kan tolkes som unormale, og om opplæring i tolking fører til bedre fødselsutfall.

Det vil også være nyttig med forskning som tar utgangspunkt i andre perspektiver enn de rent ”nyttemedisinske”. Målet med slik forskning må være å forstå hele spekteret av betydninger obstetrisk teknologi kan ha, spesielt de kulturelle og sosiale betydningene, og hvordan dette virker inn på forståelse, tolking og bruk.

5.3 Anbefalinger

Ettersom CTG (og CTG + ST-analyse) regnes som nyttig ved risikofødsler, er det viktig at fødselshjelperne kan tolke det og kjenner til metodens begrensninger. Det er prinsipielt ingen forskjell i tolking av innkomst-CTG og andre typer CTG. Det bør tilbys opplæring til nyutdanna jordmødre for å hindre at de føler seg utrygge på grunn av manglende kunnskap. Eventuelt må denne opplæringa gjøres i løpet av jordmorutdanninga. Det bør være regelmessig opplæring i bruk og tolking av CTG til alle fødselshjelpere. Fødeinstitusjonene bør ha skriftlige rutiner for hvem som tolker og signerer, og om hvordan journalføring og arkivering skal utføres.

TAKK

Takk til Helse Nord som har finansiert hoveddelen av dette arbeidet. Helse Finnmark har bidratt med lønns- og driftsmidler, og Nordiska skolan för folkhälsovetenskap (NHV) har bidratt med lønnsmidler. Sandviks bokforlag har gitt økonomisk støtte til gjennomføring av studie V.

Min veileder Pål Øian, Universitetssykehuset Nord-Norge/Universitetet i Tromsø, har vært en god støtte underveis, og har alltid tatt seg tid til god og grundig veiledning. Tusen takk! Takk til biveilederne Lennart S. Öhlund, Högskolan i Gävle, og Bo Eriksson, NHV. Pål Øian er også medforfatter på studie I-IV, og Lennart S. Öhlund på studie V.

Flere har bidratt til de fem studiene som inngår i avhandlingen. Takk til medforfatterne Karen Sofie Koss, Sykehuset Asker og Bærum, Oddvar Sviggum, Ålesund sykehus, Atle Klovning, Universitetet i Bergen/Bergen kommune og en spesiell takk til Liv Merete Reinart, Nasjonalt kunnskapssenter for helsetjenesten. Flere ved Institutt for samfunnsmedisin, Universitetet i Tromsø, har gitt gode råd og hjulpet til med statistikk og metode. Takk til Merethe Kumle, Tormod Brenn og Bjørn Straume. Jørgen Hilden, Københavns universitet, gav gode råd da vi strevde som verst med studie IV, og Doris T. Kristoffersen, Nasjonalt kunnskapssenter for helsetjenesten, hjalp til med meta-analysen. Takk til Ingunn Brox og Rachel Myr som har gitt språkråd. Takk også til de åtte observatørene som tolket CTG-strimler (studie II og III), de 12 jordmødrene som ble intervjuet i studie V og de fire jordmødrene som gjorde eksternvalidering av resultatene i studie V.

Takk til Turid Bakkevoll, Universitetssykehuset Nord-Norge for lay-out, bearbeiding av manuskript til trykking og mange gode råd. Takk til jordmor Anne Sagen som stilte opp som fotomodell sammen med Linda Merete Jensen som hadde Sander Schjølberg i magen. Avdelingsjordmor Siv-Heidi Pettersen ved føde/gyn-avdelinga, Hammerfest sykehus har alltid ordnet det slik at jeg har kunnet reise på kurs og få permisjoner når jeg trengte det. Tusen takk!

Mine foreldre, Gunvor og Knut Blix, har passet barn og dyr når det har trengtes. Sist, men ikke minst, takk til Torben, Esben, Anna, Emil og Sigurd. De fire siste er fullstendig uinteresserte i inntøst-CTG, og det er kanskje like greit.

LITTERATUR

- (1) Acheson D. Public health in England: the report of the committee of inquiry into the future development of the public health function. London: Great Britain Secretary of State for Social Services, 1988.
- (2) Svanström L. En introduktion till folkhälsovetenskap. Lund: Studentlitteratur, 2002.
- (3) Sosial- og helsedirektoratet. Retningslinjer for svangerskapsomsorgen. Oslo: Sosial- og helsedirektoratet, 2005.
- (4) Bergsjø P, Brunstad A, Maltau JM. Overvåking og fødselshjelp ved normal fødsel. I: Bergsjø P, Maltau JM, Molne K, Nesheim B-I, (red). Obstetrikk og gynekologi. Oslo: Gyldendal akademisk, 2004.
- (5) Steer P. The mechanisms and management of normal labour. I: Chamberlain G, Steer P, (red). Turnbull's obstetrics. London: Churchill Livingstone, 2001.
- (6) Whittle MJ, Martin WL. Fetal monitoring in labour. I: Chamberlain G, Steer P, (red). Turnbull's obstetrics. London: Churchill Livingstone, 2001.
- (7) Ingemarsson I, Arulkumaran S, Ingemarsson E, Tambyraja RL, Ratnam SS. Admission test: a screening test for fetal distress in labor. Obstet Gynecol 1986; 68:800-806.
- (8) Arney WR. Power and the profession of obstetrics. Chicago: The University of Chicago Press, 1982.
- (9) Royal College of Obstetricians and Gynaecologists. The use of electronic fetal monitoring. London: RCOG Press, 2001.
- (10) Sundström A-K, Rosén D, Rosén KG. Fosterovervåking. Göteborg: Neoventa AB, 2000.
- (11) Tucker SM. Fetal monitoring and assessment. St. Louis: Mosby, 2000.
- (12) FIGO News. Guidelines for the use of fetal monitoring. Int J Gynaecol Obstet 1987; 25:159-167.
- (13) Menihan CA, Zottoli EK. Electronic fetal monitoring. Concepts and applications. Philadelphia: Lippincott Williams & Wilkins, 2001.
- (14) Thacker SB, Stroup D, Chang M. Continuous electronic heart rate monitoring for fetal assessment during labor. Cochrane Database Syst Rev 2001:CD000063.
- (15) Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. Cochrane Database Syst Rev 2003:CD000116.
- (16) Nielsen PV, Stigsby B, Nickelsen C, Nim J. Intra- and inter-observer variability in the assessment of intrapartum cardiotocograms. Acta Obstet Gynecol Scand 1987; 66:421-424.

- (17) Murphy KW, Johnson P, Moorcraft J, Pattinson R, Russell V, Turnbull A. Birth asphyxia and the intrapartum cardiotocograph. *Br J Obstet Gynaecol* 1990; 97:470-479.
- (18) Ayres-De-Campos D, Bernardes J, Costa-Pereira A, Pereira-Leite L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. *Br J Obstet Gynaecol* 1999; 106:1307-1310.
- (19) Donker DK, van Geijn HP, Hasman A. Interobserver variation in the assessment of fetal heart rate recordings. *Eur J Obstet Gynecol Reprod Biol* 1993; 52:21-28.
- (20) Williams B, Arulkumaran S. Cardiotocography and medicolegal issues. *Best Pract Res Clin Obstet Gynaecol* 2004; 18:457-466.
- (21) Symon A. The importance of cardiotocographs. *British Journal of Midwifery* 1997; 5:192-194.
- (22) Gibb D, Arulkumaran S. Fetal monitoring in practice. Oxford: Butterworth Heinemann, 1997.
- (23) Banta HD, Thacker SB. Electronic fetal monitoring. Lessons from a formative case of health technology assessment. *Int J Technol Assess Health Care* 2002; 18:762-770.
- (24) Banta HD, Thacker SB. Assessing the costs and benefits of electronic fetal monitoring. *Obstet Gynecol Surv* 1979;34:627-642.
- (25) Banta HD, Thacker SB. Historical controversy in health technology assessment: the case of electronic fetal monitoring. *Obstet Gynecol Surv* 2001; 56:709-719.
- (26) MacDonald D, Grant A, Sheridan-Pereira M, Boylan P, Chalmers I. The Dublin randomized controlled trial of intrapartum fetal heart rate monitoring. *Am J Obstet Gynecol* 1985; 152:524-539.
- (27) Grant A, O'Brien N, Joy MT, Hennessy E, MacDonald D. Cerebral palsy among children born during the Dublin randomised trial of intrapartum monitoring. *Lancet* 1989; 2:1233-1236.
- (28) Berglund S, Nordström L. National survey (Sweden) of routines for intrapartum fetal surveillance. *Acta Obstet Gynecol Scand* 2001; 80:552-553.
- (29) Devane D, Lalor JG, Daly S, McGuire W. Cardiotocography versus intermittent auscultation of fetal heart on admission for assessment of fetal wellbeing. *The Cochrane Database of Systematic Reviews* 2005; Issue 1, Art No.:CD005122.
- (30) Statens helsetilsyn. Faglige krav til fødeinstitusjoner. Utredningsserie 1-97. Oslo: Statens Helsetilsyn, 1997.
- (31) Dalaker K (red). Veileder i fødselshjelp. Skriftserie for leger: Utdanning og kvalitetssikring. Oslo: Den norske lægeforening, 1998.

- (32) Eikeland T, Blix E, Øian P, Henriksen T, Yli BM, Haugeberg B, Hjelle S. CTG i graviditet og innkomsttest. I: Sand S, Henriksen T, Øian P, (red). Veileder i fødselshjelp 2006. Oslo: Den norske lægeforening, 2006.
- (33) Umstad MP. The predictive value of abnormal fetal heart rate patterns in early labour. *Aust N Z J Obstet Gynaecol* 1993; 33:145-149.
- (34) Kulkarni AA, Shrotri AN. Admission test: a predictive test for fetal distress in high risk labour. *J Obstet Gynaecol Res* 1998; 24:255-259.
- (35) Nordström L, Waldenström U. Handläggning av normal födsel. Stockholm: Socialstyrelsen, 2001.
- (36) Phelan JP. Labor admission test. *Clin Perinatol* 1994; 21:879-885.
- (37) Phelan JP, Ahn MO. Perinatal observations in forty-eight neurologically impaired term infants. *Am J Obstet Gynecol* 1994; 171:424-431.
- (38) Phelan JP. The nonstress test: a review of 3,000 tests. *Am J Obstet Gynecol* 1981; 139:7-10.
- (39) Krebs HB, Petres RE, Dunn LJ, Smith PJ. Intrapartum fetal heart rate monitoring. VI. Prognostic significance of accelerations. *Am J Obstet Gynecol* 1982; 142:297-305.
- (40) Ducey J, Guzman E, Schulman H, Farmakedis G, Karmin I. Value of a screening fetal heart rate tracing in the latent phase of labor. *J Reprod Med* 1990; 35:899-900.
- (41) Sarno AP, Jr., Phelan JP, Ahn MO. Relationship of early intrapartum fetal heart rate patterns to subsequent patterns and fetal outcome. *J Reprod Med* 1990; 35:239-242.
- (42) Farrell T, Mires GJ, Owen P, Patel NB. The influence of interpretation on the value of routine labour admission cardiotocography in a 'low risk' obstetric population. *J Obstet Gynaecol* 1995; 15:161-164.
- (43) Munro J, Ford H, Scott A, Furnival E, Andrews S, Grayson A. Action research project responding to midwives' views of different methods of fetal monitoring in labour. *MIDIRS* 2002; 12:495-498.
- (44) Meberg A, Broch H, Irgens LM. Cerebral parese som nyfødtdedisinsk kvalitetsindikator *Tidsskr Nor Lægeforen* 2001; 121:2917-2922.
- (45) Meberg A, Broch H. Etiology of cerebral palsy. *J Perinat Med* 2004; 32:434-439.
- (46) MacLennan A. A template for defining a causal relation between acute intrapartum events and cerebral palsy: international consensus statement. *BMJ* 1999; 319:1054-1059.
- (47) Jepson HA, Talashek ML, Tichy AM. The Apgar score: evolution, limitations, and scoring guidelines. *Birth* 1991; 18:83-92.

- (48) Medisinsk fødselregister. Årstabeller 1967-1998.
<http://www.uib.no/mfr/tabellverk/aarsmelding/tabelloversikt.html> (05.12.05).
- (49) Medisinsk fødselregister. Årstabeller.
<http://www.uib.no/mfr/tabellverk/aarsmelding/2002/tabelloversikt.html> (02.12.05).
- (50) Moster D. Fødselens epidemiologi. I: Bergsjø P, Maltau JM, Molne K, Nesheim B-I (red). Obstetikk og gynekologi. Oslo: Gyldendal Akademisk, 2004.
- (51) WHO. Perinatal and neonatal mortality: global, regional and country estimates. Geneva: WHO, Department of reproductive Health and Research, 2001.
- (52) WHO. Maternal and newborn health in the European Region: the challenges and the way forward. Fact sheet EURO/03/05. Copenhagen: World Health Organization, 2005.
- (53) Lawson JB, Harrison KA, Bergström S. Maternity care in developing countries. London: RCOG Press, 2001.
- (54) Medisinsk fødselregister. Årstabeller 2000. <http://www.uib.no/mfr/tabellverk/aarsmelding/2000/tabelloversikt.html> (05.12.05) .
- (55) Medisinsk fødselregister. Årstabeller 2001. <http://www.uib.no/mfr/tabellverk/aarsmelding/2001/tabelloversikt.html> (05.12.05).
- (56) Medisinsk fødselregister. Fakta om perinatal dødelighet. Bergen: MFR, 2005.
www.fhi.no (31.03.06).
- (57) WHO. Making pregnancy safer. Geneva: WHO, 1998. http://w3.who.org/EN/Section13/Section36/Section129/Section396_1445.htm (24.11.05).
- (58) Morrison AS. Screening. I: Rothman K, Greenland S (red). Modern epidemiology. Philadelphia: Lippincott-Raven, 1998.
- (59) Last JM. A dictionary of epidemiology. Oxford: Oxford University Press, 1995.
- (60) Knottnerus JA (red). The evidence base of clinical diagnosis. London: BMJ Books, 2002.
- (61) Waldenström U. Det medicinskt möjliga gör oss fartsblinda. Lakartidningen 2005; 102:1882-4, 1886.
- (62) Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology. A basic science for clinical medicine. Boston: Little, Brown and Company, 1991.
- (63) Guyatt G, Sackett D, Haynes B. Evaluating diagnostic tests. I: Haynes B, Sackett D, Guyatt G, Tugwell P (red). Clinical epidemiology. How to do clinical practice research. Philadelphia: Lippincott Williams & Wilkins, 2006. Philadelphia: Lippincott Williams & Wilkins, 2006.
- (64) Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004; 329:168-169.

- (65) Egger M, Smith GD, Altman GD. Systematic reviews in health care. London: BMJ Books, 2001.
- (66) Bergsjø P. Fødselshjelpens organisering og evaluering i Norge. I: Bergsjø P, Maltau JM, Molne K, Nesheim B-I (red). Obstetikk og gynekologi. Oslo: Gyldendal Akademisk, 2004.
- (67) Schmidt N, Abelsen B, Eide B, Øian P. Fødestuer i Norge. Tidsskr Nor Lægeforen 1997; 117:823-826.
- (68) Kjølørød L. Verstefallstenkningen i fødselsomsorgen. Tidsskr f samfunnsforskning 1985; 26:133-157.
- (69) Innst.S nr 300 (2000-2001). Instilling fra Sosialkomiteen om akuttmedisinsk beredskap. 2001.
- (70) Moster D, Lie RT, Markestad T. Relation between size of delivery unit and neonatal death in low risk deliveries: population based study. Arch Dis Child Fetal Neonatal Ed 1999; 80:221-225.
- (71) Schmidt N, Abelsen B, Øian P. Deliveries in maternity homes in Norway: results from a 2-year prospective study. Acta Obstet Gynecol Scand 2002; 81:731-737.
- (72) Declercq E, Menacker F, MacDorman M. Rise in "no indicated risk" primary caesareans in the United States, 1991-2001: cross sectional analysis. BMJ 2005; 330:71-72.
- (73) Douglas M, Wildavsky A. Risk and culture. Berkeley: University of California Press, 1983.
- (74) Teigen KH. Hvordan bedømmes risiko og sjanser i hverdagen? I: Thelle DS, Breivik G, Enebakk V, Skolbekken J-A, Teigen KH, (red). På den usikre siden. Risiko som forestilling, atferd og rettesnor. Oslo: Cappelen akademisk forlag, 2001.
- (75) Bjørndal A, Flottorp S, Klovning A. Medisinsk kunnskapshåndtering. Oslo: Gyldendal Akademisk, 2000.
- (76) Eggebø TM, Gjessing LK. Blødning etter vaginal fødsel. Tidsskr Nor Lægeforen 2000; 120:2860-2863.
- (77) Enkin M, Keirse MJNC, Neilson J, Crowther C, Duley L, Hodnett , Hofmeyr J. A guide to effective care in pregnancy and childbirth. Oxford: Oxford University Press, 2000.
- (78) Haug B. Modernitet - epoke eller erfaring? En analyse av endringene i fødselsomsorgen. Hovedoppgave i sosiologi/sosialpolitikk. Tromsø: Universitetet i Tromsø, 1997.
- (79) Sandvik GB. Moderskap og fødselsarbeid: Diskurser i reproduktivt arbeid. Bergen: Fagbokforlaget, 1997.
- (80) Wagner M. Pursuing the birth machine. Camperdown: ACE Graphics, 1994.

- (81) Walsh D. Feminism and intrapartum care: a quest for holistic birth. I: Stewart M (red). Pregnancy, birth and maternity care. Feminist perspectives. Edinburgh: Books for Midwives, 2004.
- (82) Helman CG. Culture, health and illness. Oxford: Butterworth Heinemann, 1994.
- (83) Backe B, Øian P, Eide B, Nilsen ST, Pedersen SP, Schmidt N, Schou E, Taraldsen S. Desentralisert og differensiert fødselshjelp - et paradigmeskifte. Tidsskr Nor Lægeforen 2005; 125:606-607.
- (84) Moster D, Lie RT, Markestad T. Fornuft og følelser i fødselsomsorgen. Tidsskr Nor Lægeforen 2005; 125:2818-2820.
- (85) Backe B, Øian P. Fornuft og følelser om små tall. Tidsskr Nor Lægeforen 2005; 125:3150.
- (86) WHO. Care in normal birth: a practical guide. Geneva: World health organization, 1996.
- (87) Bergsjø P. Fødselsdynamikk. I: Bergsjø P, Maltau JM, Molne K, Nesheim B-I (red). Obstetikk og gynekologi. Oslo: Gyldendal akademisk, 2004.
- (88) WHO. Maternal mortality in 2000: Estimates developed by WHO, UNICEF and UNFPA. Geneva: World Health Organization, 2004.
- (89) Diniz SG, Chacham AS. "The cut above" and "the cut below": the abuse of caesareans and episiotomy in Sao Paulo, Brazil. Reprod Health Matters 2004; 12:100-110.
- (90) WHO. Research on reproductive health at WHO: biennial report 2000-2001. Geneva: UNDP/UNFPA/WHO/World bank special programme of research development and training in human reproduction, 2004.
- (91) Afghan ministry of public health, UNICEF, Center for disease control and prevention. Maternal mortality in Afghanistan: magnitude, causes, risk factors and preventability. Kabul: Afghan Ministry of Public Health, 2002.
- (92) Thomson A. Is there evidence for the medicalisation of maternity care? MIDIRS 2000; 10:416-420.
- (93) Kolås T, Hofoss D, Daltveit AK, Nilsen ST, Henriksen T, Häger R, Ingemarsson I, Øian P. Indications for cesarean deliveries in Norway. Am J Obstet Gynecol 2003; 188:864-870.
- (94) WHO. Having a baby in Europe. Copenhagen: WHO Regional Office for Europe, 1985.
- (95) Chalmers B, Mangiaterra V, Porter R. WHO principles of perinatal care: the essential antenatal, perinatal, and postpartum care course. Birth 2001; 28:202-207.

- (96) Jordan B. Authoritative knowledge and its construction. I: Davis-Floyd RE, Sargent CF (red). Childbirth and authoritative knowledge. Berkely: University of California Press, 1997.
- (97) Blåka G. Grunnlagstenking i et kvinnefag. Bergen: Fagbokforlaget, 2002.
- (98) Hofmann B. The technological invention of disease - on disease, technology and values. Oslo: Unipub AS, 2002.
- (99) Hofmann B. On the value-ladenness of technology in medicine. Med Health Care Philos 2001; 4:335-346.
- (100) Vollebæk LE. Teknologi og magi. Tidsskr Sykepl 2004; 93:44-45.
- (101) Fraser DM, Cooper MA (red). Myles textbook for midwives. Edinburgh: Churchill Livingstone, 2003.
- (102) Faxelid E, Hogg B, Kaplan A, Nissen E. Lärobok för barnmorskor. Lund: Studentlitteratur, 2001.
- (103) Page L. The midwife's role in pregnancy and labour. I: Chamberlain G, Steer P (red). Turnbull's obstetrics. London: Churchill Livingstone, 2001
- (104) Statens helsetilsyn. Legkvinnekonferanse om fødsels- og barselomsorgen. Oslo: Statens helsetilsyn, 1999.
- (105) ICM. Internasjonal definisjon av jordmor. Den Haag: International Confederation of Midwives, 2005. www.internationalmidwives.org. (10.11.05)
- (106) Helse- og omsorgsdepartementet. Lov om helsepersonell. Oslo: Helse- og omsorgsdepartementet, 1999.
- (107) Council directive of 21.January 1980 concerning the coordination of provisions laid down by law, regulation or administrative action relating to the taking up and pursuit of the activities of midwives. Rdir. 1980/115/EØF ed. 1980.
- (108) Ohnstad B. Jordmoransvaret. Oslo: Den norske jordmorforening, 2004.
- (109) Hodnett ED, Gates S, Hofmeyr GJ, Sakala C. Continuous support for women during childbirth. Cochrane Database Syst Rev 2003:CD003766.
- (110) AbouZahr C. Maternal mortality overview. In: Murray CJI, Lopez A, editors. Health dimensions og sex and reproduction; the global burden of sexually transmitted diseases, HIV, maternal conditions, perinatal disorders, and congenital anomalies. Cambridge: Harvard School of Public Health, 1998.
- (111) Mander R, Fleming V. Failure to progress. The contraction of the midwifery profession. London: Routledge, 2002.
- (112) Heptinstall T, Lee G. Place of birth. I: Henderson C, Macdonald S (red). Mayes' midwifery. Edinburgh: Ballière Tindall, 2004.

- (113) Mires G, Williams F, Howie P. Randomised controlled trial of cardiotocography versus Doppler auscultation of fetal heart at admission in labour in low risk obstetric population. *BMJ* 2001; 322:1457-1460.
- (114) Impey L, Reynolds M, MacQuillan K, Gates S, Murphy J, Sheil O. Admission cardiotocography: a randomised controlled trial. *Lancet* 2003; 361:465-470.
- (115) Cheyne H, Dunlop A, Shields N, Mathers AM. A randomised controlled trial of admission electronic fetal monitoring in normal labour. *Midwifery* 2003; 19:221-229.
- (116) Grant JM. Editor's Choice. *BJOG: An International Journal of Obstetrics and Gynaecology* 2003; 110:xv-xvii.
- (117) Ingemarsson I, Ingemarsson E. Elektronisk fosterövervakning. Lund: Studentlitteratur, 1987.
- (118) Khan KS, Chien PF. Evaluation of a clinical test. I: assessment of reliability. *BJOG* 2001; 108:562-567.
- (119) Rothman KJ, Greenland S. Modern epidemiology. Philadelphia: Lippincott-Raven, 1998.
- (120) Fleiss JL. Statistical methods for rates and proportions. New York: Wiley, 1981.
- (121) Grant JM. The fetal heart rate trace is normal, isn't it? Observer agreement of categorical assessments. *Lancet* 1991; 337:215-218.
- (122) Donker DK, Hasman A, van Geijn HP. Interpretation of low kappa values. *Int J Biomed Comput* 1993; 33:55-64.
- (123) Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; 357:1191-1194.
- (124) Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3:25.
- (125) Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006; 6:9.
- (126) Malterud K. Kvalitative metoder i medisinsk forskning. Oslo: Universitetsforlaget, 2003.
- (127) Kvale S. Det kvalitative forskningsintervju. Oslo: Gyldendal Akademisk, 2001.
- (128) Strauss A, Corbin Juliet. Basics of qualitative research. Thousand Oaks: Sage Publications, 1998.

- (129) Taylor SJ, Bogdan R. Introduction to qualitative research methods. New York: John Wiley & Sons, 1998.
- (130) Delleve L, Abrahamsson KH, Trulsson U, Hallberg L R-M. Grounded theory in public health research. I: Hallberg L R-M (red). Qualitative methods in public health research. Lund: Studentlitteratur, 2002.
- (131) QRS NVivo 2. Melbourne: QRS International, 2002.
- (132) Mays N, Pope C. Rigour and qualitative research. *BMJ* 1995; 311:109-112.
- (133) Golditch BD, Ahn MO, Phelan JP. The fetal admission test and intrapartum fetal death. *Am J Perinatol* 1998; 15:273-276.
- (134) Kushtagi P, Naragoni S. Labour admission test-an effective risk screening tool. *J Indian Med Assoc* 2002; 100:234-236.
- (135) Villarreal EC, Alcalà FM, Sevilla RM, Alvarez NC. Cardiotocografia al ingreso en pacientes de riesgo normal. *Ginec Obstet Mèx* 1978; 44:175-181.
- (136) Feinkind L, Abulafia O, Delke I, Feldman J, Minkoff H. Screening with Doppler velocimetry in labor. *Am J Obstet Gynecol* 1989; 161:765-770.
- (137) Tannirandorn Y, Wacharaprechanont T, Phaosavasdi S. Fetal acoustic stimulation for rapid intrapartum assessment of fetal well-being. *J Med Assoc Thai* 1993; 76:606-612.
- (138) Blix E, Øian P. Labor admission test: an assessment of the test's value as screening for fetal distress in labor. *Acta Obstet Gynecol Scand* 2001; 80:738-743.
- (139) Chan FY, Lam C, Lam YH, To WK, Pun TC, Lee CP. Umbilical artery Doppler velocimetry compared with fetal heart rate monitoring as a labor admission test. *Eur J Obstet Gynecol Reprod Biol* 1994; 54:1-6.
- (140) Kunz J, Schmid J, Bader P. Aufnahme-amnioskopie und aufnahme-cardiotokogramm. *Z Geburtshilfe Perinatol* 1974; 178:389-392.
- (141) Rossavik IK. The predictive value of cardiotocography. *Ann Chir Gynaecol* 1980; 69:75-78.
- (142) Chua S, Arulkumaran S, Kurup A, Anandakumar C, Selemat N, Ratnam SS. Search for the most predictive tests of fetal well-being in early labor. *J Perinat Med* 1996; 24:199-206.
- (143) Ingemarsson I, Arulkumaran S, Paul RH, Ingemarsson E, Tambyraja RL, Ratnam SS. Fetal acoustic stimulation in early labor in patients screened with the admission test. *Am J Obstet Gynecol* 1988; 158:70-74.
- (144) Farrell T, Seaton L, Owen P. Evaluation of fetal movements as an early labour admission test in low-risk pregnancies. *Clin Exp Obstet Gynecol* 1998; 25:23-25.

- (145) Somerset DA, Murrills AJ, Wheeler T. Screening for fetal distress in labour using the umbilical artery blood velocity waveform. *Br J Obstet Gynaecol* 1993; 100:55-59.
- (146) Sarno AP, Jr., Ahn MO, Brar HS, Phelan JP, Platt LD. Intrapartum Doppler velocimetry, amniotic fluid volume, and fetal heart rate as predictors of subsequent fetal distress. I. An initial report. *Am J Obstet Gynecol* 1989; 161:1508-1514.
- (147) Elimian A, Lawlor P, Figueroa R, Wiencek V, Garry D, Quirk JG. Intrapartum assessment of fetal well-being: any role for a fetal admission test? *J Matern Fetal Neonatal Med* 2003; 13:408-413.
- (148) Blix E, Sviggum O, Koss KS, Øian P. Inter-observer variation in assessment of 845 labour admission tests: comparison between midwives and obstetricians in the clinical setting and two experts. *BJOG* 2003; 110:1-5.
- (149) Petrie A, Sabin C. *Medical statistics at a glance*. Oxford: Blackwell Science, 2000.
- (150) Herbst A, Ingemarsson I. Intermittent versus continuous electronic monitoring in labour: a randomised study. *Br J Obstet Gynaecol* 1994; 101:663-668.
- (151) Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *J Clin Epidemiol* 2000; 53:499-503.
- (152) Bernardes J, Costa-Pereira A, van Geijn H, Pereira-Leite L. A more objective fetal heart rate baseline estimation. *Br J Obstet Gynaecol* 1996; 103:714-715.
- (153) Bernardes J, Costa-Pereira A, Ayres-De-Campos D, van Geijn HP, Pereira-Leite L. Evaluation of interobserver agreement of cardiotocograms. *Int J Gynaecol Obstet* 1997; 57:33-37.
- (154) Ayres-De-Campos D, Bernardes J. Early, variable and late decelerations: can a consensus be reached in their identification? *Int J Gynaecol Obstet* 1999; 65:305-306.
- (155) Devane D, Lalor J. Midwives' visual interpretation of intrapartum cardiotocographs: intra- and inter-observer agreement. *J Adv Nurs* 2005; 52:133-141.
- (156) Ayres-De-Campos D, Bernardes J, Marsal K, Nickelsen C, Makarainen L, Banfield P, Xavier P, Campos I. Can the reproducibility of fetal heart rate baseline estimation be improved? *Eur J Obstet Gynecol Reprod Biol* 2004; 112:49-54.
- (157) Beckley S, Stenhouse E, Greene K. The development and evaluation of a computer-assisted teaching programme for intrapartum fetal monitoring. *BJOG* 2000; 107:1138-1144.
- (158) Trepanier MJ, Niday P, Davies B, Sprague A, Nimrod C, Dulberg C, Watters N. Evaluation of a fetal monitoring education program. *J Obstet Gynecol Neonatal Nurs* 1996; 25:137-144.

- (159) Young P, Hamilton R, Hodgett S, Moss M, Rigby C, Jones P, Johanson R. Reducing risk by improving standards of intrapartum fetal care. *J R Soc Med* 2001; 94:226-231.
- (160) Bohra U, Donnelly J, O'Connell MP, Geary MP, MacQuillan K, Keane DP. Active management of labour revisited: the first 1000 primiparous labours in 2000. *J Obstet Gynaecol* 2003; 23:118-120.
- (161) Fraser WD, Turcot L, Krauss I, Brisson-Carrol G. Amniotomy for shortening spontaneous labour. *Cochrane Database Syst Rev* 2000: CD000015.
- (162) Sinclair M. Midwives' attitudes to the use of the cardiotocograph machine. *J Adv Nurs* 2001; 35:599-606.
- (163) Sinclair M. Midwifery managers' perspectives on midwives' use of birth technology. *All Ireland Journal of Nursing & Midwifery* 2001; 1:213-219.
- (164) Woodward VM. Professional caring: a contradiction in terms? *J Adv Nurs* 1997; 26:999-1004.
- (165) Sinclair M, Gardner J. Midwives' perceptions of the use of technology in assisting childbirth in Northern Ireland. *J Adv Nurs* 2001; 36:229-236.
- (166) Smythe E. Uncovering the meaning of 'being safe' in practice. *Contemp Nurse* 2003; 14:196-204.
- (167) Henderson S. Power imbalance between nurses and patients: a potential inhibitor of partnership in care. *J Clin Nurs* 2003; 12:501-508.
- (168) Halldorsdottir S, Karlsdottir SI. Empowerment or discouragement: women's experience of caring and uncaring encounters during childbirth. *Health Care Women Int* 1996; 17:361-379.
- (169) Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, Bossuyt PMM. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282:1061-1066.
- (170) Khan KS, Kunz R, Kleijnen J, Antes G. Systematic reviews to support evidence-based medicine. London: The Royal Society of Medicine Press Ltd., 2003.
- (171) Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet* 2001; 358:483-488.
- (172) Moen MS, Holmen M, Tollefsrud S, Rolland R. Lavrisikofødende ved en kvinneklinikk - hvordan føder de? *Tidskr Nor Lægeforen* 2005; 125:2635-2637.
- (173) Häger RM, Daltveit AK, Hofoss D, Nilsen ST, Kolaas T, Øian P, Henriksen T. Complications of cesarean deliveries: rates and risk factors. *Am J Obstet Gynecol* 2004; 190:428-434.
- (174) Wen SW, Rusen ID, Walker M, Liston R, Kramer MS, Baskett T, Heaman M, Liu S. Comparison of maternal mortality and morbidity between trial of labor and elective

- cesarean section among women with previous cesarean delivery. *Am J Obstet Gynecol* 2004; 191:1263-1269.
- (175) Liu S, Heaman M, Joseph KS, Liston RM, Huang L, Sauve , Kramer MS. Risk of maternal postpartum readmission associated with mode of delivery. *Obstet Gynecol* 2005; 105:836-842.
 - (176) Barratt-Due A, Hagen I, Dahl V. Smertelindring til fødende i norske sykehus. *Tidsskr Nor Lægeforen* 2005; 125:2504-2506.
 - (177) Blix E, Pettersen SH, Eriksen H, Røyset B, Pedersen EH, Øian P. Bruk av oksytocin som ristimulerende medikament etter spontan fødselsstart. *Tidsskr Nor Lægeforen* 2002; 122:1359-1362.
 - (178) Klein MC. Obstetrician's Fear of Childbirth: How Did It Happen? *Birth* 2005; 32:207-209.
 - (179) Reime B, Klein MC, Kelly A, Duxbury N, Saxell L, Liston R, Prompers FJ, Entjes RS, Wong V. Do maternity care provider groups have different attitudes towards birth? *BJOG* 2004; 111:1388-1393.
 - (180) Gottvall K, Grunewald C, Waldenstrom U. Safety of birth centre care: perinatal mortality over a 10-year period. *BJOG* 2004; 111:71-78.
 - (181) Amer-Wahlin I, Hellsten C, Noren H, Hagberg H, Herbst A, Kjellmer I, Lindoff C, Månsson M, Mårtensson L, Olofsson P, Sundström A, Marsal K. Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. *Lancet* 2001; 358:534-538.
 - (182) Johanson R, Newburn M, Macfarlane A. Has the medicalisation of childbirth gone too far? *BMJ* 2002; 324:892-895.

SUMMARY IN ENGLISH

The dissertation is based on the following papers

- Paper I Blix E, Øian P. Labor admission test: An assessment of the test's value as screening for fetal distress in labor. *Acta Obstet Gynecol Scand* 2001;80:738-43.
- Paper II Blix E, Sviggum O, Koss KS, Øian P. Inter-observer agreement in assessment of 845 labour admission tests: comparison between midwives/obstetricians in the clinical setting and two experts. *BJOG* 2003;110:1-5.
- Paper III Blix E, Øian P. Interobserver agreements in assessing 549 labor admission tests after a standardized training program. *Acta Obstet Gynecol Scand* 2005;84:1087-92.
- Paper IV Blix E, Reinar LM, Klovning A, Øian P. Prognostic value of the labour admission test and its effectiveness compared with auscultation only: a systematic review. *BJOG* 2005;112:1595-1604.
- Paper V Blix E, Öhlund LS. Norwegian midwives' perception of the labour admission test. *Midwifery* 2006 (Epub ahead of print).

The papers have been reprinted by kind permission of the publishers.

Background

Public health is the science and art of preventing disease, prolonging life through and promoting health organised efforts of society¹. Surveillance of the fetus during labour is preventive care; the aim of fetal surveillance and other forms of labour management is to ensure the delivery of a healthy baby in good condition with the minimum of intervention².

Cardiotocography (CTG) is a complex technology applied for surveillance of the fetus during pregnancy, labour and delivery. The fetal heart rate and the frequency of the uterine contractions can be continuously monitored. There are two methods of monitoring the fetal heart rate: the external method, in which an ultrasound transducer is placed on the mother's abdomen by an elastic belt, and the internal method, in which electrocardiographic leads are attached directly to the fetus after rupture of the membranes. The uterine activity is monitored by an external tocotransducer or by an intra-uterine catheter; the transducers or electrodes are connected to a cardiotocograph where the fetal heart rate and the frequency and duration of the contractions are displayed. Both are recorded in a continuous tracing on paper³.

The labour admission test is a CTG of 20-30 minutes' duration, carried out when the woman is admitted to the labour ward. The test is a part of the labour admission assessment, which, among other things, includes assessing the size, position and presentation of the fetus, the frequency and strength of the contractions and how the woman is coping with the contractions. The test was introduced as a risk screening in early labour, to detect the compromised fetus on admission and to select the women in need of continuous electronic fetal monitoring during labour^{4; 5}. The test is commonly used for screening in many countries^{6; 7}, and was introduced at a time when scientific knowledge about the test was scant, but trust in the benefits of electronic fetal monitoring was great. No systematic assessments or evaluations were performed before the test was taken into widespread use. Screening is the presumptive identification of unrecognised disease by the application of tests, examinations or other procedures. Screening tests sort out persons without symptoms who probably have a disease from those who probably do not⁸.

British guidelines published in 2001⁹ do not recommend the labour admission test in low risk women, while Swedish guidelines published the same year¹⁰ recommend the test in all women. The British recommendations were based upon three studies^{4; 11; 12}, and the Swedish upon seven studies^{4; 5; 13-17}.

Aim of the dissertation

The aim of the present dissertation was to assess to what degree the labour admission test can predict adverse outcomes, the effectiveness of the test in preventing adverse outcomes compared with auscultation only, the inter-observer variability (or the test's reliability) and to explore what information and knowledge the labour admission test is perceived to provide, and what meaning the test carries in the daily work of practicing midwives.

Material and methods

The patient records of all women (n=1639) who gave birth at Hammerfest Hospital in 1996, 1997 and 1998 were retrospectively reviewed. Variables like indication for admission, how the labour admission test was classified, time span from admission to delivery, interventions, medicaments, mode of delivery, indication for operative delivery, Apgar score, resuscitation of the infant, admission to neonatal unit and other maternal and neonatal outcomes were registered. Copies were taken of all labour admission tracings. Samples from these data were used in Studies I-III.

Paper I: Women who were admitted because of labour, were in the first stage of labour on admission, had a singleton pregnancy and a gestational age of more than 196 days and gave birth within 24 hours after the labour admission test were included. In all, 932 women, both high-risk and low-risk, met the inclusion criteria. Fetal distress was defined as operative delivery for the indication changed/ominous fetal heart rate, or Apgar score less than 7 five minutes after spontaneous delivery (or after operative delivery for other reasons than fetal distress)⁴. Sensitivity, specificity, positive and negative predictive value (PPV and NPV) and likelihood ratio for positive test result (LR+) and for negative test result (LR-) for fetal distress in labour were calculated at different cut-off levels in the whole population and in subgroups of high- and low-risk women, in women who delivered within six hours after the labour admission test and in women who delivered between six and 24 hours after the test.

Paper II: Women who were admitted because of labour, were in the first stage of labour on admission, had a singleton pregnancy and a gestational age of more than 196 days, gave birth within 24 hours after the labour admission test and where the labour admission tracing was present in the woman's file were included. In all, 845 women, both high-risk and low-risk, met the inclusion criteria. Two experts were given copies of the CTG tracings for assessment. The expert assessment was carried out by two consultant obstetricians from other institutions. Inter-observer agreement (weighted kappa (κ_w) and proportion of agreement (Pa)) between the two experts and between the experts and the midwives/obstetricians who had assessed the tracings in the clinical setting were calculated. Sensitivity, specificity, PPV, NPV, LR+ and LR- for fetal distress in labour were calculated for the two experts.

Paper III: Women who gave birth within 24 hours after the labour admission test, had a singleton pregnancy and a gestational age of more than 196 days, where the labour admission tracing was present in the woman's file and had a labour admission test of minimum 15 minutes duration were included. In all, 1199 women met the inclusion criteria. A sample of 549 CTG tracings was included with the purpose to include as many abnormal tests as possible. The sample included both high-risk and low-risk women. The tests were assessed by three midwives and three obstetricians from other institutions who had completed a standard training program¹⁸. Inter-observer agreement (κ_w) and Pa) were calculated. Sensitivity, specificity, PPV, NPV, LR+ and LR- for fetal distress in labour was calculated for each of the six observers.

Paper IV: To assess the effectiveness of the labour admission test in preventing adverse outcomes compared with auscultation only and to assess the test's ability to predict adverse outcomes, a systematic review was performed. Three randomised controlled trials and

11 observational studies were included. The randomised studies comprised 11 259 women and the observational studies 5831 women. Main outcome measures were obstetric interventions (augmentation of labour, continuous electronic fetal monitoring, epidural analgesia, fetal blood sampling and operative deliveries) and neonatal outcomes (perinatal mortality, Apgar score, seizures, resuscitation and admission to neonatal unit). Relative risks (RR) for the outcome measures and sensitivity, specificity, PPV, NPV, LR+ and LR- for different adverse outcomes (operative deliveries for fetal distress, Apgar score less than seven five minutes after delivery, fetal distress⁴, resuscitation of the infant and admission to neonatal unit) were calculated.

Paper V: This is a qualitative study where a theoretical sample of 12 practising midwives from four different institutions were interviewed in order to explore how the labour admission test is perceived. The in-depth interviews were transcribed verbatim and analysed using the grounded theory technique¹⁹.

Findings

When the tests were assessed by midwives and obstetricians in the clinical setting (Paper I), sensitivity in predicting fetal distress was 15, specificity, 95, PPV 16, NPV 95, LR+ 3.2 and LR- 0.9 in the entire study population. In the subgroups of low- and high-risk women, who delivered within six hours or between six and 24 hours after the labour admission test, sensitivity varied between 0 and 36, and PPV varied between 0 and 27. Specificity varied between 92 and 96, and NPV between 89 and 97. When the tests were assessed by experts in a non-clinical setting (Papers II and III), sensitivity varied between 22 and 48, specificity 72 and 70, PPV 11 and 15, NPV 94 and 95, LR+ 2.3 and 1.6 and LR- 0.8 and 0.9. For the 11 observational studies include in the systematic review (Paper IV), the values were generally low except for specificities and positive predictive values: sensitivity was above 50 in four of 28 single outcomes, LR+ was above 10 in two outcomes and between 5 and 10 in six outcomes. LR- was 0.2 in one single outcome, and between 0.5 and 0.2 in three of the 28 single outcomes.

When the labour admission tests were classified by midwives/obstetricians in the clinical setting and two experts in a non-clinical setting (Paper II), κ varied between 0.25 and 0.38 between the three pairs of observers. Pa for a normal test varied between 0.85 and 0.89 and Pa for an equivocal/ominous test between 0.18 and 0.33. When the tests were assessed by six observers who had completed a standardised training program and in a non-clinical setting, κ varied between 0.57 and 0.75 in the 15 pairs of observers. Pa for a normal

test varied between 0.78 and 0.88, while Pa for intermediary/abnormal test varied between 0.56 and 0.69.

Meta-analysis of the controlled trials found that women randomised to the labour admission test were more likely to have minor obstetric interventions like epidural analgesia (RR 1.2; 95 % confidence interval (CI) 1.1-1.4), continuous electronic fetal monitoring (RR 1.3; 95 % CI 1.2-1.5) and fetal blood sampling (RR 1.3; 95 % CI 1.1-1.5) compared with women randomised to auscultation on admission. There were trends towards more operative deliveries for all reasons (caesarean section, vacuum and forceps) (RR 1.1; 95 % CI 1.0-1.3), operative deliveries for fetal distress (RR 1.1; 95 % CI 1.0-1.3) and caesarean sections for all reasons (RR 1.2; 95 % CI 1.0-1.4) among the women randomised to the labour admission test, although these differences did not reach statistical significance.

The analyses of the twelve interviews identified the core category “experiencing contradictions”, which explains what meanings the labour admission test can carry for practising midwives. The core category indicated that the midwives found conflicting interests within themselves or between themselves and others when using the labour admission test. They experienced contradictions between professional identity and the increasing use of obstetric technology, between feeling safe and feeling unsafe and between having power and being powerless.

Conclusions

There is no scientific evidence of benefit from the labour admission test in low risk women. The test performs poorly in preventing adverse outcomes and is a poor predictor of these adverse outcomes. There is not enough research available to draw any conclusions about its use in high risk women. Midwives and obstetricians who had completed a standardised training program achieved good levels of agreements in assessing labour admission tests. Medical staff taking care of labouring women should be aware that there is a certain disagreement in interpreting labour admission tests, especially in tests assessed not normal. Interviews with 12 midwives found that interpreting CTG tracings could be difficult, especially for newly qualified midwives. The hierarchy of power in the labour ward influenced the use and interpretation of the labour admission test, and some of the midwives felt their professional identity threatened and that midwives in general are losing their traditional skills because of the increasing use of obstetric technology.

Recommendations for practice

Low-risk women should not routinely be offered the labour admission test. Midwives and obstetricians should be aware that there are inter-observer variations in the assessment of the labour admission test, especially in tests assessed as intermediary/abnormal. Pregnant women should be informed about why the test is not recommended anymore.

Recommendations for future research

Future research should emphasise the most appropriate method of fetal surveillance in high-risk labours and if training programs in assessing CTG are associated with improved outcomes of labour. Finally, the effect of obstetric technology on traditional midwifery skills, should be explored.

Key words

Cardiotocography; labour admission test; screening; inter-observer agreement; systematic review; midwifery practice.

Literature

- (1) Acheson D. Public health in England: the report of the committee of inquiry into the future development of the public health function. London: Great Britain Secretary of State for Social Services, 1988.
- (2) Whittle MJ, Martin WL. Fetal monitoring in labour. In: Chamberlain G, Steer P, editors. Turnbull's obstetrics. London: Churchill Livingstone, 2001.
- (3) Tucker SM. Fetal monitoring and assessment. St. Louis: Mosby, 2000.
- (4) Ingemarsson I, Arulkumaran S, Ingemarsson E, Tambyraja RL, Ratnam SS. Admission test: a screening test for fetal distress in labor. *Obstet Gynecol* 1986; 68:800-806.
- (5) Phelan JP. Labor admission test. *Clin Perinatol* 1994; 21:879-885.
- (6) Berglund S, Nordström L. National survey (Sweden) of routines for intrapartum fetal surveillance. *Acta Obstet Gynecol Scand* 2001; 80:552-553.

- (7) Devane D, Lalor JG, Daly S, McGuire W. Cardiotocography versus intermittent auscultation of fetal heart on admission for assessment of fetal wellbeing. The Cochrane Database of Systematic Reviews 2005; Issue 1, Art No.:CD005122.
- (8) Last JM. A dictionary of epidemiology. Oxford: Oxford University Press, 1995.
- (9) Royal College of Obstetricians and Gynaecologists. The use of electronic fetal monitoring. London: RCOG Press, 2001.
- (10) Nordström L, Waldenström U. Handläggning av normal födsel (Management of normal labour). Stockholm: Socialstyrelsen, 2001.
- (11) Umstad MP. The predictive value of abnormal fetal heart rate patterns in early labour. Aust N Z J Obstet Gynaecol 1993; 33:145-149.
- (12) Kulkarni AA, Shrotri AN. Admission test: a predictive test for fetal distress in high risk labour. J Obstet Gynaecol Res 1998; 24:255-259.
- (13) Phelan JP, Ahn MO. Perinatal observations in forty-eight neurologically impaired term infants. Am J Obstet Gynecol 1994; 171:424-431.
- (14) Phelan JP. The nonstress test: a review of 3,000 tests. Am J Obstet Gynecol 1981; 139:7-10.
- (15) Krebs HB, Petres RE, Dunn LJ, Smith PJ. Intrapartum fetal heart rate monitoring. VI. Prognostic significance of accelerations. Am J Obstet Gynecol 1982; 142:297-305.
- (16) Ducey J, Guzman E, Schulman H, Farmakedis G, Karmin I. Value of a screening fetal heart rate tracing in the latent phase of labor. J Reprod Med 1990; 35:899-900.
- (17) Sarno AP, Jr., Phelan JP, Ahn MO. Relationship of early intrapartum fetal heart rate patterns to subsequent patterns and fetal outcome. J Reprod Med 1990; 35:239-242.
- (18) Sundström A-K, Rosén D, Rosén KG. Fetal surveillance. Göteborg: Neoventa AB, 2000.
- (19) Strauss A, Corbin Juliet. Basics of qualitative research. Thousand Oaks: Sage Publications, 1998.

ERRATA

Studie I: I Table VI og VII er det regnefeil i konfidensintervallene for sensitivitet og PPV.

Riktige verdier står i Tabell 5, side 42 i avhandlingen.

Studie II: I Table 3 og 4 er det regnefeil i konfidensintervallene for sensitivitet og PPV.

Riktige verdier står i Tabell 6, side 43.

Avhandlingen vurderer nytteverdien av rutinemessig kardiotokografi (CTG) ved innkomst i fødeavdeling. Testen er i utbredt bruk i den vestlige verden, og er rutine ved de fleste fødeinstitusjoner i Norge.

Det ble undersøkt i hvilken grad testen kan forutsi uønska fødselsutfall, og om den kan forebygge disse utfallene. Det ble også undersøkt om forskjellige observatører var enige om tolking av innkomst-CTG, og hvilken betydning testen kan ha for jordmødre i deres daglige arbeid.

Journalene til alle kvinner som fødte ved Hammerfest sykehus i årene 1996-1998 ble gjennomgått. Det ble registrert hvordan innkomst-CTG ble tolket, inngrep i fødselsforløpet og utfall for mor og barn. Åtte særlig erfarne jordmødre og leger tolket CTG-utskrifter uavhengig av hverandre, og det ble gjort en systematisk oversikt over forskningsresultater om innkomst-CTG. Det ble gjort kvalitative intervjuer med 12 jordmødre.

Avhandlingen konkluderer at det ikke finnes vitenskapelig kunnskap som tyder på at innkomst-CTG er nyttig for lavrisiko fødende. Testen forebygger ikke uønska utfall og kan heller ikke forutsi disse utfallene. Det finnes for lite forskning til at man kan konkludere om testen er nyttig hos risikofødende. Jordmødre og leger må være klar over at det vil være en viss uenighet i tolking av innkomst-CTG, spesielt når det gjelder unormale tester. Intervjuene med de 12 jordmødrene fant at makthierarkiet i fødeavdelingen påvirket bruk og tolking av innkomst-CTG, og at informantene erfarte at jordmødre generelt ble dårligere i tradisjonelle kliniske ferdigheter på grunn av økende bruk av obstetrisk teknologi.



Nordisk ministerråd