

# Eesti Meditsiinilise Sünniregistri andmete kvaliteet

**Sigrid Vorobjov<sup>1,2</sup>, Mati Rahu<sup>1,2</sup>,  
Kaja Rahu<sup>1,2</sup>, Helle Karro<sup>3</sup> – <sup>1</sup>Tervise  
Arengu Instituudi epidemioloogia ja  
biostatistika osakond, <sup>2</sup>Eesti Käitumis- ja  
Terviseteaduste Keskus, <sup>3</sup>TÜ naistekliinik**

**Võtmesõnad:** meditsiiniline sünniregister, andmed, tunnused, andmekvaliteet, kooskõlamäär, kooskõlakordaja

**Uuringu eesmärk oli hinnata Eesti Meditsiinilises Sünniregistris registreeritud andmete kvaliteeti. 1997. a sündidest (12 593 sündi) moodustati valim (666 sündi), mille kohta täideti haigla originaaldokumentide põhjal uued sünnikaardid (641 kaarti). Uuesti täidetud sünnikaartide ja registri andmete vastavuse võrdlemiseks kasutati kooskõlamäära ja kooskõlakordajat (95% usaldusvahemikuga). Sotsiaaldemograafilised andmed, rasedust iseloomustavad tunnused, sünnituseelsete nõuandlakülastuste arv, viimase menstruatsiooni kuupäev, lapse sugu, sünnikaal, sünnipikkus ning haiglasse sisse- ja sealt väljakirjutamise kuupäevad olid usaldusväärsed. Esimest nõuandlakülastust, ema rasedusaegset suitsetamist ning raseduse ja sünnituse kestust käsitlevad andmed olid kehvema kvaliteediga. Kui raseduse ja sünnituse kestuse võrdlemisel kasutati leebemaid tingimusi, siis andmete kooskõlamäär paranes. Suuri puudujääke esines raseduse ja sünnitusega seotud protseduuride ja diagnooside registreerimisel.**

**Uuring näitas, et paljud tunnused olid hea kvaliteediga, kuid kehvema kvaliteediga tunnuste kasutamisel teadustöodes tuleb olla ettevaatlik.**

Põhjamaade pikaajaline kogemus näitab (1–4), et tervise- ja sotsiaalregistrite andmeid kasutatakse edukalt teadustöös. Registreeritud andmestiku alusel tehakse rahvastikupõhiseid uuringuid odavalt ja kiiresti. Saab jälgida harva esinevaid sündmusi või käsitleda selliseid probleeme, mida teistsuguste meetoditega ei suudetaks mingil põhjusel teha.

Meditsiiniliste sünniregistrite andmete põhjal saab eelkõige uurida raseduse kulgu, sünnitusabi ja sündimust. Nende uuringute eeltingimuseks on andmete hea kvaliteet, määravat osa etendab andmekogumise viis ja hõlmatus. Registreeritud andmed peavad olema valiidsed, s.t vastama tegelikku- sele, muidu osutuvad uuringutulemused kallutatuks. Näiteks mõjutavad puudulikud andmed surnultsündide kohta riigistatistika õigsust (5). Samuti on andmete võrreldavuse eesmärgil vajalik tagada kodeerimise ühtsus.

Meditsiiniliste sünniregistrite andmekvaliteeti on põhjalikult uuritud Rootsis (6, 7), Soomes (8–11), Taanis (12) ja Norras (13, 14) ning sünnitunnistuste andmekvaliteeti Ameerika Ühendriikides (15–20). Uuringutest selgub, et mittemeditsiinilised andmed on enamasti hea kvaliteediga, ent ülejäänud tunnuste kvaliteet kõigub suurel määral.

Eesti Meditsiiniline Sünniregister (EMSR) loodi 1991. aastal, et mõõta sündimust Eesti Vabariigis. EMSRi andmebaas sisaldab ühtse meetodikaga kogutud rah-

vastakupõhiseid andmeid ja on seetõttu väärtuslik eelkõige perinataalepidemioloogia (21, 22) ja elukäigu-epidemioloogia (ingl *life course epidemiology*) (23, 24) jaoks. Registris on andmed ema, raseduse, sünnituse ja vastsündinu kohta. Isa kohta registri algaastail teavet ei kogutud. Käesoleva töö **eesmärk** oli analüüsida EMSRi andmete kvaliteeti.

## MATERJAL JA MEETODID

Kõigis Eesti sünnitusabi osutavates haiglates täidetakse iga sünni kohta sünnikaart, mis saadetakse registrisse (25). Puuduliku või vigase andmestikuga sünnikaardid saadetakse registrit tagasi haiglasse ülekontrollimiseks. 1997. a sündidest (12 593 sündi) moodustati nn mugavusvalim. Valimisse kaasati 10% Lääne-Tallinna Keskhaigla Pelgulinna sünnitusmajas ja Tartu Ülikooli Kliinikumi naistekliinikus toimunud sündidest ning 20% Rakvere ja Pärnu haiglas aset leidnud sündidest (n = 666 sündi). Vastsündinu isikukoodi, sünnitusloo numbri ja

sünnikuupäeva alusel otsiti ajavahemikul 03.–10.1999 haiglatest välja algdokumendid (sünnituslugu, antenataalne kaart ja vastsündinulugu), mille põhjal täideti sünnikaardid, tagasi saadi 641 kaarti. Kvaliteedi hindamiseks võrreldi uuesti täidetud sünnikaarte registris leiduva infoga.

Iga tunnuse kvaliteedi mõõtmiseks kasutati kooskõlamäära (ingl *agreement proportion*), s.t leiti kokkulangevate väärtustega kirjete osakaal kirjete koguarvust. Kooskõlamäär hinnati heaks ( $\geq 99,0\%$ ), rahuldavaks ( $90,0-98,9\%$ ) või halvaks ( $\leq 89,9\%$ ) (10). Kooskõla statistilist olulisust hinnati Coheni kapa-kordaja (ingl *Cohen's kappa coefficient*) ehk kooskõlakordaja (ingl *coefficient of agreement*) – koos 95% usaldusvahemikuga (UV) (26) – alusel. Kapa väärtused on esitatud kategooriatena hea (0,81–1,00), rahuldav (0,61–0,80) ja halb ( $\leq 0,60$ ) (14). Pidevate tunnuste kvaliteedi hindamiseks kasutati kahe andmebaasi vaheliste erinevuste keskmist ja standardhälvet (27). Analüüs tehti andmetöötluspaketiga Stata 9 (28).

**Tabel 1.** EMSRi ja sünnikaartide andmete kooskõla (v.a rasedus- ja sünnituspuhused protseduurid ning diagnoosid) aastal 1997

Tunnus	Sündide arv		Kooskõlamäär %	Kapa	95% usaldusvahemik
	EMSR	Sünnikaart			
Ema isikukood	619	622	98,6	0,986	0,980–0,997
Ema elukoht	641	638	93,5	0,929	0,912–0,953
Ema rahvus	641	635	99,7	0,990	0,973–1,000
Ema haridus	638	634	94,7	0,927	0,899–0,947
Perekonnaseis	641	641	99,8	0,997	0,989–1,000
Varasemate raseduste arv	641	641	96,3	0,952	0,931–0,966
Varasemate elussündide arv	641	641	98,8	0,980	0,953–0,990
Varasemate surnultsündide arv	641	641	99,4	0,775	0,560–0,989
Spontaanabortide arv	641	641	96,6	0,833	0,764–0,913
Indutseeritud abortide arv	641	641	96,6	0,939	0,916–0,959
Naistenõuandla külastuste arv	641	641	98,6	0,873	0,848–0,901
Esimene nõuandlaskülastuse kuupäev	626	604	71,4	0,713	0,696–0,721
Suitsetamine raseduse ajal	634	594	81,6	0,351	0,328–0,396
Viimase menstruatsiooni algus	633	625	92,8	0,927	0,905–0,943
Raseduse kestus (päevad)	641	635	70,9	-0,547*	-0,856–0,747
Sünnitamiskiis	641	624	98,6	0,929	0,894–0,971
Sünnituse kestus (tunnid, minutid)	625	608	52,3	0,160*	-3,732–3,411
Lapse isikukood	633	637	98,1	0,981	0,972–0,992
Sünniaeg (kellaeg)	641	639	98,6	0,986	0,980–0,991
Sugu	641	641	99,5	0,991	0,980–1,000
Sünnikaal (g)	641	641	98,1	0,349*	-33,627–34,325
Sünnipikkus (cm)	640	640	99,1	0,002*	-0,305–0,308
Haiglasse saabumise kuupäev	641	640	95,6	0,956	0,943–0,978
Haiglast väljakirjutamise kuupäev	640	640	97,7	0,976	0,962–0,986

\*Pidevale tunnusele arvatati kahe andmestiku vaheline keskmine erinevus.

**TULEMUSED**

Üldjuhul saadi sünnikaartide täitmiseks vajalik info algdokumentidest. Kõige puudulikumate andmetega tunnused olid “esimese arstliku läbivaatuse kuupäev” (37 sündi), “sünnituse kestus” (33 sündi), “sünnitusviis” (17 sündi) ning “viimane menstruatsioon” (16 sündi). Seitsme tunnuse puhul esines 1–7 juhtu, kus algandmed puudusid. Valimi esinduslikkuse hindamiseks võrreldi selle andmete jaotust kõigi 1997. a sündidega. Võrdlusest ilmselt, et ükski uuritud tunnuste jaotustest (raseduse kestus, sünnituse kestus, lapse sugu ja sünnikaal) ei näidanud statistiliselt olulist erinevust.

Kooskõla kahe andmebaasi vahel oli eriti hea sotsiaaldemograafiliste tunnuste, nt ema perekonnaseis, rahvus ja isikukood osas (vt tabel 1).

Rasedust hõlmavad tunnused (varasemad elussünnid, raseduse katkemine, esilekutsutud abordid, rasedusaegsete nõuandlakülastuste arv ja viimase menstruatsiooni kuupäev) jäid peaaegu heale tasemele. Probleeme esines varasemate surnultsündide, ema rasedusaegse suitsetamise, esimese nõuandlakülastuse kuupäeva ning raseduse ja sünnituse kestuse registreerimisel. Kõrge kooskõlamäär varasemate surnultsündide registreerimisel näitab suurt kokkulangemist neil, kellel ei olnud surnultsünde; väike kapa väärtus koos laia usaldusvahemikuga osutab sellele, et esitatud vähesed juhud (9 juhtu) olid halvas kooskõlas. Registris esines puudulikke andmeid emade rasedusaegse suitsetamise kohta: uuesti täidetud sünnikaartide põhjal oli suitsetavaid naisi (82 naist) tunduvalt rohkem kui registris (47 naist). Samas puudus sünnikaartidel andmeid (7%; 47 juhtu) sagedamini kui registris (1%; 7 juhtu).

Raseduse kestus oli sünnikaartidel märkimata jäänud kuuel korral, EMSRis oli see aga alati näidatud. Kui raseduse kestust oli mõõdetud päevades, oli kooskõlamäär vaid 71%; täisnädalate korral suurenes kooskõlamäär 90%-ni (kapa = 0,868;

95% UV 0,842–0,894). Nii enneaegsete (< 37 nädalat) kui ka ülekantud (> 42 nädalat) laste puhul ilmselt hea vastavus (kooskõlamäär 99–100%).

Sünnitust hõlmavate tunnuste kooskõla oli hea, välja arvatud sünnituse kestus. Kuigi keskmine erinevus väga suurt kõikumist ei näidanud, ulatus lahknevus siiski mõnel juhul üle 10 tunni. Kui kooskõlamäära arvutamisel lubati kõikumist vahemikus ± 30 minutit, suurenes kooskõlamäär 52%-lt 91%-ni. Kõige probleemsemaks osutus keisrilõikega sünnituste kestuse registreerimine, mis puudus 49%-l juhtudest sünnikaardil ja 29%-l registris. Keisrilõikega sünnituste analüüsisist väljajätmine ei parandanud kahe andmekogu vahelist kooskõla (52% vs 53%).

Lapse sünnikaalu ja -pikkuse andmed näitasid head kooskõla, vastavalt 98% ja 99%. Kahe andmestiku keskmine sünnikaal erines 34 grammi. Erinevused vastasündinu sünnipikkuses olid väga väikesed, seda iseloomustab suur kooskõlamäär (vt tabel 1).

Isikukood on asendamatu ühendusliku andmebaaside linkimisel. Uuringus ilmselt, et 3%-l (22 naist) esines segadusi isikukoodiga. Neist 19-l puudusid koodi viimased neli numbrit. Laste isikukoodid olid valed 1,3%-l (8 juhtu), sealhulgas kahel korral oli vale sugu ja sajandit tähistav number.

Suuri puudujääke esines rasedus- ja sünnituspuhuste protseduuride ning diagnooside registreerimisel. Juhud, mis olid registreeritud EMSRis, puudusid sünnikaartide andmestikust ja vastupidi. Üldiselt oli sünnikaartidel mainitud rohkem diagnoose kui EMSRis, vastavalt 129 ja 98. Kooskõlas diagnooside ja protseduuride arv jäi väga väikeseks. Kooskõla osutus mitterahuldavaks isegi selliste tõsiste diagnooside puhul nagu ähvardav raseduse katkemine või enneaegse sünnituse oht (vt tabel 2). Nende tunnuste suhteliselt suur kooskõlamäär viitab rahuldavale kooskõlale diagnoosimata juhtudel, kuid väike kapa näitab vähest kokkulangemist diagnoositud juhtude vahel.

**Tabel 2.** EMSRi ja sünnikaartide andmete kooskõla viie kõige sagedamini registreeritud rasedus- ja sünnituspuhuse protseduuri ja diagnoosi korral aastal 1997

Tunnus	Sündide arv		Kooskõla-määr %	Kapa	95% usaldus-vahemik
	EMSR	Sünnikaart			
<b>Rasedusega kaasnevad riskitegurid:</b>					
varasem keisrilõige	13	16	98,9	0,753	0,576–0,930
emakakaela õmblus	5	5	99,4	0,597	0,236–0,958
lootevee uuring enne 25. nädalat	5	3	99,1	0,246	–0,152–0,644
ultrahelidiagnostika enne 21. nädalat	462	470	78,8	0,466	0,389–0,542
aneemia	138	124	84,1	0,511	0,429–0,594
<b>Haiglaravi vajanud komplikatsioonid:</b>					
kõrgenenud vererõhk	26	33	96,4	0,592	0,440–0,743
ähvardav enneaegne sünnitus	39	44	94,9	0,575	0,444–0,706
neeruhaigus	31	35	97,5	0,744	0,624–0,865
südamehaigus	4	2	99,1	–0,004	–0,010–0,002
ähvardav raseduse katkemine	38	48	95,0	0,602	0,476–0,727
<b>Rasedus- ja sünnituspuhused diagnoosid:</b>					
platsenta enneaegne irdumine	4	6	99,7	0,798	0,525–1,000
asfüksia	22	34	96,9	0,627	0,477–0,778
tuharseis	15	13	99,1	0,781	0,610–0,951
gestoos	40	55	94,5	0,603	0,483–0,722
<b>Toimingud sünnitusel:</b>					
indutseerimine	34	60	91,9	0,407	0,277–0,536
lootepöie avamine	241	186	86,7	0,704	0,647–0,761
stimulatsioon (medikamentoosne)	136	152	90,3	0,723	0,658–0,787
episio-, perineotoomia	138	138	91,9	0,760	0,698–0,822
platsenta manuaalne eemaldamine	16	13	98,3	0,612	0,401–0,823
emakaõõne abrasioon	5	6	99,5	0,725	0,426–1,000
loote elektrooniline jälgimine	409	489	75,7	0,431	0,358–0,503
<b>Lapse diagnoosid:</b>					
sepsis	1	2	99,5	–0,002	–0,006–0,002
RH-immunisatsioon	2	2	100,0		
respiatoorne distress-sündroom	5	19	96,9	0,156	–0,044–0,357
väärareng	15	7	98,1	0,446	0,186–0,706
sünnitrauma	15	19	98,8	0,758	0,597–0,920
<b>Lapsega seotud toimingud:</b>					
respiraatori kasutamine	1	14	98,0	0,131	–0,100–0,362
ravi jälgimisosakonnas	17	23	95,3	0,226	0,046–0,407
ravi intensiivosakonnas	48	55	95,8	0,715	0,613–0,817
antibiootikumravi	9	9	98,4	0,437	0,148–0,725
ultraviolettkiiritusravi	3	1	99,4	–0,002	–0,007–0,002

Sünnituspuhustest diagnoosidest esines suurim kooskõlamäär (99%) tuharseisu korral, kuigi kapa väärtus jäi suhteliselt väikseks (kapa = 0,781; 95% UV 0,610–0,951). Vastsündinule esimesel elunädalal pandud diagnooside puhul näitas enamik kooskõlakordaja väärtustest madalat kooskõla, kõige paremini registreeriti sünnitraumat (vt tabel 2).

Samalaadi olukord esines vastsündinule tehtud protseduuride korral, kus vastavuse tase osutus halvaks (kapa < 0,60). Mõnevõrra parem kokkulangevus ilmnis tunnuses „ravi intensiivosakonnas” (kooskõlamäär 96%; kapa = 0,715; 95% UV 0,613–0,817). Samas tunnused „ravi vastsündinute osakonnas” ja „ravi mujal haiglas” näitasid

märkimisväärset erinevust andmekogude vahel. Isegi selliste rutiinsete ja sagedaste protseduuride puhul nagu loote elektrooniline jälgimine ja ultraheliuuring enne 21. rasedusnädalat osutus andmekvaliteet mittehuldavaks (vt tabel 2). Mõnevõrra parem andmekvaliteet iseloomustas episiootomiat ja lootepöie avamist (kapa > 0,70).

#### ARUTELU

Käesolev uuring tehti juba aastal 1999, ent tulemuste analüüs ja üllitamine jäi pikaks ajaks soiku muudele töödele eelisõiguse andmise tõttu. Arvestades aga, et Eesti epidemioloogiale/rahvatervishoiule huvi pakuvate registreeritud kvaliteeti on ülimalt vähe

uuritud (29, 30), annab iga selleteemaline kirjutus vajalikku teavet. Registriandmete kasutaja peab teadma nende kvaliteeti, millest suuresti sõltub epidemioloogilise uuringu tegemise võimalikkus ja tase. Meie tööst selgus, et paljude tunnuste kvaliteet EMSRis oli küll hea, ent mitme tunnuse puhul osutus siiski mitterahuldavaks.

Rasedusaegsete nõuandlakülastuste koguarv oli heas vastavuses, ent esimese visiidi kuupäev lahknes tihti. Osaliselt võis see olla mõjutatud sellest, millist visiiti esimeseks peeti. Kui analüüsimisel kasutati ainult esimese visiidi kuud ja aastat, suurenes kooskõlamäär 71%-lt 81%-ni. Sama kooskõlamäär Soome Meditsiinilises Sünniregistris ulatus veidi üle 95%, kui analüüsis lubati kõikumist vahemikus  $\pm 30$  päeva (10).

Enam rasedusaegset suitsetamist puudutavate andmete kehv kvaliteet on üldisem probleem. Sellele osutavad ühtviisi nii sünnitunnistuste andmete (20, 31) kui ka Rootsi Meditsiinilise Sünniregistri uuring (7). Mõnevõrra parem kvaliteet tuli ilmsiks Soome uuringutes (9, 10). Käesolevas töös tuvastasime, et sünnikaartidel oli suitsetajatena kirjas rohkem emasid kui EMSRis. Et registreisse saadab andmeid mitu inimest, tekib rohkem veavõimalusi. Mõningat kallutatust võis põhjustada asjaolu, et uuringu jaoks täitsid sünnikaarte meedikud, kes tõenäoliselt pööravad enam tähelepanu last ootavate emade kahjulikele harjumustele (31). Eri haiglates on täitmine korraldatud erinevalt, enamasti on täitjaks meedik, kuid osa andmeid täidab sekretär. Samas on üldtuntud (32), et rasedad eelistavad oma suitsetamisharjumuste kohta esitada väärteavet.

Raseduse kestus on tugevalt seotud teiste rasedust ja sünnitust iseloomustavate tunnustega ning seetõttu võib halb andmekvaliteet siingi moonutada uurimistulemusi (9, 12, 17, 19). Meie töös võis raseduse kestust puudutavate andmete puudulikkusele avaldada mõju see, et uuringu korraldamise ajal arvutati haiglates raseduse kestust eri moodi. Leebemate tingimuste lubamine

(raseduse kestuse mõõtmine täisnädalates) suurendas kooskõlamäära 95%-ni. Rõõmustava tulemusena osutus andmekvaliteet väga heaks, kui analüüsi eraldi suure riskiga, enneaegseid ja ülekanatud lapsi.

Kui EMSRi andmestiku põhjal teha raseduse kestust käsitlevaid uuringuid, oleks võimalikuks lahenduseks kasutada kestuse arvutamisel viimase menstruatsiooni kuupäeva või täisnädalaid. Kindlasti oleks vaja teha meie kvaliteediuuringule järgnevat uuringuid, et välja selgitada, kas nimetatud probleem on praegu aktuaalne või pigem seotud registri algusaastatega.

Rasedus- ja sünnituspuhuseid diagnoosiv ning protseduure käsitlevate andmete kvaliteet registris oli kehv. Ka varasemad uuringud (9, 12, 16–20) on osutanud sellelaadsete andmete nigelale kvaliteedile. Isegi taolisi sagedasi riskitegureid nagu aneemia või rasedusaegne hüpertensioon registreeritakse tihti puudulikult (17, 19). Tõenäoselt on tegemist ühtsete diagnoosikriteeriumide puudumise või nende mitmeti tõlgendamisega. Samuti võib juhtuda, et sünnikaarte täitvatele isikutele edastatakse lünklik info (10). Üldjuhul kantakse EMSRi teatisel protseduuri või diagnoosi juurde kuuluvasse ruutu „linnuk“. Kasulikuks võiks osutada variandi „andmed puuduvad“ või „ei esinenud“ lisamine. Harva esinevate tunnuste analüüsimisel tuleb teada, et nende väärtuste juhuhajuvus on suur.

Sünnikaardi täitmise juhend peab olema selge ja üheselt mõistetav. Kvaliteedi juhtimisel aitab suuresti sünnikaarti täitvate inimeste väljaõpetamine ja tagasiside andmine haiglatele, sh teavitamine, miks haiglast saadetud andmeid vajatakse, mida nendega tehakse, mida kasulikuks võiksid haiglad saada sünniregistrist (32). Abiks oleks eri haiglate andmete võrdlus, mis näitaks, kas teatud toimingute tegemine erineb tõesti või on tegemist teistmoodi registreerimisega ja teiste kriteeriumide kasutamisega. Samuti aitab andmeid korras hoida pidev kvaliteedikontroll (33), mis hõlmab ka andmesisesust. Hoolikamat jälgimist vajab kuupäeva-

de järgnevuse testimine, seda eriti puuduva päeva ja kuu puhul. Plaanilise keisrilõike korral ei tohiks arvutiprogramm lubada registreerida sünnituse indutseerimist, mis üldjuhul on ebatõenäone ja vajab täiendavat kontrollimist. Meie uuringus oli näha, et ebatõenäosed juhud on võimalikud.

Andmete sagedas kasutamine teadustöös aitab andmekvaliteeti parandada ning hea kvaliteet muudab omakorda andmed uurijatele ahvatlevaks ja väärtuslikuks (1). Käesolev töö on võrdlusalus teiste samalaadsete kvaliteediuuringute jaoks. Analüüs näitas,

et paljude tunnuste kvaliteet oli hea ja et kehvema kvaliteediga tunnuste kasutamisel teadusuuringutes tuleb olla ettevaatlik ning puudustega arvestada.

## TÄNU

Autorid on tänulikud Eesti Meditsiinilise Sünniregistri ja koostööd teinud haiglate töötajatele. Uuringut toetasid Eesti Teadusfond (grandid 2978 ja 3899) ning Haridus- ja Teadusministeerium (sihtfinantseerimine 01921112s02 ja SF0940026s07).

*sigrid.vorobjov@tai.ee*

## KIRJANDUS

- Gissler M, Haukka J. Soome tervise- ja sotsiaalregistrid ning epidemioloogilised uuringud. *Eesti Arst* 2004;83:543–52.
- Gravseth HM, Bjerkedal T, Irgens LM, et al. Life course determinants for early disability pension: a follow-up of Norwegian men and women born 1967–1976. *Eur J Epidemiol* 2007;22:533–43.
- Morken NH, Källen K, Jacobsson B. Outcomes of preterm children according to type of delivery onset: a nationwide population-based study. *Paediatr Perinat Epidemiol* 2007;21:458–64.
- Sun Y, Vestergaard M, Christensen J, et al. Prenatal exposure to maternal infections and epilepsy in childhood: a population-based cohort study. *Pediatrics* 2008;121:1100–7.
- Lydon-Rochelle MT, Cardenas V, Nelson JL, et al. Validity of maternal and perinatal risk factors reported on fetal death certificates. *Am J Public Health* 2005;95:1948–51.
- Cnattingius S, Ericson A, Gunnarskog J, et al. A quality study of a medical birth registry. *Scand J Soc Med* 1990;18:143–8.
- The Swedish Medical Birth Register – a summary of content and quality. Stockholm: National Board of Health and Welfare, Centre for Epidemiology; 2003. <http://www.sos.se/fulltext/112/2003-112-3/2003-112-3.pdf>
- Hemminki E, Teperi J, Tuominen K. Need for and influence of feedback from the Finnish Birth Register to data providers. *Int J Qual Health Care* 1992;4:133–9.
- Teperi J. Multi method approach to the assessment of data quality in the Finnish Medical Birth Registry. *J Epidemiol Community Health* 1993;47:242–7.
- Gissler M, Teperi J, Hemminki E, et al. Data quality after restructuring a national medical registry. *Scand J Soc Med* 1995;23:75–80.
- Gissler M, Shelley J. Quality of data on subsequent events in a routine Medical Birth Register. *Med Inform Internet Med* 2002;27:33–8.
- Kristensen J, Langhoff-Roos J, Skovgaard LT, et al. Validation of the Danish Birth Registration. *J Clin Epidemiol* 1996;49:893–7.
- Skomsvoll JF, Østensen M, Baste V, et al. Validity of a rheumatic disease diagnosis in the Medical Birth Registry of Norway. *Acta Obstet Gynecol Scand* 2002;81:831–4.
- Rasmussen S, Albrechtsen S, Irgens LM, et al. Unexplained antepartum fetal death in Norway, 1985–97: diagnostic validation and some epidemiologic aspects. *Acta Obstet Gynecol Scand* 2003;82:109–15.
- Northam S, Knapp TR. The reliability and validity of birth certificates. *J Obstet Gynecol Neonatal Nurs* 2006;35:3–12.
- Buescher PA, Taylor KP, Davis MH, et al. The quality of the new birth certificate data: a validation study in North Carolina. *Am J Public Health* 1993;83:1163–5.
- Piper JM, Mitchel EF Jr, Snowden M, et al. Validation of 1989 Tennessee birth certificates using maternal and newborn hospital records. *Am J Epidemiol* 1993;137:758–68.
- Dobie SA, Baldwin LM, Rosenblatt RA, et al. How well do birth certificates describe the pregnancies they report? The Washington State experience with low-risk pregnancies. *Matern Child Health J* 1998;2:145–54.
- Reichman NE, Hade EM. Validation of birth certificate data. A study of women in New Jersey's HealthStart program. *Ann Epidemiol* 2001;11:186–93.
- Zollinger TW, Przybylski M, Gamache RE. Reliability of Indiana birth certificate data compared to medical records. *Ann Epidemiol* 2006;16:1–10.
- Nguyen RH, Wilcox AJ. Terms in reproductive and perinatal epidemiology: 2. Perinatal terms. *J Epidemiol Community Health* 2005;59:1019–21.
- Paneth M. Perinatal epidemiology. New York: Oxford University Press; 2008.
- Kuh D, Ben-Shlomo Y, Lynch J, et al. Life course epidemiology. *J Epidemiol Community Health* 2003;57:778–83.
- Kuh D, Shlomo YB, eds. A life course approach to chronic diseases epidemiology. 2nd ed. New York: Oxford University Press; 2004.
- Karro H, Rahu M, Gornoi K, et al. Estonian Medical Birth Registry 1992–1994: association of risk factors with perinatal mortality. *Eur J Obstet Gynecol Reprod Biol* 1998;80:151–6.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20:37–46.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- StataCorp. 2005. Stata Statistical Software: Release 9. College Station, TX: Stata Corp LP.
- Lang K, Mägi M, Aareleid T. Study of completeness of registration at the Estonian cancer registry. *Eur J Cancer Prev* 2003;12:153–6.
- Rahu M, Rahu K, Baburin A. Eesti surmaandmestiku kvaliteedianalüüs: Statistikaameti surmaregister. Lõpparuanne (30. november 2005). Tallinn: Eesti Demograafia Instituut; 2005.
- Dietz PM, Adams MM, Kendrick JS, et al. PRAMS Working Group. Completeness of ascertainment of prenatal smoking using birth certificates and confidential questionnaires: variations by maternal attributes and infant birth weight. *Am J Epidemiol* 1998;148:1048–54.
- Hemminki E, Teperi J, Tuominen K. Need for and influence of feedback from the Finnish birth register to data providers. *Qual Assur Health Care* 1992;4:133–9.
- Croen LA, Shaw GM, Jensvold NG, et al. Birth defects monitoring in California: a resource for epidemiological research. *Paediatr Perinat Epidemiol* 1991;5:423–7.

**SUMMARY**

---

**Data quality in the Estonian Medical Birth Registry**

---

The **objective** of the study was to estimate the quality of the data recorded in the Estonian Medical Birth Registry.

**METHOD.** A sample was drawn from all births (12,593 births) in 1997. Copies of the original hospital records were requested (666 births) and received for 641 births. The records were re-abstracted by trained gynecologists at clinics. The level of agreement between the hospital records and the records of the Estonian Medical Birth Registry was measured using agreement proportion and Cohen's kappa (with its 95% confidence interval).

**RESULTS.** Most of the sociodemographic data, variables in the history of pregnancy,

number of antenatal visits, date of the last menstrual period, child's gender, birth weight, birth length, and dates of hospitalization and discharge were reliable. The date of the first antenatal visit, maternal smoking, gestational age (in days) and duration of delivery were at an unsatisfactory level. Concerning gestational age, the agreement reached a satisfactory level when full weeks were compared. Data quality in the pregnancy- or delivery-related diagnoses and procedures was the poorest.

**CONCLUSION.** This study showed that the registry data were accurate for many variables, however, poorly recorded variables should be used with caution in research.