

Teabevihik

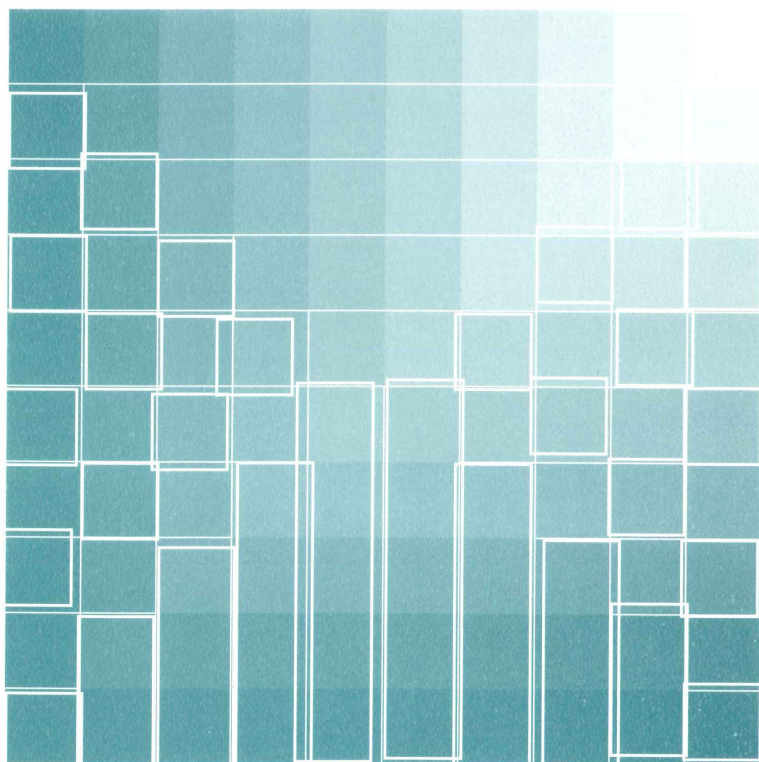
Tallinn 2001



12

EESTI STATISTIKASELTS

RAHVASTIKUSTATISTIKA JA REGISTRID



12

Eesti Statistikeseltsi teabevihik

RAHVASTIKUSTATISTIKA ja REGISTRID

Tallinn 2001

Koostanud Ebu Tamm

Kaane kujundanud Michael Walsh

ISSN 1406-314X

ISBN 9985-9168-3-2

© Eesti Statistika Selts, 2001

© Statistikaamet, 2001

SISUKORD

EESTI STATISTIKASELTSI 13. KONVERENTS	5
Ebu Tamm. Rahvaloenduse andmete kvaliteedi hindamisest	8
Anne Herm. Rahvastiku pidevstatistika olukord ja hetkeväljavaated	15
Mare Ainsaar. Rahvastikuandmed Eestis — kellele ja kui vabalt	20
Peeter Kääriks, Ragnar Õun. Eesti rahvastikuregistri olevik ja tulevik	29
Tarmo Sumeri. Rahvastikuregistri seis ja staatus	36
Alvi Tellmann. Eesti meditsiiniline sünniregister ja selle staatus	42
Tiiu Vallner. Rahvaloenduse andmekaitse	47
Ene-Margit Tiit. Eesti rahvaarvust ja selle hindamisest	53
Anneli Kukk, Kaja Sõstra, Imbi Traat. Rahvastikuregistri kasutamisest Eesti leibkonna- ja tööjõu-uuringus	63
Kaja Rahu. Isikuandmete kvaliteedist meditsiiniregistrates ehk linkimisuuringute võimalikkusest Eestis	75
Mihkel Servinski, Kaja Sõstra. Rahvastikuandmed Eesti regionaalarengu andmebaasis	79
Mati Rahu. Epidemioloogia ja registripõhised uuringud	85
Ene-Margit Tiit, Ene Käärik, Meelis Käärik. Eesti naiste reproduktiiv- käitumine aastail 1994–1998 (sünni- ja abordiregistri andmetel)	90

KROONIKA	101
Doktoritöö biostatistikast	101
Doktroritöö tõenäosusteooriast	102
Kaitstud bakalaureusetööd 1999–2000	104
Kaitstud magistritööd 1999–2000	106
Kalev Pärna. 6. Tartu mitmemõõtmelise statistika konverents	107
Eest Statistikaseltsi üldkoosolek	109

Eesti Statistikeseltsi 13. konverents

8. ja 9. novembril 2000 Tallinnas Rahvusraamatukogu konverentsisaalis

RAHVASTIKUSTATISTIKA JA REGISTRID

Kava

I sektsioon. Eesti rahvastikustatistika hetkeseis ja rahvaloenduse tulemused

8. nov 2000 10.00–12.30

Juhataja Ü. Randaru (Statistikaamet)

1. Avasõna. E.-M. Tiit (TÜ) 10.00–10.10
2. 2000. a rahva- ja eluruumide loendus ja selle esialgsed tulemused. M. Sundja (Statistikaamet) 10.10–10.40
3. Rahvaloenduse andmete kvaliteedi hindamisest. E. Tamm (Statistikaamet) 10.40–11.10
4. Rahvastiku pidevstatistika olukord ja hetkeväljavaated. A. Herm (Statistikaamet) 11.10–11.30
5. Arvestuslik rahvaarv ja rahvaloendus. A. Maamägi (Statistikaamet) 11.30–11.50
6. Rahvaloendustest Eestis. L. Tepp (Statistikaamet) 11.50–12.10
Läbirääkimised 12.10–12.30

Lõuna 12.30–13.30

II sektsioon. Rahvastikuandmeid sisaldavate registrite seis ja staatus Eestis

8. nov 2000 13.30–16.00

Juhataja E.-M. Tiit (TÜ)

7. Registriandmete kasutamine riiklike statistiliste vaatluste korraldamisel. Ü. Randaru (Statistikaamet) 13.30–14.00
8. Rahvastikuandmed Eestis — kellele ja kui vabalt. M. Ainsaar (Rahvastikuminstri büroo) 14.00–14.30
9. Rahvastikuregistrist. P. Kүүts, R. Öun (Siseministeerium) 14.30–15.00

10. AS Andmevara, Eesti rahvastikuregistri volitatud töötaja.
T. Sumeri (AS Andmevara) 15.00–15.20
11. Haigekassade ravikindlustatute andmebaas.
M. Thetloff (Keskhaigekassa) 15.20–15.50
12. Eesti meditsiiniline sünniregister ja selle staatus.
A. Tellmann (EKMI) 15.50–16.10
13. Riikliku pensionikindlustuse registri kasutuselevõtt.
I. Kressa (Sotsiaalkindlustusamet) 16.10–16.30

Kohvipaus 16.30–17.00

Eesti Statistikeseltsi üldkoosolek 17.00–18.30

1. Juhatuse tegevuse aruanne
2. Juhatuse ja revisjonikomisjoni valimised

Õhtusöök 19.00

III sektsioon. Andmed ja andmekaitse tänapäeva Eestis

9. november 2000 9.30–11.00

Juhataja E. Tamm (Statistikaamet)

14. Andmete kaitsest. V. Praust (Andmekaitse Inspeksioon) 9.30–10.00
15. Riigiasutuste andmeturbe olukorrast.
U. Puus (Küberneetika AS) 10.00–10.20
16. Rahvaloenduse andmekaitse.
T. Vallner (Statistikaamet) 10.20–10.40
17. Rahvastikunäitajate hinnangud rahvaloenduse andmete põhjal.
E.-M. Tiit (TÜ) 10.40–11.00

Kohvipaus 11.00–11.30

IV sektsioon. Rahvastikuandmete ja registrite kasutamise rakenduslikke küsimusi

9. november 2000 11.30–13.00

Juhataja V. Tamm (TÜ)

18. Rahvastiku andmebaasi kasutamisest Eesti tööjõu- ja leibkonna-
uringutel. I. Traat, K. Sõstra (TÜ) 11.30–12.00

19. Isikuandmete kvaliteedist meditsiiniregistrites.
K. Rahu (EKMI) 12.00–12.30
20. Rahvastikuandmed Eesti regionaalarengu andmebaasis.
M. Servinski, K. Sõstra (Statistikaamet) 12.30–13.00

Lõuna 13.00–14.00

**V sektsioon. Rahvastikuandmete ja registrite kasutamise
teoreetilisi ja meetoodilisi küsimusi**

9. november 2000 14.00–16.00

Juhataja K. Pärna (TÜ)

21. Epidemioloogia ja registripõhised uuringud.
M. Rahu (EKMI) 14.00–14.30
22. Eesti naiste reproduktiivkäitumise monitooring aastail 1994–1998 sünni- ja abordiregistri andmetel.
E.-M. Tiit, E. Käärik, M. Käärik (TÜ) 14.30–14.50
23. Eesti suremustrendist.
T. Kollo, N. Gluškova (TÜ) 14.50–15.10
24. Andmeaudit andmelaonduse ülesannete puhul. AS Resta kogemus.
M. Vähi (TÜ), T. Tambaum (Resta) 15.10–15.30
25. Siserände registreerimine Eestis.
T. Tammaru (TÜ) 15.30–15.50

Kohvipaus 16.00–16.30

**Ümarlaud. Registrid, riiklik statistika, andmekaitse ja tarbijad
— kas vastuolude hulknurk või ühishuvid?**

9. november 2000 16.30–18.00

Juhatajad T. Kollo ja K. Ääremaa (TÜ)

RAHVALOENDUSE ANDMETE KVALITEEDI HINDAMISEST

Ebu Tamm
Statistikaamet

On iseenesest mõistetav, et igasuguse lõpetatud töö tulemuste kvaliteeti on vaja ühel või teisel viisil hinnata. Sealhulgas ja eriti kehtib see rahva- ja eluruumide loenduse kui suure, tähtsa ja kalli projekti kohta. Olenevalt võimalikest võrdlusallikatest, ressurssidest ja seadustest on selleks kasutatud mitmesuguseid meetodeid. Põhilised loendustulemuste hindamise meetodid, mida kasutasid Lääne-Euroopa riigid 1990. ja 1991. aasta loenduste järel, on tabelis 1 [Begeot 1993].

Tabel 1

Riik	Võrdlus eelmise loenduse tulemustega	Võrdlus rahvastiku -registriga	Võrdlus tööhõive uuringuga	Valikuline järelküsitlus	Muud võrdlus- meetodid
Belgia		■			■
Kreeka				■	
Hispaania		■	■	■	
Prantsusmaa				■	
Iirimaa	■				■
Itaalia		■		■	
Luksemburg	■	■			
Portugal	■			■	■
Suurbritannia	■			■	■
Austria	■				■
Norra				■	
Rootsi			■	■	
Šveits	■		■		■

Nagu tabelist 1 näha, on iga riik kasutanud üht või mitut võrdlusmeetodit. Loenduse kvaliteeti ainult ühel viisil on hinnanud kolm riiki: Kreeka, Prantsusmaa ja Norra ning kõigil kolmel juhul on olnud selleks valikuline järelküsitus. Rahvaloenduse järelküsitus on üldjuhul töenäosuslik valikuuring, millel on kaks eesmärki: 1) hinnata loendatud ja samal ajal ka loendamata jäänud isikute arvu (kaetust) ja 2) hinnata loendusega kogutud andmete kvaliteeti. Järelküsitus peab olema loendusest sõltumatu. Sõltumatus tagatakse järgmiste tingimuste täitmisega:

- 1) kõik loenduslehed peavad olema laekunud enne järelküsitlust;
- 2) järelküsitluse välitöid ei tohi teha needsamad loendajad, kes loendasid põhiloendusel (loendajaid võib kasutada, kuid mitte neis piirkondades, kus nad loendasid);
- 3) järelküsitlejad ei tohi teada põhiloenduse tulemusi selle osa kohta, mida nad järelküsitlevad;
- 4) loendajad ei tohi teada neid piirkondi, mida hiljem järelküsitletakse;
- 5) järelküsitluse töötlus tuleb hoida lahus loenduse töötlusest.

Eesti 2000. aasta rahva ja eluruumide loenduse järelküsitluse välitööd tehti 14.–19. aprillil, st 5.–10. päeval pärast rahva ja eluruumide loenduse välitööde lõppu. Olemasolevad ressursid võimaldasid teist korda loendada ligikaudu ühe protsendi kogu elanikkonnast. Kui oleks olnud vaja hinnata ainult loendusega kogutud andmete kvaliteeti, siis oleks võinud võtta juhusliku valimi isikutest, näiteks igast loendusjaoskonnast kolm isikut, ja küsitleda neid teist korda. Selline valik aga ei oleks võimaldanud hinnata loendatud (ja vahelejäänute) määra ehk teiste sõnadega: saada hinnanguid kaetuse kohta. Kaetuse hindamiseks tuli juhuslik valim võtta mingist elanikke sisaldavate objektide ehk klastrite hulgast. Teatavasti nõuab töenäosuslik valikuuringute meetoodika, et valimi võtmisel oleks olemas üldkogumi nimekiri ehk valimi freim. Rahva- ja eluruumide loenduse valiku objekt oli loendusjaoskond ja valimi freim loendusjaoskondade nimekiri, millest oli eemaldatud institutsioone (vanadekodud, lastekodud jms) katvad loendusjaoskonnad. Selles nimekirjas oli 4917 loendusjaoskonda, neist 3069 linnas ja 1848 maal. Et linnades asuvad loendusjaoskonnad katsid igaüks keskmiselt 332 ja kokku 1 017 000 inimest ning maal asuvates jaoskondades olid need arvud vastavalt 242 ja 448 000, siis 1%-se isikuvalimi saamiseks nii linnas kui ka maal tuli linnajaoskondadest valida juhuslikult 31

loendusjaoskonda ja maalt 19. Enne valimi võtmist jagati loendusjaoskondade üldkogum kahte kihti (linnajaoskonnad ja maaajaoskonnad) ja seejärel võeti kummastki kihist sõltumatu lihtne juhuslik valim. Tabelis 2 on esitatud mõlema kihi üldkogumi maht, kihi valimi maht ja kihi laiendustegur.

Tabel 2

Kiht	Üldkogumi maht N	Valimi maht n	Laiendustegur N/n
Linnades asuvad loendusjaoskonnad	3 069	31	99
Maal asuvad loendusjaoskonnad	1 848	19	97,26
KOKKU	4 917	50	

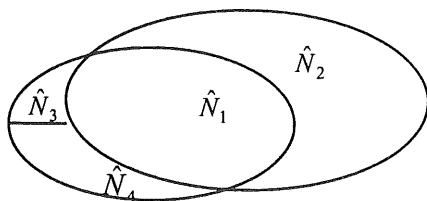
Valimisse sattunud loendusjaoskonnad küsitleti teist korda ja küsitlejateks olid loendajate juhendajad. Järelküsitluse küsimustikena kasutati rahva- ja eluruumide loenduse küsimustikke vähendatud mahus. Isikulehelt küsiti järelküsitluse käigus 16 (põhiloendusel 31) ja eluruumilehelt 6 (põhiloendusel 12) küsimust.

Kaetuse hindamine

Järelküsitlusele kuulunud 50 juhuslikult valitud loendusjaoskonnast koguti kaks andmestikku: kõigepealt põhiloenduse andmestik ja seejärel järelküsitluse andmestik. Rahva- ja eluruumide loenduse kaetust hinnatakse nende kahe andmestiku võrdlemisega, millest arvutatakse järgmised suurused:

- 1) nii loendatute kui ka järelküsitletute laiendatud arv \hat{N}_1 ,
- 2) järelküsitletute, kuid loendusel vahelejäanute laiendatud arv \hat{N}_2 ,
- 3) loendatute, kuid järelküsitlusel vahelejäanute laiendatud arv \hat{N}_3 ,

4) valesti loendatute laiendatud arv \hat{N}_4 .



Järelküsitletud 50 loendusjaoskonna mõlemaid andmestikke võib käsitleda kui valikuuringuga kogutud andmeid, mille põhjal võib teha üldkogumi ehk käesoleval juhul rahvastiku kohta mitmesuguseid järeldusi.

Järelküsitleluse andmestiku põhjal saadud rahvaarvu hinnang on

$$\hat{N}_{(L)} = \hat{N}_1 + \hat{N}_2, \quad (1)$$

millest loendusel vahelejäänute määr on

$$\hat{N}_2 / (\hat{N}_1 + \hat{N}_2) \quad (2)$$

ja kaetuse määr

$$\hat{N}_1 / (\hat{N}_1 + \hat{N}_2) . \quad (3)$$

Nendesamade 50 järelküsitletud loendusjaoskonna loendusandmete põhjal saame rahvaarvule teise hinnangu

$$\hat{N}_{(L)} = \hat{N}_1 + \hat{N}_3 . \quad (4)$$

Kui lõplikke loendustulemusi veel ei ole, saab rahvaarvule anda nn esialgse hinnangu eespool leitud hinnangute $\hat{N}_{(JL)}$ ja $\hat{N}_{(L)}$ kaudu:

$$\hat{N}_{\text{esialgne}} = \hat{N}_{(L)} \cdot \frac{\hat{N}_{(JL)}}{\hat{N}_1} = (\hat{N}_1 + \hat{N}_3) \frac{\hat{N}_1 + \hat{N}_2}{\hat{N}_1}, \quad (5)$$

kus

$$\frac{\hat{N}_1 + \hat{N}_2}{\hat{N}_1} = \frac{1}{\text{kaetuse määr}} = f$$

on parandustegur.

Netoviga arvutatakse valemiga

$$\text{netoviga} = \hat{N}_{\text{esialgne}} - \hat{N}_{(L)} \quad (\geq 0) \quad (6)$$

ja brutoviga valemiga

$$\text{brutoviga} = \hat{N}_2 + \hat{N}_4 \quad (\geq 0) . \quad (7)$$

Kui loenduse lõplik tulemus N_{loendus} on leitud, saab anda tegelikule rahvaarvule lõpliku hinnangu

$$\hat{N}_{\text{lõplik}} = f \cdot N_{\text{loendus}} . \quad (8)$$

Hinnanguid (1)–(8) on võimalik arvutada ka rahvastiku mitmesuguste alamkogumite kohta. Üks osa tabeleid võiks sisaldada teavet kaetuse määrast (vt tabel 3).

Tabel 3

	Loendatud ja järelküsitletud isikute laiendatud arv N_1	Järelküsitletute, kuid loendamata isikute laiendatud arv N_2	Kaetuse määr (%) $\frac{N_1}{N_1 + N_2} 100$
Linn			
mehed			
naised			
Maa			
mehed			
naised			
KOKKU			

Teine tabelite rühm võiks sisaldada mitmesuguseid hinnanguid rahvaarvu kohta (vt tabel 4).

Tabel 4

	Rahvaarvu hinnang jä- relküsitluse põhjal $\hat{N}_1 + \hat{N}_2$	Rahvaarvu hinnang loenduse põhjal $\hat{N}_1 + \hat{N}_3$	Rahvaarvu esialgne hinnang $\hat{N}_{esialgne} =$ $(\hat{N}_1 + \hat{N}_3)f$	Netoviga $\hat{N}_{estalgne} -$ $\hat{N}_1 + \hat{N}_3$	Brutoviga $\hat{N}_2 + \hat{N}_4$
Linn					
mehed					
naised					
Maa					
mehed					
naised					
KOKKU					

Tabelitega 3 ja 4 sarnaseid tabeleid võib koostada ka sünniaasta ja/või vanuserühma, kodakondsuse, perekonnaseisu jne järgi. Tuleb ainult meeles pidada, et kõik arvud nendes tabelites on valikuviga sisaldavad hinnangud ja seetõttu tuleb enne järelküsitluse tulemuste avaldamist arvutada kõik standardvead, et oleks võimalik otsustada, milliseid tulemusi avaldada ja milliseid mitte.

Kasutatud kirjandus

Begeot, F., Smith, L. 1993. First Results in Western European Censuses. Population Trends, No. 74, .

Design and Implementation of a Post-Enumeration Survey: Developing Country Example. 1993. International Statistical Programs Center. U.S. Bureau of the Census.

Evaluating Censuses of Population and Housing. 1985. U.S. Department of Commerce. Bureau of the Census.

Särndal, C.-E., Svensson, B., Wretman, J. 1992. Model Assisted Sample Surveys. Springer-Verlag. New York, Inc.

RAHVASTIKU PIDEVSTATISTIKA OLUKORD JA HETKEVÄLJAVAATED

Anne Herm
Statistikaamet

Rahvaloendusi korraldatakse harva ja suurem osa nendega kogutud andmetest iseloomustab rahvastikku loendusmomendil. Pidev ehk jooksev rahvastikustatistika jälgib muutusi rahvastikus. Pidev rahvastikustatistika kogub rahvastikusündmuste andmed kokku võimalikult kiiresti nende toimumise järel. Kogutud andmete alusel arvutatakse rahvaarvu ja rahvastiku koosseisu andmed iga aasta kohta ja tehakse rahvastiku perioodnäitajate arvutusi.

Pidevalt ja kõikselt kogutakse andmeid sündmuste registreerimise kaudu. Uuritavad sündmused on sünnid, surmad, abiellumised, abielulahutused ja elukohavahetused.

Euroopa Liidu Statistikaamet nõuab lisaks demograafiliste sündmuste andmetele andmeid ka teist laadi sündmuste kohta, nagu kodakondsusmuutused, asüüli- ja põgeniku staatuse taotlused, ning andmeid välismaalaste arvu ja koosseisu kohta. Selles valdkonnas on pidevstatistika väljaarendamisel. Ka perekonnastatistika pidevstatistikanähtena senini puudub: rahvastikusündmuste statistika ei võimalda selle kohta täielikku ülevaadet.

Suurem osa Eesti rahvastiku pidevstatistikast on niisiis sündmustekeskne: tegeldakse andmetega iga sündmuseliigi kohta eraldi. Vaatlusperiood on kalendriaasta, andmeallikas on rahvastikusündmuste järjepidev registreerimine kohalikes omavalitsustes ja perekonnaseisuasutustes.

Väga pikka aega kasutati statistiliste andmete saamiseks perekonnaseisuakti. Seoses perekonnaseisuaktide muutmisega võeti alates 1994. aastast sündide, surmade, abielude ja abielulahutuste andmete kogumiseks kasutusele statistilise arvestuse erivormid. Sellise otsusega kindlustati nende sündmuste andmete pidevus ka olukorras, kus registreerimisakti statistilisi andmeid enam ei sisalda.

Andmeid sündide ja surmade kohta kogutakse aga ka juba enne ametlikku registreerimist: nimelt sünnid raviaasutuses, kus laps sündis, surmad aga arsti poolt, kes surma tuvastas.

Sündide ja surmade registreerimisega suuri probleeme ei ole ja need allikad annavad rahvusvahelises plaanis üsna hästi võrreldava statistika. See ei tähenda muidugi, et puudusi ei olegi, nad ei ole aga praegu nii teravad kui mõnes teises valdkonnas.

Sünnistatistika sisaldab praegu sündide, mille kohta on perekonnaseisuasutustes, sealhulgas Eesti välisesindustes koostatud sünniakt. On aga ka veel teine andmeallikas: haiglates lapse sünni kohta täidetav sünnikaart. See on meditsiinilise sünniregistri alusdokument.

Sünnikaart sisaldab sünnistatistika jaoks olulisi andmeid, mida ei ole võimalik saada sünni registreerimise käigus. Peale selle võimaldab nende kahe andmekogumi võrdlus selgitada välja lahknevused, parandada vigu ja täiendada andmeid. Statistikaametil ja meditsiinilisel sünniregistril on selle koostöö kogemus.

Nagu eespool öeldud, tugineb ametlik sünnistatistika siiski registreerimisandmetele. Perekonnaseisuasutustest saadav andmestik sisaldab mõnevõrra rohkem sündide, kui haiglates lapsi sünnib, sest registreeritakse ka kodus ja välismaal sündinud Eesti elanike lapsed. Samas aga jääb osa Eesti haiglates sündinud lapsi perekonnaseisuasutuste registreerimissüsteemist välja või registreeritakse nad nii palju hiljem, et pole võimalik käimasoleva aasta statistikas kajastada. Sünnikaardi ja registreerimisandmete võrdlus näitab, et ligikaudu üks protsent Eesti haiglates sündinutest jäävad aasta jooksul õigeaegselt registreerimata ja seega jäävad registreerimisandmetele tuginevast sünnistatistikast välja.

Nende kahe andmeallika — haiglates täidetava sünnikaardi ja registreerimisel täidetava statistilise lehe — kooskasutamine võimaldaks saada täielikumat sünnistatistikat. Kui täiendada registreerimisandmeid haiglates kogutavate andmetega, saaks vältida registreerimise hilinemise või mitteregistreerimise tõttu tekkivat alakaetust.

Surmaandmed kogutakse arstliku surmatõendiga. Statistilise andmekogumise eesmärki täitev surmaleht kujunes tegelikult alusdokumendiks surmaakti täitmisele, mistõttu selle nimi ja tähendus muutusid. Arstliku surmatõendi edastavad surnu sugulased perekonnaseisuasutusse, kus surmaakti koostamisel seda täiendatakse.

Surmaakti koostamiseks ei oleks vaja nii põhjalikku teavet surmapõhjustest ja asjaoludest, kui seda on vaja surmapõhjuste statistikas. Piisaks surnu isiku- ja peamise surmapõhjuse andmetest.

Mitmetes riikides, eriti seal, kus sündmuste andmed saadakse rahvastiku-registri vahendusel, kogutakse surmaandmeid paralleelselt kahte teed pidi: sündmuse fakti kohta registreerimise süsteemi kaudu ja surmapõhjused otse raviastutustest. Võibolla oleks ka meil otstarbekas välja kujundada skeem, kus surmapõhjuste üksikasjad edastavad meedikud otse Statistikaametile. Surmapõhjusi täpsustatakse sageli surmaakti koostamisest palju hiljem ja aktile neid parandusi üldiselt ei kanta. Selline ümberkorraldus tooks statistikutele mõnevõrra suurema töökoormuse. Andmete edastamine otse arstilt Statistikaametile võimaldaks ära hoida andmemoonutusi. Näiteks on mõnikord ilmset tahtliku enesekahjustuse tagajärge esitatud õnnetusjuhtumina.

Abielude ja abielulahutuste andmed kogutakse samuti perekonnaseisuastutustes ja selleks on vastava perekonnaseisuaktiga samaaegselt täidetav spetsiaalne statistiline abielu- või abielulahutuse leht.

Varem jõustusid abielulahutused alles siis, kui nad olid registreeritud perekonnaseisuastutuses. 1995. aastast aga vormistatakse osa abielulahutusi juba lõplikult kohtus. Ka nende lahutuste kohta täidetakse kohtutes statistilised lehed.

Abielude ja abielulahutuste registreerimine katab ainult osa ametlikke perekonnaseisu muutusi — abiellu astumist ja selle lahutamist, kuid ei kajasta abielu lõppemist ühe abikaasa surma tõttu. Niisiis ei ole võimalik saada ülevaadet kogu rahvastikust perekonnaseisu järgi. 2000. aasta rahvaloendus annab andmeid rahvaloenduse aja kohta, kuid pidevaid andmeridu ei ole ikkagi.

Et saada andmeid leseksjäänute kohta, oleks vaja registrit, mis sisaldaks perekonnaliikmete seoseid. Teine võimalus oleks, et surmaandmete kogumisel lisatakse surnu perekonnaseisu andmetele andmed tema abikaasa kohta. Selline andmete kogumine ei olekski nii väga erandlik, sest vähemalt kolmes Euroopa riigis — Prantsusmaal, Luksemburgis ja Suurbritannias — on nii tehtud.

Rahvastiku tegeliku abielulisuse osas ei ole võimalik tugineda registreerimisandmetele. Mõnes riigis saadakse tegelike kooselude andmed rahvastiku-registri ja need määratakse kindlate valitud tunnuste kombinatsioonina. Võimalikuks andmeallikaks võivad olla ka pidevad küsitlusuuringud. Sobi-

vaks võiks kujuneda leibkondade sissetuleku ja kulutuste uuring, mille perepilt sisaldab liikmetevahelisi suhteid ja ei piira küsitletavate hulka vanusega. Küsitlusuuringute puhul aga jääb alati probleemiks see, et ei ole võimalik saada madalama haldusüksuse andmeid.

Rändeandmed kogutakse registreerimisteatega, mille täidab kohaliku omavalitsuse ametnik, kui isik teatab elukohast lahkumisest või saabumisest elukohta. Omavalitsuse ametnik registreerib nii riigi sise- kui ka välisrändesündmusi ja andmed antakse üle maakonna rahvastikuregistrile, kust need omakorda edastatakse Siseministeeriumi rahvastiku arvestuse andmebaasi. Statistikaamet saab andmed sellest andmebaasist AS Andmevara vahendusel.

Rändestatistikat peetakse vähem usaldusväärseks võrreldes teiste rahvastikusündmuste statistikaga. Nii on see ka teistes riikides, sest täpseid andmeid on raske koguda. 1990. aastatel on Eesti rändeandmete kvaliteet järjepidevalt halvenenud. Kvaliteedi oluline langus algas nõukogudeaegse rahvastikuarvestuse süsteemi lagunemisega 1990. aastate alguses, kui kaotati senine rahvastikuarvestuse alus — kohustuslik elukoha sissekirjutus passi. Et uut arvestussüsteemi ei olnud, jätkasid omavalitsused nende elanike elukohamuutuste registreerimist, kes selleks soovi avaldasid.

Andmekvaliteet langes kahes suunas: kõiki elukohavahetusi enam ei registreeritud ja need, mis registreeriti, ei pruukinud vastata tegelikele elukohamuutustele. Seda tõendavad nii statistiliste valikuuringute käigus kui ka viimasel rahvaloendusel selgunud lahknevused mingil aadressil elavate ja seal samas elavatenas registreeritud isikute puhul.

Rahvastikuregistri seadusega pandi alus uuele rahvastikuarvestuse süsteemile. 2001. aasta jaanuarist läheb käiku uus elukoha registreerimise kord. Paraku ei toeta rahvastikuregistri seaduse sätted statistikat. Vabatahtlik elukohast teatamine olukorras, kus isiku vajadus ja huvi tegeliku ja registreeritud elukoha kokkulangevuse järele puudub, ei garanteeri tõepärast elukohajärgset arvestust. Niisiis ei ole loota, et elukoha registreerimise andmed hakkaksid lähemal ajal peegeldama tegelikku rännet, moonutused tulevad nii ajalises kui ka ruumilises mõõtmes.

Vajalikus mahus rändeandmete kogumist jooksva aasta kohta on aga raske ette kujutada ilma kõiki elukohamuutusi registreerimata. Statistikas võiks ju kasutada mitut sellist andmeallikat, kuhu laekuvad isikute elukoha muutused aasta jooksul. Selleks on vaja aga andmeid ühendada isikuandmete tasandil, seega peab rakendama andmete riskasutust.

Pideva rahvastikustatistika hulka on seni kuulunud ka iga-aastane statistika rahvaarvu ja rahvastiku koosseisu kohta, mis on ka igati loogiline, sest tugineb ju rahvastikusündmuste andmetele.

1990. aastatel saadi 1. jaanuari rahvaarv ja rahvastiku koosseisu andmed nii, et 1989. aasta rahvaloenduse soo-vanuskoosseisu andmete tabelid korrigeeriti iga järgmise aasta sündmuste (sündide, surmade, sisse- ja väljarände) andmetega. Praegu ei oleks otstarbekas rahvaarvu arvutusi jätkata senisel viisil, s.o koondandmetena. On elektroonilised kirjed 2000. aasta rahvaloenduse andmetega, pidevalt kogutakse sündmuste andmeid ja need kõik salvestatakse üksikkirjetena. Otstarbekas oleks rahvastikumutuste andmed lisada baasandmetele juba üksikkirje tasandil ja nii lahendada andmete lahkumise probleemid kohe, vältides vigu, mis koondandmete tasandil võivad tekkida.

Statistiliste andmete valikut laiendada, lünki täita ja andmekvaliteeti tagada saab, kui andmeid saadakse paljudest allikatest ja et need andmed on võimalik üksühesesse vastavusse viia. Peab olema võimalik minna üksikkirjeni, ilma milleta ei saa olla kindel, et andmed esindavad just seda nähtust, mida neid arvatakse esindavat. Ilma isikutasandil andmeid võrdlemata ei saa välistada ka, et uuritav sündmus ei ole mitu korda arvestatud.

On täiesti selge, et statistika ei saa oma funktsioone hästi täita ilma andmeid isikuandmete tasandil analüüsivata ja erinevatest andmeallikatest andmeid ühendamata. On vaja kasutada nii Statistikaameti enda kui ka teiste ametkondade kogutud andmeid ja need ühendada kirje tasandil. See ei ole võimalik isikukoodi või teiste identifitseerivate isikutunnusteta.

Kahjuks ei ole praegu võimalik isikuandmeid sisaldavaid ametkondlikke registreid ja andmekogusid laialdaselt statistikas kasutada, sest realselt on andmete kättesaadavus takistatud. Üldisem põhjus tundub olevat, et ei mõisteta isikutunnuste vajadust statistika tegemisel.

Lõpetuseks rõhutagem veelkord, et rahvastiku pidevstatistika tugineb peamiselt rahvastikusündmuste registreerimise riiklikule süsteemile. Statistika usaldusväärsus sõltub seega väga tugevasti sellest, kas registreerimissüsteem on üles ehitatud tegelikkuse kajastamise põhimõttel või ei. Et aga statistilisi andmeid saada ei ole selle süsteemi peamisi eesmärke, siis on nüüdisajal tõepäraste statistiliste andmete kogumine raskendatud.

RAHVASTIKUANDMED EESTIS — KELLELE JA KUI VABALT

Mare Ainsaar
Rahvastikuministri büroo

Artikli eesmärgiks on anda ülevaade Eesti andmehalduses esinevatest probleemidest andmeandjate ja -kasutajate vahel. Eelkõige keskendutakse kahele küsimusele:

1. Millised probleemid kerkivad esile andmete tarbija ja tootja vahel ning tootjate omavahelises koostöös?

2. Kes, mis tingimustel ja kellele peaks andmeid andma? Ja veel täpsemalt — kas andmete edastamine peaks olema tasuline?

Ettekanne, mille alusel artikkel on valminud, esitati enne avaliku teabe seaduse ja Statistikaameti tegevust reguleerivate muudatuste jõustumist 2001. aastal. Käesolevas artiklis on arvestatud juba muudatustega. Avaliku teabe seadusega määratakse kindlaks paljud üldised põhimõtted, mille alusel peaks teavet levitama. Artiklist selgub, et olulisemad andmehaldust käsitlevad seadused sätestavad tootjate ja tarbijate õigusi erinevalt.

Sissejuhatus

1992. aastal pidas Eesti Statistikalts esimese aastakonverentsi, mis oli pühendatud rahvastikuandmetele. Kaheksa aastaga on Eesti andmehalduses toimunud suured muutused.

1. On selgeks saanud Eesti kindel suundumus registritepõhisele andmehaldusele. Praegune Eesti andmehaldus on segu statistilisest ja halduslikust andmehaldusest. Tähtsamateks eeskujudeks selles valdkonnas on olnud eelkõige Soome ja Rootsi. On välja kujunenud registrite süsteem, vastu võetud andmekogude seadus ja loodud andmekogude register. Järgmiseks olulisemaks sammuks oleks soodustada registrite ristikasutust.

2. Poolsalastatud ja ainult ametkondlikuks kasutamiseks mõeldud rahvastiku-statistikast on saanud üks kõige kasutatavamaid statistika valdkondi. Lisaks paber kandjatele on võimalik andmeid saada ka elektrooniliselt ja muudel infokandjatel. Mitmekesistunud on andmete allikad (valikuuringud, registrid).

3. Paralleelselt andmete kogumise ja kasutamise arenemisega on alus pandud andmekaitsele. 1999. aastal hakkas jõulisemalt tegutsema andmete järelevalvega tegelev Andmekaitse Inspektsioon.

4. Aastaid on vaieldud andmehaldust juhtiva institutsiooni vajalikkuse ja funktsioonide üle. 2000. aastast pretendeerivad sellele kaks asutust — Statistikaamet ja algselt Riigikantselei ning alates 2000. aasta suvest Teede- ja Sideminiisteeriumi juures asuv riigi infosüsteemide osakond. 2000. aasta juulist jõustunud statistikaseaduse muudatusega kohustati kõiki andmekogusid kooskõlastama Statistikaametiga andmete struktuur, definitsioonid ja kasutatavad klassifikaatorid. Infotehnoloogilise ülesehituse ja arengu ühtlustamine on olnud erinevate tööühikute ülesanne.

5. Vastu on võetud olulised andmehaldust reguleerivad seadused. Rahvastiku-andmeid puudutavad on neist tähtsamad:

- isikuandmete kaitse seadus (1996),
- andmekogude seadus (1997),
- riikliku statistika seadus (1997),
- rahvastikuregistri seadus (2000),
- avaliku teabe seadus (2000) [Vt ka Oone 1999].

Probleemid

1998. aastal PHARE regionaalarengu programmi raames tehtud intervjuude käigus selgusid Eesti andmehalduses järgmised probleemid, mis on aktuaalsed ka praegu.

Info tootjad

Info tootjatena on meie huvisfääris eelkõige need tootjad, kes tegelevad info-ga avaliku teenistuse või avaliku sektori pakutava teenuse raames (statistika, registrid ja teised andmebaasid, omavalitsused, organisatsioonid ja eraisikud, kes on avaliku info töötleja vastutava täitja ülesannetes).

Peamised probleemid info tootjate poolt nähtuna on järgmised:

- ebakindlus oma tegevuse vajalikkuses (vahel tundub, et kogutakse rohkem andmeid kui vaja);
- sageli nähakse täiendavates infotarbijates ja infopäringutes pigem lisatööd, mis nõuab lisavahendeid;
- pidev mure finantseerimise pärast;
- pole võimalik olla pidevalt kursis kõige uuemate suundumustega metoodikas, uute definitsioonide ning andmete jaotamise põhimõtetega (klassifikaatorid);
- probleemid tarbijate asjatundmatu andmete tõlgendamisega. Kohati on levimas suhtumine, et parem juba andmeid üldse mitte anda;
- kartus eksida isikuandmete kaitse seaduse vastu. Ollakse väga ettevaatlikud ka juhtudel, kui oleks õigus andmeid anda. 1999. aastast on probleemiks ka Andmekaitse Inspektsiooni liigrange tegevus;
- puudus kogu andmehaldust koordineeriva asutuse või kogu järele (eriti maakondlikul tasandil);
- probleemiks on andmebaaside omavaheline koostöö, mida raskendab ka isikuandmete kaitse ja andmekogude seadus.

Info tarbijad

- ei tea, kust infot saada;
- aeg-ajalt tunnetavad andmete hulga ebapiisavust;
- soovivad saada ainult absoluutselt õigeid andmeid;
- sõltuvad info tootjate heast tahtest.

6. oktoobril 1999. aastal peeti samadele küsimustele lahenduse otsimiseks rahvastikuministri büroo korraldatud seminar, kuhu olid kutsutud nii andmete tootjate kui tarbijate esindajad. Seminaril tehti ettepanekuid, mis aitaksid lahendada eespool nimetatud probleeme. Tehti järgmised järeldused.

1. Iga ministeerium võiks ise olla vastutav oma valdkonna metaandmestiku eest, sealhulgas klassifikaatorite levitamise ja väljatöötamise eest. Seda tööd võiks teha koordineeritult ja koostöös klassifikaatorikeskusega.

2. On vaja aktiivsemalt jagada infot isikuandmete edastamise ja kõigi sellealaste õigusaktide kohta.

3. Kõik andmebaasid muuta avatumaks, kohustuslik võiks olla veebileht, kus oleks üleval ka hinnakiri ja andmete saamise ja kogumise põhimõtted.

Kellele ja kuidas infot anda?

Paralleelselt registrite ja muude andmebaaside arenemisega on tekkinud vajadus reguleerida täpsemalt andmete väljastamise korda. Seadusi analüüsidest ilmneb selge tendents, et uuemates õigusaktides on juurdepääs andmetele palju täpsemalt reguleeritud kui varasemates. Kui esialgu sõltus (ja sõltub osaliselt praeguseni) andmete väljastamine andmete omaja ja küsija omavahelise usalduse astmest, siis koos seadusandluse arenemisega on seadustesse kirja pandud erinevad regulatsioonid.

Riikliku statistika seadus

Seadus käib riiklike statistiliste vaatlustega kogutava info kohta ja käsitleb Statistikaameti, ka Eesti Panga ja teiste statistikaga seotud institutsioonide tegevust.

- Riikliku statistilise vaatlusega kogutud andmeid tohib kasutada ainult statistilistel eesmärkidel, kusjuures tuleb tagada andmekaitse. Ometi ei selgu seadusest, mida statistilised eesmärgid täpselt hõlmavad.
- Statistiliste vaatlustega kogutavat infot ei tohi kasutada kontrolli, maksumise ega ühelgi teisel mittestatistilisel otstarbel, v.a juhtudel, kui seda ei nõua mõni teine seadus.
- Kogutud andmeid võib riikliku statistilise vaatluse korraldaja edastada ja avaldada ainult sellisel kujul, mis **väljastab andmeesitaja otsese või kaudse identifitseerimise võimaluse**. Et otsustada, kas andmeesitaja on identifitseeritav, võetakse arvesse kõik vahendid, mida kolmas isik võib põhjendatult kasutada nimetatud andmeesitaja identifitseerimiseks. Andmeid avaldatakse ja edastatakse üldistatult, rühmitatult ja ilma identifitseerimist võimaldavate tunnusteta, grupeeritult vähemalt kolme kaupa, kus ühegi isiku andmete osatähtsus koondandmetes ei ole suurem

kui 50 protsenti. Konfidentsiaalseid andmeid avaldatakse või edastatakse ainult paragrahvi 9 lõike 2 kohaselt. Erandina nähakse ette vaid teadusuuringuteks kasutatavad andmed, kus isik võib olla identifitseeritav, või andmed, mille edastamiseks on andmeesitajalt luba küsitud.

- Andmeesitajal on õigus üks kord aastas tutvuda tema kohta kogutud andmetega.

Isikuandmete kaitse seadus

Seadust ei kohaldata: 1) füüsilise isiku enda kogutud isikuandmete töötlemisel isiklikuks otstarbeks ega 2) riigisaladust sisaldavate isikuandmete puhul. Üldiselt sätestab seadus, et isikuandmed on kõik andmed üheselt tuvastatud füüsilise isiku kohta. Seaduses tehakse vahet mittedelikaatsetel ja delikaatsetel isikuandmetel. (Delikaatsed isikuandmed on poliitilised vaated, usulised ja maailmavaatelised veendumused, etniline päritolu ja rassiline kuuluvus, tervislik seisund, pärilikkus, seksuaalelu ja kriminaalmenetluse käigu andmed enne avalikku kohtuistungit. Delikaatsetele isikuandmetele on kehtestatud eripiirangud.)

Üldjoontes ütleb seadus, et andmeid tohib töödelda

- ilma isiku nõusolekuta, kui seda tehakse **seadusega ettenähtud kohustuse** täitmiseks;
- avalikku või üldist huvi silmas pidava ülesande täitmiseks, mis on **seadusega pandud** vastutavale töötlejale või kolmandale isikule, kellele andmed üle antakse.

Andmekogude seadus

Seadus kehtib riigi, kohaliku omavalitsuse ja eraõigusliku isiku andmekogu kohta.

Seadus ei käsitle andmekogusid, mida peetakse 1) raamatupidamise seaduse alusel; 2) ainult teaduslikul, pedagoogilisel ja kodu-uurimuslikul eesmärgil; 3) organisatsiooni operatiivse juhtimise vajaduseks õigusaktidest tulenevate või põhikirjaliste ülesannete täitmisel, kui neist kogudest ei väljastata andmeid muudeks vajadusteks ja teistele isikutele.

Avalik-õiguslikul isikul on õigus tema peetavasse andmekogusse koguda ainult andmekogu asutamise seaduses nimetatud ülesannete täitmiseks vajalikke andmeid.

Seadus sätestab väga üldised põhimõtted ja jätab andmete väljastamise teiste seaduste otsustamissfääri. Riigi ja kohaliku omavalitsuse andmekogudes peetavad andmed on avalikud ja igal Eesti kodanikul on õigus nendega seaduses sätestatud korras tutvuda ning neist ärakirju saada, välja arvatud juhul, kui seadusega on andmetega tutvumine või andmete väljastamine keelatud või need on ette nähtud ainult ametialaseks kasutamiseks.

Üheselt on määratud vaid, et isikul on õigus saada teavet tema enda kohta käivate andmete kohta. Isiku soovil teavitab vastutav töötaja teda tema kohta käivate isikuandmete olemasolust või nende puudumisest.

Avaliku teabe seadus

Seadus reguleerib selle info levitamist, mis tekib riigi või kohaliku omavalitsuse ülesande täitmise käigus.

Teavet saavad:

- riigi või kohaliku omavalitsuse asutuse töötaja — tööülesannete täitmiseks;
- isik — eraelulisi isikuandmeid kolmanda isiku kohta, juhul kui infopäring on põhjendatud.

Isikuandmeteks, millele kehtib juurdepääsupiirang ja mis on ette nähtud asutusesiseseks kasutamiseks, loetakse delikaatseid ja eraelulisi andmeid (perekonnaelu üksikasjad, sotsiaalabi või sotsiaalteenuste osutamise taotlemine, isiksuse omadusi, võimeid või muid isiku iseloomuomadusi kirjeldavad, vaimseid või füüsilisi kannatusi kirjeldavad andmed; maksustamisega kogutud teave, välja arvatud maksuvõlgnevuste kohta ja tasutud tulumaksu kogusumma).

Rahvastikuregistri seadus

Seadus sätestab, et andmed on avalikuks kasutamiseks üksnes isiku enda nõusolekul või riigi- ja kohaliku omavalitsuse asutusele või juriidilisele või füüsilisele isikule avaliku ülesande täitmiseks.

Esmakordselt on seaduses püütud reguleerida andmete usaldusväärsuse probleemi, sätestades andmete töötaja kohustuse informeerida andmesaajat võimalikest vigadest.

Volitatud töötaja on koguni kohustatud sulgema juurdepääsu mittetäielikele või ebaõigetele rahvastikuregistri objekti andmetele või andmetele, mille

suhtes on tekkinud kahtlus, kas need on õiged. Andmesaaja, kes kasutab andmeid, on kohustatud volitatud töötlejale viivitamata teatama avastatud ebaõigetest, kaheldavatest või mittetäielikest andmetest.

Seega on õigus andmeid saada:

- täisealisel isikul enda ja oma alaealiste laste ning eestkostetavate kohta;
- riigi- ja kohaliku omavalitsuse asutusel ning juriidilisel või füüsilisel isikul avaliku ülesande täitmiseks;
- füüsilisel ja juriidilisel isikul õigustatud huvi korral. Juhul kui õigustatud huvi ületab 100 isikut aastas, on selleks vaja Andmekaitse Inspektsiooni nõusolekut;
- välisriigi asutusel ja isikul, kui see õigus on sätestatud välislepinguga või vastutava töötleja korraldusel.

Eraldi on seaduses märgitud, et teaduse või statistika tarbeks andmete edastamisel tuleb andmetele anda kuju, mis ei võimalda üheselt tuvastada rahvastikuregistri objektiks olevat isikut. Kasutamise järel tuleb rahvastikuregistri objekti ühest tuvastamist võimaldavad andmed arhiveerida või hävitada.

Kui palju maksab info?

Võrreldes andmesaajate ringi kindlaksmääramisega on info maksumus tunduvalt nõrgemalt reguleeritud õigusaktide tasandil. Samas hakkab info hind tulevikus enam määrama selle tegeliku leviku võimalusi, st kas riikliku andmehalduse käigus kogutu on kättesaadav kõigile. Tendentsi andmehalduse kommertsialiseerumisele kiirendab paljude vastutava töötleja funktsioonide andmine eraettevõtetele, kes on otseselt huvitatud mitte niivõrd andmete levikust kui andmehalduse majanduslikust tasuvusest.

Oluliseks üldisema arutluse objektiks võiks kujuneda aksioom, et andmehaldus on avaliku teenuse osa (registrid, statistika, uurimused jm) ja enamasti finantseeritakse andmehaldust juba niigi riigieelarvest. Seega on küsimus, kui suur võiks olla andmete väljastamise hind.

Seni on andmete hind sageli kujunenud küsija finantsvõimaluste põhjal. Pikaajalisem andmete väljastamise kogemus on Statistikaametil ja AS Andmevaral. Kogemused on aga erinevad. Kui Statistikaametis on järgitud põhimõtet, et andmete väljastamisel maksab tarbija enamasti ainult and-

mete otsese edastamisega seotud kulud (paber, ketas, tööaeg jne), siis AS Andmevara on sisuliselt kehtestanud hinna infole kui sellisele.

Andmekogude seaduse järgi tuleb tulevikus kõigi registrite asutamise määru- ses tuua andmekogust tasuta infot saavate asutuste nimekirja.

Siinjuures on esitatud ülevaade sellest, kuidas mõned seadused reguleerivad info hinda.

Riikliku statistika seadus ei määra andmete väljastamise hinda. Seaduses on igaks juhuks küll sätestatud, et vaatluste korraldajale esitavad riigiasutused, omavalitsused ja juriidilised organid andmeid kohustuslikult ja tasuta. And- mete mitteesitamise või hilinemisega esitamise puhul võib määrata raha- trahvi.

Andmekogude seaduses sätestatud juhtudel on riigi või kohaliku omavalit- suse andmekogus olevate andmetega tutvumine, andmete väljastamine ja neist ärakirjade saamine tasuta.

Isikuandmete seaduses ei räägita info hinnast. Isikul on õigus saada vastu- tavalt töötlejalt enda kohta andmeid üks kord aastas tasuta, teist korda juba tasu eest. Tasu ei tohi ületada andmete väljastamisega seotud kulusid.

Avaliku teabe seaduse järgi on juurdepääs teabele üldjuhul tasuta. Tasu võib võtta vaid väljastamise otseste kulutuste eest. Seaduses soovitatakse teadus- töö, oma õiguste ja kohustuste jaoks vajatav info vabastada ka otseste kulu- tuste maksmisest.

Rahvastikuregistri seaduse kohaselt on täisealisel isikul enda, oma ala- ealiste laste ja eestkostetavate kohta õigus saada tasuta infot üks kord kalend- riaastas.

Seaduse järgi tasuvad andmesaajad andmete väljastamise kulud kinnitatud hindade alusel. Rahvastikuregistri pidamise tulude ja kulude eelarve kinnitab siseminister.

Samal ajal näeb seadus ette, et andmeandja poolt rahvastikuregistrisse and- mete andmisel tehtud kulutused kaetakse riigieelarvest andmeandjale selle ülesande täitmiseks eraldatud vahenditest ehk andmeandjate oma eelarvest. Nähakse ette ka võimalus, et neid kulutusi saab katta andmete töötlemise lepingu raames Siseministeeriumi rahvastikuregistri pidamiseks ettenähtud vahenditest, kuid see on ainult võimalus ja mitte kohustus. Ka riigi- ja koha-

liku omavalitsuse asutuse andmetele juurdepääsu kulud kaetakse riigiasutuse või kohaliku omavalitsusüksuse eelarvest.

Kokkuvõte

Viie peamise rahvastikundmete haldust reguleeriva seaduse analüüs näitas, et

- 1) seadustes on andmete edastamine reguleeritud erinevalt, kuid üldiselt on ajaliselt hilisemates seadustes andmete edastamine tunduvalt põhjalikumalt reguleeritud. Kahtlemata on see mõjutatud andmete edastamise ja tarbimise aktuaalsemaks muutumisest;
- 2) andmete edastamisel erinevatele tarbijate rühmadele tehakse tavaliselt vahet avaliku ülesande täitjatel ja erahuvides tegutsejatel;
- 3) andmete hind ja selle reguleerimine on Eestis veel kujunemisjärgus.

Kasutatud kirjandus

- Einberg, U. 1999. Andmekogudest aastal 1999. Odrats, I. (koost). Infotenoloogia haldusjuhtimises. Aastaraamat 1999, <http://www.eik.ee/it99/pealeht.htm> (17.12.2000). Eesti Informaatikakeskus.
- Oone, K. 1999. Aasta seadusloome poole pealt. Odrats, I. (koost). Infotenoloogia haldusjuhtimises. Aastaraamat 1999, <http://www.eik.ee/it99/pealeht.htm> (17.12.2000). Eesti Informaatikakeskus.

EESTI RAHVASTIKUREGISTRI OLEVİK JA TULEVİK

Peeter K uuts, Ragnar  un
Siseministeerium

P arast Eesti taasiseseisvumist 1991. aastal oli vaja kiiresti lahendada rahvastiku arvestusega seotud k usimused (valijate nimekirjade koostamine, isiku identifitseerimiskoodide v aljastamine jne). Nii loodigi 1992. aasta rahavahe- tuse nimekirjade p ohjal andmebaas, mis sisaldas 1 520 490 isikukirjet mini- maalsete andmetega (isikukood, s unniaeg, ees- ja perekonnanimi, elukoht). 1993. aastal t aiendati andmebaasi andmetega, mis saadi elukoha registreeri- mise teadetest ja Eesti Vabariigi kodaniku passi taotlustest (isanimi, s unni- koht, kodakondsus, emakeel, elukoha muutused). 1995. aasta II poolaastal v ottis valitsus vastu rahvastiku arvestuse andmebaasi pidamise p ohim aaruse. 1996. aastast on rahvastiku arvestuse andmebaasi vastutav t ootleja Sise- ministeerium.

1997. aasta 12. m artsil vastuv oetud andmekogude seadus m aarab, et rahvas- tikuregister on Eesti p ohiregister. Rahvastikuregistri seadus v oeti Riigikogus vastu 2000. aasta 30. mail. Seadusega s atestati, et rahvastikuregister v oetakse kasutusele 2002. aasta 1. jaanuaril. Praegu on kasutusel kolm andmebaasi:

- rahvastiku arvestuse andmebaas,
- Eesti h aale iguslike kodanike riiklik register ja
- Eesti h aale iguslike v alismaalaste riiklik register.

Rahvastikuregistri seaduse kohaselt tuleb rahvastikuregister moodustada eelmainitud kolme andmebaasi p ohjal.

2000. aasta oli peale rahvastikuregistri seaduse vastuv otmise oluline ka selle poolest, et algas haldusreform. Rahvastikuregistriga on see reform seotud eelk oige selle kaudu, et rahvastikuregistri seaduses on inimeste elukoha  ule arvestuse pidamine j aetud kohalikele omavalitsustele ja nende suurus on ot- seselt seotud nende tehtud t oo kvaliteediga.  Uhelt poolt, mida v aiksem on omavalitsus, seda paremini tunnevad ametnikud seal elavaid inimesi ja nende  ule arvestuse pidamine on lihtsam ja t ापsem. Teiselt poolt on aga selliste v ai- keste omavalitsuste pidamine kulukam ja suuremad omavalitsus uksused v oimaldaksid ka kvaliteetsemat t ooj odu. Niisiis on ka rahvastiku arvestuse

poolt vaja leida optimaalne lahendus Eesti haldusjaotuses. Veel võiks 2000. aastal tehtust esile tõsta järgmist:

veebruar — koostööleping Siseministeeriumi ja Soome rahvastikuregistri-keskuse vahel;

mai — Riigikogu võttis vastu rahvastikuregistri seaduse;

oktoober — valitsus määras rahvastikuregistri volitatud töötlejaks AS Andmevara;

oktoober — peeti 10. rahvastiku arvestuse seminar.

Siseministri määrusega

- kehtestatakse andmeandjate kaupa üleantavate andmete loetelu;
- kinnitatakse elukohateate vorm ja selle täitmise juhend;
- kehtestatakse isikukoodi moodustamise ja andmise kord;
- sätestatakse rahvastikuregistri andmete töötlejate atesteerimisnõuded ja atesteerimise kord;
- kinnitatakse andmete edastamise vormid ja vajaduse korral nende täitmise juhendid.

Infotehnoloogia ja andmetöötuse poolt olid aasta 2000 ülesanneteks ja märksõnadeks rahvastikuregistri tarkvara koostamine rahvastiku arvestuse andmebaasi tarkvara täiendamiseks, elukohateadete töötlemise tarkvara koostamine ja kasutuselevõtt, kohalike omavalitsuste töötlejate koolitus elukohateadete kasutuselevõtuks, *WEB*-tehnoloogia ja *on-line*-töökohtade kasutuselevõtt vähemalt pooltes linnades ja valdades. Rahvastikuregistri töötajatel on selge, et registri andmekvaliteedi parandamise kõige lihtsam ja kasulikum moodus on võimalikult suur kasutajaskond ja nendepoolne tagasiside. Võiks öelda, et andmebaas on täpselt sama kvaliteetne, kui kvaliteetseks peab teda kõige norivam klient. Oluline on, et kasutaja saaks just ise (muidugi seadusega lubatud juhtudel) leida registrist teda huvitavad andmed, mitte ei kasutaks registripidajat vahelülina. Selleks ongi rahvastikuregistri töötlejad pannud suurt rõhku *on-line*-töökohtade arvu suurendamisele. Tabel 1 näitab, kuidas on neljal viimasel aastal suurenenud *on-line*-päringute arv ja kui palju on eri ametkondadel juba töökohti. Eriti tuleb tähelepanu juhtida AS Andmevara kui andmebaasi volitatud töötleja tehtud päringute osa tunduvalt vähenemisele kogu päringute hulgas. See on just see, et lastakse klientidel pöörduda andmete saamiseks otse andmebaasi poole, mitte aga volitatud või vastutava

Tabel 1

	Töö- ko- had	Ka- sut.	Päringuid 2000. a		1999	1998	1997
			I pa arv	võrreldes 1999. a, %			
KOKKU	768	802	38 4770	58	659 395	341 073	99 594
Kohal. oma- valitsused	120	117	11 445	74	15 376	11 955	4 233
Tallinna lin- navalitsus	11	11	1 121	42	2 641	240	55
Elukoha re- gistr. talitused	97	88	98 154	42	235 531	127 677	6 391
Tallinnas	43	35	42 966	39	111 523	58 519	2 348
mujal Eestis	54	53	55 188	45	124 008	69 158	4 043
AS Andme- vara	21	21	65 675	37	176 647	153 932	87 598
KMA	43	38	17 794	59	30 051	9 277	182
Tallinnas	15	10	15 012	50	29 817	8 705	49
mujal Eestis	7	7	2 782	1 188	234	572	133
Sisemin. rahv. arv. osak.	3	3	4 949	155	3 188	-	-
Sisemin. rahv. toim. osakond	4	4	604	-	0	-	-
Piirivalve	51	53	9 988	158	6 329	4 150	405
Tallinnas	25	28	7 763	265	2 924	2 747	405
mujal Eestis	26	25	2 225	65	3 405	1 403	-
Politseiamet	284	323	94 180	148	63 468	20 353	483
Tallinnas	131	138	54 757	185	29 593	11 080	483
mujal Eestis	153	185	39 423	116	33 875	9 273	-
Päästeamet	3	8	211	60	354	626	-
AKI	1	1	31	76	41	20	-
Tolliamet	31	29	4 053	68	5 989	5 581	302
Tallinnas	18	17	2 234	65	3 430	3 457	302
mujal Eestis	13	12	1 819	71	2 559	2 124	-
Statistikaamet	1	1	2 059	-	0	0	0
Maksuamet	1	1	4 018	200	2 003	-	-

	Töö- ko- had	Ka- sut.	Päringuid 2000. a		1999	1998	Järg
			I pa arv	võrreldes 1999. a, %			1997
Kaitsej. peastaap	9	10	3 190	129	2 481	307	-
Harju ja Tal- linna RKO	5	6	4 645	57	8 115	5 587	-
Välismin.	4	4	1 419	692	205		-
Sotsiaalkind- lustusamet	5	5	9 549	61	15 519	-	-
Pensionia- ametid	5	5	21 537	25	87 421	282	-
Tallinnas	7	7	5 364	11	48 589	282	-
mujal Eestis	25	25	16 173	49	38 832	-	-
Haigekassa	3	3	4 128		0	-	-
Haiglad	2	2	4 752	407	1 168	-	-
ARK	5	5	1 767	116	1 528	-	-
Justiitsmin.	1	1	4	-	0	-	-
Kohtud	19	19	9 773	-	0	-	-
Keskkon- naamet	4	9	111	-	0	-	-
Muud	8	8	2 386	60	4 006	1 197	-

töötajate poole. Kahjuks ei saa võimaldada *on-line*-töökohti kõigile soovijatele, sest piire seavad finantsilised ressursid. Kõigile klientidele, kes need töökohtad saavad, tuleb korraldada väljaõpe, et nad oskaksid pakutavaid võimalusi maksimaalselt ära kasutada.

Nüüd lühidalt sellest, mida tuleb teha lähitulevikus, s.o eelkõige aastal 2001, et aastal 2002 saaks rahvastikuregistri kasutusele võtta. Tuleb vastu võtta rahvastikuregistri tööks vajalikud õigusaktid. Näiteks tuleb rahvastikuregistri tarkvara väljatöötamisel ja kasutuselevõtmisel kinnitada vastutava töötaja poolt rahvastikuregistri andmete töötlemise tehnoloogiline ülesehitus, rahvastikuregistri algoritm, tarkvara kirjeldus ja tarkvara nimi ning anda luba rahvastikuregistri moodustamisel tarkvara kasutuselevõtmiseks. Tuleb sõlmida lepingud andmesaajate ja andmeandjatega. Rahvastikuregistri vastutav töötaja korraldab ka andmeandjate ja andmesaajate väljaõppe. On kavas

sõlmida andmevahetusleping Soomega, et lõpeks inimeste n-õ topeltarvestamine. Praegu on juhtumeid, kus inimene on arvel nii Soome kui ka Eesti rahvastikuregistris. Peab kaduma olukord, kus näiteks Eesti andmetel viibib Soomes kolmsada Eesti kodanikku, Soome andmetel on see arv aga kümme korda suurem.

Kuidas ettekande alguses nimetatud kolmest andmebaasist moodustatakse rahvastikuregister? Esimese sammuna kantakse rahvastiku arvestuse andmebaasist, Eesti hääleõiguslike kodanike riiklikust registrist ja Eesti hääleõiguslike välismaalaste riiklikust registrist andmed üle rahvastikuregistri arhiivi. Sealt kantakse rahvastikuregistri aktuaalsete andmete hulka nende isikute andmed, kelle kohta on järgmised andmed:

- 1) perekonnanimi või perekonnanimed;
- 2) eesnimi või eesnimed;
- 3) sünniaeg;
- 4) sünnikoht;
- 5) sugu;
- 6) isikukood;
- 7) elukoha aadress vähemalt kohaliku omavalitsusüksuse täpsusega;
- 8) Eesti kodakondsus või kehtiv elamisluba.

Mis saab siis, kui inimese kohta pole kõiki neid andmeid? Kui isiku kohta puuduvad osaliselt sätestatud andmed, siis esitab volitatud töötaja puuduvate andmete väljaselgitamiseks vähemalt üheksa kuud enne rahvastikuregistri kasutuselevõttu järelepärimise andmeandjatele. Andmeandja peab välja selgitama ja esitama õiged andmed kolme kuu jooksul järelepärimise saamisest rahvastikuregistri seaduses andmete üleandmiseks sätestatud korras. Kui andmeandjal ei ole mõne isiku kohta andmeid, siis teatab ta andmete puudumisest. Need isikud, kellel jääb osa sätestatud andmeid puuduma ka pärast andmete täpsustusetappi, jäävad rahvastikuregistri arhiivi ja aktuaalsete andmete hulka ei jõua.

Lõpetuseks väike ülevaade elanikkonnast, kes on arvel Eesti rahvastiku arvestuse andmebaasis (vt tabel 2) ja mõned viited huvilistele, kes tahaksid rohkem ja põhjalikumalt end asjaga kurssi viia.

Tabel 2

	Vanuserühm				Kokku
	0-14	15-39	40-64	65-...	
EV kodaniku passid, arv	59 922	402 755	344 516	164 598	971 791
eestlased, %	61,3	75	69,4	71,5	71,6
venelased, %	17,3	13,7	14	8,7	13,2
määramata, %	19,5	8,7	13	16,8	12,3
Alalised elamisload, arv	19 727	66 944	80 153	32 907	199 731
ühegi riigi kodakondsusega, %	50	77,1	60,8	42	62,1
kvoodi alla käivate riikide kodanikud, %	50	22,9	39,2	58	37,9
Tähtajalised elamisload, arv	15 067	26 116	23 470	11 149	75 802
ühegi riigi kodakondsusega, %	18,2	62	50,7	27,6	44,7
kvoodi alla käivate riikide kodanikud, %	81	34,6	45,9	71,5	52,7
Dokumenteerimata, arv	151 821	30714	22 051	8046	212 632
eestlased, %	69,7	21,6	8,4	9,5	54
venelased, %	9,2	12,7	17,6	9,1	10,6
määramata, %	20,1	62,3	68	78	33,4
KOKKU	246 537	526 529	470 190	216 700	1 459 956

Tabeli kommentaariks võib öelda vaid seda, et dokumenteerimata 0-15-aastastest enamik on Eesti Vabariigi kodanike lapsed, kes ei ole endale veel EV kodaniku passi võtnud (arvatavasti kõik eestlased pluss ka veel suur osa määramata rahvusega lastest, kokku seega pea 90% 151 821 lapsest). Selles tabelis sisaldub ligikaudu 70 000 inimest, kes on end oma viimasest elukohast välja registreerinud, kuid kellest rahvastiku arvestuse andmebaas midagi rohkemat ei tea. Nendest paljud peaksid kuuluma nende dokumenteerimata inimeste hulka, kes on vanemad kui 15-aastased ja kelle puhul võib eeldada, et nad on Eestist lahkunud.

Veebilehtede viidad:

Siseministeeriumi rahvastiku arvestuse osakond

<http://www.sisemin.gov.ee/rahvastik/arvestus.html>

Rahvastikuregistri volitatud töötleja AS Andmevara

<http://www.andmevara.ee>

Rahvastikuregistri seadus

<http://seadus.ibs.ee/aktid/rk.s.20000531.69.20000801.html>

Soome rahvastikuregistrikeskus

<http://www.vaestorekisterikeskus.fi>

Läänemere maade registrite konverentsi ettekanded

http://www.regno.ee/ettekanne_frame.htm

RAHVASTIKUREGISTRI SEIS JA STAATUS

Tarmo Sumeri
AS Andmevara

Ajalooline ülevaade

Tabelis on ülevaade rahvastiku arvestuse andmebaasi praeguse seisu kujunemisest.

1992	Rahvastiku arvestuse keskandmebaasi asutamine Statistikaametis. Andmevahetus Statistikaametiga. <i>Mainframe</i> -tehnoloogia. Failkommunikatsiooni teke andmeandjatega
1994	Rahvastiku arvestuse andmebaasi volitatud töötlemise algus Riigiarvutuskeskuses. <i>PC</i> -d ja <i>Novell</i> -võrk
1995	Rahvastiku arvestuse andmebaasi põhimääruse kinnitamine. <i>Unix</i> -server, <i>Progress</i> -andmebaasisüsteem
1997	Keskbaasi <i>on-line</i> kasutamine päringuteks ja andmete korrigeerimiseks
1998	<i>Off-line</i> -teeninduse väljakujunemine
1999	Hääleõiguslike valijate registreerimise põhimääruste kinnitamine
2000	Rahvastikuregistri seaduse vastuvõtmine. AS Andmevara kinnitamine rahvastikuregistri volitatud töötlejaks. Ametnike <i>on-line</i> -töökohad
2001	Rahvastikuregistri seaduse rakendusaktide väljatöötamine ja kinnitamine. Ametnike <i>on-line</i> -töökohtade lisandumine, integratsiooni arendamine

Nagu tabelist näha, asutati rahvastiku arvestuse keskandmebaas 1992. aastal. Baasi aluseks olid rahavahetusnimekirjad, mis sisestati maakondades arvutitesse ja koondati Statistikaametis ühte andmebaasi. Samal ajal hakati andmeid aktualiseerima — kohalikud omavalitsused teatasid maakonnavalitsuse kaudu failkommunikatsiooni teel kõigist registreeritud elukohavahetustest. Hiljem teatati samal meetodil ka perekonnaseisu sündmustest.

Seadusandluse areng

1995. aastal kinnitas valitsus rahvastiku arvestuse andmebaasi põhimääruse. Põhimäärus lõi seadusliku aluse andmebaasi tegutsemisele, andmeandjatele pandi kohustus edastada andmed keskregistrile. Määrati kindlaks rahvastiku arvestuse keskandmebaasist andmete väljastamise alused.

Väga oluline oli isikuandmete kaitse seaduse ja andmekogude seaduse vastuvõtmine Riigikogus vastavalt 1996. ja 1997. aastal. Nimetatud seadused löid aluse andmekogude valdkonna korrastamiseks.

1999. aastal kinnitas Vabariigi Valitsus hääleõiguslike valijate registrite põhimäärused.

2000. aasta 31. mail võeti Riigikogus vastu rahvastikuregistri seadus. Seadus jõustub kahes etapis: 1. jaanuarist 2001 jõustuvad elukoha registreerimise korda sätestavad peatükid ja 1. jaanuarist 2002 jõustub kogu seadus. Seaduse järgi moodustatakse rahvastikuregister kolmest praegu juriidiliselt eraldiseisvast registrist. Need on rahvastiku arvestuse andmebaas, hääleõiguslike kodanike riiklik register ja hääleõiguslike välismaalaste riiklik register.

Rahvastikuregistri staatus

Eespool nimetatud õigusaktide alusel peetakse praegu rahvastiku arvestuse andmebaasi. Andmebaasi vastutav töötleja on Siseministeerium ja volitatud töötleja lepingu alusel AS Andmevara. Rahvastikuregistri seaduse alusel minnakse samal ajal üle rahvastikuregistri pidamisele. Ettevalmistustööde raames kinnitas Vabariigi Valitsus 2000. aasta oktoobris rahvastikuregistri volitatud töötleja. Enne kinnitamist hindasid volitatud töötleja kandidaati sõltumatud audiitorid ja Andmekaitse Inspektsioon.

Kogu rahvastikuregistri töö tugineb õigusaktidele, kogu info, mida kogutakse, peab põhinema juriidilistel alusdokumentidel.

Tehnoloogia areng

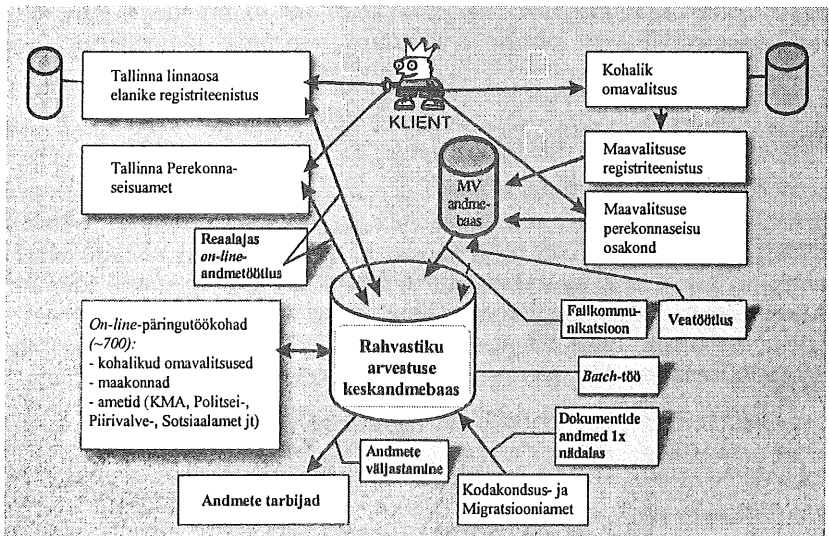
1997. aastal alustati rahvastiku arvestuse keskandmebaasiga ühendatud *on-line*-terminalide paigaldamist. Sellega tekkis võimalus lisaks senisele andmete *off-line*-tarbimisele saada keskbaasist infot isikute kohta reaalajas. Praegu on otsejuurdepääs päringute tegemiseks rohkem kui tuhandel avaliku sektori ja kohaliku omavalitsuse töötajal. Suuremad tarbijad on näiteks

politseiamet, piirivalveamet, tolliamet, kodakondsus- ja migratsiooniamet, maavalitsused, kohalikud omavalitsused, haigekassa, pensioniamet jt. Aastas teevad andmete tarbijad üle miljoni päringu.

2000. aasta sügisest on alustatud ka reaalajas andmehõivet otse ametnike töökohtadelt. Selleks on loodud perekonnaseisumetniku ja elukoha registreerimise ametniku töökohtade tarkvara, mis suhtleb otse keskandmebaasiga. Paeguseks laekub *on-line*-töökohtade kaudu üle 30% andmete uuendustest.

On-line-tehnoloogiate kasutuselevõtt on loonud eeldused operatiivse, reaalajas toimiva ja õigusaktide alusel andmete laekumise tagamiseks. Selle tulemusena on ka viimastel aastatel oluliselt paranenud andmete kvaliteet ja andmebaasi täitumus.

Rahvastiku arvestuse infosüsteemi toimimise skeem



Präegune olukord on kujutatud skeemil. Keskandmebaasi laekuvad andmete uuendused Tallinna Perekonnaseisumetist ja linnaosade elukoha registreerimise osakondadest reaalajas otse ametnike töökohtadelt. Teistest kohalikest omavalitsustest laekub info failkommunikatsiooniga üks kord näda-

las. Üks kord nädalas saabub info ka KMAst isikute dokumentide kohta. Lisaks on ametnikel otsejuurdepääs isikuandmete korrigeerimiseks keskandmebaasis. Koos rahvastikuregistri seaduse elukoha registreerimise peatükkide rakendamisega (1.01.2001) hakkab info laekuma samuti otsepöördusega *on-line*-töökohtadelt.

Andmete tarbijad saavad infot kas otsejuurdepääsuga *on-line*-terminalidelt või perioodiliste väljavõtetenä.

Andmete praegune seis rahvastiku arvestuse keskandmebaasis

Isikukirjed	
elus	1 471 480
surnud	146 503
välismaal	50 239
Isikusuhted	
vanem — laps	342 908
abikaasa	66 272
lõppenud suhteid	12 996
Isikute aadressid	
alaline, tähtajaline	1 364 679
väljaregistreeritud	77 310
muud (KMA teatud aadress, vale aadress)	29 491

Nagu tabelist näha, sisaldab rahvastiku arvestuse keskandmebaas ligikaudu 100 000 kirjet rohkem, kui andis rahvaloenduse tulemus. Erinevuse põhjused võivad olla järgmised.

1. Rahvastiku arvestuse andmebaasi salvestatakse isikusündmuste andmeid ainult juriidiliste dokumentide alusel. Kui isik on kolinud välismaale, aga sellest ei ole kohalikku omavalitsust teavitanud, ei kajastu see ka andmebaasis. Samuti esineb surmajuhtumeid, kus surnud isik jääb esialgu tuvastamata. Olukorra peaks selles osas lahendama rahvastikuregistri seadus, mis annab kohalikule omavalitsusele õiguse teatud juhtudel isikukirjeid arhiveerida.
2. Alaloendus rahvaloendusel.

3. Vigased isikukirjed rahvastiku arvestuse keskandmebaasis. Keskandmebaasi on veel jäänud mõningad rahavahetusnimikirjade isikukirjed, mille õigsust on mitmel põhjusel raske kontrollida. Olukorra lahendab *on-line*-tehnoloogia levik kohalikesse omavalitsustesse. Praegu võib öelda, et selliseid kirjeid andmebaasis ei tohiks olla rohkem kui mõni tuhat.

RR keskandmebaasi täitumus andmetega

	1996	1997	1998	1999	2000
Kirjeid andmebaasis	1 553 787	1 551 303	1 529 247	1 508 254	1 471 480
Täitumus andmetega (% kirjete arvust)					
Isanimi	39	55	66	71	81
Sünnikoht	36	54	67	75	86
Kodakondsus	34	56	71	74	80
Emakeel	31	48	62	68	78
Rahvus	38	56	68	76	85
Perekonnaseis	27	45	59	64	74
Haridus	25	39	52	57	64
Tegevusala	21	36	40	54	62

Tabelis on toodud aastati keskandmebaasi täitumus isiku lisaandmetega. 1992. aastal oli kõigi nende rekvisiitide täitumus null.

Uute andmete laekumisel on tagatud isanime, sünnikoha, kodakondsuse ja perekonnaseisu sisestamine keskandmebaasi. Emakeel, rahvus, haridus ja tegevusala sisestatakse vabatahtlikkuse alusel isiku ütluspõhiselt.

Andmete väljastamine rahvastiku arvestuse andmebaasist

Teenuste osutamine põhineb juriidiliselt rahvastiku arvestuse andmebaasi põhimäärusel ja siseministri 22.02.1999 kinnitatud määrusel nr 44 "Andmetele juurdepääsu korraldamine". Teenuste lähem tutvustus on toodud veebilehel <http://www.andmevara.ee/>.

Aastate jooksul on välja kujunenud hea koostöö mitmete andmeid tarvitavate asutustega, sealhulgas ka Statistikaameti isikustatistika teenistusega.

Teenustööd kliendi tellimuse alusel:

- juhuvalimid,
- isikuandmete võrdlus,
- kliendi andmebaasi uuendamine,
- *on-line*-teenused.

Perioodilised statistilised väljundid:

- isikute arv kohaliku omavalitsuse, soo ja sünniaasta järgi,
- andmebaasi täitumus,
- sündmuste arv,
- muudatuste arv.

Kokkuvõte

Rahvastiku arvestus põhineb õigusaktidel ja isikusündmused registreeritakse andmebaasi ainult juriidiliste dokumentide alusel. Rahvastikuregistri seaduse järgi sisestatakse ka isiku statistilised andmed isiku allkirjaga kinnitatud teate alusel. Ainult sellisel viisil on võimalik tagada korrektne ning juriidilise ja statistilise tarbimisväärtusega andmekogum.

EESTI MEDITSIINILINE SÜNNIREGISTER JA SELLE STAATUS

Alvi Tellmann
Ekspérimentaalse ja Kliinilise Meditsiini Instituut

Eesti meditsiiniline sünniregister töötab 1992. aastast [Karro 1995; Tellmann et al 2000].

Registri omanik ehk vastutav töötleja on Sotsiaalministeerium. Registri pidaja ehk volitatud töötleja on Ekspérimentaalse ja Kliinilise Meditsiini Instituut. Registrit peetakse eelnimetatud ministeeriumi ja instituudi vahel sõlmitud lepingu alusel.

Registri pidamise eesmärk on kõikide Eesti territooriumil toimunud laste sündide (nii elus- kui surnultsündide) registreerimine ning vastsündinu ja tema vanemate, raseduse ja sünnituse kohta käivate andmete töötlemine Sotsiaalministeeriumile tervishoiukorralduse seaduse ja teiste õigusaktidega pandud ülesannete täitmiseks.

Registri objekt on lapse sünd Eesti territooriumil.

Registri andmete kogumise vorm ehk andmete alusdokument on sünnikaart.

Sünnikaardile kantakse:

- haigla andmed (haigla nimi ja sünnitusloo number);
- vanemate isikuandmed (ema ja isa nimi, isikukood, elukoht, rahvus, haridus, tegevusala, amet, ema perekonnaseis, abielu registreerimise või vabaabielu alustamise aeg);
- andmed ema varasemate raseduste ja sünnituste kohta (varasemate sünnituste arv, nende lõppemine elus- või surnultsünniga, eelmise sünnituse kuupäev, varasemate abortilõppega raseduste arv aborti liigi järgi, elus olevate ja esimesel elunädalal surnud laste arv);
- andmed käesoleva raseduse kulu kohta (külüstuste arv naistearsti, perearsti ja ämmaemanda juurde, raseduskestus esimesel külüstusel, suitsetamine raseduse ajal, rasedusaegsed toimingud, riskitegurid ja tüsistused);
- sünnituse andmed (raseduskestus sünnitushetkel, sünnituspuhused ja -järgsed diagnoosid, sünnitamisviis, valutustamine, muud toimingud

sünnitusel, sünnituse ning väljutusperioodi kestus, peresünnituse olemasolu);

- vastsündinu andmed (isikukood, sündimise kellaaeg, sugu, sündinud laste arv, mitmikute puhul ka sündimise järjekord, sündimine elusalt või surnult, sünnikaal ja -pikkus, 1. ja 5. minuti Apgari indeks, sünnikoht, diagnoosid ning toimingud, lapse asukoht seitsme päeva vanuseks saamisel);
- sünnikaardi vastutava täitja andmed (nimi, arstikood ja allkiri).

Surnult sündinud või esimesel elunädalal surnud lapse puhul kantakse kaardile ka surmapõhjused koos diagnoosi koodidega. Sünnikaarti on kahel korral muudetud (1995. ja 1998. aastal), kuid need muutused ei ole olnud suured [Tellmann et al 2000].

Isikuandmete kaitse seaduse järgi kuuluvad mõned sünnikaardi andmed, mis kirjeldavad isiku etnilist päritolu, tema tervislikku seisundit või seksuaalelu, delikaatsete isikuandmete hulka. Need on: ema ja isa rahvus, ema varasemate abortilõppega raseduste arv, rasedusaegsed riskitegurid ja tüsistused, sünnituspuhused ja -järgsed diagnoosid, 1. ja 5. minuti Apgari indeks, lapse diagnoosid ning surmapõhjused.

Sünnikaardi koostamisel on võetud eeskujuks Soome meditsiinilise sünniregistri kaart. Ka Soomes ja teistes Põhjamaades põhineb perinataalstatistika meditsiiniliste sünniregistrite andmetel, kusjuures registrite ülesehitus nendes maades on sarnane Eesti omaga [Karro 1995; Gissler et al 2000].

Registrisse andmeesitajad on kõik Eestis sünnitusabi osutavad tervishoiuasutused. Andmed esitatakse sünnikaardil. Kaart täidetakse lapse seitsme päeva vanuseks saamisel või enne seda, kui laps kirjutatakse koju või sureb. Andmeesitajad toimetavad täidetud sünnikaardid kinnises ümbrikus kuude kaupa registrisse nii, et eelmise kuu kaardid laekuvad hiljemalt järgmise kuu 15. kuupäevaks (surnult sündinud või esimesel elunädalal surnud lapse sünni korral hiljemalt ülejärgmise kuu 15. kuupäevaks). Andmed saadetakse registrisse käsi- või tähitud postiga.

Registri töökorraldus. Meditsiiniline sünniregister alustas tööd 1992. aastal Eesti Meditsiinistatistika Büroo koosseisus. Alates 1998. aastast asub register Eksperimentaalse ja Kliinilise Meditsiini Instituudi juures.

Juba 1994. aastast on meditsiinilist sünniregistrit peetud koos abordinegistriga (praegune nimetus *raseduskatkestusandmekogu*). See tähendab, et samad töötajad teevad mõlema andmekogu tööd.

Eksperimentaalse ja Kliinilise Meditsiini Instituudi epidemioloogia ja biostatistika osakonna koosseisus on nende andmekogude pidamiseks moodustatud sünni- ja abordiregistri töörühm. Töörühmas on neli töötajat:

1 juhtivinsener — töörühma juhataja,

1 juhtivinsener — infosüsteemi administraator (andmetöötleja ülesannetes),

2 vanemtehnikut (andmesisestaja ülesannetes).

Meditsiinilise sünniregistri pidamiseks on arvestatud 2,6 ametikohta.

Registri pidamist finantseeritakse riigieelarvest Sotsiaalministeeriumi eelarve kaudu. Aastal 2000 eraldati meditsiinilise sünniregistri pidamiseks 350 000 krooni.

Registris andmed:

- kontrollitakse ja täpsustatakse;
- kodeeritakse (elukoht, rahvus ja amet);
- sisestatakse arvutisse;
- standardtöödeldakse ja töödeldakse päringute alusel;
- säilitatakse ja paberdokumendid arhiveeritakse.

Arvestades asjaolu, et meditsiinilises sünniregistris töödeldakse delikaatseid isikuandmeid, mis on seaduse järgi piiratud juurdepääsuga, on sünni- ja abordiregistri töörühmal lokaalne arvutivõrk ja oma serverarvuti, kus hoitakse ainult nende andmekogude andmeid. Kasutatav andmetöötlussüsteem hoiab andmeid serverarvuti andmebaasis krüpteeritud andmebaasiserveri *SQL Anywhere* abil. Süsteem on koostatud nii, et iga selle kasutaja pääseb juurde ainult temale kasutamiseks lubatud isikuandmetele (juurdepääs on reguleeritud paroolide süsteemiga).

01.11.2000 oli meditsiinilises sünniregistris andmeid enam kui 120 000 sünni kohta. Registrisse saabub ühes kuus ligikaudu 1000 sünnikaarti.

Meditsiiniline sünniregister vahendab rahvastikuregistrilt ka vastsündinute isikukoodide tabelleid sünnitusabi osutavatele haiglatele. Need tabelid on kahe poolega, millest üks jääb haiglasse ja teine saadetakse täidetult registrisse tagasi koos vastava kuu sünnikaartidega. Nii saab registris kontrollida, et kõikide laste kohta, kellele on antud isikukood, jõuab registrisse ka sünnikaart.

Kui laps sünnib väljaspool haiglat, siis kehtiva korra kohaselt peavad vanemad lapsele isikukoodi saamiseks pöörduma elukohajärgsesse sünnitusabi osutavasse haiglasse. Seal antakse isikukood, täidetakse sünnikaart ja saade-

takse registrisse. See kord tagab ka väljaspool haiglat sündinud laste registreerimise meditsiinilises sünniregistris.

Et tagada ja kontrollida kõikide laste registreerimist, tuleks meditsiinilise sünniregistri andmebaasi võrrelda ehk linkida Statistikaameti sünnilehtede ja surmatõendite andmebaasiga. Kuni 1999. aastani võrreldi meditsiinilises sünniregistris ning Statistikaameti sünnilehtede ja surmatõendite andmebaasis olevate laste kirjeid Sotsiaalministeeriumi ja Statistikaameti vahel sõlmitud koostöölepingu alusel. Uue koostöölepingu järgi (sõlmiti 14.12.1999) on eespool nimetatud andmeid lubatud võrrelda eraldi lepingu sõlmimisel pärast Andmekaitse Inspeksioonist isikuandmete riskasutuse loa saamist. 01.11.2000 seda luba ja lepingut veel ei olnud.

Andmete väljastamine. Meditsiinilisel sünniregistril on andmete töötlemisel ja väljastamisel kindlad ülesanded, mis on fikseeritud ka registri pidamiseks sõlmitud lepingus. Need ülesanded on:

- koostada standardtabelid eelmise aasta sünnikaartidel olevate andmete põhjal ja esitada need Sotsiaalministeeriumile, maakonnaarstidele ning sünnitusabi osutavatele haiglatele (tähtaeg 1. mai);
- esitada Statistikaametile andmed nende sünnilehtede andmebaasi täiendamiseks (aastal 2000 on selle ülesande täitmist takistanud see, et meditsiinilise sünniregistri ja Statistikaameti sünnilehtede andmebaasi isikuandmete riskasutuse luba puudub);
- esitada korrastatud sünnistatistika Sotsiaalministeeriumi meditsiinistatistika büroole nende aastaraamatu jaoks [Eesti tervishoiustatistika aastaraamat 1997. 1999];
- kord aastas koostada lühikogumik [Tellmann et al 2000; Tellmann, Karro 1998; Tellmann et al 1999];
- vastata päringutele.

Meditsiinilisele sünniregistrile laekub umbes kolm päringut kuus. Sagedasemad päringute esitajad on Sotsiaalministeeriumi ametnikud, tervishoiuasutuste töötajad, kõrgkoolide õppejõud ja üliõpilased, ajakirjanikud.

Andmete väljastamise üle otsustab isikuandmete korral registri vastutav töötaja Sotsiaalministeerium, muudel juhtudel volitatud töötaja.

Meditsiinilise sünniregistri andmete baasil on tehtud ja tehakse teadustööd. Registri andmeid on kasutatud eelkõige Eksperimentaalse ja Kliinilise Meditsiini Instituudi teadusprojektides, näiteks "Perinataalsete riskitegurite jao-

tumine ja raseduse lõpe Eestis”, “Reproduktiivset tervist mõjustavad tegurid ja nende seos raseduse lõpptulemusega” [Karro et al 1998]. Käesoleval ajal toetatakse Sotsiaalministeeriumi rahvatervise sihtprogrammist projekti “Eesti elanike seksuaal- ja reproduktiivkäitumine”, mis on Tartu Ülikooli ning Eksperimentaalse ja Kliinilise Meditsiini Instituudi ühisprojekt.

Registri juriidiline staatus. Kuigi meditsiiniline sünniregister töötab 1992. aastast, ei ole tema staatus veel juriidiliselt korrektselt vormistatud. Valminud on sotsiaalministri määruse eelnõu Eesti meditsiinilise sünniregistri asutamise, registri pidamise korra ja andmete arvestusvormi, st sünnikaardi kinnitamise kohta, kuid see määrus ei ole veel kinnitatud.

Kasutatud kirjandus

Eesti tervishoiustatistika aastaraamat 1997. 1999. Tallinn: Sotsiaalministeerium.

Gissler, M., Karro, H., Tellmann, A., Hemminki, E. 2000. *Births in Finland and Estonia from 1992 to 1996: convergent differences?* BJOG, 107(2).

Karro, H. 1995. Meditsiinilised sünniregistrid Põhjamaades ja Eestis. — Eesti Arst, nr 5.

Karro, H., Rahu, M., Gornoi, K., Baburin, A. 1998. *Estonian medical birth registry 1992–1994: association of risk factors with perinatal mortality.*— Eur. J. Obstet. Gynecol, No 2.

Tellmann, A., Karro, H., Serkina, V. 1999. Eesti Meditsiiniline Sünniregister 1992–1997. Eesti Abordiregister 1996–1997. Tallinn: Eksperimentaalse ja Kliinilise Meditsiini Instituut.

Tellmann, A., Karro, H., Serkina, V. 2000. Eesti Meditsiiniline Sünniregister 1992–1998. Eesti Abordiregister 1996–1998. Tallinn: Eksperimentaalse ja Kliinilise Meditsiini Instituut.

Tellmann, A., Karro, H. 1998. Sünnistatistika Eestis 1992–1996. Tallinn: Eesti Meditsiinistatistika Büroo.

RAHVALOENDUSE ANDMEKAITSE

(2000. aasta rahva ja eluruumide loendusega kogutud andmete kaitse andmete sisestamisel ja kodeerimisel)

Tiiu Vallner
Statistikaamet

Artiklis käsitletakse ainult loendusega kogutud andmete sisestamist ja kodeerimist, ei käsitleta loendusseadusest ja teistest loendusega seonduvatest õigusaktidest tulenevaid küsimusi ega anta hinnangut kasutatud turvameetmetele ja nende tõhususele.

Kogu rahvaloendus on pikaajaline protsess, mis vältab ligikaudu 10 aastat (1993 kuni 2002). Andmetöötluse see osa, mis hõlmab andmete sisestamist ja kodeerimist, kestis umbes viis kuud, maist oktoobrini aastal 2000.

Andmed töödeldi Statistikaameti ruumides (nn andmetöötlusruumid), mis olid selleks ette valmistatud (remonditud jne). Andmeid sisestasid ja kodeerisid Statistikaametisse ajutiselt tööle võetud 135 arvutioperaatorit. Töötati tööpäevadel kahes vahetuses kahe võimsa skanneriga, märgituvastamiseks oli 15 töökohta ning andmete kodeerimiseks 50 töökohta. Andmetöötlusega olid seotud peale juhtkonna veel turvajuht, vahetusejuhid, arhiivitöötajad, andme- ja süsteemihaldurid, meetodikud ja konsultandid, lepingupartnerite tugiisikud.

Loenduse ajal täideti ja toodi Statistikaametisse 6500 loendusportfelli, mis sisaldasid loendaja märkmikke, loendajate, nende juhendajate ja ringkonnajuhatajate täidetud vorme (nn saatelehed) ning loenduslehti (3,5 miljonit A4 formaadis mõlemalt küljelt täidetud lehte), seejuures isikulehti kaks korda 1,4 miljonit ja eluruumilehti 0,7 miljonit.

Loenduslehtede andmed sisaldasid ka delikaatseid isikuandmeid, sest küsiti isiku tervisliku seisundi ja/või usuliste tõekspidamiste, samuti töökohaga seotud andmeid.

Delikaatsete isikuandmete töötlemisel tuleb isikuandmete kaitse seaduse järgi registreerida isikuandmete töötlemine Andmekaitse Inspeksioonis (AKI) ja nende kaitsmiseks tuleb võtta tarvitusele organisatsioonilised ja tehnilised abinõud. Peavad olema täidetud isikuandmete kaitse seaduse

paragrahvi 12 lõikes 2 toodud 9 punkti (kohustust), mida AKI turvameetmete rakendamise kontrollimisel aluseks võtab.

Andmekaitse tagamiseks tuleb kindlaks teha turvavajadused, määratleda turvanõuded, kehtestada turvaklass.

Kogu andmeturve jagatakse neljaks haruks, mida klassifitseeritakse eraldi:

- terviklus (T);
- aegkriitilisus (K);
- hilineamise tagajärgede kaalukus (R);
- konfidentsiaalsus (S).

Kui turvaklass on määratud, siis igale turvaklassi komponendile seatakse vastavusse etalonmeetmed, mis omakorda jagunevad turvameetmete liigi (organisatsioonilised, infrastruktuuriga seotud, personaliga seotud, tark- ja riistvaraga seotud, kommunikatsioonid, talitluse katkematus) ja turvameetmete toime (ennetavad, avastavad ja taastemeetmed) järgi. Siis seatakse vastavusse olemasolevad ja vajalikud turvameetmed, hinnatakse turvameetmete rakendamise kulutusi ja tõhusust, kavandatakse tegevus jne. Kuivõrd turvameetmete liike on palju, neil on hulga parameetreid, nende mõjud kattuvad, siis turvameetmete kavandamiseks töötatakse välja eriprogramme. Niisugust metoodikat saab loodetavasti lähitulevikus kasutada. Statistikaametis veel konkreetseid turvaklasse ei määratud, etalonmetoodikat ja spetsiaalseid arvutiprogramme ei kasutatud.

Rahvaloenduse turvapoliitika väljatöötamist alustati Statistikaametis juba 1997. aastal koostöös AS Andmevaraga ja jätkati pärast prooviloendust koostöös AS Küberneetikaga. 1999. aastal valmis rahvaloenduse turvapoliitika uuendatud versioon AS Küberneetika infotehnoloogia osakonnas. Aruandes esitati ülevaade rahvaloenduse infosüsteemiga seonduvast varast, esitati rahvaloenduse infosüsteemi riskianalüüs, kus loetleti ohud loendusmaterjalile, digitaalandmetele, riistvarale ja seadmetele, tarkvarale, ruumidele, personalile, sideteenustele ning hinnati ohtude kriitilisust.

Suurimateks riskideks peeti:

- loendusmaterjali käideldavuse häireid, seejuures hävimist suures ulatuses;
- isikulehtede, samuti isikuandmekogu konfidentsiaalsuse kaitset — kogu andmebaasi sattumist "võõrastesse kättesse";
- infotehnoloogiliste vahendite, samuti andmete tõrkeid või hävimist, mis põhjustab loendusandmete töötlemise hilinemist, topelttööd jms.

Rahvaloenduse turvapolitiitika töötati välja standardi EVS ISO/IEC 13335 soovitude alusel. Turvapolitiitika sätestab üldised eesmärgid ja kavandatavad tegevusmallid järgmistes valdkondades:

- turvaeesmärgid ja -põhimõtted;
- turbe organisatsioon ja infrastruktuur;
- infotöötluse turva- ja riskianalüüs ning riskihalduse strateegia;
- info tundlikkus ja riskid;
- riist- ja tarkvara turve;
- sideturve;
- füüsiline turve;
- personaliturve;
- dokumentide ja andmekandjate turve;
- avariijärgse taaste ja talitluse pidevuse tagamine.

Turvapolitiitika uuendatud versioon andis üldistatud lähenemise ja nõuded rahvaloenduse turvameetmete väljatöötamiseks. Selle alusel vormistati rahvaloenduse andmetöötluse turvaeeskiri.

Rahvaloenduse turvaeeskirjas on eespool nimetatud valdkondade kaupa käsitletud kõiki andmetöötluses rakendatud turvameetmeid. Turvaeeskirja üldosas on viited rahvaloendusega seotud õigusaktidele ja Statistikaametis kehtivatele juhenditele ja korrale. Turvaeeskirja põhiosa on rahvaloenduse andmetöötluse turvet tagavad juhendid ja kord, mis vormistati turvaeeskirja lisadena.

Lisa 1. Infosüsteemi turvajuhend

Lisa 2. Füüsilised turvanõuded

Lisa 3. Infosüsteemi pääsu reguleerimise kord

Lisa 4. Tarkvara hanke ja kasutamise eeskiri

Lisa 5. Infovarade arvestuse kord

Lisa 6. Varundamise eeskiri

Lisa 7. Parasiittarkvara tõrje eeskiri

Lisa 8. Avarii- ja taasteprotseduuride reeglid

Lisa 9. Tegevuskava ohtlikus olukorras

Kinnitati rahvaloenduse teenistuse sisekorraeeskiri, turvanõuded rahvaloenduse teenistujate ja arvutioperaatorite ametijuhendites ning täiendati lepingupartnerite lepinguid turvanõuetega.

Turvalisuse tagamiseks rakendati järgmisi meetmeid:

— infotehnoloogilised turvameetmed — peavad kaitsma infot infotehnoloogilistest vahenditest tulenevate ohtude eest (turvaeeskirja lisad 1, 3 ja 7);

— füüsilised turvameetmed — peavad kaitsma infot keskkonnast ja kõrvalistelt isikutelt lähtuvate ohtude eest (turvaeeskirja lisa 2);

— organisatsioonilised turvameetmed — s.o töökorraldus, mis vähendab inimeste tegevusest lähtuvaid ohte (sisekorraeeskiri, turvaeeskirja lisad 3 kuni 9, töötajate ametijuhendid, lepingupartnerite spetsialistidele esitatavad turvanõuded).

Infotehnoloogilised (IT) turvameetmed:

— andmed töödeldi lokaalvõrgus, st ei olnud füüsilist ühendust kogu Statistikaameti võrgu või välisvõrguga, sh Internetiga;

— andmetöötamise rakendustarkvara töötati välja koostöös tunnustatud firma (AS *AboBase Systems*) spetsialistidega;

— andmetöötamise tarkvara põhineb *Oracle*'i andmebaasi ohjesüsteemil, millel on suured võimalused andmete kaitsmiseks, näiteks mitmesugused kasutajate õigused, juurdepääsud, paroolid, tegevuse jälgimine jne;

— töötati välja eriprogramm, nn turvapäevik, mis registreeris kõikide kasutajate toimingud;

— andmetöötamise tarkvara võimaldab andmeid krüpteerida, mida kasutatakse isikut identifitseerida võimaldavate andmete hoidmisel;

— süsteemsete logide salvestamine ja nende jälgimise võimalus;

— varukoopiaid tehti regulaarselt (iga päev) ja neid säilitati turvakappides varundamise eeskirja järgi; RL-2000 (*Oracle*'i baasi) andmebaasi varundamise varukoopiaid kontrolliti taastatavuse suhtes;

— IT vahendeid hooldasid regulaarselt lepingupartnerite AS *AboBase*, *Microlink* Süsteemide AS ja OÜ *ICL Invia* spetsialistid;

— infosüsteemis vormistati aruandeid iga vahetuse põhiliste tööloikude kohta;

— vormistati IT seadmete loetelu koos tehniliste andmetega (infovara arvestuse kord).

Füüsilised turvameetmed:

- remonditi andmetöötlusruumid;
- komplekteeriti arvutioperaatorite töökohad;
- paigaldati ventilatsiooni- ja kliimaseadmed, akendele rulood;
- paigaldati turvauksed ja -lukud, sh avariilukud, koodiga lukud, magnetkaardiga avatavad lukud;
- rakendati kaardisüsteem, mis andis mitmesuguseid õigusi uste avamiseks ja lukustamiseks, võimaldas jälgida liikumist;
- paigaldati ja ühendati signalisatsiooniseadmed, valvekaamerad, sissepääsu juurde pandi videotelefon;
- muretseti mitme võimsuse ja võimalusega tulekustutid, sh eriti suure võimsusega gaasikustutid dokumendihoidlatesse;
- dokumendihoidlatesse ja serveriruumi paigutati lisameetmetena signalisatsiooniantureid (vee-, temperatuuri-, suitsuandurid, klaasipurunemisandurid jms);
- muretseti tulekindlad turvakapid;
- elektritoite tõrgete vastu kasutati reservtoiteallikaid (eraldi UPS andmetöötlusruumide tehnika jaoks ja üldine diisलगeneraator);
- nähti ette varuruumid avariolukorra puhuks.

Organisatsioonilised turvameetmed:

- infoturbe haldus ja poliitika, sh õigusaktide väljatöötamine ja rakendamine;
- turbe järgimise kontroll — määrati ametisse rahvaloenduse turvajuht ja kinnitati tema ametijuhend, mille kohaselt turvajahi põhikohustus on turvanõuetest kinnipidamist kontrollida, turvapäevikud läbi vaadata ja analüüsida, esitada regulaarsed turvaaruanded peadirektorile;
- kontrollimisel olid tõhusaks toeks AKI inspektorid;
- personali puudutavad meetmed — arvutioperaatorite valik, koolitus (nii sisuline kui ka tehniline), ümberõpe, vahetusejuhtide rakendamine, töövaldkondadel põhinev aruandlus ja alluvus, ametijuhendites juurdepääsuõiguste sätestamine töövaldkonna järgi jms;
- taaste- ja avariolukordade, samuti mitmesuguste tõrgete vältimise, teatamise ja tegevuse kavandamine eriolukorra puhul (teatamise järjekord, telefoninumbritega varustatus jms);
- rahvaloenduse sisekorraeeskirja järgimine.

Rahvaloenduse teenistuse sisekorraeeskiri sätestas:

- töökorralduse kahe vahetusega töörežiimil;
- ruumidesse pääsu ja seal viibimise tingimused;
- kõikide andmetöötlusruumides viibijate registreerimise, nimesiltide kandmise kohustuse;
- külastajate piiramise, külastuste registreerimise;
- sündmuste registreerimise vahetusejuhi päevikus, samuti haldurite päevikus;
- puhkepauside reguleerimise;
- võtmete haldamise korra ja selle järgimise tingimused;
- vahetusejuhi — koordinaatori — ameti loomise (vahetusejuhi roll aitas oluliselt kaasa sisekorraeeskirja ja turvameetmete toimimisele).

Kõiki meetmeid rakendati selleks, et kaitsta vara, et kindlustada andmete käideldavus, terviklus ja konfidentsiaalsus, et tagada isikuandmete kaitse seaduse nõuete täitmine.

Statistikaamet andis rahvaloenduse turvapoliitika väljatöötamiseks ja turvameetmete rakendamiseks mitmesuguseid vahendeid: sõlmiti lepingud Eesti parimate spetsialistidega (AS Küberneetika, AS Andmevara); eraldati ja remonditi andmetöötlusruumid, kus rakendati tugevdatud organisatsioonilisi, füüsilisi ja infotehnoloogilisi turvameetmeid; andmetöötlusruumides rakendati range kodukord, millest kinnipidamist jälgisid vahetusejuhid ja rahvaloenduse turvajuht.

EESTI RAHVAARVUST JA SELLE HINDAMISEST

Ene-Margit Tiit
Statistikaamet

1. Eesti 2000. aasta rahvaloendus ja rahvaarv

2000. aasta rahvaloenduse esialgsete tulemuste avaldamisele järgnes šokk: osutus, et rahvaloendusel saadud rahvaarv oli 62 500 võrra väiksem kui Statistikaameti jooksva statistika käigus eelmise rahvaloenduse baasil arvu-
tatud rahvaarv (vt ridu 2 ja 3 tabelis 1).

Tabel 1

Jrk nr		Rahvaarv
1	2000. a loenduse faktiline rahvastik	1 370 500
2	2000. a loenduse alaline rahvastik	1 376 700
3	Alaline rahvastik 1.01.2000 (1989. a baasil)	1 439 197
4	Alaline rahvastik 1.01.1998 (1989. a baasil)	1 453 844
5	<i>Uus arvestuslik rahvaarv 1.01.1998</i>	<i>1 535 300</i>
6	1989. a loenduse faktiline rahvastik	1 572 900

Allikad: [Eesti statistika...2000; Katus 2000; Rahvaloenduse esialgsed...]

Šokki teravdas asjaolu, et mõned demograafid olid varem avaldanud artik-
leid, kus oletati, et faktiline rahvaarv on Statistikaameti rahvastikuarvestuse
hinnangust enam kui 80 000 võrra suurem (vt [Katus 2000; Katus 1998],
viimasest kirjutisest pärineb rida 5 tabelis 1). Vallandus küsimuste laviin. Kui
palju meid siiski on? Kas tõesti ei oska ega suuda statistikud inimesi korrali-
kult kokku lugeda? Sel eesmärgil eraldati ju suur hulk raha! Ajakirjandusse
sattus teravaid väljendusi: rahvaloendus on ebaõnnestunud, asjatu, sellele
kulutatud raha on raisatud.

Omaette küsimus seostub ka rahvastiku suure kaoga — 11 aasta jooksul on
Eesti rahvaarv vähenenud enam kui 200 000 inimese võrra (read 1 ja 6 tabe-
lis 1). Käesoleva artikli eesmärgiks on vaagida mõningaid selles seoses tek-
kinud küsimusi.

- Kas võimalikud vead rahvaloenduse tulemustes on viimase rahvaloenduse eripära või on neid esinenud ka varem ja mujal?
- Mis põhjustab vigu rahvaloendustes?
- Mida saab ja tuleb teha rahvaloenduse andmetega, mis sisaldavad vigu?

Neile ja teistele küsimustele vastamisel läheneme rahvaloendusele matemaatiku-statistiku seisukohast.

2. Mõõtmise ja mõõtmisviga statistiku pilguga

Mingi suuruse kindlakstegemiseks kasutatakse üldiselt *mõõtmist*, kusjuures mõõtmismeetodite hulka kuulub ka *loendamine*. Igasuguse mõõtmisega kaasneb *mõõtmisviga*. Kuigi mõõtmistäpsuse suurendamiseks tehakse eri valdkondades suuri pingutusi, jäävad juhuslikkuse seaduspärasustele alluvad mõõtmisvead siiski alles. Nende mõju võetakse arvesse ja/või kõrvaldatakse mõõtmistulemuste statistilise analüüsi käigus.

Mõõtmisvigade teooria arenes kõigepealt füüsikaliste mõõtmiste valdkonnas, kusjuures olulise panuse andsid siin astronoomilised mõõtmised, mida tehti instrumentide täpsuse piiiril, mistõttu vead olid eriti tuntavad. Kui lisaks füüsikalistele mõõtmistele hakati tegema ka psühholoogilisi mõõtmisi, selgus, et siin on olukord veelgi hullem, sest mõõtmistäpsust mõjustas palju segavaid tegureid ja sageli ei õnnestunud uuritavat nähtust üldse vahetult mõõta. Masiliste valikküsitluste kui sotsioloogilise uurimismeetodi arenedes selgus asjaolu, et kõiki uuringusse kavandatud isikuid ei õnnestu küsitluse käigus kätte saada. Sellega seoses sõnastati *ala-* ja *ülekaetuse* mõisted ning töötati välja hindamismeetodid, et kaetusprobleemidest tingitud mõõtmisvigu vältida ja kompenseerida.

Praegusajal on jõutud arusaamiseni, et

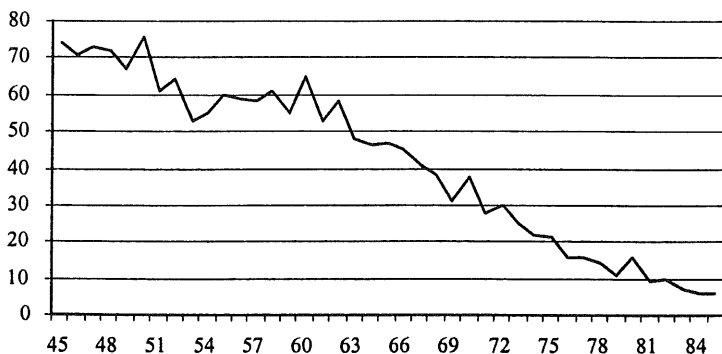
- *mõõtmisvead* on igasuguste mõõtmiste paratamatud saatjad, kusjuures vastavalt mõõtmistäpsuse suurenemisele mõõtmisvigade keskmine suurus ja selle tõttu ka mõju üldjuhul väheneb;
- mõõtmistulemus ei ole kunagi adekvaatne *tegeliku suurusega*, vaid mõõtmistulemuste põhjal leitakse mõõdetavale suurusele *hinnang*;
- tegelikule suurusele saadava *hinnangu täpsus* sõltub niihästi *mõõtmistäpsusest* kui ka *hindamismeetodist*;
- hinnang on korrektne vaid siis, kui ta koos uuritava suuruse hinnanguga annab teavet ka *hindamistäpsuse (hinnanguvea)* kohta.

3. Traditsioonilised rahvaloenduse vigade allikad

Küllaltki pikka aega ei peetud vajalikuks rahvaloenduste puhul mõõtmisvigu arvutada, sest oli ju tegemist kõikse uuringuga, mille puhul juhuslikku (valimiga seonduvat) viga ei tohiks olla. Ometi sai selgeks, et ka rahvaloenduse andmed ei ole veatud. Et rahvaloenduse andmeid kogutakse isikutevahelise suhtlemise käigus (rahvaloendaja ja loendatav), siis mõjustavad rahvaloenduse tulemusi niihästi mõõtja (loendaja) põhjustatud inimlikud vead, mõõdetava (loendatava) koostöövalmidus ja informeeritus, võimalikud küsitluslehe mittemõistetavusest tingitud süstemaatilised vead, eelteabe puudulikkusest tulenev alakaetus ja lõpuks veel mitmesugused juhuslikud vead, mis tekivad andmete kogumise, töötlemise, säilitamise ja analüüsi protsessis.

Näiteks varasemate rahvaloenduste tüüpiline viga on vanusehinnangud: üldiselt on teada, et nulliga lõppevate vanuste esinemissagedus on suurem nende naabervanuste esinemissagedusest, kusjuures see nähtus ilmneb eriti vanemate inimeste puhul. Ilmselt on siin põhjuseks tõsiasi, et vanemad inimesed ei tea endi vanust täpselt (täisaastates) ja annavad oma vanuse kohta *ümar-datud hinnangu* (vt joonis, kus on kasutatud Eesti 1922. aasta rahvaloenduse andmeid, vt ka [Statistiline Album...1925]).

45–85-aastaste naiste vanuskoosseis
1922. aasta rahvaloenduse andmetel



Hinnatavate suuruste puhul võivad rahvaloenduse tulemusi mõjustada ka elanikkonna hoiakud, üks selliseid on näiteks keelteoskus.

4. Loendamisel kasutatavate mõistete määratlemine

Vigu ja kooskõla puudumist rahvaloenduse tulemustes võivad põhjustada ka defineeritavad mõisted. Hoolimata definitsioonide aina suurendatud täpsusest, on nendes äärmiselt raske võtta arvesse elu paljutahulisust.

Üks selliseid raskeid küsimusi on leibkonna koosseis. Missugusest hetkest loetakse teise linna õppima siirdunud nooruk omaette leibkonda kuuluvaks? Esialgu elab ta tavaliselt vanemate toetusel ja külastab kodu tihti ja nii vanemad kui nooruk ise on arvamusel, et ta kuulub vanemate leibkonda. Muutused tulevad samm-sammult: nooruk hangib lisisissetulekuallikaid (õppelaen), soetab esialgu ajutise, hiljem püsiva eluaseme, külastused vanematekoju hõrenevad, leitakse elukaaslane ja nii ongi moodustunud omaette leibkond. Missugune on otsustav hetk? Kas aasta pärast vanematekodust lahkumist? Kuid koolivaheaegadel viibib nooruk ju taas vanemate juures, seega pole ta üle aasta eemal olnudki. Kas alates omaette eluruumi soetamisest? See sündmus pole aga enamasti kuidagi fikseeritud. Kas alates kooselust elukaaslasega? Ka kooselu võib olla muutuv, partneridki võivad vahetuda. Ka leibkonnaliikmetelt küsimine ei aita alati selgust tuua, sest võib juhtuda, et nooruk ja tema vanemad on eri arvamusel, nende arvamuste kõrvutamine aga on tehniliselt raske.

Niisama keeruline küsimus on (tegelik) perekonnaseis (vt tabel 2).

Tabel 2

Aasta	Abielus mehi	Abielus naisi	Erinevus
1959	242 768	259 487	16 719
1970	305 953	309 934	3 981
1979	342 329	342 121	-208
1989	362 825	364 633	1 808

Allikas: [Eesti rahvastik...1996]

Kui naised väärtustavad abieluseisust meestest märksa kõrgemalt, on rahvaloenduste tulemusena abielunaisi rohkem kui -mehi. Kui see üldine hoiak kaob, läheneb abielus meeste ja naiste arv teineteisele.

Tänapäeval on kõigis maades küsimuseks üha laiemalt leviva vabaabielu määratlus. Kui pikka kooselu võib juba vabaabieluks pidada? Kas vabaabielu eeldab abikaasade ühist eluruumi? Skandinaavia maade sotsioloogid uurivad juba kümnekonna aasta jooksul intensiivselt ka lahuselavate abikaa-

sadega perekonda. (Tegemist võib olla nii seadusliku kui ka registreerimata partnerlusega.)

Isegi laste arv peres pole tänapäeval üheselt määratud suurus(!). Lahutatud vanemad lepidvad sageli kokku selles, et laps on osa aega ühe, osa aega teise vanema peres. Kumbki leibkond peab last oma liikmeks. Kas tuleks siis lugeda laps poolenisti ühte, poolenisti teise peresse kuuluvaks?

Määratluste mitmeti mõistetavus mõjustab mitte ainult loendatavate näitajate (vanus, perekonnaseis, leibkondlik kuuluvus, elukoht) väärtusi, vaid ka rahvaarvu. Näiteks noor inimene läheb välismaale õppima kindla kavatsusega tagasi tulla. Õpingute edenedes tekib töövõimalus ja kodumaale naasmine nihkub tulevikku. Missugusest hetkest peale seda inimest enam ei loeta Eesti alalise rahvastiku hulka? Mida öelda meremeeste kohta, kes alustasid karjääri laeval, mille kodusadam on Eestis, kuid vahetasid laeva, ja kellel on igas sadamas kallim, kuid püsivat peret ei kusagil?

Tegelikult on peaaegu igale erijuhule täpselt fikseeritud reeglid. Enamasti on need teatavas mõttes bürokraatlikud, st ei jäli mitte niivõrd inimese enese liikumist, vaid temaga seotud dokumente. Enamikus riikides eeldatakse, et igal inimesel on üks püsielukoht (tema enese valikul), mille ta on riiklikult registreerinud (mis paikneb kas isiku- või aadressiregistris) ja mille kaudu ta suhtleb riigiga. Loomulikult võib inimesel lisaks püsielukohale olla veel teisi elukohti (näiteks suvekodu). *Püsielukoht on kogu leibkonnal ühine ja see ongi leibkonna määratluse üks alustugesid* (mõnes riigis koguni ainus leibkonda määrav kriteerium).

Loendades kokku need inimesed, kelle püsielukoht on mingil territooriumil, saadakse selle territooriumi (riigi, linna) *alaline rahvastik*. Kui riigis on olemas toimiv (pidevalt aktualiseeritav) rahvastikuregister, siis on alalise rahvastiku arvu, soo-vanusstruktuuri, geograafilist jaotust jne võimalik arvatada registri põhjal, st *teha rahvaloendus rahvastikuregistri põhjal* (nagu seda tehakse Soomes ja Taanis). Muidugi sõltub sellise loenduse alusel saadud teabe maht rahvastikuregistris ja sellega seotud registrites sisalduva teabe mahust. Samal põhimõttel käib ka rahvastiku jooksev arvestus: rahvaloenduse ajal kogutud rahvastikuandmeid täiendatakse (arvutatakse ümber) jooksvalt registreeritud sündmusstatistika põhjal.

Probleemiks on siin see, kuidas võtta arvesse inimesi, kellel ei ole püsieluruumi? Lisaks kodututele võivad ilma püsieluruumita olla ka näiteks vana-inimene, kes elab kord ühe, kord teise sugulase juures, mingitel asjaoludel

kodu kaotanud ja/või senisest elukohast lahkunud inimesed, kellel on üksnes ajutine elukoht jne. Neid asjaolusid arvestades toimib registripõhine loendus üksnes niisugustes riikides, kus ühelt poolt on olemas kvaliteetne register, kuid teiselt poolt on ka valdaval osal elanikkonnast olemas püsieluruum, st neis maades, kus n-õ kuuse all, telgis või paadi all elada pole kliimaolude ja sotsiaalsete hoiakute tõttu mõeldav.

5. Rahvaloenduse kaetus

Kui soovitakse teada saada, kui palju inimesi tegelikult *loenduse kriitilisel hetkel* riigi territooriumil viibis, siis tuleb korraldada *rahvaloendus*, kus küsitletakse nimeliselt ja üksahaaval kõiki isikuid, märgitakse üles ka kõik rongides, laevades, hotellides, lennujaamades viibivad inimesed (välja arvatud välisriikide diplomaadid ja sõjaväelased ning nende pereliikmed). Saadakse *faktiline rahvaarv*, mis sisaldab peale alalise rahvastiku ka parajasti riigis viibivaid külalisi, kuid ei sisalda ajutiselt välismaal viibivaid inimesi. Fakti- list rahvaarvu ei saa määrata registri põhjal. Inimeste küsitlemise korral on aga alati kaks ohtu:

- osa inimesi jääb küsitlemisel vahele ja
- osa satub küsitlusse korduvalt.

Korduvalt küsitlusse sattumine pole eriti suur mure. Selle vea saab kõrvaldada, kui kõik loendatud isikud üle kontrollida. See töö on tülikas siis, kui isikuid tuleb identifitseerida ainult nime ja sünniaja kaudu, kuid isikukoodide abil on see samm suhteliselt lihtne.

Rahvaloenduste suurim probleem on aga *alakaetus*, st loendusest väljajäänud isikud. Väljajäämise põhjusi on mitmesuguseid:

- rahvaloendajad ei leia üles selle isiku elukohta (eelinformatsioon elukohatade kohta on puudulik või isik elab mitte-eluruumis, nt suvilas, pooleli ehitises vm);
- isikut ei saada kätte, sest ta viibib loendajate külastuse ajal kodust eemal ja keegi teine ei esita tema andmeid;
- isik hoidub teadlikult või keeldub loendamisest;
- loendaja töötab hooletult ja jätab osa isikuid loendamata.

Kõige tõenäosemalt jäävad loendusest välja ajutiselt oma põhielukohast eemal, sh välismaal viibivad isikud.

Et põhiline rahvaloenduse viga on alakaetus, siis jäeldub siit, et rahvaloenduse andmed enamasti kipuvad *faktilist rahvaarvu alahindama*.

6. Rahvaloenduse alakaetuse globaalsed põhjused

Hoolimata tehnika arengust (mis küll oluliselt lihtsustab rahvaloenduse andmete töötlemist ja analüüsi) ei ole viimasel ajal rahvaloenduste täpsus suurenenud. Põhjusi on siin mitu, loetleme olulisemad neist.

- Rahvastiku suurenev mobiilsus, mida stimuleerib viisarežiimi kaotamine paljude riikide vahel, vaba tööjõu liikumine, õpilaste ja üliõpilaste vahetusprogrammid jne.
- Üldine liberaliseerumine, millega seostub ka elukoha registreerimise reeglite nõrgenemine (asjaolu, mis tugevasti mõjub Eestis, kus elukoha registreerimine ei olnud rahvaloenduse toimumise ajal ja sellele eelnenud 10 aasta jooksul kohustuslik).
- Üksikisikute suurem enesekesksus, mille tulemusena ei soovita oma andmeid esitada ja hoidutakse pigem loendusest eemale. Seda võimendavad arutlused isikuandmete kaitse üle meedias.
- Rahvaloenduse seostamine poliitiliste sümpaatiatega ja lojaalsusega riigile ja/või valitsevale parteile.

7. Kas rahvaloenduse andmeid saaks ja tuleks *parandada*?

Igasuguste mõõtmiste puhul on mõõtmistulemuste *tagantjärele parandamine* üpriski problemaatiline protseduur, mida üldjuhul tuleks vältida. *Parandada tohib ilmseid vigu üksnes siis, kui õige väärtus on teada*. Kui inimese sünniaastaks on märgitud 1199, aga isikukoodilt jäeldub, et sünniaasta peaks olema 1919, siis on selge, et eksitus on sünniaasta märkimisel. Samas ei ole suur osa andmeid dubleeritud ja seega saab loogiliste kontrollimeetoditega jämedaid vigu küll avastada, aga enamasti pole neid võimalik parandada. Vigane väärtus lihtsalt eemaldatakse.

Mida teha aga rahvaloenduse alakaetusega?

On selge, et esimene ja kõige tähtsam ülesanne on teha selgeks, kui suur osa kahe rahvastikuhinnangu vahest on tingitud tegelikust, kuid seni arvestamata jäänud rahvaarvu muutusest (sh ka eelmise rahvaloenduse veast), ja kui suur osa rahvaloenduse alakaetusest.

Missugused andmed saavad statistikuid selle juures aidata?
Kõigepealt andmed, mis selgitavad rahvaarvu vähenemist, sh

- mitteametlik eelmise rahvaloenduse täpsushinnang;
- hinnangud militaarrände arvestuste ebatäpsuse kohta;
- naaberriikide registrite andmed sisserännanud (ka ajutiselt kohalviibivate) elanike kohta;
- oletatavad vead jooksvas rahvastikustatistikas (arvestamata jäänud surmajuhud jne).

Teiseks on andmed, mis võimaldavad hinnata alakaetuse suurust.

- Rahvaloenduse järelkontrolli andmete analüüs (valikuuring), et anda hinnang rahvaloenduse kaetusele (soo-vanuserühmade kaupa).
- Võrdlus haigekassa andmebaasiga, mis võimaldab identifitseerida haigekassas kindlustatud isikuid, keda rahvaloenduse käigus ei loendatud, ja lisada nende andmed olulisematesse jaotustesse (soo-vanusjaotus). Haigekassa andmebaasi tugevaks küljeks on selle pidev aktualiseeritus (suurem osa inimesi külastab arsti vähemalt kord aastas, mistõttu on andmed suhteliselt värsked). Selle baasi puuduseks on, et ta ei kata kogu elanikkonda (mittekindlustatud), samuti see, et haigekassa geograafiline paiknevus ei ole üksüheselt seotud isiku elukohaga.
- Võrdlus rahvastikuregistri andmebaasiga: saab identifitseerida rahvastikuregistris paiknevaid isikuid, keda ei loendatud. Siin on probleemiks, et rahvastikuregistri andmete aktualiseerituse tase on ilmselt madalam kui haigekassa andmetel, eriti vananenud on elukohti puudutav teave. Võib oletada ülekaetust, sest registris leidub kirjeid Eestist teistesse riikidesse lahkunud inimeste kohta, keda ei ole ametlikult välja registreeritud.
- Võrdlus sotsiaalregistriga, kus on olemas andmed mitme rahvastikurühma (pensionärid, lapsed jt) kohta.
- Võrdlus maksuameti andmebaasiga. Seda võrdlusallikat kasutatakse paljudes riikides, Eestis pole seni veel kasutatud. Loomulikult ei paku võrdlusel rahvaloenduse andmetega huvi isikute tulud, vaid lihtsalt isikute olemasolu. Võib oletada, et teatav hulk tööelistest tervetest ini-

mestest puudub haigekassa registrist, kuid on olemas maksuameti registris.

- Lisateavet annavad pikka aega korraldatud valikuuringud (tööjõu-uuring, leibkonna sissetuleku ja kulutuste uuring, eluolu-uuring), mis kõik baseeruvad rahvastikuregistri andmebaasile ja mille tulemusi saab kasutada hindamaks rahvastiku soo-vanusjaotust ning selle erinevust rahvastikuregistri ja jooksva rahvastikuarvestuse jaotustest.

Rahvaloenduse andmete ja muude registriandmete ühitamine saab olla edukas üksnes järgmistel tingimustel.

- Kõik asjaosalised, sh ka Andmekaitse Inspektsioon, teadvustavad enesele, et tegemist on riigi seisukohalt olulise, mitte üksikute ametkondade või üksikisikute huvides tegevusega.
- Eesmärgiks ei ole mitte rahvaloenduse andmete parandamine või täpsustamine, vaid Eesti elanikkonna soo-vanuselise struktuuri ja elukohtade jaotuse hindamine kõigi olemasolevate andmekogude baasil.
- Nimetatud samme saab edukalt teha identifitseeritud isikuandmete abil. Selleks tuleb kehtestada isikuandmete kaitse reeglid. Ilmselt tegeleb andmetega suhteliselt väike, vajaliku ettevalmistuse saanud ja kõrge vastutustundega isikute ring. Pärast soo-vanuse-elukoha-struktuuri hinnangute leidmist (mis peavad olema kooskõlas rahvastiku vähenemise andmetega) ei ole vaja identifitseeritud andmeid kasutada. Leitud hinnangud muutuvad üldkasutatavateks rahvastikuhinnanguteks ja need tuleb üldistel alustel publitseerida.
- Kasutades leitud soo-vanuse ja elukohtade struktuuri jaoks optimaalseid hinnanguid, saab valikuuringute teooria meetoditega laiendada kogu elanikkonnale ka kõigi teiste rahvaloenduses küsitud tunnuste andmeid.

Pärast niisugust lisaanalüüsi, mis hinnalt on palju kordi odavam rahvaloendusest, on võimalik kinnitada, et Eesti rahvastiku koosseis on hinnatud nii täpselt, kui see üldse võimalik oli, ja lisaks muule on teada ka hinnangute täpsusehinnangud.

Kasutatud kirjandus

Eesti rahvastik rahvaloenduste andmetel. 1996. Tallinn: Statistikaamet.

Eesti statistika aastaraamat. 2000. *Statistical Yearbook of Estonia*. 2000. Tallinn: Statistikaamet.

Katus, K. 2000. Rahvastiku areng. — Eesti uue aastatuhande lävel. Tallinn: TEA Kirjastus.

Katus, K. 1998. Rahvastikuareng. — Sotsiaaltrendid. Tallinn: Statistikaamet.

Rahvaloenduse esialgsed tulemused...

http://www.stat.ee/wwwstat/content/E_U_PR_0074/1.html

Statistiline Album. 1925. Vihik I: Maa ja Rahvas. Tallinn: Riigi Statistika Keskbüroo.

RAHVASTIKUREGISTRI KASUTAMISEST EESTI LEIBKONNA- JA TÖÖJÕU-UURINGUS

Anneli Kukk, Kaja Sõstra, Imbi Traat
Statistikaamet, Tartu Ülikool

Sissejuhatus

Statistikaameti tellimisel ja Maailmapanga finantseerimisel korraldati aastatel 1998–1999 Eesti leibkonna eelarve uuringu (LEU) ümberdisainimine. Neli töögruppi ja töö hilisemas faasis lisandunud viies töögrupp analüüsisid uuringu eri etappe ja töötasid välja täiustatud meetodika, mida rakendatakse uuringus 2000. aasta jaanuarist. Samast ajast läks uuele meetodikale üle ka Eesti tööjõu-uuring, kasutades valimi võtmise teatud etappides leibkonna-uuringuga sarnaseid ideid.

Käesoleva artikli autorid kuulusid valimi töögruppi, kus töötati valiku- ja hindamismeetodite täiustamisega [Traat 1999]. Et leibkondade juhuslik valik tehakse rahvastiku andmebaasist ehk rahvastikuregistrist (edaspidi *register*) (registrit peab AS Andmevara) ja ümberdisainitud uuringus toetutakse veelgi tugevamalt registris sisalduvale infole, siis omandasid registri teatud näitajate kvaliteediküsimused erilise tähtsuse. Selles artiklis tuleb juttu nendest registriga seonduvatest aspektidest, mis on tähtsad juhusliku valiku tegemisel ja nihketa hinnangute arvutamisel. Seejärel kirjeldatakse uuele meetodikale üleminekuga saavutatud tulemusi nii leibkonna eelarve uuringus kui ka tööjõu-uuringus.

1. Rahvastikuregister

Rahvastikuregister kujutab endast andmebaasi, milles on iga Eesti kodaniku ja elamisloaga isiku kohta isikuandmetega kirje. Kirje sisaldab isiku nime, isikukoodi, aadressi (mitme komponendiga) jm infot.

1.1. Leibkondade valik registrist

Leibkondade juhuslik valik registrist tehakse süstemaatilise valikumeetodi järgi. Tulemuseks on kirjete hulk isikunimedega, aadresside jm infoga. Klassikaline meetod on, et iga valitud isik toob valimisse oma leibkonna. Tagajär-

jeks on suuremate leibkondade üleesindatus valimis (võrreldes leibkondade jaotusega üldkogumis). Leibkonna kaasamistõenäosuseks (lihtsustatult) on

$$\pi_i = \frac{p_i n}{N}, \quad (1)$$

kus n on valitavate kirjete arv, N kõigi kirjete arv registris ja p_i valitud kirje-ga seonduva leibkonna liikmete arv. Nihketa hinnangute saamiseks niisugu-sest valimist on kaks levinumat meetodit, mida kasutatakse ka Statistikaameti uuringutes:

- tagasikaalumine hindamisetapil (LEU), kaaluks on kaasamistõenäosuse pöördväärtus;
- osa leibkondade valimist väljaviskamine, nii et üleesindatus kaob (ETU), kaalud võrdsustuvad.

Mõlemal juhul on äärmiselt tähtis teada **valitud** leibkonna täpne suurus p_i . Probleemiks osutus see mõlemas uuringus: tööjõu-uuringus seepärast, et seda oli vaja teada enne leibkonnaga kontakteerumist (osa leibkondade väljavis-kamiseks); leibkonnauuringus seepärast, et aadressil elas sageli hoopis teine leibkond (14 protsenti juhtudest). Ajapuudusel polnud võimalik valitud leib-konda jälitada ja valimisse kaasati valikusse sattunud aadressil elav leibkond. Selle leibkonna suurus kasutati valemis (1) tema kaasamistõenäosuse ja see-ga ka kaalu määramiseks, mis oli meetodiliselt ebakorrekne. Kuigi väikesed ebakorrektsused kaaludes tulemustele erilist mõju ei avaldanud, otsustati siin meetodikat parandada. Tekkis idee rakendada kogu uuringus **aadressvalikut**, kus registrist valitud kirje aadressosa määrab leibkonna (pole tähtis, missu-gune leibkond elab sellel aadressil). Kaasamistõenäosuse määrab sellisel ju-hul aadressi esinemissagedus registris.

1.2. Registri analüüs

Üleminekul aadressvalikule uuriti, kas registri kvaliteet on selleks piisav:

- kas aadressid on unifitseeritud;
- kas need võimaldavad määrata leibkonna;
- kas on võimalik leida aadresside esinemissagedusi registris;
- kuidas on olukord aadresside üle- ja alakaetusega.

Kui esimesele ja kolmandale küsimusele sai vastata jaatavalt, siis teise küsimuse puhul selgus, et täielik aadressvalik pole võimalik. Maapiirkondades on sageli aadressiks vaid küla nimi ja seega ei saa aadressi järgi ühte konkreetset leibkonda kindlaks määrata. Küll aga saab külas otsida üles valitud isiku ja lülitada uuringusse tema leibkonna. Osutus, et registris oli 89 protsenti (98-08-19) niisuguseid kirjeid, mis võimaldasid identifitseerida leibkonna (täpne aadressosa oli olemas või oli küla täpsuseni), 5 protsendi kirjete aadressosa oli täiesti tühi ja ülejäänud olid nn väljakirjutusega kirjed.

Eraldi uurimise alla võeti väljakirjutusega kirjed (kaasa arvatud ka nn ebasaldatava aadressiga kirjed), sest neid Andmevaras senini valimi võtmiseks ei kasutatud. Kirjete jaotus maakonniti registri osas 15-aastased ja vanemad isikud (RR15+) on toodud tabelis 1.

Tabel 1. Väljakirjutuste sagedus, RR15+ (30.09.1999)

Maakond	Isikute arv	Väljakirjutused	Sissekirjutused	Väljakirjutuste %
Tallinn	361 822	32 928	328 894	9,1
Harju	102 717	4 133	98 584	4,0
Hiiumaa	9 195	120	9 075	1,3
Ida-Viru	172 769	4 195	168 571	2,4
Jõgeva	32 972	791	32 181	2,4
Järva	34 180	1 230	32 926	3,6
Lääne	26 268	857	25 410	3,3
Lääne-Viru	61 327	3 966	57 335	6,5
Põlva	28 439	677	27 762	2,4
Pärnu	79 869	3 154	76 715	3,9
Rapla	31 762	1 010	30 748	3,2
Saaremaa	31 850	753	31 097	2,4
Tartu	124 313	7 321	116 992	5,9
Valga	31 483	692	30 789	2,2
Viljandi	50 911	1 805	49 106	3,5
Võru	35 630	1 539	34 090	4,3
KOKKU	1 215 507	65 171	1 150 275	5,4

Nagu näha, 5,4 protsenti kirjetest ei kasutata valimi võtmisel. Need inimesed saavad valimisse sattuda ainult mõne pereliikme kaudu, kes on registris sissekirjutusega. Analüüs näitas, et registris oli 15 839 tühja aadressi, kuhu polnud ühtegi isikut sisse kirjutatud. Et see moodustas 2,5 protsenti Eesti elamufondist (622 100, Eesti Statistika aastaraamat. 1997. *Statistical Yearbook of Estonia*, lk 263) ja suur osa nendest olid tõenäoliselt mitteleluumid, siis nende kirjete väljajätmine valiku tegemisel olulist alakaetust ei põhjusta.

Tähtis on, missugune on aadresside esinemissagedus registris ja kas neid on võimalik leida. Osutus, et on ja seda iseloomustab tabel 2. Aadressi esinemissagedus registris näitab sellele aadressile sissekirjutatute arvu. Välja on jäetud kirjed, mille aadressosa oli tühi või millel oli väljakirjutuskood.

Tabel 2. Aadresside esinemissagedus (98-08-13)

Sagedus	Isikute arv	Aadresside arv	%	Sagedus	Isikute arv	Aadresside arv	%
1	180 398	180 398	35,62	11	2 398	218	0,04
2	267 912	133 956	26,45	12	1 680	140	0,03
3	271 536	90 512	17,87	13	1 534	118	0,02
4	243 144	60 786	12,00	14	1 288	92	0,02
5	116 265	23 253	4,59	15	885	59	0,01
6	51 612	8 602	1,70	16	1 056	66	0,01
7	23 086	3 298	0,65	17	1 173	69	0,01
8	11 712	1 464	0,29	18	900	50	0,01
9	5 697	633	0,12	19	1 007	53	0,01
10	3 390	339	0,07	20 ->	173 667	2 362	0,47
KOKKU					1 360 340	506 468	100,00

35,62 protsenti aadressidest on ühe elanikuga (sissekirjutusega), 26,45 protsenti kahe elanikuga jne. Aadresside jaotus elanike arvu järgi langes hästi kokku leibkonnauuringust saadud hinnangutega leibkonna suuruse kohta, mis kinnitas veendumust, et üleminek uuele kaalusüsteemile (rajatud registri sissekirjutuste arvule) ei põhjusta hüppeid hinnangutes.

Põhjalikult uuriti registri alakaetust aadresside osas. On teada, et mitte kõik elukohad ei esine aadressina registris. Aadressivaliku korral jäävad nendes elukohtades elavad leibkonnad valimist välja. Võrreldes rahvaloenduse prooviloendusel (20.–29.03.1998) leitud elukohti registriga, hinnati registri alakaetuseks aadresside osas 5–6 protsenti [Traat 1999]. Sama analüüs Viljandi maakonnas andis tööjõu-uuringu aadresside alakaetuseks 3 protsenti.

1.3. Leibkonna valik ümberdisainitud uuringus

Leibkonna lõplik valik tehakse kas aadressreegli või isikureegli järgi. Aadressreegli järgi võetakse valimisse aadressil elav leibkond (kui on mitu leibkonda, siis kõik). Isikureegli kohaselt otsitakse üles valimisse tulnud isik ja uuringusse kaasatakse ainult selle isiku leibkond.

Statistikaamet saab AS Andmevaralt juhuslikult valitud kirjed (täpsemalt vt [Traat 1999]) ja koos iga kirjega sellele aadressile registreeritud isikute arvu. Kui see arv on väiksem kui 8, rakendatakse selle kirje korral aadressreeglit, muidu isikureeglit. Osutus, et nii klassifitseerub 14 protsenti kirjetest isikureegli alla. Kaasamistõenäosuste arvutamisel valemiga (1) on suurus p_i võrdne aadressile sissekirjutatud isikute arvuga registris aadressreegli leibkonna jaoks ja on võrdne leibkonna suurusega isikureegli leibkonna jaoks. Sedasama skeemi kasutatakse modifitseeritud kujul tööjõu-uuringus, mida kirjeldatakse allpool.

2. Ümberdisainimise mõju LEU hinnangutele

Punktis 1 kirjeldati rahvastikuregistriga seotud aspekte juhusliku valiku tegemisel ja hinnangute arvutamisel. Hindamisetapil rakendati leibkonnauuringus veel teisi uudseid meetodeid. Lisaks muudetud kihistamisele ja järelkihistamisele lülitati juurde hinnangute kalibreerimine.

Kalibreerimise mõju hindamiseks vaadeldi hinnanguid enne (s.o pärast järelkihistamist) ja pärast kalibreerimist. Selgus, et kalibreerimine suurendas nii netosissetulekut kui ka väljaminekut 0,8 protsenti. Üksikute tulu- ja kululiikide muutused olid muidugi suuremad. Kalibreerimise mõju sissetuleku liiki-dele on toodud tabelis 3.

Tabel 3. Leibkonnaliikme kuukeskmise netosissetulek kroonides, I kvartal 2000

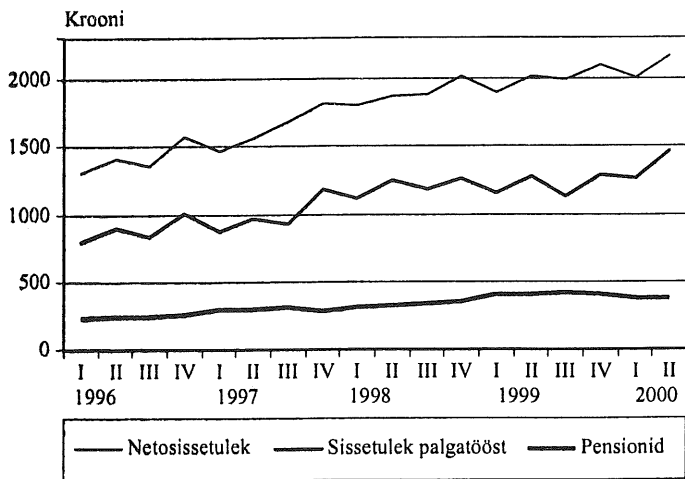
	Järel- kihistamine	Kalibree- rimine	Kalibreeri- mise mõju, %
Netosissetulek	1 990,4	2 006,7	0,8
palgatööst	1 265,9	1 305,1	3,1
individuaalsest teisest tegevusest	101,2	102,8	1,6
siirded	591,1	566,1	-4,2
pensionid	407,8	380,8	-6,6
muu sissetulek	32,3	32,7	1,5

Teoorias eeldatud tulemus leidis ka praktikas kinnitust, st kõige rohkem suurenes sissetulek palgatööst ja vähenes sissetulek pensionidest. Saadud hinnangute usaldusväärsuse kinnituseks võrreldi LEU hinnanguid sotsiaalkindlustusametist saadud arvudega. Ilmnes, et sotsiaalkindlustusameti I kvartali andmetel oli ühe Eesti elaniku (elanike arv rahvastikustatistika andmetel) kohta pensionideks makstav summa ühes kuus 369,7 krooni. Leibkonna eelarve uuringu andmete põhjal leitud hinnang oli 370 krooni, mis ei sisalda välisriikide makstavaid pensione (erinevalt tabelis 3 toodud arvust). Kalibreerimiseta hinnang (kaalutud üksnes järelkihistamisega) andis tulemuseks 395,4 krooni, mis on 7 protsenti kõrgem haldusandmete alusel saadust. 1999. aasta IV kvartalis oli ühe inimese kohta väljamakstav pensionisumma 370,4 krooni, sama perioodi uuringu hinnang oli 402,1 krooni, mis on rohkem kui 8 protsenti kõrgem. Võrreldes väljamakstud lastetoetust inimese kohta uuringu hinnanguga, oli 1999. aasta IV kvartalis uuringu hinnang 2,9 protsenti ja 2000. aasta I kvartalis 1,3 protsenti madalam haldusallikate infost.

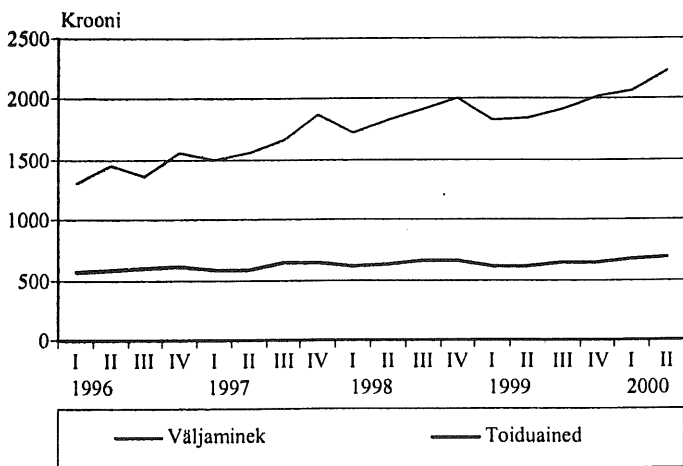
Väljaminekute poolelt suurendas kalibreerimine väljaspool kodu söömisele tehtavaid kulutusi 9 protsenti, kulutusi alkoholile ja tubakale 6 protsenti ja kulutusi toidule 4 protsenti.

Vaadeldes sissetulekute ja kulutuste muutust ajas (joonised 1 ja 2), ei ole ümberdisainimise mõju neile märgata. Väljaminekute 2000. aasta II kvartali tavapärasest suurem tõus oli peamiselt tingitud transpordikulude suurenemisest (135 krooni I kvartalis ja 228 krooni II kvartalis), mille omakorda on põhjustanud kütusehindade järjepidev kasv.

**Joonis 1. Leibkonnaliikme kuukeskmise sissetulek,
I kvartal 1996 – II kvartal 2000**



**Joonis 2. Leibkonnaliikme kuukeskmise väljaminek,
I kvartal 1996 – II kvartal 2000**



3. Eesti tööjõu-uuringust

Esimene Eesti tööjõu-uuringu (ETU) küsitlus korraldati jaanuarist aprillini 1995. Valimi aluseks oli 1989. aasta rahvaloenduse andmebaas, kust võeti juhuvalim 15–74-aastaste isikute hulgast. Küsitleti valimisse sattunud isikuid.

Aastatel 1997–1999 oli ETU küsitlus II kvartalis. Valimi alus oli rahvastiku arvestuse andmebaas, kust võeti juhuvalim 15–74-aastastest isikutest. Välitööde käigus intervjueriti kõiki valimisse sattunud isiku leibkonna tööelisi (15–74-aastasi) liikmeid.

2000. aasta jaanuaris alustati pidevuuringuga, kus valim on ühtlaselt jaotatud 52 uuringunädalale. Küsitlustulemused avaldatakse kord kvartalis.

3.1. Valimi moodustamine 1997–1999

Valim moodustati mitmes etapis. Esiteks telliti rahvastiku andmebaasist juhuslikult valitud 15–74-aastaste isikute nimekiri.

Teiseks täpsustasid küsitlusvõrgu maakondade koordinaatorid nimekirja sattunud isikute tööeliste leibkonnaliikmete arvu. Täpsustamiseks kasutati maal vallavalitsuste abi, linnades sageli rahvastikuregistri andmeid, mis ei olnud eriti täpsed. Sellega tehti juba valimi moodustamise teisel etapil palju vigu. Hilisemal võrdlemisel selgus, et täpsustatud liikmete arv langes küsitluse käigus selgunud tegeliku arvuga kokku vaid veidi üle pooltel juhtudest.

Kolmandaks moodustati saadud andmete põhjal lõplik valim, kuhu jäid kõik ühe tööelise liikmega leibkonnad, pooled kahe tööelise liikmega leibkondadest, kolmandik kolme tööelise liikmega leibkondadest jne. Välitööde käigus otsiti üles valimisse sattunud isik ja küsitleti temaga samas leibkonnas elavaid tööelisi isikuid.

Eelkirjeldatud protseduur tagas leibkondade ja isikute võrdse valimisse sattumise tõenäosuse. Kõiki leibkondi käsitleti ühtse eeskirja järgi, asendusi ei lubatud. Välitööde käigus tulid kasuks eelteadmised leibkonna suuruse kohta: sai paremini plaanida küsitlusmaterjali vajadust, ette oli teada küsitluse keel.

Meetodi puuduseks olid suur ebatäpsus valimi moodustamise teisel etapil, mis tekitas nihkeid leibkondade kaasamistõenäosuses; tekkis raskusi nooremate ja liikuvamate inimeste leidmisel. Hilisem küsitlute soo-vanusjaotuse analüüs näitas 20–30-aastaste alakatust. Lihtsam oli olukord siis, kui mujal

iseseisvat elu alustanud isik oli registreeritud oma vanemate juurde, kes enamasti teadsid oma lapse uut aadressi. Tihti peale aga oli korter müüdud ja uued omanikud ei teadnud endisest perest midagi.

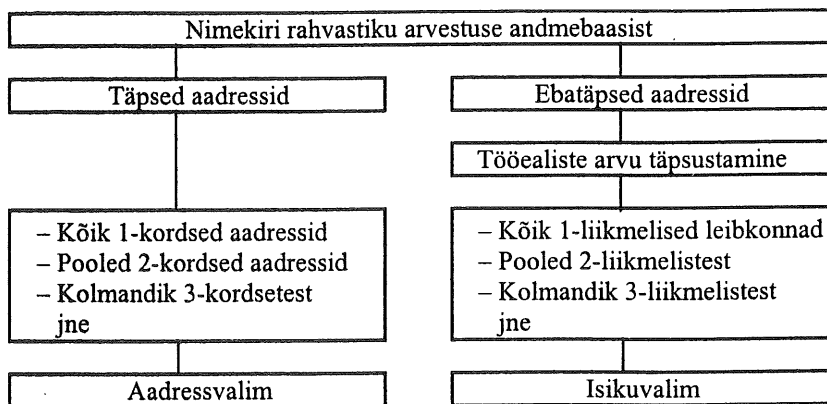
Küsitlustulemuste analüüsimisel selgus, et kao põhjuste hulgas oli kõigil aastatel esikohal see, et küsitletavat ei leitud, sest ta ei elanud registris näidatud aadressil. Kao põhjustena järgnesid keeldumine ja leibkonnaga kontaktteerumise ebaõnnestumine.

3.2. Üleminek pidevuuringule 2000. aastal

Pidevuuringut ette valmistades tuli mõelda ka valimiga seotud probleemidele. Raskused leibkonnaliikmete arvu täpsustamisel tekivad igas kvartalis, sest see on küllaltki tülikas ja aeganõudev. Samal ajal on selle töö tulemused küsitava väärtusega.

Paremate hinnangute saamiseks on kvartali valim jaotatud ühtlaselt 13 uurin-
gunädalale ja küsituleja peab minema intervjuud tegema vahetult uurin-
gunädalale järgneval nädalal. Erandjuhul on lubatud küsitleda ka hiljem (kuni viie
nädala jooksul). Ajaliste piirangute tõttu pole enam võimalik isikuid üle Eesti
taga otsida.

Lahenduseks otsustati kasutusele võtta leibkonna eelarve uuringu ümberdi-
sainimise käigus väljatöötatud meetodika, mida kohandati ETU nõudmistele.
Valim moodustati järgmise skeemi järgi:



Rahvastiku andmebaasist tellitud nimekiri sisaldab lisaks isikuandmetele samal aadressil registreeritud üle 15-aastaste isikute arvu (aadressi esinemissagedus). See on täpseteks ja ebatäpseteks aadressideks jaotamise alus.

Aadressid esinemissagedusega 1–6 loetakse täpseteks ja üle 10 ebatäpseteks, vahepealne nimekirjaosa kontrollitakse ridahaaval ja otsustatakse aadressi täpsuse üle vastavalt toodud andmetele. Täpne aadress määrab eluruumi üheselt (linnas: tänav, maja, korter; maal: küla, talu nimi). Ebatäpsed aadressid esinevad maapiirkonnas, kus aadressiks on ainult küla nimi.

Esinemissageduse järgi lülitatakse aadressivalimi ossa kõik üks kord esinevad aadressid, pooled kaks korda esinevatest jne. Ebatäpsete aadressidega tehakse läbi sama protseduur mis eelmistel aastatel.

Aadressivalimi osa oli umbes 85 protsenti valimist. Registri kvaliteedi paranedes peaks aja jooksul aadressivalimi osatähtsus pidevalt kasvama. Piirkonniti on aadressivalimi osatähtsus praegu väga erinev: ligi 100 protsenti Tallinnas, Saare- ja Viljandimaal, aga alla 50 protsenti Põlva- ja Võrumaal.

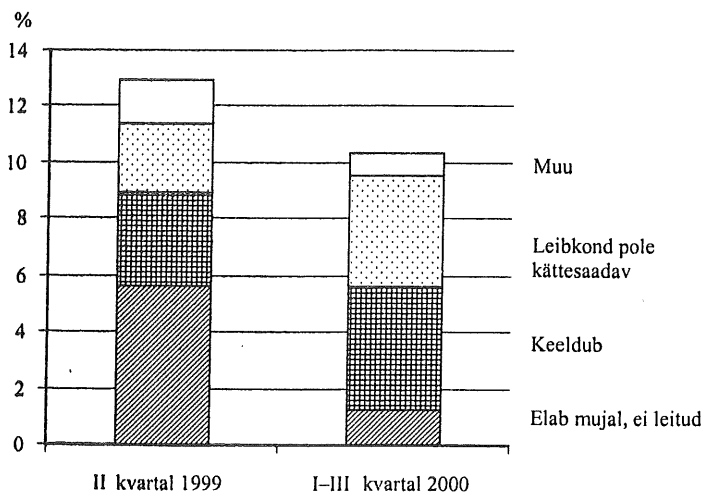
Küsitlejatele jagatud valimi nimekirjas on märgitud igal real käsitusviis. Aadressivalimi korral peab küsitaja minema märgitud aadressile ja intervjuerima sellel aadressil elavaid tööelisi liikmeid. Aadress on aluseks isegi siis, kui on teada, et valimis toodud isikut seal ei ela ja tema õige elukoht on teada. Kui aadress osutab rohkem kui ühele eluruumile (2–3-korteriga elamu), siis minnakse üle isikuvalimile. Isikuvalimi korral otsitakse isikut algul valimis osutatud vallast ja külast, kui teda seal ei ela, siis maakonna piires.

Uuel käsitusviisil on mitmeid eeliseid. Täpsustamist vajab vaid väike osa algnimekirjast. Niisiis väheneb ajakulu ja paraneb täpsustamise kvaliteet. Valim on isekaaluv, aadressivalimis on aadressid ja isikuvalimis leibkonnad võrdse kaasamistõenäosusega. Väheneb ajakulu küsitletavate otsimisel, valimit ei anta üle ühest maakonnast teise.

Puudusteks on valimi vigade osatähtsuse kasv, ebatäpsed aadressid, mille järgi pole võimalik eluruumi leida, keerulisem valimi käsitlemise eeskiri, probleemid kontaktkirjade kohaletoimetamisel. Valimi vigadeks on aadressid, mis osutavad asustamata eluruumile või mitteeluruumile, üle 74-aastastega asustatud eluruumile jms). Mõnes piirkonnas on majade numeratsiooni ja tänavate nimesid põhjalikult muudetud, aga registris on endised andmed. Küsitlajatel oli esialgu raske aru saada, et nad peavad minema küsitlema aadressi, mitte inimese järgi. Sageli juhtus, et kontaktkirjad saadeti avamata tagasi või viis kirjakandja kirja õigele inimesele, mitte märgitud aadressile.

Küsitlustulemus näitas, et 2000. aasta I–III kvartali jooksul moodustasid valimi vead 8 protsenti ja kadu 10 protsenti esialgsest valimist. Võrreldes varasema perioodiga on kadu vähenenud.

Joonis 3. Kao põhjused, II kvartal 1999, I–III kvartal 2000



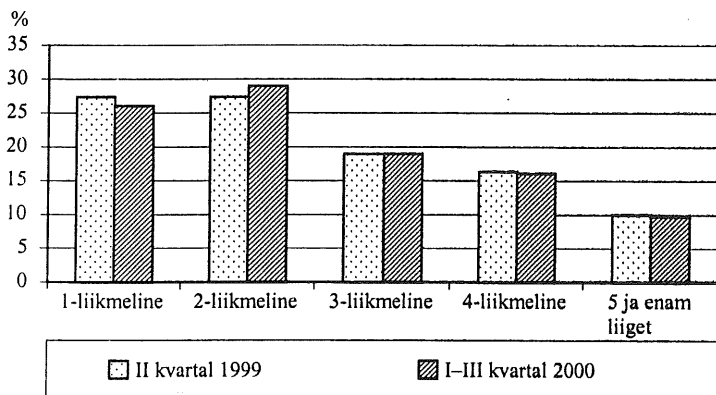
Mitteleidmise osatähtsus on tunduvalt vähenenud, suurenenud on keeldumine ja kontakteerumise ebaõnnestumine tingituna küsitluse ajalisest piirangust.

3.3. Tulemuste kvaliteedi hindamine

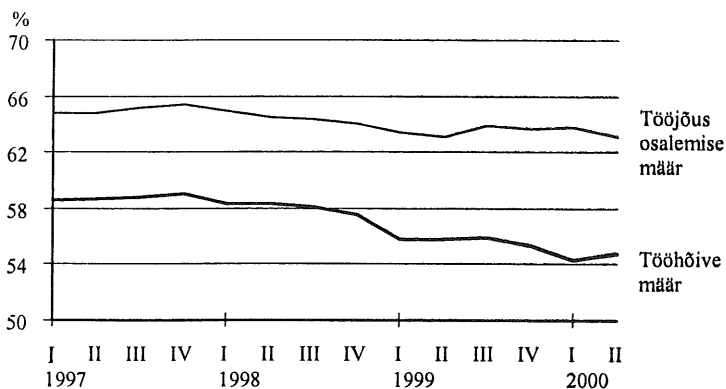
Uuringu meetodika muutmisel kerkib alati küsimus, kuidas mõjutab see saadavaid hinnanguid. Kontrollimaks, kas 2000. aastal küsitletud leibkonnad erinevad eelmistel aastatel valimisse sattunud leibkondadest, on võrreldud valimi leibkondade struktuuri. Joonis 4 näitab, et väikesed erinevused on 1- ja 2-liikmeliste leibkondade puhul, üldiselt aga leibkondade struktuur valimisel langeb erinevate valimi võtmise meetodikate korral üsna hästi kokku.

Veelgi olulisem meetodika muutuste korral on eri aastatel saadud hinnangute võrreldavus. Tööjõus osalemise määra ja tööhõive määra jooniselt on näha, et aegridades ei ole toimunud suuri hüppeid — seega on tulemused võrreldavad.

Joonis 4. Leibkondade struktuur ETU valimis, II kvartal 1999, I–III kvartal 2000



Joonis 5. Töajõus osalemise määr ja tööhõive määr, I kvartal 1997 – II kvartal 2000



Kasutatud kirjandus

Traat, I. 1999. Leibkonna eelarve uuringu ümberdisainimine. *Valimi töörühma lõpparuanne*. Tallinn: Eesti leibkonna eelarve uuringu ümberdisainimise teadusnõukogu.

ISIKUANDMETE KVALITEEDIST MEDITSIINI- REGISTRITES ehk LINKIMISUURINGUTE VÕIMALIKKUSEST EESTIS

Kaja Rahu
Ekspimentaalse ja Kliinilise Meditsiini Instituut

Linkimisuuringute puhul kasutatakse mingi hüpoteesi kontrollimiseks varem andmebaasidesse kogutud andmeid, mis isiku tasandil kokku viiakse. Meditsiiniliste andmekogudena tulevad kõne alla eelkõige meditsiiniregistrid. Näiteks mingi haigusregistri linkimine surmajuhtude registriga on eeltingimuseks, et hinnata elulemust selle haiguse puhul. Linkimisuuringud võtavad vähe aega ja on odavad, kuid eeldavad kvaliteetsete isikuandmetega andmebaaside olemasolu.

Linkimisel tehakse kahesuguseid vigu:

- pannakse kokku vale paar isikuid;
- õige paar jääb leidmata ja andmed ühendamata.

Isikukood teeb linkimise küll imelihtsaks, kuid ei lahenda probleemi täielikult, sest:

- isikukood tekkis alles koos rahvastikuregistriga 1992. aastal, seega ei aita see meid vanemate andmete korral;
- isikukoodi väli võib andmebaasis olla täitmata;
- kontrollsümbol ei vasta alati valemile;
- eri isikutel on sama isikukood;
- samal isikul on eri isikukoodid;
- naisel on mehe isikukood ja vastupidi;
- kontrollsümbol on võltsitud (ise arvutatud);
- isikukoodist tuletatud sünniaeg ei ole reaalne kuupäev.

Tabel 1 näitab, et isikukoodile saab toetuda järjest rohkem, sest mõne aastaga on olukord suuresti paranenud.

Tabel 1. Isikukoodid Eesti kolmes meditsiiniregistris, 1994 ja 1997

Aasta/register	Kirjete arv (%)			
	isikukood vastab valemile	isikukood puudub	isikukood ei vasta valemile	kokku
1994				
Vähiregister	1 685 (30,5)	3 833 (69,4)	3 (0,0)	5 521 (100)
Sünniregister				
lapsed	14 156 (100,0)	0 (0,0)	0 (0,0)	14 156 (100)
emad	5 494 (38,8)	8 590 (60,7)	72 (0,5)	14 156 (100)
isad	229 (1,6)	13 927 (98,4)	0 (0,0)	14 156 (100)
Surmade andmebaas	6 836 (30,9)	15 307 (69,1)	7 (0,0)	22 150 (100)
1997				
Vähiregister	5 272 (88,6)	667 (11,2)	14 (0,2)	5 953 (100)
Sünniregister				
lapsed	12 591 (100,0)	2 (0,0)	0 (0,0)	12 593 (100)
emad	11 332 (90,0)	1 213 (9,6)	48 (0,4)	12 593 (100)
isad	2 074 (16,5)	10 519 (83,5)	0 (0,0)	12 593 (100)
Surmade andmebaas	15 587 (83,9)	2 977 (16,0)	4 (0,0)	18 568 (100)

Et isikukoodi ei saa alati usaldada ja mõnikord seda polegi, tuleb linkimisel aluseks võtta nimi ja sünniaeg. Abiks on sugu, elukoht, rahvus või mis tahes muu tunnus olenevalt andmete koosseisust. Nimed võivad olla paberdokumendile kirjutatud raskesti loetava käekirjaga või kuulmise järgi, mistõttu nime kuju võib andmebaasiti erineda (tabel 2).

Haiguseregistrites jälgitakse kindlasti, et diagnoos ja sugu poleks vastuolus, kuid hoopis harvemini kontrollitakse nime ja soo vastavust. Võib juhtuda, et tüüpilisele mehenimele vastab sugu *naine* või vastupidi. Sel korral peab kontrollima algandmeid, sest sugu ei pruugi olla valesti kodeeritud, äkki on viga hoopiski nimes.

Tabel 2. Isikute paarid (reaalsed andmed on muudetud)

Nimi	Sünniaeg	Surmaaeg	Elukoht
Rein Kaupo	21.07.1928		Järvamaa
Kaupo Rein	21.07.1929		Järvamaa
Üpits Boriss	. .1908	12.10.1970	Jõhvi
Jupits Boris	. .1908	12.10.1970	Jõhvi
Veerma Maria	.07.1912	09.08.1986	Võru
Vurma Maria	12.07.1912	08.09.1986	Võru
Riemer Elsa-Anna	01.01.1943		Tallinn
Rimer Anna-Elsa	12.08.1943		Tallinn
Gjamjaljainen Vera	17.05.1935		Paldiski
Hämäläinen Veera	17.05.1935		Paldiski
Alexeev Semyon	12.10.1951		Paide
Aleksejev Semjon	01.11.1955		Paide
Hemmak Mihkel	25.12.1929	17.05.1995	Saaremaa
Nettak Mihkel	25.12.1929	17.05.1995	Saaremaa

Et linkimine oleks edukas, tuleb olla eriti tähelepanelik andmete testimisel ja parandamisel:

- perekonnanimi ja eesnimi seisku eraldi väljades;
- puuduvate kuupäevade asemel ei tohi kasutada fiktiivseid kuupäevi (nt 01.01; 01.07; tuleneb *date*-formaadi nõuetest);
- tuleb kontrollida lubamatuid koode ja loogikavigu;
- klassifikaatorid olgu eri andmebaasides samad, siis on lihtne võrrelda;
- topeltkirjed tuleb leida ja kustutada;
- testimiseks peab kasutama kogu infot, mida andmebaas sisaldab (ka kirjetevahelised testid).

Linkimine ei õnnestu veatult, kui isikuandmetele pärast sisestust tähelepanu ei pöörata (tabel 3).

Tabel 3. Näiteid nimekujudest ühes ja samas registris

Jjudmilla	Al-der
Kjudmilla	Alekander
Kudmila	Alekder
Liudmilla	Aleksaander
Ljudmila	Aleksadr
Ljudmilla	Aleksand
Ljüdmilla	Aleksander
Lljudmilla	Aleksandr
Ludmila	Aleksanr
Ludmilla	Aleksanser
Lumilla	Alesander
Ljudmiöa	Alksander

Kui suur võib olla ühe leidmata või valesti lingitud isiku kaal järelduse tegemisel, sõltub uuritava seose tugevusest ja andmebaasi suuruselt. Näitena sobib hüpotees Tšernobõli veteranide uuringust: enesetappe esineb nende meeste hulgas sagedamini kui Eesti meesrahvastikus üldse. Hüpoteesi kontrollimiseks lingiti veteranide andmebaas surmajuhtude andmebaasiga aastatel 1986–1997:

leitud enesetappude arv 48;
 eeldatav enesetappude arv 35,74;
 standarditud suremusmäär 1,34;
 95% usaldusvahemik 0,99–1,78.

Järeldus — veteranide enesetappude arv pole oluliselt suurem kui Eesti meesrahvastikus kokku.

Nüüd oletame, et leiame veel ühe enesetapujuhu lisaks:

leitud enesetappude arv 49;
 eeldatav enesetappude arv 35,74;
 standarditud suremusmäär 1,37;
 95% usaldusvahemik 1,01–1,81.

Järeldus — veteranide enesetappude arv on oluliselt suurem kui Eesti meesrahvastikus kokku.

Et saaksime anda oma panuse teadusesse kiirelt tehtavate, vähe ressursse nõudvate ja kvaliteetsete linkimisuuringute kaudu, tuleb isikuandmeid kohelda suure lugupidamisega.

RAHVASTIKUANDMED EESTI REGIONAALARENGU ANDMEBAASIS

Mihkel Servinski, Kaja Sõstra
Statistikaamet

Regionaalarengu andmebaasist

Eesti regionaalarengu andmebaasi loomist alustati 1997. aasta augustis, kui käivitus Phare projekt "Euroopa Liidu toetus Eesti regionaalarengule". Nimetatud projekti peäülesanneteks oli välja töötada Eesti regionaalarengu strateegia ja luua toimiv regionaalarengu andmebaas.

See andmebaas luuakse regionaalarengu seire kindlustamiseks — nii on kirjas Eesti regionaalarengu strateegias, mis kiideti heaks valitsuse istungil 16. novembril 1999. Andmebaas sisaldab riiklike statistiliste vaatlustega kogutud ja riigiasutustes väljaspool riiklikke statistilisi vaatlusi kogutud andmeid.

Regionaalsete andmete hetkeolukord on vastuoluline. Ühelt poolt tundub, et regionaalseid andmeid on küllaltki palju, teiselt poolt on andmete tarbijatel raskusi regioonide kohta andmete saamisega. Olemasolevad andmed on hajakutud — puudub ülevaade sellest, milliseid regionaalseid andmeid, mis tingimustel ja kust on võimalik saada. Andmed on halvasti kirjeldatud. Tarbijal puudub võimalus saada asjatundlikku abi andmete hankimisel ja kasutamisel. Eriti keeruline on olukord andmetega, mida koguvad riigiasutused väljaspool riiklike statistiliste vaatluste kava.

Eesti regionaalarengu andmebaas muudab olemasolevate regionaalsete andmete tarbijatele kättesaamise lihtsamaks, tagab asjatundliku abi regionaalsete andmete hankimisel ja kasutamisel.

Andmeid esitatakse regioonide kohta, milleks ideaalsel juhul on kohaliku omavalitsuse üksus, aga ka maakond või nende rühm.

Andmebaasi põhitarbijad on kohalikud omavalitsused, maavalitsused ja ministereeriumide regionaalküsimustega tegelevad üksused.

Andmebaasi sisu

Andmebaasi andmeid kogutakse etapiviisi, st et tulemusteni (andmete tarbijatele kättesaadavaks tegemiseni) jõutakse üksikute andmevaldkondadega järjestikuse tegelemise kaudu. Valdkondadega tegelemise järjekord määratakse tarbijate vajadusi ja andmete kättesaadavuse võimalusi arvestades. Esmalt kogutakse valdkonna kõige olulisemaid andmeid.

Esmased valdkonnad, mille kohta hakatakse andmeid koguma:

- rahvastik,
- tööhõive,
- elanike sissetulekud,
- kohalike omavalitsuste eelarved,
- ettevõtlus.

Andmebaasi loomist alustati olemasolevate andmete kogumise, süstematiseerimise ja kättesaadavaks tegemisega veebilehel <http://www.stat.vil.ee/pks/>

Sellelt lehelt võib praegu leida infot järgmiste valdkondade kohta:

- regionaalne sisemajanduse koguprodukt,
- rahvastik,
- sotsiaalkindlustus,
- õigus,
- leibkonna sissetulek ja väljaminek,
- palk,
- tööturg,
- rahandus,
- ettevõtlus,
- tööstus,
- kinnisvara ja ehitustegevus,
- väliskaubandus.

Rahvastikuandmed andmebaasis

Praegu on eespool nimetatud veebilehel kättesaadavad ligi 1200 andmetabelit. Ligikaudu pooled andmetest on rahvastikustatistika andmed:

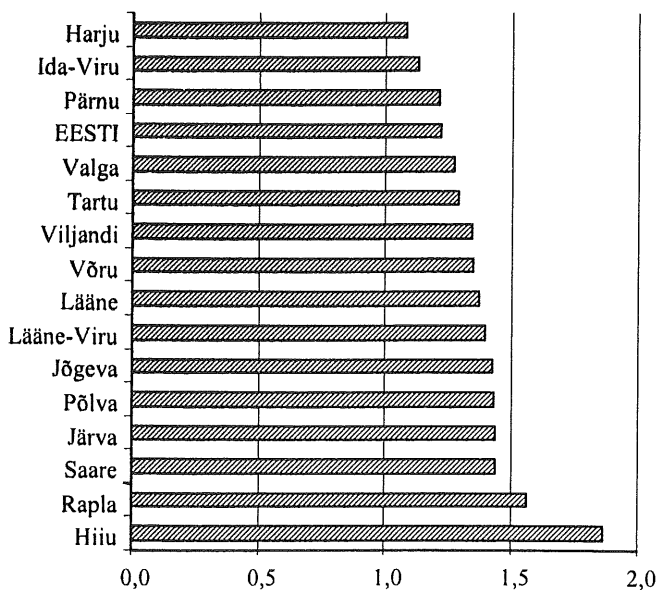
- rahvaarv,
- rahvastik soo ja valitud vanusegruppide järgi,
- pindala ja rahvastiku asustustihedus,
- rahvastikusündmused,

- surmapõhjused,
- ränne,
- ülalpeetavate määr,
- demograafilise töötururive indeks,
- rahvastiku rahvuskoosseis.

Enamik rahvastikuandmeid on esitatud omavalitsusüksuste kaupa. Lisaks absoluutarvudele on toodud ka indeksid, et erineva suurusega piirkondi omavahel võrrelda.

Demograafilise töötururive indeks võimaldab hinnata rahvastiku vanusstruktuuri tasakaalustatust, piirkondade demograafilist taastootmispotentsiaali ja nende muutusi. Indeks arvutatakse 5–14-aastaste ja 55–94-aastaste suhtena.

Demograafilise töötururive indeks, 1.01.2000



Regionaalarengu strateegia kesksed näitajad

1999. aasta 16. novembril Vabariigi Valitsuse kinnitatud Eesti regionaalarengu strateegia arendab edasi 1994. aastal valitsuses heaks kiidetud regionaalpoliitika käsitust.

Eesti regionaalarengu strateegia määrab kindlaks regionaalarengu soovitatavad suundumused, riigi regionaalpoliitika alused ja harupoliitikate regionaalse mõju suunamise lähtekohad Euroopa Liiduga liitumise eelsel perioodil.

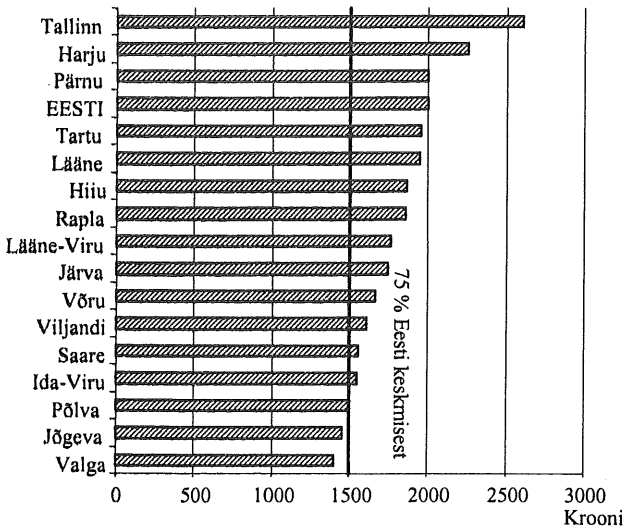
Eesti regionaalpoliitika on suunatud kõigis regioonides elanike stabiilse kõrge elukvaliteedi (turvaliste ja heade elutingimuste, sissetulekute, eneseteostuse võimaluste) saavutamisele.

Eesti regionaalpoliitika vahetu eesmärk on riigi regionaalareng tasakaalustada kohalike arengueelduste tugevdamise ja nende maksimaalse ärakasutamise teel. Sellega aidatakse kaasa riigi üldisele makromajanduslikule ja sotsiaalsele arengule.

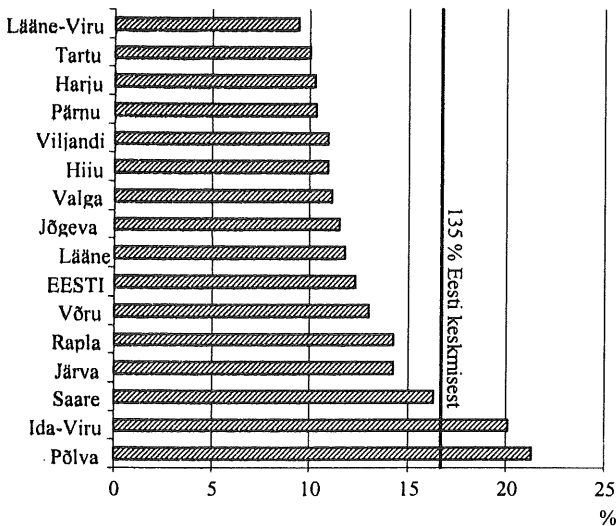
Eesmärgi saavutamise hindamisel on kesksed indikaatorid elanike keskmine sissetulek, töötuse määr ja maksutulude laekumine. Lisaks kasutatakse piirkonna spetsiifikast tulenevalt muid näitajaid. Eesti regionaalpoliitika rakendamise oodatav tulemus on regionaalarengu tasakaalustatus aastaks 2003 sedavõrd, et ühegi maakonna:

- keskmine elatustase (mõõdetuna leibkonnaliikme keskmise sissetulekuna) ei oleks madalam kui 75 protsenti Eesti keskmisest;
- tööpuudus (mõõdetuna töötuse määrana Rahvusvahelise Tööorganisatsiooni metoodika järgi) ei ületaks enam kui 35 protsenti Eesti keskmisest;
- kohalike omavalitsuste maksutulude laekumine (mõõdetuna kohalike omavalitsuste eelarvesse laekuva üksikisiku tulumaksuna) ei oleks madalam kui 75 protsenti Eesti keskmisest, millest on välja arvatud Tallinn.

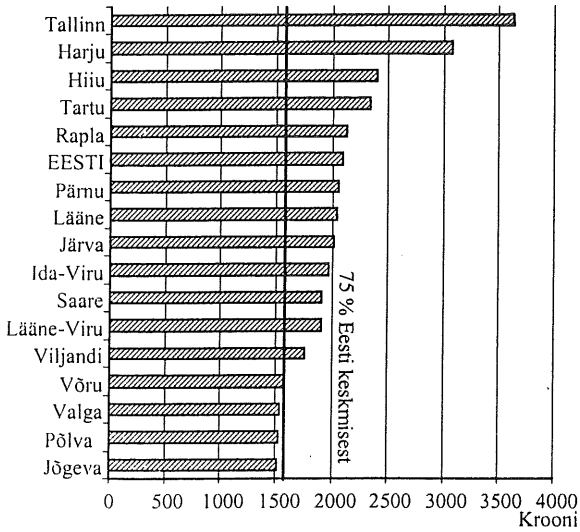
Leibkonnaliikme keskmine sissetulek, 1999



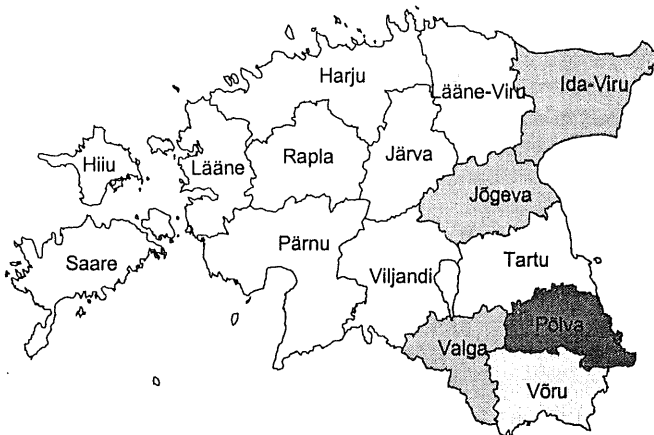
Töötuse määr, 1999



Maksutulude laekumine, 1999



Probleemsed maakonnad, 1999



EPIDEMIOLOOGIA JA REGISTRIPÕHISED UURINGUD

Mati Rahu

Ekspérimentaalse ja Kliinilise Meditsiini Instituut

Epidemioloogia jahib haiguste — täpsemalt tervisega seotud seisundite ja sündmuste — ja ekspositsioonide (tegurite, mõjurite) vahelisi seoseid. Epidemioloogid mõõdavad haiguste ja ekspositsioonide esinemist rahvastikus, samuti nendevaheliste seoste tugevust. Epidemioloogilistes uuringutes leitu põhjal võidakse välja pakkuda mooduseid haiguste sageduse vähendamiseks, s.o soovitusi rahvatervishoiupoliitikaks.

Epidemioloogia edusammud viimastel kümnenditel seostuvad suuresti kahe võimsa hoovaga — bio- ja infotehnoloogilise progressiga.

Biotehnoloogiline progress on molekulaarbioloogia ja -epidemioloogia kaudu andnud võimaluse avastada haigust võimalikult vara (haiguse markerid), mõõta ekspositsiooni jälgi (ekspositsiooni markerid) ja hinnata indiviidi eeldusi haigestuda (tundlikkuse markerid). Siinkohal huvitab meid biotehnoloogiline progress niivõrd, kui võrd temaga kaasub bioloogilise materjali — vereseerumi, DNA, küünte, juuste jms — registre (pankade) loomine ja kasutamine epidemioloogilisteks uuringuteks. Eestis on biotehnoloogia edusammud pärvinud avalikkuse tähelepanu eriti johtuvalt kavandatud geenivaramuprojektist. Seda projekti võib õigustatult vaadelda kui laiaulatuslikku epidemioloogilist uuringut, milles üht ekspositsiooniosist — geneetilisi mõjureid — käsitletakse tavapärasest süvendatumalt.

Infotehnoloogilise progressi ajal hakati seni paberdokumentidel (nimistud, ankeedid, sissekanderaamatud, kartoteegid) hoitud (tervise)andmeid säilitama elektroonselt. Elektroonsete andmekogude loomise tõttu suurenes märgatavalt teadustöö viljakus, sest terviseandmete linkimine muutus hõlpsamaks.

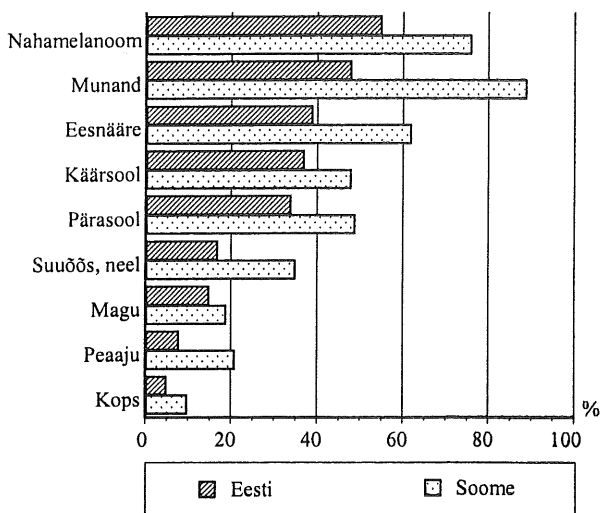
Teatavasti on dokumendilinkimine (*record linkage*) protseduur, mille korral andmed nt ühe ja sama isiku kohta kogutakse kokku eri andmekogudest (registritest). Mõiste *dokumendilinkimine* löi 1946. aastal Halbert Dunn (USA). Tema ideede edasiarendajatest on tuntuimad Howard Newcombe (1950. aastaist terviseandmete linkimine Kanadas), Donald Acheson (1962 esimene terviseandmete linkimise süsteem Inglismaal Oxfordi piirkonnas)

ning Ivan Fellegi ja Alan Sunter (1969 kirjutus dokumendilinkimise teooria kohta) [Gill et al 1993].

Epidemioloogia vaatevinklist tähendab registripõhine uuring (ehk register-uuring) haiguste esinemise ja/või neid tingivate ekspositsioonide kindlakstegemist ühe või mitme registri alusel. Seega registripõhine uuring aitab mõõta rahvastikus selliseid protsesse nagu haigestumus, suremus, levimus ja elulemus ning tuvastada haiguse ja ekspositsiooni vahelisi seoseid.

Vaatleme arvjoonist, mille tegemiseks vajalik andmestik pärineb Euroopa riikide ühisprojekti EURO CARE II kokkuvõttest [Coebergh et al 1998].

Viie aasta suhteline elulemusmäär vähipaikmeti Eestis ja Soomes, mehed (haigus diagnoositud 1985–1989)



Jooniselt näeme võrdlevalt Eesti ja Soome kohta, mitu protsenti mingisse vähki (nt kopsuvähki, maovähki) haigestunud meestest elab veel viis aastat pärast haiguse diagnoosimist. Näiliselt lihtsa lintdiagrammi konstrueerimise eelnes tohtu töö — andmete ettevalmistamine ja analüüs. Igas riigis, kes projektis osales, pidi olema alaline ja heal tasemel vähiregister, iga vähki haigestunu eluseisundit (antud kuupäeva seisuga elus, surnud või riigist lahkunud) tuli korralikult jälgida riigis loodud andmebaaside abil, elulemusarvu-

tusteks olid vajalikud rahvastiku elutabelid ja arvutiprogrammid. Need, kellele meeldib kasutada sõna *analüüs*, võivad olla kindlad, et selline arvjoonis saab valmida üksnes põhjaliku analüüsi tulemusena. Kuidas elulemusanalüüsi tuleneva joonisega või tabeli andmestikuga edasi toimida, sõltub paljuski juba muudest lülidest. Kas on oskajaid, kes mõõtmistulemuse edastavad kliinitsistidele? Või neid, keda leitu ergutab tervisedendusprojekti algatama? Kas riigil jätkub huvi ja jõudu luua vähitõrjeprogramm, milles elulemusnäitajategi najal plaanida tulevikutegemisi?

Ükskõik kui innukalt ja mitmekülgselt soovime registripõhise uuringu tulemusi panna teenima edasisi eesmäärke, oleneb nii uuringu enda kui ka rakenduste tase registri andmete kvaliteedist. Haiguseregistri korral tähendab kvaliteetsete andmete kogumine *tingimatut* linkimist tervishoiuasutuste andmebaaside, rahvastikuregistri ja surmajuhtude registriga. Niisugusest linkimisest (riskasutamisest) peavad teadlikud olema isikud, kes registreid kavandavad, loovad ja peavad, kes registreite õigusakte koostavad ja kes on kooskõlastaja rollis. Tegelikult ei tohiks Eestis tekkida haiguseregistreid, mille tava-tegevust kajastavates õigusaktides puudub range nõue andmete linkimiseks mitmete teiste andmekogudega ja mis on seetõttu loomise hetkest alates isoleeritud kvaliteedi tagamise protseduurist. *Täiesti kurioossele rajale astume vastutustundetult siis, kui registri asutamise õigusakti eelnõus väldime sihilikult igasuguse linkimise mainimist "õilsal" kaalutlusel, et nii olevat hõlpsam saada asutamisluba.*

Registripõhised uuringud koos neile omase dokumendilinkimisega on tänapäeva epidemioloogia alusmüür. Riikides, kus tõsiselt tegutsetakse rahva tervise seire ja parandamise huvides, saavad epidemioloogid oma töös tugineda mitmesugustele registritele. Siinjuures on esitatud mõned näited registrite kasutamise kohta epidemioloogilistes uuringutes.

Norras kontrolliti hüpoteesi, et isa kutseekspositsioon, mis põhjustab mutatsioonid spermas, võib viia kaasasündinud väärarendite tekkimisele lastel [Irgens et al 2000]. Uuringus lingiti Norra Meditsiinilise Sünniregistri andmed 1,2 miljoni sünni kohta aastatel 1970–1993 kolme rahvaloenduse (1970, 1980, 1990) andmebaasiga, mida säilitatakse elektroonselt ja mis sisaldavad isikut tuvastavaid tunnuseid.

Soomes mõõdeti laste terviseseisundit seitsme eluaasta jooksul [Gissler et al 1999]. Selleks jälgiti 1987. aastal Soomes sündinud 60 192 last Soome Meditsiinilise Sünniregistri andmestiku linkimise teel rahvastikuregistri, surmapõhjuste registri, haiglast lahkunute registri, ravimite andmebaasi, tervise-

hoolduse andmebaasi ja 18 piirkondliku registriga intellektipuudega laste kohta.

Kanadas sooviti teada saada, kas rinnavähiskriiningus on kliiniline läbivaatus koos mammograafiaga eelistatavam kui ainult kliiniline läbivaatus [Miller et al 2000]. Epidemioloogilise katsena korraldatud uuringus osales 39 405 naist, kes randomiseeriti kahte rühma ja kelle terviseseisundit seirati keskmiselt 13 aastat Kanada Riiklikust Vähiregistrist ja Kanada Riiklikust Suremuse Andmebaasist saadud teabe abil.

Registripõhised linkimisuuringud on juurdunud Põhjamaades. Sealsetest eriilmelistest töödest tuleks mainida järgmisi: vähihaigete enesetapud (avaldatud 1992); juuksurite vähihaigestumus (1992); kõrgepingeliinide läheduses elavate laste vähihaigestumus (1993); siirdatud neeruga isikute vähihaigestumus (1995); emakaelavähi kui suguhaiguse riskitegurid; abort ja väike sünnikaal järgmise raseduse korral; skisofreenia kui rasedusaegsete tüsistuste tulem; tippsportlaste vähihaigestumus (kõik 2000).

Lõviosa nimetatud uuringutest kuulub meie jaoks lahtrisse "hapud viinamarjad" ehk "helesinised unistused". Ja põhjusel, et *me ei pea vajalikuks süüvida haiguste registreerimise teooriasse, et meil puuduvad registreerimise standardid, et aeglus õigusaktide koostamisel ja muud asjaolud on teinud registrid kergesti haavatavaks; lisaks on pidurdunud registripõhine teadustöö, olematu on teadlaste ühisrinne põhjendamaks üldsusele ja ametnikele vajadust isikuandmetega meditsiiniregistrite järele*. Ometigi ei tule eeskuju otsimisel ja võtmisel kaugele rännata — Põhjamaade registritest ja registripõhistest uuringutest on siiani õppust võetud ja seda saab jätkata tulevikuski. Kas ei vääriks süvatähelepanu Taani arvukate meditsiiniregistrite töökorraldus või Rootsi 40 riikliku tervisehoolduse kvaliteediregistri eesmärgid [National Health... 2000]? Soomest saame aimu, mida tähendab kodanikkonna *registrateadvuse* arendamine: et registrid pole loodud inimeste eraelus tuhnimiseks, vaid ühiskonnale elulist huvi pakkuvate küsimuste lahendamiseks; et inimesed peavad teadma, missugused registrid on olemas ja mida nad sisaldavad; et registriandmetele juurdepääs ja nende töötlemine käib õigusaktides sätestatu alusel; et registreid rajatakse selleks, et neid kõigekülgselt kasutada.

Aga ükskõik kui hästi ühiskond tervikuna mõistab vajadust registrite järele, leidub alati isikuid, kes soovivad osa saada hüvedest, ise midagi vastu andmata. Sellekohase isekuse näite toob rahvusvaheline tervisepoliitika konsultant William Lowrance [Lowrance 1997]:

Arst: *Palun, siin on ravim, mis Teid aitab.*

Patsient: *Kuidas Te seda teate?*

Arst: *Uuring 10 000 inimese ravi tulemuste kohta näitas, et 9247 neist paranesid.*

Patsient: *Hästi, ma hakkas seda ravimit võtma. Kuid ärge teatage mitte kellelegi, kas ma sain tervemaks.*

Kasutatud kirjandus

Coebergh, J. W. W. et al. *Survival of Adult Cancer Patients in Europe Diagnosed from 1978–1989: The EUROCARE II Study.* — *Eur J Cancer*, Vol 34, No 14, 1998, pp 2137–2278.

Gill, L. et al. *Computerised Linking of Medical Records: Methodological Guidelines.* — *J Epidemiol Community Health*, Vol 47, No 4, 1993, pp 316–319.

Gissler, M. et al. *Can Children's Health Be Predicted by Perinatal Health?* — *Int J Epidemiol*, Vol 28, No 2, 1999, pp 276–280.

Irgens, A. et al. *Birth Defects and Paternal Occupational Exposure. Hypotheses Tested in a Record Linkage Based Dataset.* — *Acta Obstet Gynecol Scand*, Vol 79, No 6, 2000, pp 465–470.

Lowrance, W. W. *Privacy and Health Research.* U.S. Department of Health and Human Services. Washington, 1997, p 12.

Miller, A. B. et al. *Canadian National Breast Screening Study-2: 13-Year Results of a Randomized Trial in Women Aged 50–59 Years.* — *JNCI*, Vol 92, No 2, 2000, pp 1490–1499.

National Health Care Quality Registries in Sweden 1999. Federation of Swedish County Councils, and National Board of Health and Welfare. Stockholm, 2000.

EESTI NAISTE REPRODUKTIIVKÄITUMINE AASTAIL 1994–1998

(sünni- ja abordiregistri andmetel)

Ene-Margit Tiit, Ene Käärik, Meelis Käärik
Tartu Ülikool

1. Ülesande püstitus

Käesoleva uuringu aluseks on Eesti meditsiinilise sünniregistri ja aastail 1994–1998 eksisteerinud Eesti abordiregistri andmestik nimetatud viieaastast perioodist, kokku 65 660 kirjet sünniregistrist ja 96 891 kirjet abordiregistrist. Andmestikud ühendati ja korrastati ning korrastatud andmestik sisaldab 45 699 kirjet, mille põhjal on tehtud alljärgnev analüüs.

Uuringu eesmärgiks on modelleerida eestlaste seksuaal- ja reproduktiivkäitumist, sealjuures eriti selgitada neid tegureid, mis mõjustavad rasestunud naise otsust sünnituse või raseduse katkestamise kasuks ja samuti teha oletusi rasedusvastaste vahendite kasutamise kohta. Eriti pakub huvi:

- hinnata tõenäosusi, et raseduse katkestamisele järgneb sünnitus ja sünnitusele järgneb sünnitus (ning vastupidi),
- leida mudel selgitamiseks sünnitamise otsuse langetamise tõenäosust sõltuvalt raseda varasemast reproduktiivkäitumisest ning tema demograafilistest ja sotsiaal-majanduslikest karakteristikutest,
- uurida ja hinnata rasestumistevaheliste ajavahemike jaotust tegemaks oletusi rasestumisvastaste vahendite kasutamise kohta.

2. Reproduktiivsündmuste jadad

Vaadeldav andmestik piirdub ainult viie aastaga, mil abordiregister tegutses registrina; pärast selle reorganiseerimist *raseduse katkestuste andmekoguks* kahjuks kadus võimalus reproduktiivkäitumist iseloomustavate jadade analüüsimiseks. Algselt oli andmestik esitatud standardse statistilise andmekogu kujul, kus iga kirje vastas üksiksündmusele (kas sünnitusele või raseduse katkestusele). Püstitatud ülesande lahendamiseks oli andmestik tarvis teisen-dada nii, et objektiks kujuneks sündmuste jada, mis koosnes sündmustest A ja S, kusjuures olid fikseeritud ka sündmustevahelised ajavahemikud (vt tabel 1).

Arvestades ka jadade osajadasid (näiteks SAS sisaldab jadasid SA ja AS), on tabeli 1 põhjal võimalik arvutada järgmised tõenäosuste hinnangud:

pärast S toimumist on S tõenäosus 48,92 ja A tõenäosus 51,08;
 pärast A toimumist on S tõenäosus 40,85 ja A tõenäosus 59,15;
 pärast SS toimumist on S tõenäosus 41,48 ja A tõenäosus 58,52;
 pärast SA toimumist on S tõenäosus 31,31 ja A tõenäosus 68,69;
 pärast AS toimumist on S tõenäosus 22,56 ja A tõenäosus 77,44;
 pärast AA toimumist on S tõenäosus 38,16 ja A tõenäosus 61,84.

Tabel 1. Suurema sagedusega reproduktiivsündmuste jadad

Kahe reproduktiivsündmuse jadad								
Jada	SS	SA	AS	AA				
Sagedus	4 163	4 346	2 368	3 429				
Kolme reproduktiivsündmuse jadad								
Jada	SSS	SSA	SAS	SAA	ASS	ASA	AAS	AAA
Sagedus	341	481	387	849	111	381	280	666
Nelja reproduktiivsündmuse jadad								
Jada	SSSS	SSSA	SSAS	SSAA	SASS	SASA	SAAS	SAAA
Sagedus	29	27	19	150	17	94	86	314
Jada	ASSS	ASSA	ASAS	ASAA	AASS	AASA	AAAS	AAAA
Sagedus	4	17	21	96	16	79	81	161

Pikemate jadade puhul ei ole tõenäosuse hinnangud enam nii usaldusväärsed, lisaks sellele seab piirid vaatlusperioodi lühidus (kokku ligikaudu 1800 päeva). Üldiselt ilmnevad järgmised tendentsid.

- Kõige suurema tõenäosusega langetavad otsuse sünnituse kasuks need naised, kes pole varem ühtki rasedust katkestanud — ka neljas laps (viie aasta jooksul) sünnib tõenäosusega 0,52.
- Tõsiasi, et naise reproduktiivsündmuste jada on alanud abordiga, mõjustab teda suurema tõenäosusega langetama otsust aborti kasuks.
- Suurema tõenäosusega katkestavad raseduse need naised, kes on sedasama teinud ka eelmise rasedusega.
- Ühe- ja kahelapselistes peredes ei mõjusta olemasolevate laste arv sünnitamise otsust oluliselt.

3. Sünnitamise otsust mõjustavad demograafilised ja sotsiaal-majanduslikud tegurid

3.1. Mudel ja oletatavad argumendid

Eelmises punktis esitatud mudelid ei anna vastust küsimusele, mis mõjustab rasestunud naist langetama otsuse sünnitamise kasuks. Selgitamaks seda konstrueeriti logistilised mudelid, mis hindavad eraldi *esmasünnitaja ja kor-duvsünnitaja* puhul tõenäosust langetada otsus sünnitamise kasuks. Logistilises mudelis on prognoositavaks suuruseks uuritava sündmuse tõe-näosuse P funktsioon $P/(1-P)$, nn *šanss*. Mudel ise on multiplikatiivne, mis saavutatakse sel teel, et leitakse lineaarne regressioonmudel šansi logaritmile

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \sum_{i=1}^k \beta_i X_i \quad (1)$$

Oletatavad argumendid, mis sünnitamise otsuse tõenäosust mõjustavad, on järgmised.

- Perekonnaseis, millel on tegelikult kaks binaarset tunnust (indikaatorit): 1) kas on seaduslikus abielus ja 2) kas on vabaabielus. Kui mõlema tun-nuse väärtuseks on null, tähendab see kolmandat võimalust, st ei ole ei seaduslikus ega vabaabielus.
- Elukoht, mis koosneb samuti kahest binaarsest tunnusest — *elab Tallin-nas* ja *elab linnas*. Mõlema tunnuse nullväärtusele vastab ülejääv või-malus *elab maal*.
- Tegevus (tööalane staatus), mis koosneb kolmest binaarsest tunnusest: *on töötaja*, *on töötu*, *on õppur* (sh nii õpilane kui ka üliõpilane); neljas, ülejääv variant (eespool mainitud kolme tunnuse väärtuseks on null) on see, et *isik on kodune*.
- Haridus, mida käsitletakse arvtunnusena (mõõdab läbitud haridus-astmeid).
- Vanus, mille mõju pole oletatavasti lineaarne, mistõttu on lisaargumen-diks loetud ka vanuse ruut.
- Rahvus, vaadeldud binaarse tunnusena (eestlane = 1, muu rahvus = 2).
- Varasem abortide arv.
- Varasem sünnituste arv.
- Viimasest reproduktiivsündmusest (sünnitusest/abordist) möödunud päe-vade arv. See päevade arv on viidud võrreldavale skaalale teisendusega

$k = p - C$, kus konstandi C väärtuseks on Tartu Ülikooli Naistekliinikus määratud eksperthinnang. Et selle teguri mõju ei tarvitse olla lineaarne, on argumendiks võetud ka tema ruut.

- Eelmine sündmus (1 — sünnitus, 0 — abort).

3.2. Esmasünnitamise otsuse langetamise tõenäosuse mudel

Veel mitte sünnitanud naiste andmestik sisaldas 13 029 kirjet (9658 naise andmed, kellel eelmiste sünnituste arv on null, aga võib olla esinenud aborte), kusjuures sünnitamise tõenäosuse hinnang oli 0,64 ja vastavalt sünnitamise šanss 1,80. Valemi (1) kordajate hinnangud on esitatud tabelis 2 (teises veerus β). Kolmandas veerus *Kordaja* on arvtunnuste korral teises veerus esitatud suuruse eksponentfunktsiooni väärtus. See näitab, kui palju kordi suureneb sünnitamise šanss, kui vastav argument suureneb ühe ühiku võrra. Seega mõõdab see kordaja šansside suhet vastavalt argumenti muutusele ühe ühiku võrra. Kui aga tunnus on esitatud mitme binaartunnuse kaudu, näitab kordaja šansi muutumist üksikute binaartunnuste kaupa.

Tabel 2. Esmasünnitaja sünnitamise tõenäosuse mudel

Argument	β	Kordaja	Liikme olulisus	R^2
Vabaliige	-3,1653		<0,0001	0
Perekonnaseis (abielus)	1,5649	37,06	<0,0001	
Perekonnaseis (vabaabielus)	0,4827	12,557	<0,0001	0,4298
Tegevus (töötaja)	-0,3364	0,564	<0,0001	
Tegevus (töötu)	0,8846	1,914	<0,0001	
Tegevus (õppur)	-0,7840	0,361	<0,0001	0,4489
Elukoht (Tallinn)	0,1476	0,892	0,0002	
Elukoht (muu linn)	-0,4099	0,511	<0,0001	0,4651
Varasemate abortide arv	-0,3880	0,678	<0,0001	0,4772
Rahvus	-0,4962	0,609	<0,0001	0,4806
Vanus	0,4313		<0,0001	0,4819
Haridus	-0,3117	0,732	<0,0001	0,4853
Vanuse ruut	-0,00724		<0,0001	0,4889

Tabelist 2 on näha, et kõige tugevama toimega tegur sünnitamise otsuse langetamisel on naise perekonnaseis: seaduslikus abielus oleva naise puhul on sünnitamise šanss üle 30 korra, vabaabielus naisel üle 10 korra suurem kui vallalisel, üksi elaval naisel. Determinatsioonikordaja väärtus R^2 näitab, et perekonnaseis kirjeldab esmasünnitaja sünnitamise otsuse varieeruvusest

enam kui 42 protsenti. Kõik hiljem mudelisse liidetavad argumendid on küll statistiliselt olulised, kuid parandavad ühtekokku mudelit vaid umbes 5 protsendi võrra.

Järgneb **tegevus (tööalane staatus)**. Õppuril on šanss sünnitada ligi kolm korda väiksem kui koduperenaisel. Ka töötajal on koduperenaisega võrreldes ligi poole väiksem šanss langetada otsus sünnitamise kasuks, seevastu aga töötu naine langetab sünnitamise otsuse suurema tõenäosusega.

Elukoha mõju ilmneb selles, et maal elav naine langetab suurema tõenäosusega sünnitamise otsuse kui linnas elav naine. Erinevus on suurem väiksemate linnade puhul, kus sünnitamise otsuse šanss on ligi poole väiksem kui maal. Seevastu Tallinnas on sünnitamise tõenäosus rasestumise korral peaaegu sama suur kui maal.

Varasemate abortide (siin peetakse silmas üksnes legaalseid, st naise omal valikul tehtud aborte) arv vähendab arvestatavalt sünnitamise šanssi — iga abordiga kahaneb šanss varasema seisuga võrreldes umbes 70 protsendini.

Ligikaudu samasugune toime on **rahvusel ja haridusel** — see, kui rase ei ole põhirahvusest, samuti ka iga järgnev hariduseaste vähendab sünnitamise otsuse langetamise šanssi.

Vanuse mõju ei ole lineaarne. Kuni 30. eluaastani suureneb vanuse suurenedes tõenäosus, et naine langetab otsuse sünnitamise kasuks, pärast seda hakkab sünnitamise tõenäosus vähenema.

Ühtekokku kirjeldab esitatud mudel sünnitamise otsuse langetamise tõenäosuse varieerumist peaaegu 49 protsendi ulatuses.

Varieerisime mudelit, vaadeldes ainult **esmarasedaid**, kelle kohta oli andmestik 7330 kirjet (isikut). Nende puhul oli sünnitamise tõenäosuse hinnang 0,69 ja vastavalt sünnitamise šanss 2,26. Sünnitamise tõenäosust prognoosiv mudel osutus üsna sarnaseks eelkirjeldatuga, kusjuures perekonnaseisu mõju sünnitamise otsusele oli veelgi tugevam, ulatudes 45 protsendini (determinatsioonikordaja mõttes), mudeli kirjeldatuse tase oli 50 protsenti.

3.3. Korduvsünnitajate sünnitamise otsust mõjustavad tegurid

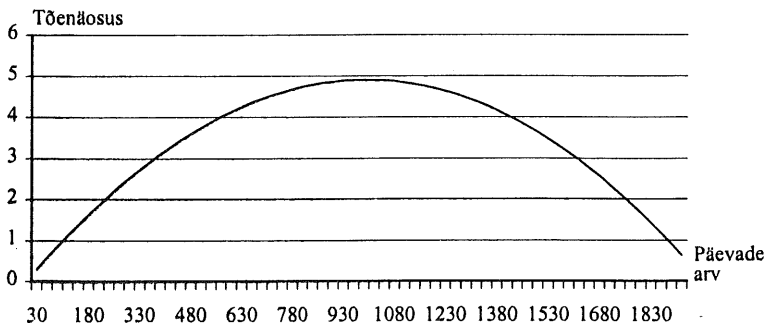
Korduvsünnitajateks loeti kõik need isikud, kellel oli varem registreeritud vähemalt üks sünnitus. Selle andmestiku suurus on 22 981 kirjet (16 717 naise andmed, kellel on reproduktiivsündmuste jadas vähemalt kaks sünnitust). Mudel konstrueeriti eelmises punktis leitud esmasünnitaja mudeli sarnaselt.

Osutus, et kõige olulisemaks argumendiks korduvsünnitaja mudelis on eelmisest reproduktiivsündmusest möödunud aeg — soodsaim on umbes 1000-päevane ajavahemik (vt joonis 1). See tunnus kirjeldabki üle veerandi sünnitamise otsuse tõenäosuse hajuvusest.

Tabel 3. Korduvsünnitaja sünnitamise tõenäosuse mudel

Argument	β	Kordaja	Liikme olulisus	R^2
Vabaliige	-5,6238		<0,0001	0
Päevade arv	0,0099	1,0099	<0,0001	0,1895
Päevade arvu ruut	-4,97E ⁻⁶		<0,0001	0,2760
Elukoht (Tallinn)	0,9166	2,500	<0,0001	
Elukoht (muu linn)	-0,6667	0,513	<0,0001	0,3196
Perekonnaseis (abielus)	1,5681	4,798	<0,0001	
Perekonnaseis (vabaabielus)	1,7666	5,851	<0,0001	0,3594
Abortide arv	-0,1559	0,856	<0,0001	0,3765
Vanus	0,1017	1,107	0,0004	0,3808
Tegevus (töötaja)	0,1179	1,125	0,2729	
Tegevus (töetu)	0,5461	1,726	<0,0001	
Tegevus (õppur)	-0,0046	0,995	0,1028	0,3836
Sünnituste arv	0,1000	1,105	<0,0001	0,3849
Kordus	-0,1333	0,875	<0,0001	0,3864
Vanuse ruut	-0,0025		<0,0001	0,3875
Haridus	-0,0622	0,940	0,0033	0,3879
Belmine sündmus	-0,1040	0,901	0,0190	0,3881

Joonis 1. Sünnitamise otsuse tõenäosuse sõltuvus reproduktiivsündmustevahelisest ajavahemikust (päevades)



Tähtsuse poolest järgmisel kohal on **elukoht**: šanss langetada sünnitamise otsus on Tallinnas 2,5 korda suurem kui maal, teistes linnades aga poole väiksem kui maal. Tulemus on mõneti ootamatu, sest Tallinnas on sündimus madalam kui maal. Nähtavasti on siin põhjuseks rasedusvastaste vahendite edukam kasutamine Tallinnas võrreldes maaelanikega, mistõttu soovimatuid rasestumisi on vähem ja rasestumine viib sagedamini sünnituseni. Mis puutub ülejäänud linnadesse, siis võib arvata, et siin räägib kaasa ka suhteliselt madal sündimus Ida-Viru linnades.

Järgneb **perekonnaseis**. Perekonnaseisu mõju sünnitamise otsuse langetamise tõenäosusele on korduvrasedatel märksa väiksem kui esmarasedatel. Võrreldes üksikutega on abielunaiste šanss langetada otsus sünnitamise kasuks ligi viis korda suurem. Üllatav on see, et vabaabielunaistel on korduvraseduse puhul sünnitamise otsuse šanss üksikutega võrreldes veelgi suurem. Need kolm tunnust kirjeldavad sünnitamise otsuse hajuvusest 36 protsenti.

Tegevusala, hariduse ja vanuse mõju korduvsünnitajate sünnitamise otsusele on üldjoontes samasugune nagu esmasünnitajatelgi. Tähelepanu väärib aga see, et suurem varasemate abortide arv mõjustab naist pigem langetama otsust taas aborti kasuks, suurem sünnituste arv aga taas sünnituse kasuks. Samas mõjutab vahetult eelnenud sündmus valikut pigem vastassuunas. Kõigi nende tegurite mõju on aga väga väike ja pärast nende kaheksa tunnuse mudelisse lisandumist tõusis kirjeldatuse tase (R^2 mõttes) 39 protsendini.

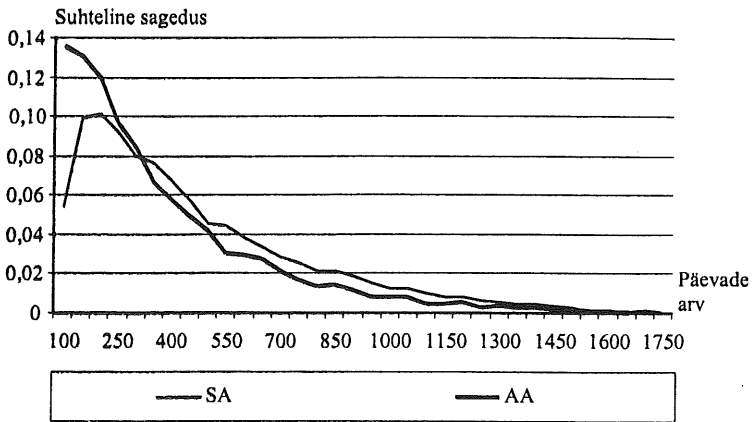
4. Reproduktiivsündmustevahelise ajavahemiku jaotus

Analüüsides reproduktiivsündmustevahelise ajavahemiku jaotust, on võimalik sõnastada hüpoteese ka selle kohta, kuidas teadlik on Eesti naiste reproduktiivkäitumine: kas rasestumiste vaheaegade jaotus on iseloomulik juhuslikult toimuvate sündmuste ajavahemike jaotusele või ilmnevad siin teatavad seaduspärasused, mis lubavad oletada, et vähemalt osa naistest väldib rasedust teadlikult ja sünnitab kavandatud ajal.

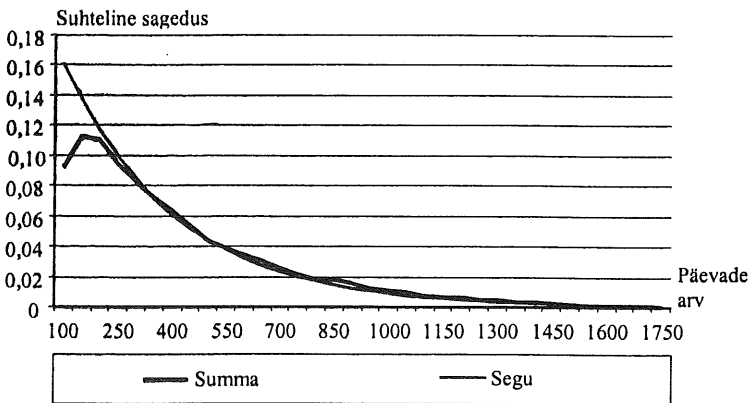
4.1. Abordi ja sellele eelnenud reproduktiivsündmustevahelise ajavahemiku jaotus

Siin oli kasutada andmeid 14 377 ajavahemiku kohta (SA 6621, AA 7756). Kõik ajavahemikud on mõõdetud päevades. Joonisel 2 on näha, et eri ajavahemike — SA ja AA — jaotused on mõnevõrra erinevad. Siiski on alust arvata, et mõlemat jaotust kirjeldab alates teisest-kolmandast mõõtepunktest üsna hästi eksponentjaotus, mis viitab sündmuste voo juhuslikkusele.

Joonis 2. Ajavahemik sünnitusest/abordist abordini



Joonis 3. Abordi toimumise aja jaotus pärast eelmist reproduktiivsündmust

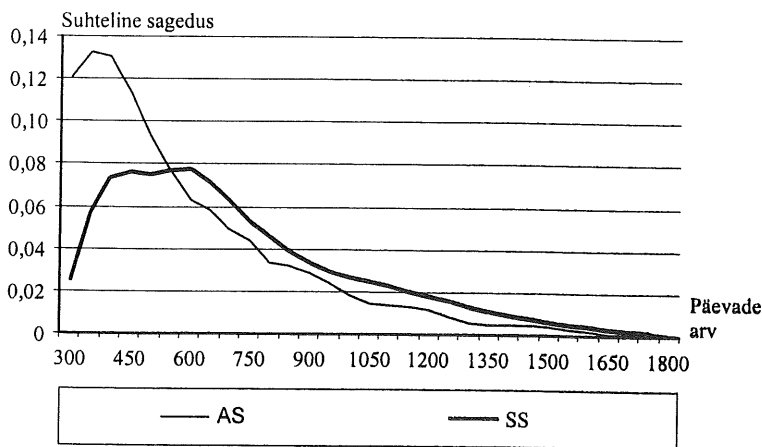


Joonisel 3 on mõlemad andmestikud ühendatud (köver *Summa*) ja neid on lähendatud kahe eksponentjaotuse seguga (kumbki jaotus on optimaalne ühe andmestiku esitamiseks, köver *Segu*). Küllalt hea kooskõla empiirilise ja teoreetilise jaotuse vahel alates kolmandast punktist lubab oletada, et kuigi sünnitusest ja abordist järgmise abordini loetud ajavahemiku jaotus on pisut erinev, on tegemist ühesuguse seaduspärasusega. Tegemist on juhusliku, kontrollimatu rasestumisega. Teoreetilise ja empiirilise jaotuse halb kooskõla joonise vasakpoolses osas on samuti sisuliselt selgitatav: sünnitanud, last rinnaga toitva naise rasestumise tõenäosus on ilmselt märksa väiksem raseduse katkestanud naise rasestumise tõenäosusest.

4.2. Sünnituse ja sellele eelnenud reproduktiivsündmuse vahelise ajavahemiku jaotus

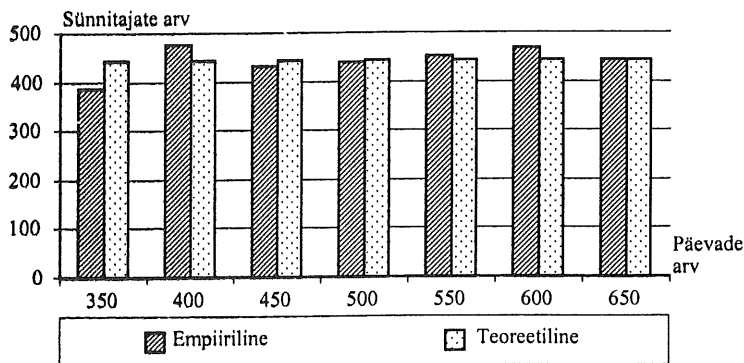
Siin analüüsiti 8427 ajavahemikku (SS — 5869, AS — 2558), nende jaotusi esitab joonis 4.

Joonis 4. Vaheaeg eelmisest reproduktiivsündmusest järgmise sünnituseni



Kahe jaotuskõvera erinevus on ilmne. Sünnitusejärgsel 250-päevasel perioodil rasedutakse üsna ühtlase jaotusega (vt joonis 5).

Joonis 5. Sünnitustevahelise ajavahe jaotus (350–650 päeva)

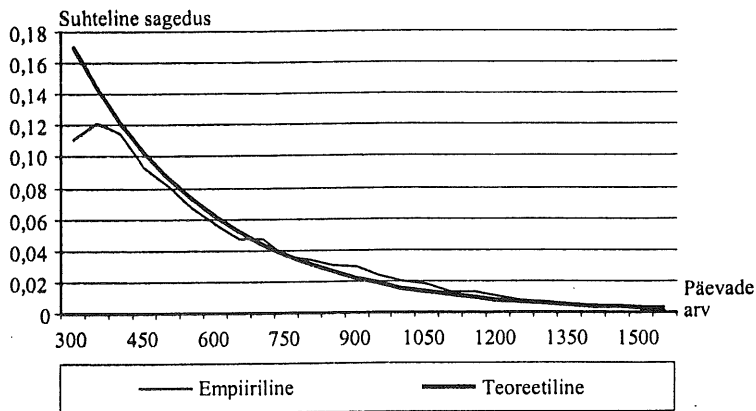


Seda nähtust saab seletada kahel viisil. Tegemist on

- 1) sihipäraselt kavandatud laste vanusevahega;
- 2) rinnaga toitvate naiste madalama rasedumise tõenäosusega.

Kõige tõenäosem on mõlema põhjuse kombineerumine.

Joonis 6. Ajavahe miks sünnituseni pärast aborti või 250 päeva möödumist eelmisest sünnitusest



Joonis 4 viib mõttele, et kahest andmestikust saadud sünnituseelsete ajavahemike jaotused saaks ühitada pärast 250-päevase nihke teostamist sünnitusjärgsetele ajavahemikele. Osutubki, et ka sellisel viisil ühendatud ajavahemike hulga jaotust kirjeldab üsna hästi eksponentjaotus (vt joonis 6).

5. Kokkuvõte

Uuringust järeldub, et Eesti naiste reproduktiivkäitumine sõltub üsna olulisel määral demograafilistest ja sotsiaalsetest teguritest.

- Kuigi praegusajal sünnib väga suur osa lapsi väljaspool abielu, on siiski abielusideme olemasolu esmasünnitaja puhul määravaks teguriks, mis mõjustab sünnitamise otsust, siinjuures on esmasünnitaja puhul seaduslik abielu märksa tugevam sünnitamise stiimul kui vabaabielu.
- Korduvsünnituse puhul on olulisim tegur see, millal sündis viimane laps — soovitav ei ole liiga väike ega ka liiga suur laste vanusevahe.
- Korduvsünnituse puhul on perekonnaseisu mõju väiksem kui esmasünnitusel; seaduslikul abielul ei ole vabaabielu ees eeliseid.
- Sünnitusega ja raseduse katkestamisega lõppevate rasedustumiste ajastuse jaotused on erinevad.
- Abordiga lõppevate raseduste ajahetkede jaotused on hästi lähendatavad eksponentjaotusega, mis viitab sündmuste voo juhuslikkusele.
- Sünnitustevaheliste ajavahemike jaotused koosnevad kahest erikujulisest komponendist. Üle 50 protsendi korduvsünnitustest on kuni eelmise lapse kaheaastaseks saamiseni, kusjuures jaotus on ühtlane. Hilisemad sünnitused, samuti kui abordile järgnevad sünnitused toimuvad eksponentjaotuse järgi.
- Varasemad abordid suurendavad naise kalduvust katkestada rasedus.
- Suhteliselt suure tõenäosusega langetavad otsuse sünnituse kasuks need naised, kes ka varem on vältinud raseduse katkestusi.

Uurimus on valminud tänu kahele grandile:

1. Rakenduslik rahvastiku-uuring "Sünnituse dünaamika Eestis" ja
2. Rahvatervise sihtprogramm "Eestlaste seksuaal- ja reproduktiivkäitumine".

Kroonika

DOKTORITÖÖ BIOSTATISTIKAST

Krista Fischer. *Struktuurse keskmise mudelid tegeliku ravi mõju analüüsil kliinilistes katsetes.* Juhendaja prof Els Goetghebeur

Tartu Ülikooli matemaatikateaduskonna nõukogu komisjoni ees kaitses 25. juunil 1999 oma doktoritööd Krista Fischer. Pärast bakalaureuseõpingute lõpetamist Tartus 1992. aastal T. Kollo juhendamisel jätkas ta õpinguid magistriõppes Belgias Limburgi Ülikoolis, kus kaitses biostatistika alal magistrikraadi prof Jim Lindsey juhendamisel 1994. aastal. Samal aastal astus ta doktorantuuri, kus jätkas töötamist kliiniliste katsete alal Genti Ülikooli prof Els Goetghebeuri käe all. Krista Fischer arenes kiiresti iseseisvaks uurijaks, kes võttis südameasjaks biostatistika ja iseäranis meditsiinistatistika arendamise Eestis ning lisaks tööle väitekirjaga valmistas ette uusi loengukursusi ja konsulteeris meedikuid nende uurimistöös. Väitekirjas uuris ta ühte võimalikku J. M. Robinsi poolt sissetoodud nn struktuurse keskmise mudelit tegeliku ravi mõjude hindamisel katstes, kus kontrollgrupile oli määratud plattsebo.

Töös tuletas ta tema poolt vaadeldava lineaarse mudeli parameetritele hinnangu maatrikskujul ja tõestas selle hinnangu asümptootilise nihketuse ja efektiivsuse. Suurt rõhku oli töös pandud meetodi rakendatavuse uurimisele: töötati välja struktuurse keskmise mudeli õigsuse kontrolli meetodid, võrreldi saadud tulemusi kirjanduses kasutatavate teiste meetoditega ja korraldati ulatuslikke simuleerimiskeskkondi. Tööst selgub, et tuntud USA statistikute B. Efroni ja D. Feldmani varem kasutusele võetud meetod jääb oma võimalustelt ja omadustelt tunduvalt alla Krista Fischeri arendatavale Robinsi meetodile. Väitekirja tulemuste eest pälvis Krista Fischer kõrge rahvusvahelise tunnustuse: 1997. aastal võitis ta Thomas Chalmersi nimelise üliõpilaspriimi Bostonis toimunud kliiniliste katsete ühingu ja rahvusvahelise kliinilise biostatistika ühingu teisel ühiskonverentsil.

Oponentide, maailma ühe juhtiva biostatistiku, Kopenhaageni Ülikooli professori Niels Keidingu ja professor Ene-Margit Tiidu arvamused olid üksmeelsed: tegemist on tööga, mis vastab igati doktoritöö nõuetele. Kaitsmiskomisjon omistas Krista Fischerile filosoofiadoktori kraadi (PhD) matemaatika alal.

DOKTORITÖÖ TÖENÄOSUSTEORIAS

Jüri Lember. *Empiiriliste k -tsentrite mõjus. Juhendaja prof K. Pärna*

9. septembril 1999 kaitses Tartu Ülikooli matemaatikateaduskonna nõukogu komisjoni ees doktoriväitekirja "Empiiriliste k -tsentrite mõjus" Jüri Lember.

Jüri Lember lõpetas TÜ matemaatikateaduskonna 1992. aastal ja astus samal aastal magistriõppesse matemaatilise statistika erialal. 1994. aastal valminud magistritöös (juhendaja K. Pärna) uuriti juhusliku suuruse hajuvuse kirjeldamist k -dispersiooni abil ja selle raames tekkis esimene kokkupuude ka tulevase doktoritöö teemaga. Kui tavaline dispersioon näitab juhusliku suuruse hajuvust ümber ühe punkti (keskväärtuse), siis k -dispersioon mõõdab hajuvust ümber k punkti. Need k punktid valitakse nii, et juhusliku suuruse kauguse ruut talle lähimast punktist oleks keskmiselt minimaalne. Sellist optimaalset k -punktilist hulka nimetatakse jaotuse k -tsentriks ja juhul kui $k = 1$, langeb see kokku tavalise keskväärtusega. Autor sai k -dispersioonide kohta huvitavaid tulemusi. Magistriõppe käigus ilmnis Jüri Lemberi huvi matemaatiliste probleemide vastu ja suurepärase anne neid lahendada. Üsnagi ootuspäraselt võitis Jüri Lember 1995. aastal Eesti parimale noormatemaatikule määratud Arnold Humala nimelise preemia.

Doktorantuuris asus Jüri Lember uurima k -tsentritega seotud küsimusi meetrilistes ruumides ja Banachi ruumides. Doktoritöös vaadeldakse olukorda, kus juhusliku suuruse teoreetiline jaotus P pole teada ja k -tsenter tuleb määrata jaotusest P võetud juhusliku valimi abil. Peatähelepanu on seejuures pööratud *mõjususe* küsimusele, st kas valimi põhjal leitud nn empiirilised k -keskmised koonduvad (Hausdorffi mõttes) jaotuse P enda k -tsentriteks, kui valimi maht piiramatult kasvab. Seega on tegemist hästituntud suurte arvude seaduse üldistusega. Autor suutis tõestada empiiriliste k -tsentrite mõjususe Jefimov-Stečkini ruumides ja suhteliselt nõrkadel eeldustel jaotuse kandja kohta ka üldisemates ruumides. Oluline saavutus on ka see, et ta näitas k -tsentrite ülesande seost juba põhjalikult läbitöötatud parima lähenduse teooriaga. Saadud tulemused lisavad senituntule palju uut ja huvitavat oma üldisuse ja samas ka sügavuse poolest.

Doktoritöö kaitsmine kulges edukalt ja oponentid prof Alfredas Račkauskas Vilniuse Ülikoolist ja prof Riho Lepp Tallinna Tehnikaülikoolist leidsid üksmeelselt, et Jüri Lember väärib matemaatikadoktori kraadi.

Jüri Lember on osalenud mitmesugustes pikemates koolitusprogrammides ja seminaridel Hispaanias, Prantsusmaal, Itaalias, Norras ja Soomes, pälvides mitmel korral eraldi tunnustust ürituste korraldajatelt. Oma teadustöö tulemusi on ta tutvustanud konverentsidel ja seminaridel Leedus, Ungaris, Hollandis, Eestis. Tema sulest on ilmunud 10 teaduspublikatsiooni. Praegu töötab Jüri Lember Tartu Ülikooli matemaatilise statistika instituudis teadurina, kuid juba lähiajal siirdub ta kaheks aastaks tööle Hollandisse. Nimelt valiti Jüri Lember Eindhovenis asuva uurimisinstituudi EURANDOM kaastöötajaks, kusjuures osa tema tööülesannetest on seotud ka Philipsi firmaga.

KAITSTUD BAKALAUREUSETÖÖD 1999–2000

Kaitstud bakalaureusetööd 1999

1. **Jelena Boldõreva.** *Dispersioonanalüüsi segamudelite uurimine statistilise modelleerimise abil.* Juhendaja T. Nahtman.
2. **Marina Jagolnikova.** *Maksimaalse entroopia printsiip ja mõned alumised tõkked entroopiale.* Juhendaja K. Pärna.
3. **Inga Jakoreva.** *Reliaabluse hindamine pakettidega STATISTICA, SAS ja SPSS.* Juhendaja E. Käärik.
4. **Natalja Jurevitš.** *Klasteranalüüsi hierarhilised meetodid.* Juhendaja A.-M. Parring.
5. **Ergo Jõepere.** *Kõrguse ja rinnasdiameetri päritavuse uurimine harilikul männil.* Juhendaja T. Möls.
6. **Andres Kütt.** *Eesti aktsiaturu statistiline analüüs: börsikrahi mõjud.* Juhendaja K. Pärna.
7. **Margit Nerman.** *Statistikat EXCEL-i diagrammide abil.* Juhendaja A.-M. Parring.
8. **Artur Novek.** *Rahvastikustatistika sõnastik.* Juhendaja E.-M. Tiit.
9. **Tatjana Obidina.** *Valikudisaini sõltuvuskarakteristikud.* Juhendaja I. Traat.
10. **Žanna Petritševa.** *Juhuslike elementide ruumid.* Juhendaja J. Lember.
11. **Margus Pihlak.** *Mitmemõõtmeliste jaotusfunktsioonide lähendamine.* Juhendaja T. Kollo.
12. **Heti Pisarev.** *Erinevate meetodite võrdlus puude aastarõngaste laiuse prognoosimisel.* Juhendaja K. Fischer.
13. **Virgi Puusepp.** *Sõltuvusstruktuuride analüüs nulle sisaldavate korrelatsioonimaatriksite abil.* Juhendaja E.-M. Tiit.
14. **Annika Randmets.** *Patsiendi koostöö kliinilises katses ja puuduvad andmed.* Juhendaja K. Fischer.
15. **Inge Ringmets.** *Segu komponentide määramine graafiliste meetoditega.* Juhendaja T. Kollo.

16. **Katre Seema.** *Regressioonikordajate asümptootiline jaotus.* Juhendaja A.-M. Parring.
17. **Siret Tamm.** *Diskriminantanalüüs. Segamudelid.* Juhendaja S. Koskel.
18. **Jelena Tšernõšova.** *Täielike ja osaliste järjestuste korrespondentsanalüüs.* Juhendaja K. Pärna.
19. **Eva Tühis.** *Black-Scholes'i mudeli kasutatavusest optsioonide hindamisel Eesti tingimustes.* Juhendaja R. Kangro.
20. **Ragnar Õun.** *Eesti rahvastiku muutused XX sajandi esimesel poolel ja mittetäpselt teadaolevate rahvastikusündmuste hindamine.* Juhendaja E.-M. Tiit.

Kaitstud bakalaureusetööd 2000

1. **Svetlana Bizjajeva.** *Kaks-korda-kaks sagedustabelite analüüsivõimalused.* Juhendaja E. Käärik.
2. **Julia Bõtšinskaja.** *Muutuva intensiivsusega järjekorrasüsteemi optimeerimine.* Juhendaja K. Pärna.
3. **Natalja Jedomskihh.** *Ettevõtete finantsnäitajate prognoosimine.* Juhendaja A.-M. Parring.
4. **Liile Jõgi.** *Eesti aktsiaturu parameetrite hindamisest.* Juhendaja R. Kangro.
5. **Eva Kessler.** *Ravimi efektiivsuse hindamine sagedustabelite põhjal.* Juhendaja K. Fischer.
6. **Tatjana Kirilina.** *Olekuruumi mudelid ja protseduur STATESPACE (vene keeles).* Juhendaja M. Viil.
7. **Kersti Kriisk.** *Gini kordaja ja tema hinnangu jaotus.* Juhendaja E.-M. Tiit.
8. **Annika Krutto.** *Pensionikindlustuse matemaatilised mudelid.* Juhendaja T. Kollo.
9. **Jüri Kuzmin.** *Üldistatud lineaarsed mudelid. Rakenduslik näide sotsioloogilisele andmestikule.* Juhendaja E.-M. Tiit.
10. **Kristi Lehto.** *Eesti laste antropoloogiliste mõõtude kasvumudelid esimese 13 elukuu vältel.* Juhendaja E.-M. Tiit.

11. **Aimar Puru.** *Statistika seletav sõnastik.* Juhendaja M. Vähi.
12. **Kristiina Rajaleid.** *Järjestusvalik.* Juhendaja I. Traat.
13. **Eneli Reimets.** *Tehingusignaalide statistiline analüüs.* Juhendaja K. Pärna.
14. **Anne Reitsak.** *Jaotuse keskring.* Juhendaja K. Pärna.
15. **Anne-Liis Rämson.** *Aktsiahindade päevasisene prognoosimine.* Juhendaja K. Pärna.
16. **Dmitri Šestopalov.** *Programm ARLK-protsesside ja ülekandefunktsioonimudelite korrelatsioonifunktsioonide arvutamiseks (vene keeles).* Juhendaja M. Viil.
17. **Harry-Anton Talvik.** *Simulatsioonprogramm "KENO-LOTO".* Juhendaja K. Pärna.
18. **Marek Tuul.** *Lüklilike rändevaatluste statistiline analüüs hallhane (Anser anser) näitel.* Juhendaja T. Möls.
19. **Monika Unt.** *Hinnangud tagasipanekuga valikudisaini korral.* Juhendaja I. Traat.

KAITSTUD MAGISTRITÖÖD 1999–2000

Kaitstud magistritööd 1999

1. **Agne Soome.** *Rahvastikustatistika õpik.* Juhendaja E.-M. Tiit.
2. **Birgit Strikholm.** *Ajas muutuva sesoonsuse modelleerimine kvartaalsetes tööstustoodanguindeksi ridades.* Juhendajad T. Teräsvirta (Stockholm) ja M. Viil.

Kaitstud magistritööd 2000

1. **Merle Kurvits.** *Normaalne Gaussi pöördjaotus.* Juhendaja T. Kollo.
2. **Meelis Käärrik.** *Jaotuse lähendamine tõkestatud hulkadega.* Juhendaja K. Pärna.
3. **Kaire Ruul.** *Korrelatsioonimaatriksi jaotuse empiiriline analüüs.* Juhendaja T. Kollo.

6. TARTU MITMEMÕÕTMELISE STATISTIKA KONVERENTS

Kalev Pärna
Tartu Ülikool

6. Tartu mitmemõõtmelise statistika konverents peeti 19.–23. augustini 1999. Konverentsil käsitleti niihästi matemaatilise statistika teoreetilisi küsimusi kui ka statistika rakendusi majandusuuringutes, sotsiaal- ja eluteadustes, sh kasvukõverate analüüsimisel, kliinilistes uuringutes ja aretusmudelite väljatöötamisel. Ettekannetega osavõtjaid oli 45, sealhulgas 30 väliskülalist 18 riigist. Konverentsi peakorraldaja oli Tartu Ülikooli matemaatilise statistika instituut (MSI).

Tartu mitmemõõtmelise statistika konverentside seeriale pandi alus 1977. aastal. Sellest peale on konverentse peetud 4–5-aastase vaheaja järel, algul üleliiduliste konverentsidena, alates 1994. aastast rahvusvahelistena. Seekord tekkis idee siduda 6. Tartu konverents Helsingis toimuva ISI (*International Statistical Institute*) 52. sessiooniga (10.–18. august 1999). See võimaldaski paljudel tuntud statistikutel tulla Helsingist Tartusse. Kutsutud esinejate ettekanded olid erakordselt kõrgetasemelised. Järgnevas teemegi juttu eelkõige tellitud ettekannetest, mis annab teatava läbilõike ka konverentsi temaatikast.

Eriti esinduslik oli ettekandjate koosseis aegridade valdkonnas. Kointegratsioon (ehk mittestatsionaarsete aegridade ühendamine statsionaarseks sobiva lineaarkombinatsiooni abil) on aegridade uurimisel üks keskseid meetodeid. Mitmemõõtmelise analüüsi klassik Theodore W. Anderson (Stanford) kõneles kointegreeritud aegridade statistilise analüüsi ja kanoonilise korrelatsiooni vahelistest seostest. Soren Johanseni (Firenze) ettekanne oli pühendatud kointegreerivate seoste arvu testimisele. James Durbini (London) teemaks oli aga tema enda väljaarendatud olekuruumi (*state space*) mudeli kasutamine mitmemõõtmeliste aegridade analüüsimiseks mitmesugustes olukordades.

Klassifitseerimise vallas oli pikem ettekanne Helmuth Strasserilt (Viin), kes tutvustas ideed, mille kohaselt suurte andmekogude puhul võiks enne statistiliste otsuste tegemist andmed kvantida, st asendada väiksema arvu lähisväärtustega. Muni Srivastava (Toronto) käsitles regressioonanalüüsi prob-

leeme, mis tekivad juhul, kui sõltuv tunnus on mitmemõõtmeline. Juhuslike maatriksitega seotud arvutusi demonstreeris Heinz Neudecker (Amsterdam). Konverentsi viimasel päeval tutvustas Kai-Tai Fang (Hongkong) kuulajatele katseplaane, kus katsepunktid täidavad katsepiirkonna ühtlaselt.

Kümmekond ettekannet tartlastelt peegeldasid nii siin tehtavat mitmekesisist teoreetilist uurimistööd statistika ja tõenäosusteooria vallas kui ka statistika rakendusi eri elualadel (K. Fischer, T. Kollo, E. Käärik, M. Vähi, J. Lember, A.-M. Parring, K. Pärna, E.-M. Tiit, I. Traat jt).

Ülevaate tehtud ettekannetest saab äsjailmunud kogumikust kirjastustelt TEV/VSP seerias *New Trends in Probability and Statistics: Multivariate Statistics, Proceedings of the 6th Tartu Conference, Tartu, Estonia, 19–22 August 1999*. Eds.: T. Kollo, E.-M. Tiit, M. Srivastava. *TEV/VSP, Vilnius/Utrecht, 2000, pp 223*.

Konverents toimus kahe ISI alamorganisatsiooni — Bernoulli Ühingu ja IASC egiidi all. Programmikomitee koosseis: T. Teräsvirta (Stockholm) — esimees Bernoulli Ühingu esindajana, T. Kollo (Tartu) — aseesimees, H.-H. Bock (Aachen), H. Caussin (Prantsusmaa), D. von Rosen (Uppsala), A. Satorra (Barcelona). Programmikomitee sisuline töö jäi põhiliselt T. Kollo kanda, kes pidas läbirääkimisi ka kutsutud esinejatega. Konverentsi korraldustoimkonna esimeheks oli K. Pärna ja sekretäri kohuseid täitis edukalt E. Puhang. Korraldustoimkonnas olid kõik MSI töötajad ning ülesandeid jätkus ka kraadiõppuritele ja üliõpilastele. Konverentsi korraldamise kogemusi said juurde ligikaudu 30 inimest.

Rahvusvaheliste statistikakonverentside seeriat on kavas edaspidi kindlasti jätkata, sest koostöösidemed teiste maade uurijatega on teaduses väga olulised.

EESTI STATISTIKASELTSI ÜLDKOOSOLEK

Eesti StatistikaSeltsi (ESS) üldkoosolek peeti 2000. aasta 8. novembril konverentsi "Rahvastikustatistika ja registrid" esimese tööpäeva lõpul Tallinnas Rahvusraamatukogu konverentsisaalis algusega kell 16.30.

Üldkoosoleku päevakord

1. ESSi presidendi ülevaade ESSi tegevusest aastail 1997–2000.
2. ESSi põhikirja muutmine osalusmaksu osas.
3. ESSi uue juhatuse ja presidendi valimine.

Koosolekut juhatas Tõnu Kollo, protokollis Mare Vähi.

Üldkoosolekul esines ülevaatega senitehtust seltsi senine president E.-M. Tiit. Ajavahemikul 1997–2000 toimus Eesti StatistikaSeltsi korraldusel neli vabariiklikku ja üks rahvusvaheline konverents ning avaldati seitse trükist, neist kolm ESSi teabevihikud. Kuni 1998. aastani suudeti konverentse korraldada regulaarselt ja ka avaldada teabevihikuid. Aruandeaaja viimasel perioodil on seltsi tegevus olnud suhteliselt väheaktiivne. Esialgu aastaks 1999 kavandatud konverents lükkus 2000. aastasse. Ühelt poolt oli selle põhjuseks rahvusvaheline konverents Tartus ja teiselt poolt rahvaloenduse ettevalmistamine Tallinnas. Tartus toimunud keskkonnateemalise konverentsi materjalid on senini trükis avaldamata (on küll valminud Interneti-versioon), oluliseks põhjuseks on siin ressursside nappus.

Aruandeperioodi jooksul selts registreeriti ümber ja kinnitati uus põhikirja. (Põhikirja viidi vastavusse Eesti seadustega.)

Peale selle asutati aruandeperioodil ESSi veebileht ja *list* info levitamiseks.

Presidendi ettekandele järgnes arutelu.

Põhikirja muudatuse suhtes võttis sõna professor Mati Rahu, kes tegi ettepaneku suurendada senist ESSi liikme 5-kroonist aastamaksu. Ettepanek:

- regulaarsete liikmete maks — 25 krooni aastas;
- üliõpilaskliikmete maks (sh ka magistrandid ja doktorandid) — 5 krooni aastas;
- sisseastumismaks on võrdne aastamaksuga, sisseastumisaastal aastamaksu ei võeta.

Ettepanek võeti vastu.

EESTI STATISTIKASELTSI ÜLDKOOSOLEK

Eesti Statistika Seltsi (ESS) üldkoosolek peeti 2000. aasta 8. novembril konverentsi "Rahvastikustatistika ja registrid" esimese tööpäeva lõpul Tallinnas Rahvusraamatukogu konverentsisaalis algusega kell 16.30.

Üldkoosoleku päevakord

1. ESSi presidendi ülevaade ESSi tegevusest aastail 1997–2000.
2. ESSi põhikirja muutmine osalusmaksu osas.
3. ESSi uue juhatuse ja presidendi valimine.

Koosolekut juhatas Tõnu Kollo, protokollis Mare Vähi.

Üldkoosolekul esines ülevaatega senitehtust seltsi senine president E.-M. Tiit. Ajavahemikul 1997–2000 toimus Eesti Statistika Seltsi korraldusel neli vabariiklikku ja üks rahvusvaheline konverents ning avaldati seitse trükist, neist kolm ESSi teabevihikud. Kuni 1998. aastani suudeti konverentse korraldada regulaarselt ja ka avaldada teabevihikuid. Aruandeaaja viimasel perioodil on seltsi tegevus olnud suhteliselt väheaktiivne. Esialgu aastaks 1999 kavandatud konverents lükkus 2000. aastasse. Ühelt poolt oli selle põhjuseks rahvusvaheline konverents Tartus ja teiselt poolt rahvaloenduse ettevalmistamine Tallinnas. Tartus toimunud keskkonnateemalise konverentsi materjalid on senini trükis avaldamata (on küll valminud Interneti-versioon), oluliseks põhjuseks on siin ressursside nappus.

Aruandeperioodi jooksul selts registreeriti ümber ja kinnitati uus põhikirja. (Põhikirja viidi vastavusse Eesti seadustega.)

Peale selle asutati aruandeperioodil ESSi veebileht ja *list* info levitamiseks.

Presidendi ettekandele järgnes arutelu.

Põhikirja muudatuse suhtes võttis sõna professor Mati Rahu, kes tegi ettepaneku suurendada senist ESSi liikme 5-kroonist aastamaksu. Ettepanek:

- regulaarsete liikmete maks — 25 krooni aastas;
- üliõpilaskliikmete maks (sh ka magistrandid ja doktorandid) — 5 krooni aastas;
- sisseastumismaks on võrdne aastamaksuga, sisseastumisaastal aastamaksu ei võeta.

Ettepanek võeti vastu.

Juhatus otsustati valida 6-liikmeline avalikul hääletamisel ja üheksast kandidaadist osutused valituteks (tähestikjärjestuses):

- Jarno Habicht, Tartu Ülikool,
- Tõnu Kollo, Tartu Ülikool,
- Ülo Randaru, Statistikaamet,
- Ebu Tamm, Statistikaamet,
- Ene-Margit Tiit, Tartu Ülikool,
- Ragnar Õun, Siseministeerium.

Eesti Statistikaamet presidendiks valiti Tõnu Kollo, kes edestas napilt Ebu Tamme.

Koosolek lõppes kell 17.45.

Kirjastanud	Statistikaamet, Endla 15, 15174 Tallinn
Trükkunud	OÜ Multico, Endla 15, 15174 Tallinn
Tellimuse nr	Aprill 2001 176

RAHVASTIKUSTATISTIKA JA REGISTRID

Mida saime/saame teada rahvaloendusest?
Kuidas teha rahvastikustatistikat pidevalt?
Kuidas pidada ja kasutada registreid?
Isikuandmed Eestis – kellele kuuluvad
ja kui vabalt saab neid kasutada?

*Neid ja paljusid teisi küsimusi käsitleti konverentsil
"Rahvastikustatistika ja registrid" 8.–9. novembril
Rahvusraamatukogus, kõigest sellest loete
ka käesolevast kogumikust.*