

Teavevihik

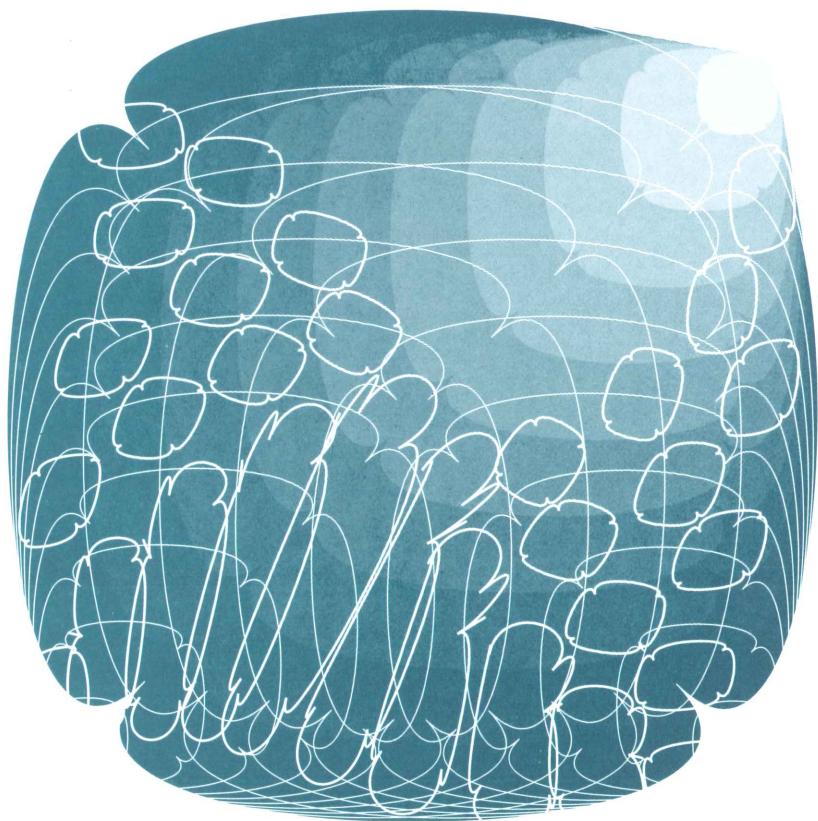
Tartu 2003

13



EESTI STATISTIKASELTS

STATISTIKAMEETODID ELUTEADUSTES





Eesti Statistika Seltsi Teabevihik

**STATISTIKAMEETODID
ELUTEADUSTES**

Tartu 2003

Toimetaja: Ene-Margit Tiit

Kaanekujundus: Michael Walsh

ISSN 1406-314X

ISBN 9985-9168-5-9

© Eesti Statistikeselts, autorid, 2003

Sisukord

Eesti Statistikaltsi 14. konverents.....	5
Tiia Bakler, Andres Kink, Rein Teesalu. Nüüdisaegne meditsiiniliste andmete kogumise süsteem (müokardiinfarktiregistri näol)	8
Agu Tamm. Laboratoorsete andmete alusel tehtavate otsustuste tüübid.....	14
Ülle Kirsimägi, Aleksander Lõhmus, Jaanus Kahu, Madis Ilmoja, Tiia Arro. Neerusiirdamine Eestis 1972–2002.....	23
Kaja Rahu. Kohortuuring epidemioloogias: teooria ja praktika ..	32
Helje Kaarma, Säde Koskel. Tartu noorte meeste ja naiste kehastruktuur.....	40
Ene Käärrik, Ene-Margit Tiit. Sündimuskõverate lähendamisest	50
Mart Kals, Eda Merisalu, Krista Fischer. Eesti koolijuhitide tööstress ja läbipõlemine.....	57
Triin Umbleja, Kersti Pärna, Krista Fischer. Rahvusvahelise koolinoorte suitsetamisuuringu mitmetasandilised mudelid....	67
Artur Nilson. Regressioonivõrrandite lineariseerimine.....	77
Artur Nilson. Uus asümptootiline funktsioon.....	87
Kalle Remm. Näidiste abil klassifitseerimine ja prognoos	95
Mati Rahu. Ühe eluteaduse lainetel: konverentsijärgseid mõttemõlgutusi.....	107
Ene-Margit Tiit. Eesti statistikaltsi kümme tegevusaastat....	118
Kroonika.....	120

SAATEKS

Kaante vahele on saamas järjekordne ESS Teabevihik, sedapuhku järjekorranumbriga 13. Võrreldes esimese vihikuga on kümne võrra suurenenud aastaarv ja märksa soliidsemaks muutunud nii kujundus kui ka välimus. Jätkunud on aga traditsioon – teabevihiku põhisisu moodustavad ESS konverentsi ettekanded. Käesolev vihik on pühendatud 2002. aasta hilissügisel Otepää külje all toimunud konverentsile ”Statistika-meetodid eluteadustes”. Eluteadusi mõistsime kõige laiemas tähenduses – konverentsil kõneldi nii meditsiinist ja käitumisprobleemidest kui ka metsa kasvust, arutati sündimus- ja geneetikamudeleid. Olulise koha ettekannete seas hõivasid arutelud, mis on seotud meditsiini-andmestiku kasutamise ja teadusuuringutes ja praktilistes otsustustes.

Ettekannete arvu ja taseme poolest oli see konverents üks sisutihedamaid statistikaselti ajaloos. Kahest tööpäevast kippus nappima, nii et seksioonide juhatajatel tuli ettekandjate suhtes karmi kätt rakendada – hoolimata sellest, et suurema osa ettekannete puhul rakendati nüüdisaegset tehnoloogiat ning tahvlile kirjutamiseks ja selle pühkimiseks aega ei kulunud. Marguse spordibaasi perenaised imestasid – mis seltskond see küll on, kes aina ettekandeid peab ja kuulab, ei mingit pidu ...

Käesolevas kogumikus on esindatud vaid kolmandik konverentsil peetud ettekannetest. Et ESS Teabevihikut (veel!) *Current Content*'is ei refereerita, jäi autoritel üks stiimul kirjutamiseks ära – ”linnuke”, mis siin avaldatud artikli eest kirja saadakse, pole eriti rasvane. ESS Teabevihik on mõeldud ju ainult eesti lugejale. Kui aga samast teemast oli eesti keeles juba kirjutatud – ajakirju ja muid väljaandeid ilmub Eestis praegu palju – ei olnud ESS toimetus kordusartiklist huvitatud.

Peale otseselt konverentsil peetud ettekannete üllatasid mõned konverentsil osalenud autorid toimetust sümfaatsete lisapaladega – need on Artur Nilsoni teine artikkel ja Mati Rahu suleosav mõttemõlgutus. Traditsiooniliselt jätkab teabevihik ka statistika kroonika esitamist, käesolevas numbris on kajastatud Eesti statistikaelu aastail 2001–2002.

E.-M. Tiit

EESTI STATISTIKASELTSI 14. KONVERENTS

STATISTIKAMEETODID ELUTEADUSTES

1.–2. nov. 2002 EPMÜ Marguse spordibaasis Valgamaal
Nüpli külas

Konverentsi programm

Reede, 1. november

11.00–12.00. *Saabumine Marguse spordibaasi, registreerimine*

12.00–13.30. Avaplenaaristung. Juhataja: E. Tamm, ESA

12.00. *Avasõna* — T. Kollo, ESSi president, TÜ

12.10. Tšernobõl ja tervis: hirmud, kuuldused, tõde — M. Rahu, EKMI

12.40. Tervishoiustatistika hetkeseis ja arengud — J. Habicht, E. Palo,
Sotsiaalministeerium

13.05. Põhjuslikud mõjud ja nende hindamine — K. Fischer, TÜ

13.30–14.30. *Lõuna*

14.30–16.30. II istung. Statistikameetodid arstiteaduslikes uuringutes. Juhataja: T. Kollo, TÜ

14.30. Kaasaegne meditsiiniliste andmete kogumise süsteem —
A. Kink, T. Bakler, R. Teesalu, TÜ

14.45. Ansamblikeskmisest lainekujust südamegevuse signaalide
analüüsil — J. Vedru, O. Makarova, TÜ

15.00. Vahend tervisega seotud elukvaliteedi hindamiseks — A. Luk-
mann, K. Pärna, J. Maaros, TÜ

15.15. Eesti koolijuhtide tööstress ja tervis — K. Fischer, E. Merisalu,
M. Kals, TÜ

15.30. Haigestumusest Eesti Haigekassa andmetel — M. Vähi, TÜ,
M. Thetloff, Haigekassa

- 15.45. Neerusiirdamised Eestis 1972—2002. Elukestusanalüüs — Ü. Kirsimägi, T. Arro, A. Lõhmus, TÜ
- 16.00. Seos vere üldkolesterooli ning agressiivse ja hüperaktiivse käitumise vahel kooliõpilastel. — K. Rüütel, M. Harro, K. Fischer, TÜ
- 16.15. Otsustused laboratoorsete andmete baasil. — A. Tamm, TÜ

16.30–17.00. *Kohv*

17.00–19.00. III istung. Statistikameetodid antropomeetria-uringutes ja andmekvaliteet. Juhataja: M. Rahu, EKMI

- 17.00. Tütarlaste ja naiste kehaehitusstruktuur — G. Veldre, J. Peterson, TÜ
- 17.15. Kutsealuste kehastruktuur — J. Lintsi, L. Saluste, TÜ
- 17.30. Tartu noorte meeste ja naiste kehastruktuur — S. Koskel, H. Kaarma, TÜ
- 17.45. Vastsündinu kaalu prognoosimudelid — K. Ohvril
- 18.00. Loote kasvu seaduspärasused UH- mõõtmiste põhjal — K. Õun, TÜ
- 18.15. Andmekvaliteet vähiregistris — K. Lang, TÜ
- 18.30. Diskussioon

19.00–20.00. *Õhtusöök*

20.00–21.30 Ümarlaud “Millist abi ja missuguseid uusi lahendusi ootavad eluteadlased statistikutelt?” Juhatavad: Kalev Pärna, K. Fischer (TÜ)

Laupäev, 2. november

8.00–9.00. *Hommikusöök*

9.00–11.00. IV istung. Statistikameetodid geneetikas ja rahvastikuteaduses. Juhataja: Ü. Randaru, ESA

- 9.00. Parameetrilised ja mitteparameetrilised *geenijahtimismeetodid* — T Kaart, EPMÜ
- 9.15. Geneetilise progressi suurendamine veistel — E. Pärna, EPMÜ

- 9.30. Mikroevolutsiooni kompuutermodeleerimisest — M. Viikma, TÜ
 9.45. Sündimuskõverate lähendamine — E. Käärik, E.-M. Tiit, TÜ
 10.00. Kohortuuring. Teooria ja praktika — K. Rahu, EKMI
 10.15. Mitmetasandilised mudelid rahvusvahelises koolinoorte suitsetamise uuringus — T. Umbleja, Kersti Pärna, K. Fischer, TÜ
 10.30. Uuringu planeerimisest meditsiinilis-käitumuslikes uuringutes — E.-M. Tiit, TÜ
 10.45. Bioinformaatika aktuaalsetest probleemidest — J. Vilo, TÜ
 11.00–11.30. *Kohv*

11.30–13.30. V istung. Statistikameetodid looduse- ja keskkonnauuringutes. Juhataja: E.-M. Tiit, TÜ

- 11.30. Eestit kattev metsa kasvukäigu püsiproovittikkide võrgustik — A. Kiviste, EPMÜ
 11.45. Regressioonivõrrandite lineariseerimisest ja ühest asümptootilisest funktsioonist — A. Nilson, EPMÜ
 12.00. Paljumõõtmeline statistiline analüüs geobotaanikas — J. Paal, TÜ
 12.15. Mustrite leidmine taimkattes — J. Liira, TÜ
 12.30. Diskussioon

13.30–14.30. Lõuna. Konverentsi lõpp külalistele

14.30–15.30. ESS üldkoosolek

- 14.30–16.00. Rahvusvahelise Biomeetria Seltsi Balti tööühma asutamine

NÜÜDISAEGNE MEDITSIINILISTE ANDMETE KOGUMISE SÜSTEEM (MÜOKARDIINFARKTI- REGISTRI NÄOL)

Tiia Bakler, Andres Kink, Rein Teesalu
TÜ Kardioloogiakliinik, tiia.bakler@kliinikum.ee

Sissejuhatus

Eestis puudub ühtne meditsiiniliste andmete kogumise infosüsteem, mis tagaks täpse meditsiinistatistika ning kogutavate andmete efektiivse kasutamise. Nii kliinilises auditis, kliinilises praktikas kui ka tervishoiusüsteemi juhtimises vajatakse usaldusväärset süstemaatilist informatsiooni. Nüüdisaegne infotehnoloogia võimaldab tervishoiuasutustest suure hulga kliiniliste andmete kogumist, säilitamist ja ühildamist, aitab kaasa andmete kättesaadavusele ning levitamisele. Äärmiselt oluliseks peetakse meditsiiniliste andmete analüüsimist ja kohapealset kättesaadavust lühikese ajavahemiku järel pärast iga juhu sisestamist [National, 1999]. Internetipõhised programmid, mis seda kõike pakuvad, loovad hea eelduse infosüsteemi (andmebaasi) aktiivseks kasutamiseks kliinistide poolt. Elektroonsete haiguslugude probleemideks arvatakse olevat puudulikult standarditud andmete kogumissüsteem ja erisugused arvutiprogrammid, puudulik kvaliteedi kontroll, konfidentsiaalsuse ja omandiõiguse huvide puudumine [Pronovost, Angus, 2000].

Käesolevas lühiartiklis kirjeldatakse internetipõhist müokardiinfarkti-registrit (MIR), mis on loodud SA Tartu Ülikooli Kliinikumi juurde, kui üks võimalus koguda andmeid tänapäeva infotehnoloogiat kasutades.

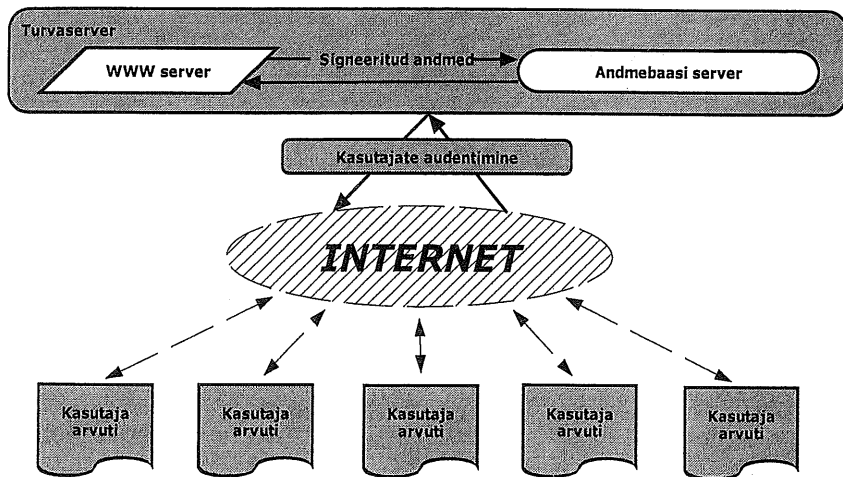
Vajadus

Südame-veresoonkonna haigused (sh müokardiinfarkt) moodustavad Eestis üle 50% kõigist surmapõhjustest [Sotsiaalministeerium, 2000]. Viimase 30 aasta jooksul on USAs ja Lääne-Euroopa riikides suurem vereringehaigustesse pidevalt langenud [Heart, 1999, Sans *et al*, 2000].

Ida-Euroopa maades, sealhulgas Eestis, niisugust tendentsi ei ilmne [Sans *et al*, 2000, Leinsalu, 1995]. Südame isheemiatõvest tingitud suremus on langenud eelkõige fataalsete ja korduvate müokardiinfarktide (MI) vähenemise tagajärjel. Need muutused tulenevad nii ägeda MI haiglaravi kui ka sekundaarse preventsiiooni meetmete (eelkõige riskitegurite modifitseerimine) efektiivsuse tõusust [Capewell *et al*, 2000, Salomaa *et al*, 1996, McGovern *et al*, 2001]. Eestis puudub informatsioon MI-ga seonduvatest aspektidest (haiguse esinemissagedus, ravi, ravitulemused jne). Seetõttu peeti vajalikuks luua müokardiinfarktiregister (MIR), mis täidaks järgmisi ülesandeid: 1) diagnoosi ühtlustamine; 2) haigestumusstatistika korraldamine; 3) ravikvaliteedi tõstmine; 4) arstide omavahelise suhtlemise tõhustamine; 5) MI-ga seonduva teadustöö edendamine. Eesmärkide täitmise tagatiseks on andmete pidev süstemaatiline tsentraliseeritud kogumine.

Metoodika

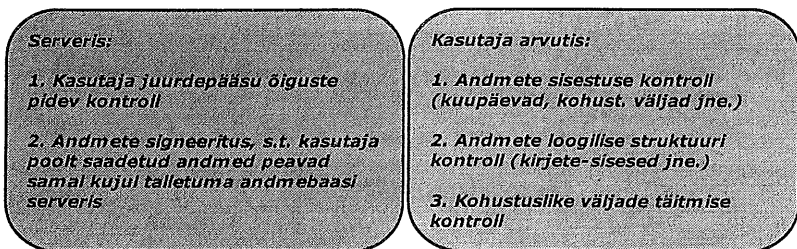
MIR on avatud kasutajaliidesega internetipõhine andmebaas, mille andmete kogumise süsteemi iseloomustab joonis 1. Andmebaasi saab siseneda igast arvutist, milles on olemas *Internet Explorer* või *Netscape*. Isikuandmete kaitseks kasutatakse nüüdisaegseid turvasüsteeme.



Joonis 1. Andmekogumissüsteem

MIRi kogutakse andmed Eesti elanikel diagnoositud MI-de (rahvusvahelise haiguste klassifikatsiooni versiooni 10 alusel I21-I22) kohta Eesti tervishoiuasutustest. Andmeid kogutakse standarditud elektroonilise vormi alusel, mis sisaldab 78 tunnust koos definitsioonidega ja on kinnitatud Eesti Kardioloogide Seltsi juurde kuuluva Eesti MIRi teadusnõukogu poolt. Definitsioonid on võimalikult lihtsad, et saada usaldusväärne üheselt mõistetav informatsioon. Tunnused hõlmavad järgmisi valdkondi: isikuandmed, riskitegurid ja eelnev kardiovaskulaarne anamnees, käsitletava atakiga seotud andmed (kulg, arstiabi kättesaamise kiirus), haiglaperioodil esinevad tüsistused, ravi ja protseduurid, laboritulemused, kliinilised diagnoosid, isiku surmatunnistuse diagnoosid ja lahanguandmed. Vormi täidab arst patsiendi väljakirjutamisel haiglast. Selle täitmine võtab aega 15–20 minutit. Kõik arstid, kes osalevad registri töös, on läbinud koolituse, kus tutvustatakse kogutavate tunnuste definitsioone ja registri organisatsiooni.

Kõigil vastavate kasutajaõigustega isikutel on võimalik igal momendil kirjeid lugeda ja lisada, kindla aja jooksul neid parandada ja täiendada. Andmete õigsuse tagamiseks rakendatavad kontrollimehhanismid (joonis 2) hõlbustavad vigaste kirjete identifitseerimist ning kvaliteetsemaks muutmist.

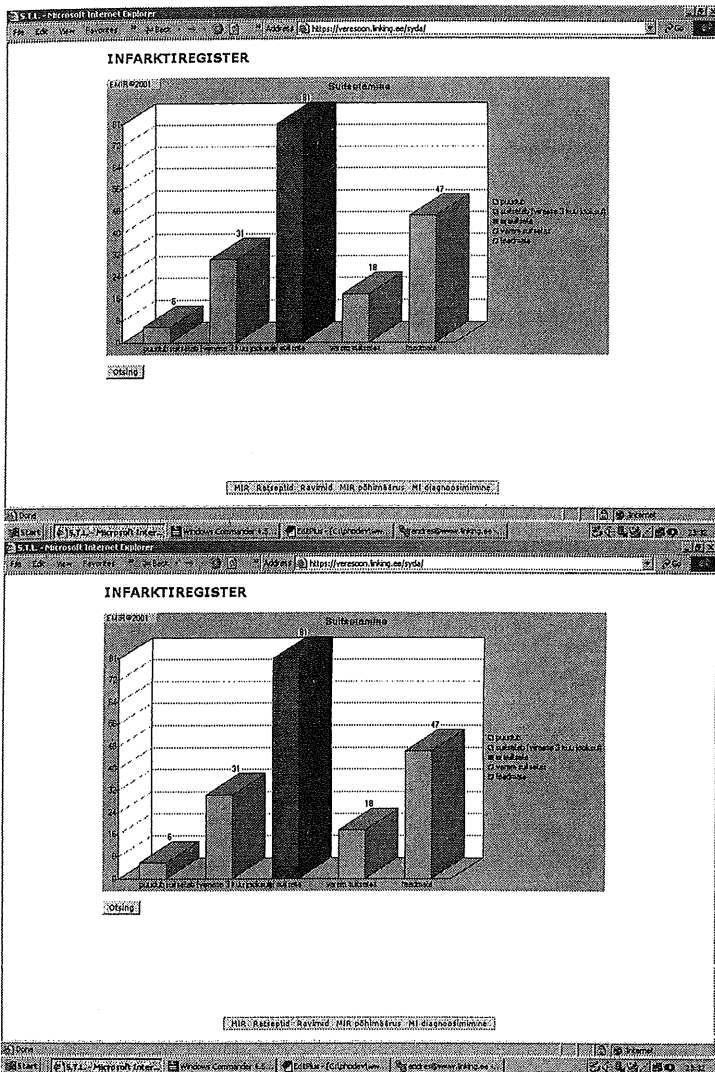


Joonis 2. Andmete õigsuse tagamiseks rakendatavad kontrollimehhanismid

Mida pakub MIR töös osalevatele institutsioonidele/isikutele?

- Kohesest juurdepääsu varem sisestatud elektroonsele infole patsiendi kohta, sõltumata raviuasutusest (on vaja teada patsiendi isikukoodi).
- Reaalajas tagasisidet graafikute näol (joonised 3a, 3b) võimalusega ise valida patsientide rühmad registreeritud tunnuste alusel ja ajavahemik.
- Võimalust kasutada andmeid aastaaruande tegemisel.

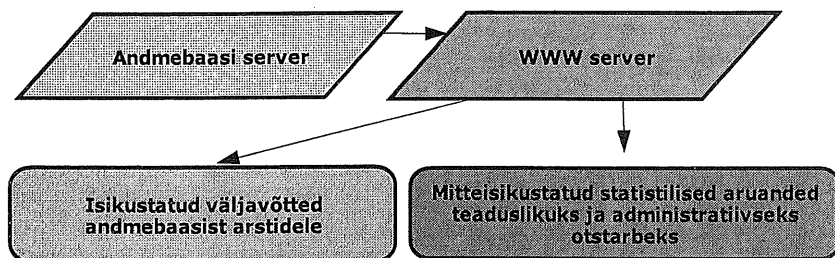
- Andmete ühisanalüüsi ja võrdlust vastavate (registri) üldkeskmistega.



Joonis 3a, 3b. Näited tagasiside moodulist

Iga kasutaja saab võrrelda oma tulemusi osakonna, haigla ja kogu registri keskmise tulemusiga, mis aitab kaasa arstil isenda töö audi-

teerimisele. Arstidele ja haiglatele on tagatud konfidentsiaalsus. Haiglad ei saa oma tulemusi võrrelda mingite teiste konkreetsete haiglate tulemustega, vaid ainult kogu registri keskmisega. Andmete käitlemise teid iseloomustab joonis 4.



Joonis 4. Andmete käitlemise teed

Üheks eeskujuks taolise infosüsteemi loomisel on olnud Rootsi kardioloogilise intensiivravi register (*The Swedish Register for Cardiac Intensive Care*), mis on alles alates 2001. aastast läinud internetipõhiseks. Sealse registri põhirõhk on ravikvaliteedi parandamisel, millele aitab kaasa väga professionaalsel tasemel väljatöötatud tagasisidesüsteem [National, 1999; Sternesand, Wallentin, 2000].

Hetkeseis ja tulevik

Infosüsteemi testimise ja andmete sisestamisega alustati jaanuaris 2001. Andmekogu vastutavaks töötajaks on SA TÜ Kliinikum ning volitatud töötajaks OÜ Linking, kes pakub elektroonilist andmetöötlusteenust. Registritöös osalevad tänase seisuga SA TÜ Kliinikum, SA Pärnu Haigla ja SA Viljandi Haigla. Kogutud kirjade arv ulatub 1000-ni. MIR sooviks tulevikus areneda üle-eestiliseks. Andmete kvaliteedi ja mitmekülgse analüüsi huvides on oluline luua lingid teiste sarnaste registritega (nt südamekirurgia, angiograafia jne) ning rahvastikuregistri, surmade andmebaasi, haigekassa andmebaasi ning haiglast lahkunute registriga.

Kokkuvõte

Käsitlev infosüsteem rahuldab väga olulisi nõudmisi andmete kättesaadavuses ja kasutatavuses. Samuti tekib infrastruktuur vajaliku infor-

matsiooni levitamiseks sihtgrupis (uued ravijuhised, uuringud jne). Meie kogemus näitas, et kogutavate andmete täielikkus ja usaldusväärsus tõusis eelkõige tänu tagasisidemoodulile. Selline infosüsteem loob hea eelduse kogu Eestit hõlmava ühtse andmebaasi loomiseks kvaliteetsete andmete kogumise eesmärgil, mida on võimalik kasutada laialdaselt ka teadustöös.

Kirjandus

1. National Health Care Quality Registries in Sweden 1999. <http://www.sos.se/mars/>
2. Pronovost, P., Angus, D. C. (2000). Using large scale databases to measure outcomes in critical care. CIMC. <http://www.spci.org/cimc2000/mesas/mr2/Angus2/databases.htm>
3. Sotsiaalministeeriumi haldusala arvudes 2000 (2000). Eesti Sotsiaalministeerium, Tallinn.
4. 2000 Heart and Stroke Statistical Update (1999). American Heart Association, Dallas, Texas.
5. Sans, S., Kesteloot, H., Kromhout, D. On behalf of the Task Force (2000). The burden of cardiovascular disease mortality in Europe. Task Force of the European Society of Cardiology on Cardiovascular Mortality and Morbidity Statistics in Europe. *Eur Heart J*, 21, 1141–51.
6. Leinsalu, M. (1995). Time trends in cause-specific mortality in Estonia from 1965 to 1989. *Int J Epid*, 24, 106–13.
7. Capewell, S., Beaglehole, R., Seddon, M., McMurray, J. (2000). Explanation for the decline in coronary heart disease mortality rates in Auckland, New Zealand, Between 1982 and 1993. *Circulation*, 102, 1511–6.
8. Salomaa, V., Miettinen, H., Kulasmaa, K. *et al.* (1996). Decline of coronary heart disease mortality in Finland during 1983 to 1992: roles of incidence, recurrence, and case fatality: The FINMONICA MI Register Study. *Circulation*, 94, 3130–7.
9. McGovern, P. G., Jacobs, D., Shahar, E. *et al.* (2001). Trends in acute coronary heart disease: mortality, morbidity, and medical care from 1985 through 1997: The Minnesota Heart Survey. *Circulation*, 104, 19–24.

LABORATOORSETE ANDMETE ALUSEL TEHTAVATE OTSUSTUSTE TÜÜBID

Agu Tamm

TÜ Sisekliinik, laboratoorse meditsiini õppetool

Raviasutuste laboratooriumide toodetavaid miljoneid arve võib arst kasutada vähemalt neljast strateegiast lähtuvalt. Nendele vastavaid lähenemisviise kirjeldab neli tüüpilist küsimust:

1. Kas tulemus on normaalne?
2. Kas tulemus erineb oluliselt sama isiku eelmise uuringu vastusest?
3. Kas selline väärtus, *resp.* patsient nõuab sekkumist, s.t ravi alustamist või muutmist?
4. Kas tulemusel on prognostilist tähendust?

Käesolevaks ajaks on kogunenud piisaval hulgal analüütilisi käsitlusi, mis võimaldavad ülaltoodud küsimustele vastata mitte ainult empiiriliselt, vaid ka lähtuvalt teoreetilistest alustest. Eeskätt on vaja selgesti määratleda/kirjeldada lähteandmeid, st baasi, millele tugineb üks või teine vastus. Siin on otstarbekas viidata ka mõningatele muutustele, mis on viimasel aastakümnel nende probleemide käsitluses aset leidnud.

Et vastata usaldusväärset uuringu tulemuse "normaalsuse" kohta, vajab hindaja usaldusväärseid *referents-* ehk *kontrollväärtusi*. Et otsustada, kas muutus sama isiku korduval uurimisel on kliiniliselt oluline, peab teadma *olulise muutuse* piiri (ingl *reference change limit*). Kas uuringu tulemuseks saadud väärtus nõuab viivitamatut sekkumist, see peab selguma korduvate kliiniliste vaatluste ja laboritulemuste kõrvutamisel. Üldistatult nimetatakse sekkumist nõudvaid väärtusi mitmeti, sealhulgas *kriitilisteks* või *tegutsemis-* või lihtsalt *otsustusväärtusteks*. Lõpuks on teada hulk laboratoorseid parameetreid, mille alusel hinnatakse suurema või väiksema usaldusväärusega haigusprotsessi kulgu või patsiendi prognoosi. Enamasti saab selliseid parameetreid kasutada siis, kui on teada teiste samasse klastrisse kuuluvate riskitegurite olemasolu.

Referentsväärtused

Juba 1980 aastatest alates soovitab Kliinilise Keemia Rahvusvaheline Föderatsioon (IFCC) mitte kasutada terminit "normaalne" kui vananenut ja ebasobivat. Sõna "normaalne" mõistavad erinevad spetsialistid vägagi erinevalt. Statistikas tähistab see termin Gaussi jaotust, kuid paljud kliinilis-laboratoorsed ja bioloogilised parameetrid ei jaotu nii: jaotus on ebasümmeetriline (ingl *skewed*), tavaliselt suuremate väärtuste poole välja venitatud, parempoolse asümmeetriaga (*raske sabaga*). Teiseks, pole alust aktsepteerida vaadet, nagu oleks otsene seos tulemuste normaalsuse ja tervise vahel. Ei tähenda ju mittenormaalne tulemus alati struktuuri ja/või funktsiooni häiret. Lõpuks, enamik seniseid "norme" on määratletud nii, et puudub adekvaatne info uuritud isikute või kasutatud määramistehnikate kohta. Seetõttu loodigi *referentsväärtuste* kontseptsioon, kus kõiki vajalikke etappe on selgelt iseloomustatud [Expert Panel, 1987].

Uue terminiga tähistatavate *referentsvahemike* puhul eeldatakse, et paljud olulised aspektid nagu referentsisikute valik, nende tervisliku seisundi määratlus, selle populatsiooni iseloomustus, keda nad esindavad, proovide kogumise süsteem, kasutatavad analüütilised tehnikad ja andmetöötlusmeetodid on selgesti defineeritud ja kirjeldatud. Sel juhul põhinevad laboratoorse parameetri kontrollpiirid tervete isikute adekvaatse suurusega rühma ühekordsel uurimisel. Kokkuleppeliselt jääb referentsvahemikku 95% uuritud populatsioonis leitud väärtustest. Seega ei välista üksiku indiviidi tulemuse paigutumine väljapoole referentsvahemikku tema kuulumist tervete hulka.

Üks tavalisi soovitusi, mis johtub ülaloodud põhimõtetest, on see, et iga labor määratleks ise oma analüütide kontrollvahemikud. Praktikas on selle soovitusi järgimine küllalt kulu- ja töömahukas, mistõttu pole sugugi harvad juhtumid, mil siiski opereeritakse kirjandusest leitud või meetodeid turustavate firmade soovitusetega. Järgnevalt kaks näidet üllatustest, mis võivad osaks saada, kui hakata võõraid allikaid kontrollima. Viimase viie aasta jooksul on allakirjutanu initsiatiivil täpsustatud Eesti populatsiooni referentsväärtusi rutiinses hematoloogias ja luuainevahetuse peamiste markerite jaoks. Leukotsüütide üldarv on kirjandusest 4×10^9 kuni 11×10^9 , oma uuringute põhjal naistel 4×10^9

kuni $9,4 \times 10^9$. Desoksüüpüridinoliini erituv hulk tervel naisel on tootja andmeil $< 7,4$ nmol/mmol kreat kohta, oma uuringute põhjal < 15 nmol/mmol kreat kohta. Esimesel juhul võimaldab meie oma tulemus diagnoosida kõrvalekallet varem, teisel juhul mitte rakendada põhjendamatu ravi.

Tulemuse (vähim) oluline muutus

Laboratoorseid tulemusi kasutataksegi raviasutuses kõige sagedamini selleks, et hinnata seisundi muutumist või püsimist samal inimesel. Sellise hinnangu andmiseks on vaja teada vähemalt määramismeetodi täpsust e reprodutseeritavust ning vastava parameetri bioloogilise varieeruvuse ulatust. Et kahe määramistulemuse erinevus oleks *analüütiliselt* 95% usaldusväarsusega veenev, peaks erinevus ületama 2,8 ($= 1,96\sqrt{2}$) korda mõõtmistulemuse (meetodi) standardhälbe SD. Et testi tulemuse muutust pidada *kliiniliselt oluliseks*, peab see samuti ületama 2,8 korda parameetri *koguvärvuse*, mida kirjeldab valem

$$SD_{KOGU} = \sqrt{SD_A^2 + SD_B^2},$$

kus SD_A ja SD_B on vastavalt vaadeldava parameetri analüütilist hajuvust ja indiviidi bioloogilist hajuvust iseloomustavad standardhälbed. Vähima olulise muutuse (VOM) mõned näited on esitatud tabelis 1.

Tabel 1. Näiteid vähima kliiniliselt olulise muutuse kohta

Analüüt	VOM	Tase	Muutus	Ühik
P-glük	+/- 20%	10	+/- 2	Mmol/L
S-K	+/- 15%	4	+/- 0,6	Mmol/L
LK	+/- 30%	9	+/- 2,7	$\times 10^9/L$
B-Hb	+/- 8%	120	+/- 10	G/L

Muutused on antud protsentides ja mõõtühikutes. Tähistused: P-glük — vereplasma glükoosi kontsentratsioon, S-K — seerumi kaaliumi kontsentratsioon, LK — leukotsüütide üldarv veres, B-Hb — täisvere hemoglobiini kontsentratsioon.

Näiteid vähima kliiniliselt olulise muutuse kohta

Samast valemist on tuletatud ka ettepanek, mis on enamasti kasutatav analüütilise töö kvaliteedi mõõdupuuna: analüütiline varieeruvus ei

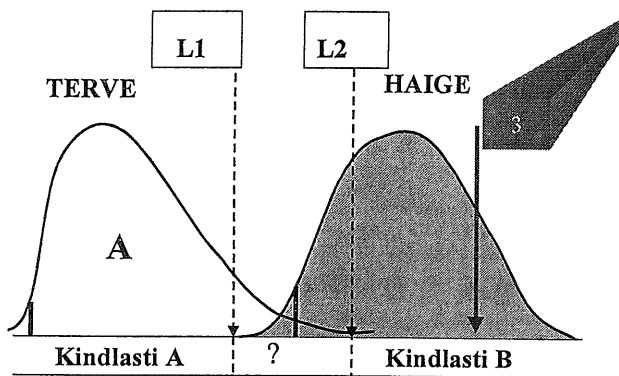
tohiks olla suurem kui pool sellesama analüüdi bioloogilise varieeruvuse ulatusest. Kui viimane tingimus on täidetud, moodustab põhiosa koguvareeruvusest bioloogilise varieeruvuse "panus".

Peamiste laboratoorsete analüütide bioloogilise varieeruvuse kohta on koostatud mitmesuguseid andmebaase. Üks, mida järjekindlalt täiendatakse ja täpsustakse (koostaja C. Ricos) on saadaval veebilehel <http://www.westgard.com/guest17.htm>.

Muud kliiniliste otsustuste piirid

Kõrvuti eespool nimetatud väärtustega vajavad ja kasutavad arstid mitmeid tasemeid, mis sõltuvalt tehtavast otsustusest võivad jääda referentsvahemikest kaugemale väljapoole või hoopis selle piiresse. Neid on nimetatud mitmeti: otsustustasemeiks, tegutsemisväärtusteks, ka kriitilisteks väärtusteks. Otsustustase terminina tundub olevat liiga lai, sest kõigi eelnimetatud, samuti järgmiste näidete puhul langetatakse mingi otsus.

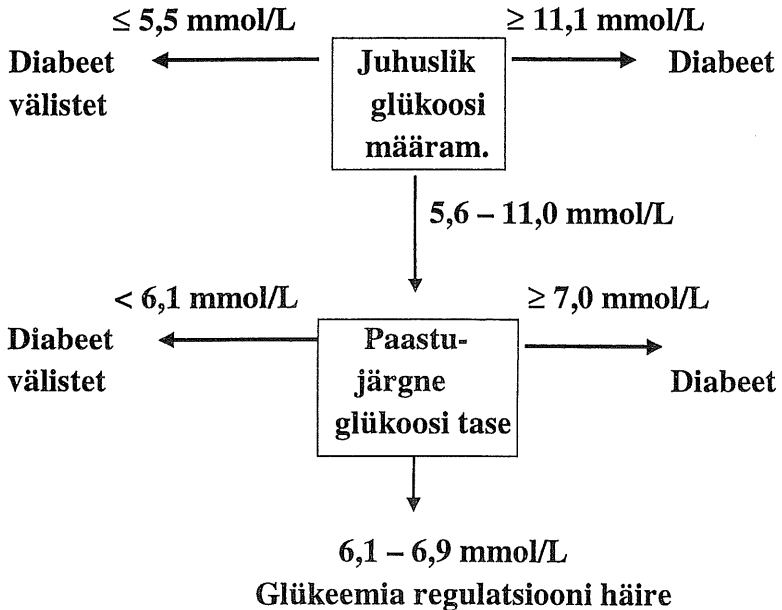
Esimene sellistest läviväärtustest on *välstav tase*, mis tähendab, et laboratoorse uuringu tulemus kinnitab teadaoleva usaldusväärsusega, et uuritud isik ei kuulu teatavasse haigete rühma (joonis 1, lävi L1).



www.westgard.com/decision.htm

Joonis 1. Otsustuste piirid (L) vastavalt isikute teoreetilisele jaotusele.
Sümbolid: L1, L2 – vastavad läviväärtused, 3 – alarmväärtus

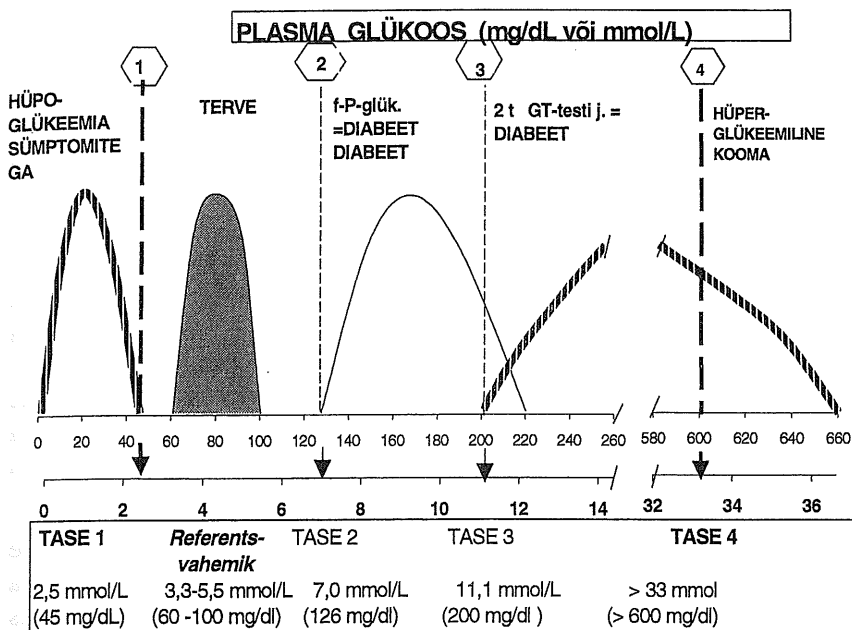
Teine väärtus oleks kinnitav tase (lävi L2), mille saavutamine/ületamine tunnistab, et patsient kuulub teatava haigusega isikute hulka. Kui kasutada näidetena veresuhkru väärtusi (referentsvahemik 3,3–5,5 mmol/L), siis käesoleval ajal peetakse juhuslikult võetud (st tingimusi määratlemata) proovi korral taset alla 5,5 mmol/L diabeeti välistavaks ja väärtust üle 11 mmol/L diabeedi diagnoosi kinnitavaks (joonis 2).



Joonis 2. Diabeedi laboratoorse diagnostika kriteeriumid

Kui uuritav isik on olnud enne proovi andmist söömata, on need väärtused mõistetavalt veidi teistsugused, võrreldes hiljuti söönud isiku omadega, vastavalt alla 6,1 ja üle 7,0 mmol/L.

Lisaks diagnoosi välistavale ja kinnitavale tasemele on paljude laboratoorsete uuringute puhul tuntud viivitamatut sekkumist nõudvad e *alarmväärtused*. Kui jätkata vere glükoosi näidetega, siis väärtused alla 2,5 mmol/L (joonis 3, lävi 1) osutavad hüpopglükeemiale, väärtused üle 33 mmol e 600 mg/dl (lävi 4) aga puudulikust ravist tingitud hüper-



Joonis 3. Vereplasma glükoosisalduse alarmväärtused (tase 1 ja 4) kõrvutatuna tervete inimeste ja diabeetikute glükoosi kontsentratsiooniga veres. Otsustuste tasemed on väljendatud kahes skaalas: mmol/L ja mg/dl

glükeemia. Esimene neist ähvardab liiga madalast glükoositasemest tingitud eluohtliku teadvuskaotusega (koomaseisundiga), viimane aga liiga kõrgest glükoositasemest tingitud komplikatsioonidega, nagu hüperosmolaarne kooma [Porte, Halter, 1981]. Seega on antud näites kohest sekkumist nõudvaiks e alarmeerivateks väärtusteks täiskasvanul glükoosi puhul < 2,5 ja > 33 mmol/L. Võrdluseks on samal joonisel toodud läviväärtused 7 ja 11 mmol/L (tase 2 ja 3), mida kasutatakse diabeedi diagnoosi kinnitamiseks varem juba nimetatud erinevates tingimustes.

Klassikaliseks allikaks ülaltooduga sarnaste otsustuspiiride kohta 100 tuntuimal laboratoorsel testil on B. E. Statlandi raamat (1987), mille

andmed on aadressil <http://www.westgard.com/decision.htm> saadaval tabeli kujul.

Prognostilised näitajad

Suurt hulka laboratoorseid teste käsitletakse teatud haiguste ohuteguritena. Multifaktoriliste haiguste korral, nagu osteoporoos või südame koronaarhaigus, on ülioluline kümnete tuhandete / miljonite ohustatute hulgast välja sõeluda need vähesed, kes saavad mõne aja möödudes vastavalt osteoporoosilise luumurru või südameinfarkti. Tähtsaimad riskitegurid on neil puhkudel vanus, sugu ja pärilikud faktorid, mida muuta ei saa. Lisateguritena aga toimivad mõned ainevahetus- ja eluviisifaktorid, mida saab mõõta ja vajadusel modifitseerida. Osteoporoosilise fraktuuri oht on pea kaks korda suurem neil isikutel, kel luustiku kollageeni resorptsioon on kiirenenud [Garnero *et al*, 2000]. Sel juhul on tegemist kergemini interpreteeritava leiuga: resorptsioon on normaalsest kiirem.

Üheks näitajaks, mida määratakse südame koronaarhaiguse süvenemise suhtes kõrge riskiga isikute leidmiseks, on plasma väikese tihedusega lipoproteiinide (LDL-kolesterooli) tase. Siin on iseäralikuks jooneks see, et riskiväärtus, nt 4,5 mmol/L, paikneb analüüdi referents- e normväärtuste (50–70-aastastel meestel 2,3–5,4) piirkonna sees. Niisuguste riskitegurite kasutamist on viimasel ajal ümber hinnatud: eraldi võetuina ei ole nad kasutatavad haiguse sõeltestina (Wald *et al*, 1999). Nende diagnostiline ja prognostiline väärtus ilmneb ainult hulga ohutegurite kollektiivse mõjuna.

Sellist lähenemist illustreerib väga hästi Münsteri ülikooli südameuuringute tulemusena konstrueeritud PROCAM-algoritm. Nimetus on tuletatud aastatel 1978–1996 ellu viidud projektist **P**rospective **C**ardiovascular **M**unster Study. See on üks suurimaid meeste südamehaiguste uurimusi Euroopas, kus 50 tuhande 40–65-aastase mehe saatust jälgiti kaheksa aasta. vältel. PROCAM-riskikalkulaator on interaktiivsel kujul kättesaadav aadressil <http://www.chd-taskforce.com>. Järgnevalt esitame näite kahest 50-aastasest mehest, kelle LDL-kolesterool on 4,5 mmol/L, kuid südameinfarkti risk täiesti erinev (tabelid 2 ja 3), tähistused: LDL-C – väikese tihedusega kolesterool, HDL-C – suure tihedusega kolesterool, TG – triglütseriidid, DM – suhkurdiabeet, MI – müokardiinfarkt.

Tabel 2. Kahe 50-aastase mehe ohutegurite võrdlus, rõhuga võrdse väärtusega LDL-kolesteroolil kui näitena toodud ohuteguril

	1. mees	2. mees
Vanus (a.)	50	50
LDL-C (mg/dl)	175	175
HDL-C (mg/dl)	70	70
TG (mg/dl)	300	300
Süst. RR (mgHg)	160	160
Suitsetamine	Ei	Jah
DM	Ei	Jah
MI perekonnas	Ei	Jah

Tabel 3. Samade meeste müokardiinfarkti riski suurus, võrrelduna sama vanade meeste keskmise riskiga (PROCAM-kalkulaatori andmetel)

	1. mees	2. mees
Vanus	50	50
LDL-C	175	175
HDL-C	70	70
TG	300	300
Süst. RR	160	160
Suitsetamine	Ei	Jah
DM	Ei	Jah
MI perekonnas	Ei	Jah
MI risk 10 a. j.	2,1%	9,8%
Keskmine MI risk tema vanuses	6,3%	6,3%

Kokkuvõte

- On vähemalt neli strateegiat, kuidas kasutada laboratoorseid uurin-
guid kliiniliste otsuste langetamisel.
- Asjad kulgevad kõige paremini, kui labor on võimeline esitama
oma referentsvahemikke ja olulise muutuse piire. Selleks on vaja
head kvaliteeti nii laborieelses kui ka analüütilises töös.
- Otsustuse piirid selgitatakse patsiendi kliinilise seisundi alusel ja
need peavad patsiendi seisundiga sobima.

- Erijuhul, kui labori tulemust esitatakse riski näitajana, vajab see veenvat tõendust.

Kirjandus

1. Expert Panel on Theory of Reference Values. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *Clin Chim Acta*, 1987; 165, 111–118.
2. Ricos, C. Use of data derived from biological variation. <http://www.westgard.com/guest17.htm>
3. Statland, B. E. (1987). Clinical Decision Levels for Lab Tests. 2nd Ed. Medical Economics Books, Oradell, NJ; vt ka <http://www.westgard.com/decision.htm>
4. Garnero, P., Sornay-Rendu, E., Claustrat, B., Delmas, P. D. (2000). Biochemical markers of bone turnover, endogenous hormones and the risk of fractures in postmenopausal women: the OFFELY study. *J Bone Miner Res*, 15, 1526–1536.
5. Porte, D. Jr., Halter, J. B. (1981). Endocrine pancreas and diabetes mellitus. In: Textbook of Endocrinology, 6th ed. R. Williams (Ed.), 80, WB Saunders Co.
6. Wald, N. J., Hackshaw, A. K., Frost, C. D. (1999). When can a risk factor be used as a worthwhile screening test? *Brit Med J*, 319, 1562–1567.
7. <http://www.chd-taskforce.com>.

NEERUSIIRDAMINE EESTIS 1972–2002

Ülle Kirsimägi¹, Aleksander Lõhmus¹, Jaanus Kahu¹,
Madis Ilmoja², Tiia Arro³

¹TÜ Kirurgiakliinik, ²Lääne-Tallinna Keskhaigla,
³TÜ matemaatilise statistika instituut,
ylle.kirsimagi@kliinikum.ee

Sissejuhatus

Neerupuudulikkus võib olla mitmete krooniliste neeruhaiguste tulemus, mille lõppstaadiumis on neerud pöördumatult kahjustunud ning patsient vajab neeruasendusravi. Erinevate maade statistilistel andmetel tuleb aastas ühe miljoni elaniku kohta 80–100 lõppstaadiumis kroonilise neerupuudulikkusega haiget [Aftan, Ruutu, 1990]. Neeruasendusravi meetoditena on kasutusel: 1) hemo- või peritoneaaldialüüs, 2) neerusiirdamine. 2001. a detsembri seisuga oli Eestis dialüüsravil 141 patsienti (hemodialüüsravil 92 ja peritoneaaldialüüsravil 49), funktsioneeriva transplantaadiga elas 191 patsienti.

Esimene edukas neerusiirdamine maailmas tehti 1954. a. USA-s, kus G. B. Murry ja J. B. Merill siirdasid elusdoonori neeru ühelt kaksikõelt teisele. Esimene neerusiirdamine Eestis toimus 1968. a. Tartus, järgnes nelja-aastane vaheaeg. Alates 1972. a. on neerusiirdamine järk-järgult juurdunud kliinilisse praktikasse ning saanud ka Eestis kroonilise neerupuudulikkuse lõppstaadiumis efektiivseimaks ravimeetodiks, mille tulemusena patsient saab tagasi täisväärtusliku elu.

Käesoleva töö eesmärgiks oli uurida 30 aasta jooksul Eestis tehtud neerusiirdamisi, hinnata patsiendi ja transplantaadi elulemust ning välja selgitada transplantaadi elulemust mõjutavad riskifaktorid.

Patsiendid

1972. aastast kuni 31. oktoobrini 2002 on Eestis toimunud 502 neerusiirdamist, neist 462 esmasiirdamist ning 40 korduvsirdamist, millest kahel patsiendil on neerud siiratud kolm korda. Patsientidest olid 55,6%

mehed (keskmise vanus 43,3) ja 44,4% naised (keskmise vanus 45,1). Kõige noorem oli kaheksa-aastane meespatsient, kõige vanem 68-aastane naispatsient. Kroonilise neerupuudulikkuse põhjusi iseloomustab tabel 1.

Tabel 1. Kroonilise neerupuudulikkuse lõppstaadiumi põhjused Eestis

Põhjus	Kõigest põhjustest (%)	Meeste osakaal
Glomerulonefriit (GN)	54,3	61,9
Püelonefriit (PN)	15,4	45,9
Polütsüstoos (PC)	13,1	42,3
Diabeet (DM)	11,9	51,1
Muud (harva esinevad)	5,3	57,1
KOKKU	100,0	55,6

Meetodid

Ajavahemikul 1972–2002 Eestis tehtud neerusiirdamiste kohta on hinnatud patsiendi ja transplantaadi elulemust. Patsiendi elulemuse hindamisel vaadeldakse ajavahemikku siirdamisest kuni patsiendi surmani, transplantaadi elulemuse korral ajavahemikku siirdamisest kuni transplantaadi funktsioneerimise lakkamiseni või patsiendi surmani.

Patsiendi/transplantaadi elulemus loetakse tsenseerituks, kui huvipakkuv sündmus vaadeldava ajahetkeni t ei ole toimunud (patsient on elus / transplantaat töötab). Kogutud andmete põhjal on hinnatud patsiendi/transplantaadi elulemusfunktsiooni Kaplan-Meieri meetodil. Elulemusfunktsioon on ajahetkel t elusoleku tõenäosus

$$\hat{S}(t) = \prod_{j:0 < t_j \leq t} \left(1 - \frac{d_j}{r_j}\right),$$

kus aluseks on ajahetkede jada $t_1 \leq t_2 \leq \dots \leq t_n$, sümbol d_j tähistab sündmuste (surm) arvu ajamomendil t_j ning r_j riskigrupi suurust vahetult enne momenti t_j ; riskigrupi moodustavad ajamomendil t_j teadaolevalt elusad patsiendid.

Neerusiirdamisprotsessi paremaks iseloomustamiseks on uuritav periood jagatud neljaks lähtuvalt olulisematest muutustest neerusiirdamise arengus Eestis:

I periood: 1972–1989 — siirdamisi on vähe;

- II periood: 1990–1994 — siirdamisi on rohkem, kirurgiline aktiivsus suureneb, sh elusdoonorite kasutamise tõttu, tsüklosporiini kasutuselevõtmine immuunosupressiooni baasravis;
- III periood: 1995–1999 — rutiinne histoloogiliste uuringute tegemise võimalus, siirdatud neeru äratõukereaktsiooni korral ravi mono- või polükloonaalsete antikehadega, mükofenoolhappe kasutuselevõtmine immuunosupressiooni baasravis;
- IV periood: 2000–2002 — ravi mono- või polükloonaalsete antikehadega siirdatud neeru äratõukereaktsiooni induktsioonravis (profülak-tika), prospektiivse elulemuse uurimuse alustamine.

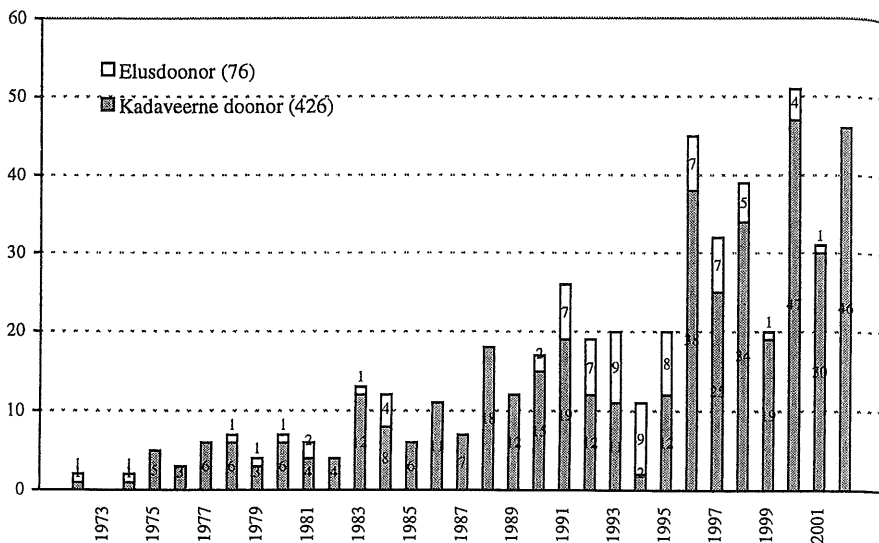
Võimalike transplantaadi elulemust mõjutavate riskifaktorite uurimiseks on vaadeldud kahe viimase perioodi andmeid, kus juhtude käsitus on sarnane, sh doonori valikuprintsiibid, transplantatsioonieelne ja -järgne ravi ning koesobivuse määramise meetodika. Andmed on kogutud retrospektiivselt haiguslugude alusel järgmiste näitajate kohta: doonori vanus, elus- või kadaveerne doonor, retsiipiendi sugu, vanus, neerupuudulikkuse põhjus, kehamassiindeks (BMI), immunoloogiline sobivus (HLA), dialüüsravi meetod, dialüüsravi kestus, üldisheemia aeg, transplantaadi funktsiooni taastumine. Arvtunnuste kohta on leitud keskmine, standardhälve (SD), minimaalne ja maksimaalne väärtus.

Käesolev uuring kestab ja andmete kogumine vaadeldava perioodi kohta ei ole veel lõppenud. Olemasoleva informatsiooni maksimaalseks kasutamiseks on Coxi võrdeliste riskide mudelit hinnatud esmalt iga võimaliku riskifaktori korral eraldi, et püstitada saadud tulemuste põhjal edasised tööhüpoteesid. Riskifaktor loetakse oluliseks $p \leq 0,05$ korral. Oluliseks osutunud riskifaktorite kohta on hinnatud ka riskide suhe ning arvutatud riskide suhte 95%-line usaldusintervall. Andmed on sisestatud neerusiirdamiste andmebaasi (MS Access) ning töödeldud statistika-paketiga StatsDirect.

Tulemused ja arutelu

Eestis tehtud neerusiirdamiste arvu kasvu iseloomustab joonis 1, kust näeme, et neerusiirdamise aktiivsus on alates 1996. aastast oluliselt suurenenud. Viimaste aastate tulemused näitavad, et Eestis on saavutatud vajalik neerusiirdamise aktiivsus, et tagada lõppstaadiumis kroonilise neerupuudulikkusega patsientidele neerusiirdamise võimalus. SA

TÜ Kliinikumi Kirurgiikliiniku andmetel on Eesti dialüüsihaigete populatsiooni arvestades neerusiirdamiste vajadus 30–40 siirdamist ühe miljoni inimese kohta aastas. Keskmise dialüüsravil oldud aeg enne neerusiirdamist oli 1995.–2002. a. andmetel alates poolest kuust kuni kuue aastani (keskmise ooteaeg $15,3 \pm 13,9$ kuud).



Joonis 1. Neerusiirdamiste arv Eestis 1972–2002

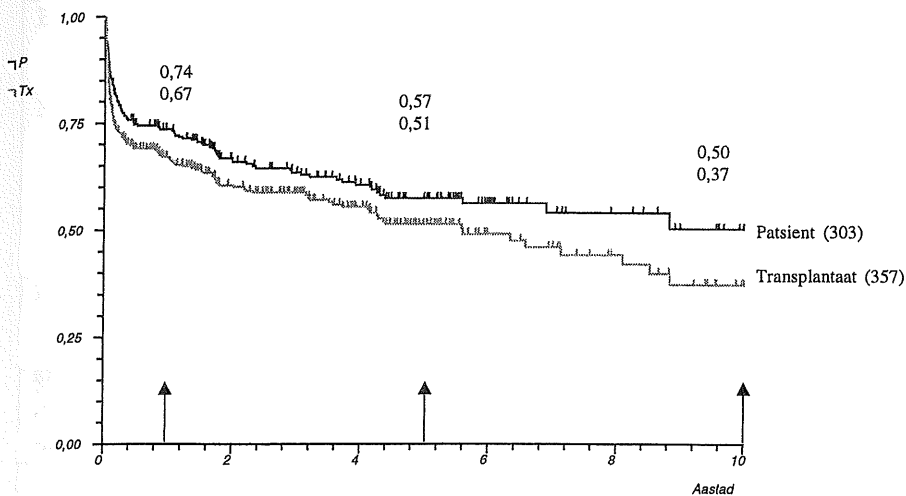
Elusdoonoreid kasutati 76 juhul (15%), kadaveerseid doonoreid 426 korral (85%). Viimastel aastatel on elusdoonorite osatähtsus veelgi langenud, olles aastail 2000, 2001 ja 2002 vastavalt 7,8%, 3,2% ja 0%.

Tabel 2. Neerusiirdamiste arv 1 mln inimese kohta kadaveersete ja elusdoonorite kasutamisel 2000. aastal

Riik	Kadaveerne doonor	Elusdoonor
Soome	36,2	1,4
Norra	28,7	17,1
Hispaania	48,4	0,5
USA	33,5	19,8
Eesti	33,6	2,9

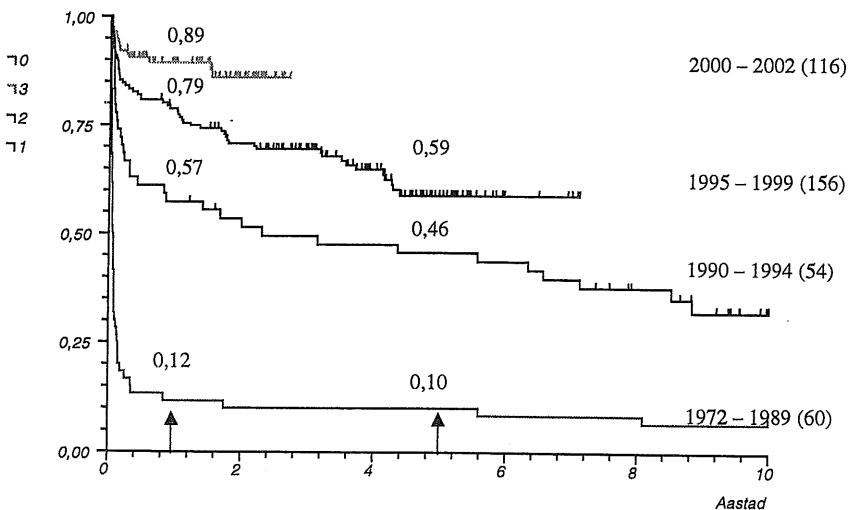
Tabelist 2 [International, 2001] ilmneb, et riigiti on elus- ja kadaveersete neerusiirdamiste arv miljoni inimese kohta erinev. Riikides, kus kadaveersete doonorite arv miljoni inimese kohta on väike, on elusdoonorite osatähtsus suurem.

Eestis on kogu vaatlusperioodi (1972–2002) kohta hinnatud patsiendi ja transplantaadi elulemust (joonis 2). Ühe aasta elulemus on patsiendil 74% ja transplantaadil 67%, viie aasta elulemus vastavalt 57% ja 51%.



Joonis 2. Patsiendi ja transplantaadi elulemus 1972–2002.
Püstkriipsud joonisel tähistavad tsenseeritud vaatlusi

Hinnatud on ka elulemuse sõltuvust sellest, kas doonor oli kadaveerne või elus. Osutus, et elusdoonorite kasutamisel on transplantaadi elulemus oluliselt parem (aasta elulemus 82%; viie aasta elulemus 69%) kui kadaveerse doonori puhul (aasta elulemus 67%, viie aasta elulemus 48%), $p = 0,0213$, riskide suhe 2,08 ning 95%-lised usalduspiirid (1,1; 3,87). Joonis 3 annab ülevaate transplantaadi elulemusest Eestis erineval neerusiirdamisperioodil. Nii patsiendi kui ka transplantaadi elulemuse hindamine näitab, et kõige kriitilisem on esimene aasta pärast siirdamist. Esile tuleb tõsta viimase kolme aasta häid tulemusi transplantaadi ühe aasta elulemuses (89%) elusdoonori (5 juhtu) ja kadaveerse doonori (123 juhtu) korral kokku.



Joonis 3. Transplantaadi elulemus erinevatel neerusiirdamisperioodidel (N=386)

Võrdluseks on toodud Inglismaa andmed, kus transplantaadi ühe aasta elulemus elusdoonori korral on 95%, kadaveerse doonori korral 88%; vastav näitaja on Soomes [Andrews, 2002, Kyllonen *et al*, 2000] kadaveerse doonori korral 91,3%.

Tabel 3 iseloomustab uuritud riskifaktoreid, mis võiksid mõjutada transplantaadi elulemust [Morris, 2001].

Mitmeid uuringuid on tehtud doonori vanuse ja BMI kui võimalike patsiendi/transplantaadi elulemust mõjutavate riskifaktorite kohta [Howard *et al*, 2001; Meier-Kriesche *et al*, 2002, Nakatani *et al*, 2001]. Doonori vanus ei ole üldjuhul riskifaktoriks, sama tulemus on saadud ka käesolevas uuringus, kus doonori vanust on vaadeldud nii pideva kui ka rühmitava tunnusena (<60; ≥60), mõlemal juhul osutus näitaja mitteoluliseks (doonori vanus vaadeldud pideva tunnusena $p = 0,3837$; rühmitava tunnusena $p = 0,1402$). Kirjanduse andmetel on nii väga madal kui ka väga kõrge BMI üks olulisemaid patsiendi/transplantaadi elulemust vähendavaid riskifaktoreid, kuigi mõned uuringud kinnitavad

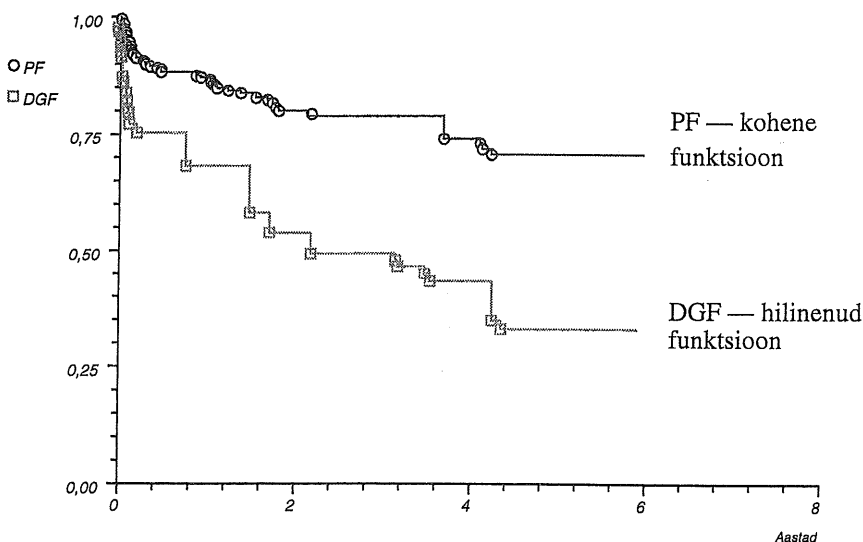
vastupidist [*Op cit*]. Kirjanduses kasutatud [Howard *et al*, 2002] BMI vahemikest lähtudes jagati patsiendid kahte gruppi: BMI<25 ja BMI≥25. Käsitletud juhtude korral ($n = 138$) osutus BMI transplantaadi elulemuse

Tabel 3. Eeldatavad riskifaktorid ja nende mõju

Riskifaktor	n (%)	Keskmine \pm sd (min; max)	p väärtus
Sugu	243		
M	145 (60)		0,6387
N	98 (40)		
Elus- või kadaverne doonor	243		
Elusdoonor	34 (14)		0,2183
Kadaverne doonor	209 (86)		
Neerupuudulikkuse põhjus	243		
Diabeet	45 (19)		0,2449
Püelonefriit	42 (17)		0,1518
Polütsüstoos	29 (12)		
Glomerulonefriit	108 (44)		0,3674
Muud (harva esinevad)	19 (8)		0,2349
Retsipiendi vanus	238	43,7 \pm 13,9 (8;68)	
≤ 55 aastat	185 (78)		0,1533
> 55 aastat	53 (22)		
Kehamassiindeks (BMI)	138	24,2 \pm 4,2 (15;39)	
< 25	92 (67)		0,2688
≥ 25	46 (33)		
Immunoloogiline sobivus (HLA)	212	4,1 \pm 1,3 (0;6)	
≤ 3	66 (31)		0,1408
> 3	146 (69)		
Transplantaadi funktsiooni taastumine	238		
PF – kohene funktsioon	189 (80)		<0,0001
DGF – hiline funktsioon	48 (20)		
Dialüüsravi meetod	237		
HD – hemodialüüs	141 (60)		0,6495
PD – peritoneaaldialüüs	96 (40)		
Üldisheemia aeg tundides	190	13,8 \pm 13,9 (1,3; 31)	0,4151
Doonori vanus	207	40,3 \pm 12,8 (14; 70)	0,3837
< 60	192 (93)		
≥ 60	15 (7)		
Dialüüsravi kestus kuudes	236	15,3 \pm 13,9 (0,5; 72)	0,6318

hindamisel statistiliselt mitteoluliseks ($p = 0,2688$), kuid saadud tulemus võib muutuda juhtude arvu suurenemisel. Samuti võib kõrgem BMI olla riskiteguriks patsiendi/transplantaadi ühe aasta elulemusele.

Vaadeldud riskifaktoritest osutus oluliseks transplantaadi funktsiooni taastumine, mis hilineanud funktsiooni korral vähendab transplantaadi elulemust oluliselt ($p < 0,0001$). Leitud riskide suhe on 2,93 ja selle 95%-lised usalduspiirid (1,76; 4,86) näitavad, et transplantaadi hilineanud funktsiooni korral on risk transplantaat kaotada ligikaudu kolm korda suurem kui transplantaadi kohese funktsiooni korral. Saadud tulemus on kooskõlas teiste autorite poolt läbiviidud uuringute tulemustega [Morris, 2001].



Joonis 4. Transplantaadi elulemus sõltuvalt transplantaadi funktsiooni taastumisest 1995-2002

Hinnatud Coxi võrdeliste riskide mudelit illustreerib joonis 4. Edaspidiste uurimustega on vaja välja selgitada transplantaadi funktsiooni hilise taastumise põhjused (riskifaktorid), millega arvestades saaks miinimumini viia transplantaadi hilise taastumise juhtude arvu, mis kokkuvõttes tagaks patsiendi/transplantaadi parema elulemuse nii lähi- kui ka kaugtulevikus.

Kokkuvõte

- Patsiendi/transplantaadi elulemuse näitajad perioodil 1972–2002:
1 aasta elulemus 74% / 67%,
5 aasta elulemus 57% / 51%,
10 aasta elulemus 50% / 37%.
- Viimasel kolmel aastal on saavutatud eelnevate neerusiirdamis-perioodidega võrreldes oluliselt paremaid tulemusi ühe aasta elulemuses (transplantaadi elulemus 89%).
- Transplantaadi elulemust mõjutavatest riskifaktoritest osutus oluliseks transplantaadi funktsiooni taastumine, mis hilinevad funktsiooni korral vähendab oluliselt transplantaadi elulemust ($p < 0,0001$).
- Käsitletud juhtude arvu korral osutus teiste faktorite (doonori vanus, elus- või kadaverne doonor, retsiptiendi sugu, vanus, neerupuudulikkuse põhjus, kehamassiindeks, immunoloogiline sobivus, dialüüsravi meetod, dialüüsravi kestus, üldisheemia aeg) mõju transplantaadi elulemusele mitteoluliseks.

Kirjandus

1. Alftan, O., Ruutu, M. (1990). *Urologia*, 393.
2. International figures on organ donation and transplantation (2000). *Newsletter Transplant*, 2001, 6, 1.
3. Andrews, P. A. (2002). Renal transplantation. *BMJ*, 324, 530–534.
4. Kyllönen, L. E. J., Salmela, K. T., Eklund, B. H., Halme, L. E. H., Höckerstedt, K. A. V., Isoniemi, H. M., Mäkisalo, H. J., Ahonen, J. (2000). Long-term results of 1047 cadaveric kidney transplantations with special emphasis on initial graft function and rejection. *Transplant Int*, 13, 122–128.
5. Morris, P. J. (2001). Results of renal transplantation. *Kidney Transplantation*, 693–715.
6. Howard, R. J., Thai, V. B., Patton, P. R., Hemming, A. W., Reed, A. I., Van der Werf, W. J., Fujita, S., Karlix, J. L., Scornic, J. C. (2002). Obesity does not portend a bad outcome for kidney transplant recipients. *Transplantation*, 73, 1, 53–55.
7. Meier-Kriesche, H. U., Arndorfer, J. A., Kaplan, B. (2002). The impact of body mass index on renal transplant outcomes: a significant independent risk factor for graft failure and patient death. – *Transplantation*, 73, 1, 70–74.
8. Nakatani, T., Kim, T., Uchida, J., Kumada, N., Tsuchida, K., Takemoto, Y., Yamamoto, K., Kishimoto, T. (2001). Clinical study of 53 cases of cadaveric kidney transplantation at Osaka City University Hospital. *Osaka City Medical Journal*, 47, 2, 95–103.

KOHORTUURING EPIDEMIOLOOGIAS: TEORIA JA PRAKTIKA

Kaja Rahu

Eksperimentaalse ja Kliinilise Meditsiini Instituut,
epidemioloogia ja biostatistika osakond, kaja@ekmi.ee

Sissejuhatus

Kohortuuringuid (*cohort study*) on Eestis tänamatult vähe tehtud. See uuringukavand (*study design*) tuleb kaalumisele siis, kui on vaja otsustada mingi ekspositsiooni efekti (*exposure effect*) üle — kas ekspositsioon suurendab või vähendab haiguseriski (*disease risk*).

Kohortuuringu selgitamiseks on alljärgnevalt kasutatud andmeid tööst, milles kontrolliti hüpoteesi, et vähihaigete hulgas on enesetapp sagedasem kui kogurahvastikus.

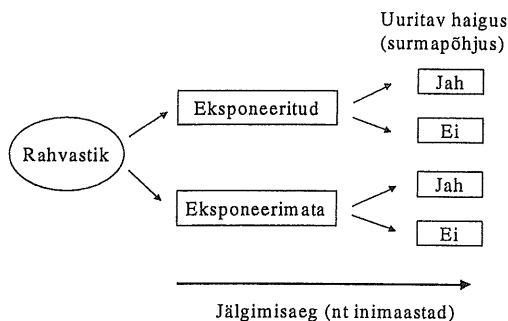
Käesolev kirjutis on omalaadne üleskutse kohortuuringute tegemiseks olukorras, kus andmete kogumisele kulutatav aeg on tühine ja võib keskenduda analüüsile.

Kohortuuringu olemus

Kohortuuringu puhul (vt joonis 1):

- eristatakse rahvastikus ekspositsiooni alusel kaks (või enam) haigusvaba kohorti (rühma); ühe kohordi moodustavad eksponeeritud ja teise eksponeerimata isikud;
- jälgitakse eksponeeritud ja eksponeerimata kohordi igat liiget uuritava haiguse/surmapõhjuse suhtes;
- leitakse riskile avatud aeg (nt inimaastad – *person-years*); kindlaks määratud ajaperioodi jooksul võib inimene olla endiselt terve, haigestuda uuritavasse haigusesse, surra või emigreeruda; et haigestumus/suremus kalendriaastati ja vanuseti muutub, tuleb inimaastate kokkulugemisel arvestada kalendriperioodide ja vanuserühmadega;

- hinnatakse suhtelist riski (*relative risk*) kohortide haigestumuse/suremuse võrdlemise teel; enamasti püütakse selgusele jõuda, kas ekspositsioon suurendab haigestumust/suremust.



Joonis 1. Kohortuuringu skeem

Kohortuuring võimaldab uurida harvaesinevaid ekspositsioone ja hinnata ekspositsiooni seost mitme haigusega/surmapõhjusega, kuid jälgimine võtab palju aega ja seoses sellega ka raha.

Kui jäädagi uskuma, et kohortuuring nõuab *alati* palju aega ja raha, siis peaksime pikemalt mõtlemata loobuma kohortuuringute plaanimisest. Kui efekt järgneb ekspositsioonile alles aastakümne(te) pärast, võib juhtuda, et meie kogutud andmeid analüüsib alles järgmine põlvkond. Tegelikult saab kohortuuringut teha väga kiiresti ja odavalt, kui kasutada varem kogutud andmeid (nt ametikohordid) ja kui võrdlusrühmana on võimalik kasutada kogurahvastikku, st uuritava haiguse/surmapõhjuse kohta on andmed meditsiiniregistritesse kantud. Sel juhul kasutatakse suhtelise riski hindamiseks:

- standarditud haigestumusmäär (standardized incidence ratio — SIR), kui mõõdame haigestumust;
- standarditud suremusmäär (standardized mortality ratio — SMR), kui mõõdame suremust.

Mõlemad näitajad arvutatakse sarnaselt:

$$\begin{aligned} \text{SIR} &= O/E, \\ \text{SMR} &= O/E, \end{aligned}$$

kus

O – tegelik juhtude arv (*observed cases*);
 E – eeldatav juhtude arv (*expected cases*).

Eeldatav juhtude arv näitab, kui palju haigus-/surmajuhte oleks eksponeeritud kohordis esinenud siis, kui haigestumus/suremus oleks olnud võrdne haigestumusega/suremusega kogurahvastikus, arvestades sugu, vanuserühma ja kalendriperioodi.

Eeldatava juhtude arvu leidmiseks peame teadma:

- haigestumus-/suremuskordajaid kogurahvastikus vanuserühmiti kalendriaasta järgi;
- kohordi inimaastaid vanuserühmiti kalendriaasta järgi.

$$E_{ij} = IR_{ij} \times PYR_{ij}$$

$$E = \sum_i \sum_j E_{ij},$$

kus

E_{ij} – eeldatav haigus-/surmajuhtude arv vanuserühmas i kalendriperioodil j ;

IR_{ij} – haigestumus-/suremuskordaja vanuserühmas i kalendriperioodil j ;

PYR_{ij} – inimaastate arv vanuserühmas i kalendriperioodil j .

Saame leida eeldatava juhtude arvu igas vanuserühmas igal kalendriperioodil eraldi. Kogu eeldatava juhtude arvu annab summeerimine üle vanuserühmade ja kalendriperioodide.

Kui:

- $\text{SIR}=1$ ($\text{SMR}=1$), siis ekspositsioon ei mõjuta haigestumust/suremust;
- $\text{SIR}>1$ ($\text{SMR}>1$), siis on tegemist riskiteguriga;
- $\text{SIR}<1$ ($\text{SMR}<1$), siis on tegemist kaitseteguriga.

Kohortuuringu teooriat kirjeldavad mitmed õpikud [Breslow, Day, 1987; Ahlbom, Norell, 1993; Rothman, Greenland, 1998] ja artiklid [Checkoway *et al*, 1989; Samet, Muñoz, 1998].

Kohortuuringu näide: vähihaigete enesetapuriski hindamine

Vähihaigete enesetapuriski uuringus kasutati Eesti vähiregistri andmeid ja moodustati kohort isikutest, kellel diagnoositi vähk aastatel 1983–1998. Uuritavaks sündmuseks oli enesetapp. Igat kohordi liiget jälgiti ajaperioodil 1983–2000. Surmaaeg ja -põhjus saadi Eesti suremuse andmebaasist ja andmed emigreerumise kohta rahvastikuregistrist. Suhelise riski hindamiseks kasutati SMRi. Võrdlusrühmaks võeti kogurahvastik, st eeldatav enesetappude arv vähihaigete hulgas arvatuti kogurahvastiku enesetapakordajate alusel.

Tabelis 1 on näha nelja isiku inimaastate kujunemine uuritava perioodil, liikumine ühest vanuserühmast teise ja ühest kalendriperioodist teise.

Tabel 1. Näide inimaastate leidmise kohta kohordis

Vanuserühm	1983–1987	1988–1992	1993–1997	1998–2000	1983–2000
0–4	–	–	–	–	–
5–9	–	–	–	–	–
10–14	–	–	–	–	–
15–19	–	1,5	–	–	1,5
20–24	–	2,9	–	–	2,9
25–29	–	–	–	–	–
30–34	–	0,3	–	–	0,3
35–39	–	3,7	1,3	–	5,0
40–44	–	–	3,7	1,3	5,0
45–49	2,6	–	–	0,4	3,0
50–54	0,5	4,5	–	–	5,0
55–59	–	0,5	4,5	–	5,0
60–64	3,1	0,9	0,5	2,0	6,4
65–69	–	4,1	0,9	–	5,0
70–74	–	–	4,1	0,9	5,0
75–79	–	–	–	2,1	2,1
80–84	–	–	–	–	–
85+	–	–	–	–	–
Kokku	6,2	18,3	15,0	6,7	46,2

Esimene isik haigestus 18-aastaselt 1988. a, surm muul põhjusel (st mitte enesetapp) 1992. a; kokku 4,4 inimaastat;

teine isik haigestus 34-aastaselt 1989. a, emigreerumine 1999; kokku 10,7 inimaastat;

kolmas isik haigestus 47-aastaselt 1984. a, enesetapp 2000; kokku 15,1 inimaastat;

neljas isik haigestus 61-aastaselt 1984. a, jälgiti perioodi lõpuni 31. 12. 2000; kokku 16,1 inimaastat.

Järgmised tabelid illustreerivad eeldatavate enesetapujuhtude leidmist meeste hulgas. Kogu kohordi meeste inimaastate jagunemine vanuserühmiti ja kalendriperiooditi on toodud tabelis 2.

Tabel 2. Inimaastate arv kohordis, mehed

Vanuserühm	1983– 1987	1988– 1992	1993– 1997	1998– 2000	1983– 2000
0–4	44,2	69,2	68,7	15,0	197,1
5–9	37,0	104,5	136,5	85,9	363,9
10–14	36,6	86,2	140,3	106,0	369,2
15–19	47,0	128,8	181,3	125,1	482,3
20–24	35,8	140,0	241,0	142,8	559,5
25–29	96,7	199,2	239,5	216,1	751,5
30–34	97,6	304,2	388,8	241,8	1032,5
35–39	153,2	389,1	504,4	338,3	1385,0
40–44	243,3	594,5	812,5	442,4	2092,6
45–49	605,3	993,3	1292,8	776,2	3667,6
50–54	942,3	1865,0	1947,3	1116,9	5871,6
55–59	1460,1	2647,4	3311,3	1810,5	9229,3
60–64	1222,1	3107,5	4257,0	2582,2	11168,8
65–69	1066,3	2628,2	4819,4	3182,6	11696,5
70–74	1061,0	1872,9	3621,0	3114,6	9669,5
75–79	846,8	1639,7	2124,8	1975,5	6586,8
80–84	322,3	928,4	1500,0	930,2	3680,9
85+	120,1	298,8	696,2	692,7	1807,9
Kokku	8437,9	17996,8	26282,8	17895,0	70612,5

Algandmetest on nüüd puudu veel Eesti meesrahvastiku enesetapukordajad samades vanuserühmades ja kalendriperioodidel. Need arvud on esitatud tabelis 3.

Tabel 3. Enesetapukordajad (100 000 in k), Eesti meesrahvastik

Vanuserühm	1983–1987	1988–1992	1993–1997	1998–2000
0–4	–	–	–	–
5–9	–	–	–	0,8
10–14	3,5	7,8	9,9	4,9
15–19	23,9	30,9	32,8	36,4
20–24	42,1	43,4	54,3	51,3
25–29	61,8	56,8	65,0	64,0
30–34	72,9	59,1	84,3	57,6
35–39	83,7	73,9	120,9	69,8
40–44	112,7	88,0	130,2	107,5
45–49	103,7	88,9	140,6	122,6
50–54	114,6	94,0	175,7	107,7
55–59	98,7	87,0	166,2	113,9
60–64	82,8	95,9	105,2	109,2
65–69	120,1	104,1	114,9	103,8
70–74	125,6	134,7	129,5	102,5
75–79	174,4	142,3	178,9	145,2
80–84	188,4	198,5	248,0	174,8
85+	299,7	322,2	263,3	277,3

Korrutades inimaastad ja enesetapukordajad omavahel vanuserühmiti ja kalendriperiooditi, saamegi eeldatava enesetappude arvu, nagu on näidatud tabelis 4, näiteks eeldatav enesetappude arv vanuserühmas 55–59 kalendriperioodil 1993–1997 on

$$3311,3 \times 166,2/100000 = 5,5.$$

Iseloomustamaks uuringu tulemusi on tabelis 5 toodud mõned standarditud suremusmäärad koos 95%-lise usaldusvahemikuga (*confidence interval* – CI).

Segamise (*confounding*) kontrollimiseks kasutatakse SMRde arvutamisel kihitamist (*stratification*) soo, vanuse, paikme (paikmete rühma) ja diagnoosimise aja järgi.

Tabel 4. Eeldatav enesetappude arv kohordis, mehed

Vanuserühm	1983– 1987	1988– 1992	1993– 1997	1998– 2000	1983– 2000
0–4	–	–	–	–	–
5–9	–	–	–	0,0	0,0
10–14	0,0	0,0	0,0	0,0	0,0
15–19	0,0	0,0	0,1	0,0	0,1
20–24	0,0	0,1	0,1	0,1	0,3
25–29	0,1	0,1	0,2	0,1	0,5
30–34	0,1	0,2	0,3	0,1	0,7
35–39	0,1	0,3	0,6	0,2	1,2
40–44	0,3	0,5	1,1	0,5	2,4
45–49	0,6	0,9	1,8	1,0	4,3
50–54	1,1	1,8	3,4	1,2	7,5
55–59	1,4	2,3	5,5	2,1	11,3
60–64	1,0	3,0	4,5	2,8	11,3
65–69	1,3	2,7	5,5	3,3	12,8
70–74	1,3	2,5	4,7	3,2	11,7
75–79	1,5	2,3	3,8	2,9	10,5
80–84	0,6	1,8	3,7	1,6	7,7
85+	0,4	1,0	1,8	1,9	5,1
Kokku	9,8	19,5	37,1	21,0	87,4

Tabel 5. Vähihaigete enesetapurisk võrreldes kogurahvastikuga

Paige	Mehed			Naised		
	O	E	SMR (95% CI)	O	E	SMR (95% CI)
Kõik paikmed	150	87,4	1,7 (1,4–2,0)	47	95,5	0,5 (0,4–0,7)
Kops	29	10,9	2,7 (1,8–3,8)	6	1,6	3,7 (1,4–8,1)

Sarnase meetodikaga kohortuuringuid on palju tehtud Põhjamaades, kus töötab aastakümneid meditsiiniregistrite süsteem, mille andmeid saab isikutasandil kokku viia teiste andmebaasidega, sh rahvastikuregistriga [Andersen *et al*, 1999, Lichtermann *et al*, 2001, Weiderpass *et al*, 2001].

Kirjandus

1. Ahlbom, A., Norell, S. (1993). *Sissejuhatus moodsasse epidemioloogiasse*. Tallinn: Huma.
2. Andersen, A., Barlow, L., Engeland, A., Kjaerheim, K., Lynge, E., Pukkala, E. (1999). Work-related cancer in the Nordic countries. *Scand J Work Environ Health* 25, 2, 1–116.
3. Breslow, N. E., Day N. E. (1987). *Statistical Methods in Cancer Research, Vol II. The Design and Analysis of Cohort Studies*. IARC Scientific Publications No 82. Lyon, IARC.
4. Checkoway, H., Pearce, N., Dement, J.M. (1989). Design and conduct of occupational epidemiology studies: I. Design aspects of cohort studies. *Am J Ind Med* 15, 4, 363–373.
5. Checkoway, H., Pearce, N., Dement, J.M. (1989). Design and conduct of occupational epidemiology studies: II. Analysis of cohort data. *Am J Ind Med* 15, 4, 375–394.
6. Lichtermann, D., Ekelund, J., Pukkala, E., Tanskanen, A., Lonnqvist, J. (2001). Incidence of cancer among persons with schizophrenia and their relatives. *Arch Gen Psychiatry* 58, 6, 579–580.
7. Rothman, K. J., Greenland S. (1998). *Modern Epidemiology*. Philadelphia, Lippincott-Raven Publishers.
8. Samet, J. M., Muñoz, A. (eds.). (1998). *Perspective: cohort studies*. *Epidemiol Rev* 20, 1, 135–136.
9. Weiderpass, E., Gridley, G., Nyren, O., Pennello, G., Landstrom, A. S., Ekblom, A. (2001). Cause-specific mortality in a cohort of patients with diabetes mellitus: a population-based study in Sweden. *J Clin Epidemiol* 54, 8, 802–809.

TARTU NOORTE MEESTE JA NAISTE KEHASTRUKTUUR

Helje Kaarma, Säde Koskel

TÜ füüsilise antropoloogia keskus, matemaatilise statistika
instituut, Antrop@ut.ee, koskel@ut.ee

Uuringu eesmärk

Nii terviseedenduses kui ka haiguste ravis omandab üha suurema tähtsuse kehaehituse iseärasuste, sh keha rasvasisalduse arvestamine. Et patsienti hinnata, on vaja teda kellegagi võrrelda, st on tarvis teada iga näitaja normaalsele kehaehitusele vastavat taset. Eestis nende näitajate normatiivid praegu puuduvad. Perearsti juures mõõdetakse küll kohustuslikult patsiendi pikkus ja kaal ning arvutatakse kehamassiindeks, kuid need pole leidnud kasutamist riigi terviseedendusprogrammides.

Meie uuringu eesmärgiks on töötada välja normatiivid eesti täiskasvanud elanike pikkuse, kaalu, kehamassiindeksi ja keha rasvasisalduse jaoks, kasutades selleks viimastel aastatel tehtud mõõtmisi.

Materjal ja meetodika

Vajalikke mõõtmisi alustas füüsilise antropoloogia keskus Tartu linnas 1998. aastal [Saluste *et al*, 2002]. Selleks seati sisse kaks antropomeetria kabinetti [Lintsi *et al*, 1999]. Neist üks töötab Maarjamõisa polikliinikus PhD Liidia Saluste juhtimisel ja teine Tartu linna polikliinikus, juhatajaks magister Mart Lintsi. Kabinettides teenindatakse inimesi tasuta, sõltumata sellest, kas patsiendid pöörduvad sinna omal algatusel või perearsti suunamisel. Patsiendil mõõdetakse pikkus (m) ja kaal (kg) ning arvutatakse valemiga

$$\text{BMI} = \text{KAAL} / \text{PIKKUS}^2$$

kehamassiindeks (BMI), kasutades vastavaid tabeleid. Pikkus mõõdetakse Martini metallantropomeetriga täpsusega 0,1 cm ja kaalutakse firma Soehnle elektronkaaluga täpsusega 50 g.

Lisaks sellele hinnatakse keha rasvasisaldus firma Omron® BF 300 uudse keharasvaanalüsaatoriga, mis töötab segmentaalse bioimpedantsi meetodil [Omron manual]. Nimetatud aparaat on võetud kasutusele USA-s ja jõudis Eestisse 1998. aastal. Aparaat hindab patsiendi keha üldrasva hulka kilodes (täpsusega 0,1 kg) ja arvutab rasvaprotsendi täpsusega 0,1%. Rasvamõõtuuri käsitlemine on väga lihtne ega vaja eriväljaõpet. Hindamise eel sisestatakse mõõturisse patsiendi algandmed: pikkus, kaal, vanus ja sugu. Mõõtmiseks hoiab seisev patsient aparaati mõlema käega täisnurga all ette sirutatult öla kõrgusel. Rasvasisalduse näitajad ilmuvad ekraanile umbes 7 sekundi jooksul [*Op cit*]. Aparaaadi tööpiirkonnaks on keha rasvasisalduse vahemik 4–50%, mis katab kogu praktilist huvi pakkuva piirkonna.

Kõik mõõtmistulemused sisestatakse arvutisse ja igale patsiendile väljastatakse blankett tema andmetega, mida ta saab vajadusel esitada perearstile või mõnele erialaarstile. Järgnevas on toodud blanketi näidis.

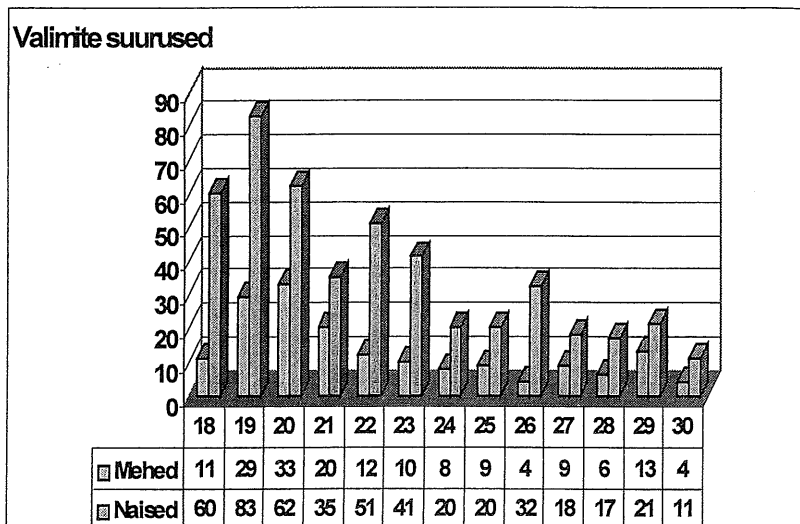
Füüsilise Antropoloogia Keskus, kab. 2031

Kuupäev:
Ees- ja perekonnanimi:
Vanus:
Sugu:
Pikkus: (cm)
Kaal: (kg)
Kehamassiindeks (BMI):
Keha üldine rasvasisaldus: (kg)
Rasva % kehakaalust:

Patsiendid võivad kabinette külastada korduvalt jälgimaks oma toitumise ratsionaalsust. Samas tehakse patsiendi soovil või perearsti soovitusel detailsem antropomeetriline mõõtmine, mõõdetakse 46 kehamõõtu (pikkus- laius-, sügavusmõõdud ja rasvavoldid) [Linti *et al*, 1997].

Töö tulemused

Käesolevas artiklis käsitletakse ajavahemikus 1998–2001 tehtud mõõtmiste tulemusi. Sellel ajavahemikul oli uuringul käinud 492 meest ja 1346 naist vanuses 18–80 aastat. Neist valiti uuringusse 18–30-aastased Tartu mehed ja naised, kokku 168 meest ja 471 naist, vt joonis 1.



Joonis 1. Uuritava kontingendi sagedusjaotus soo ja vanuse järgi

Valimite mahud vanuse kaupa kajastavad meeste ja naiste erinevat suhtumist taolistesse mõõtmistesse. Naiste osakaal oli igas vaadeldavas vanusklassis suurem. Uuritavate põhikontingendi moodustavad noored vanusevahemikus 18–23 aastat (70,5% naistest ja 68,5% meestest). Suhteliselt väikesed on aga 24–30-aastaste meeste vanuserühmad.

Esitame kõigepealt nii meeste kui ka naiste pikkuse, kaalu ja kehamassiindeksi keskmiste hinnangud vanuserühmiti (vt tabelid 1 ja 2). Tabelites on esitatud ka vastava näitaja minimaal- ja maksimaalväärtus vaadeldavas rühmas, mis on otstarbekas ekstreemsete juhtude esiletoomiseks. Tabelist 1 näeme, et vaadeldavas vanuses meeste pikkus varieerub 160 kuni 196,4 sentimeetrini, keskmine on 180,3 cm. Uuritavate vanuserühmade väiksuse tõttu ei saa teha erilisi sisulisi järeldusi kehamõõtude muutuste kohta, seega tuleb käesolevat tööd käsitleda kui metoodilist.

Tabel 1. Noorte meeste kehaehituse mõõtmistulemused vanuserühmiti

Vanus	Kaal			Pikkus			BMI		
	Min	Keskmine	Max	Min	Keskmine	Max	Min	Keskmine	Max
18	62,7	74,66	93,3	170,7	180,39	188,6	18,93	22,93	26,34
19	50,1	71,63	92,5	160,0	178,30	192,4	18,56	22,46	27,03
20	60,0	77,14	112,0	161,0	181,23	194,1	18,61	23,51	32,11
21	51,7	78,42	98,6	167,6	181,22	191,3	16,47	23,91	30,57
22	65,0	81,85	123,7	173,9	183,23	196,4	20,01	24,26	33,28
23	70,9	86,66	121,8	166,5	178,31	187,2	23,16	27,20	35,74
24	64,3	75,43	97,2	176,9	180,19	193,1	20,55	23,15	26,73
25	57,5	76,06	92,5	173,0	180,80	189,1	17,59	23,29	28,52
26	82,0	94,33	110,7	187,7	190,63	194,6	23,28	25,95	30,31
27	66,6	78,54	89,0	168,2	179,01	187,0	21,28	24,56	27,57
28	59,0	80,43	118,0	169,0	175,15	179,5	20,15	26,07	37,45
29	63,7	80,43	91,1	167,9	180,06	188,7	22,09	24,78	28,12
30	64,5	75,00	78,7	167,2	178,50	188,2	22,16	23,53	25,39
Üldine	50,1	77,75	123,7	160,0	180,29	196,4	16,47	23,89	37,45

Tabel 2. Noorte naiste kehaehituse mõõtmistulemused vanuserühmiti

Vanus	Kaal			Pikkus			BMI		
	Min	Keskmine	Max	Min	Keskmine	Max	Min	Keskmine	Max
18	48,8	59,54	74,8	158,6	169,04	180,7	17,52	20,82	25,16
19	45,8	59,09	103,6	150,2	166,84	182,3	16,93	21,19	34,18
20	45,7	59,24	83,2	148,9	167,62	181,0	16,37	21,10	27,23
21	46,7	57,81	86,4	148,0	165,06	186,4	17,99	21,18	28,68
22	43,7	60,45	85,5	144,2	166,71	180,2	16,17	21,76	28,83
23	41,5	59,03	79,7	152,9	166,81	178,3	17,75	21,14	27,47
24	47,1	59,33	77,0	157,5	167,46	176,0	18,76	21,11	25,17
25	49,0	61,11	76,8	156,6	166,25	179,5	18,13	22,10	27,81
26	49,2	63,31	95,6	154,5	166,61	178,4	17,50	22,82	33,63
27	46,2	64,53	93,4	159,2	167,41	178,1	17,69	22,98	34,10
28	53,4	66,09	99,6	158,5	169,00	191,7	18,63	23,01	32,28
29	39,2	64,09	88,4	153,5	167,85	183,1	15,64	22,60	30,02
30	53,5	62,51	71,6	152,7	166,87	177,1	19,26	22,42	24,74
Üldine	39,2	60,36	103,6	144,2	167,22	191,7	15,64	21,56	34,18

Meeste kaal muutub vahemikus 50,1–123,7 kg, keskmine on 77,75 kg. Kehamassiindeks varieerub vahemikus 16,5–37,5, keskmine on 23,9. Tabelist 2 ilmneb, et naiste pikkus jääb 144,2 ja 191,7 cm vahemikku, keskmine on 167,2 sentimeetrit. Naiste keskmine kaal varieerub 39,2 kilost kuni 103,6 kiloni, olles keskmiselt 60,4 kilo. Naiste keskmine kehamassiindeks varieerub vahemikus 15,6–34,2, keskmine on 21,56.

Keha rasvasisalduse seos kehakaalu ja kehamassiindeksiga

Teatavasti sisaldab kehakaal nii luu-, lihas- kui ka rasvkoet kaalu. Täpsemalt suudetakse tänapäeval hinnata just rasvahulga osatähtsust. Tabelitest 3 ja 4 on näha, et vaadeldavate noorte meeste puhul on keharasva üldhulk keskmiselt 10,6 kg ja rasvaprotsent 13,1. Naistel on keskmine rasvasisaldus 12,7 kg, seega 2 kg võrra suurem kui meestel, moodustades kehakaalust 20,4%. Seevastu kehamassiindeksi keskmine väärtus on naistel väiksem kui meestel.

Paljud uurijad, veelgi enam aga antropomeetriliste andmete tavatarbijad kasutavad meelsasti kehamassiindeksit keha täidluse hindamiseks, hinnates selle kaudu ühtlasi ka keha rasvasisaldust. Üldkasutatav on järgmine kehamassiindeksi klassijaotus:

1. <20 — väike kehakaal pikkuse suhtes, mis võib põhjustada terviseprobleeme;
2. 20–25 — tervise seisukohalt enamikule inimestest kõige soodsam;
3. 25–27 — algav ülekaalulisus, võib mõnel juhul põhjustada tervisehäireid;
4. 27–30 — tugev ülekaalulisus, põhjustab sageli terviseprobleeme;
5. >30 — rasvumine, põhjustab suuri terviseprobleeme.

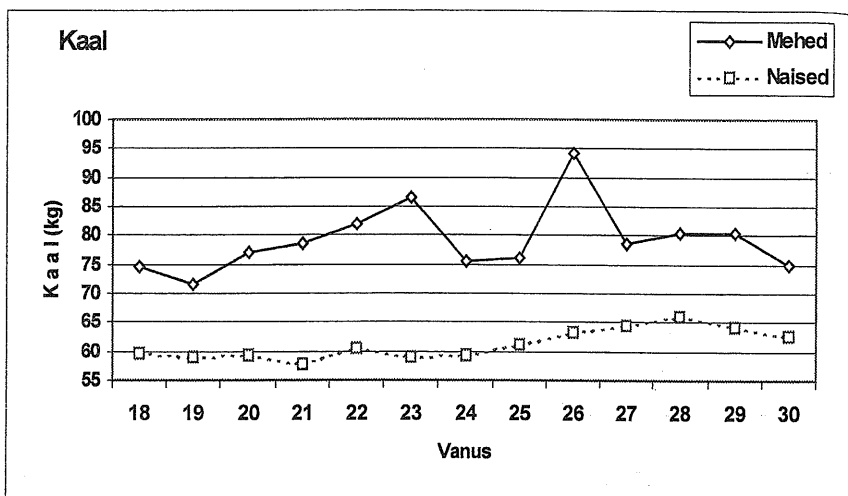
Et võrrelda kehaehituse karakteristikuid vanuserühmiti, esitame meeste ja naiste mõõtude keskmiste graafikud (vt joonised 2–4). Nendelt on näha, et meestel on nii pikkus, kaal kui ka kehamassiindeks keskmiselt suuremad kui naistel. Ettevaatusega tuleb aga suhtuda jónksudesse meeste näitude graafikul — siin on tegemist väikeste rühmadega, mistõttu rühma keskmised erinevad juhuslike mõjutuste tagajärjel üksteisest üsna palju.

Tabel 3. Noorte meeste keha rasvasisaldus vanuserühmiti

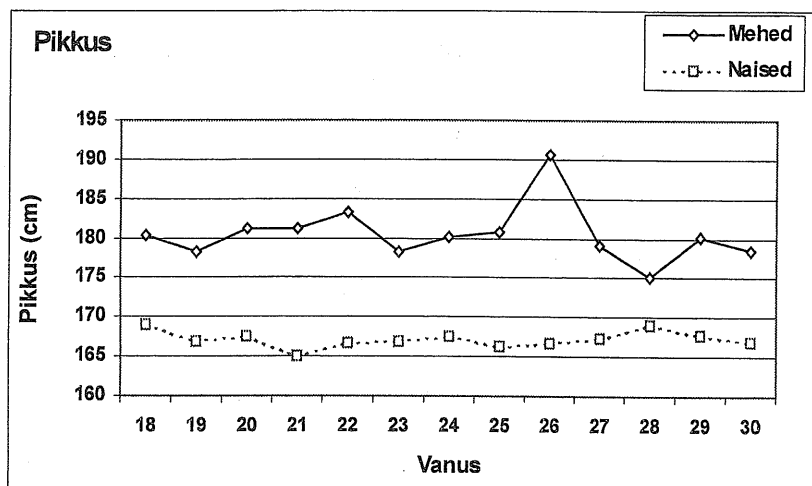
Vanus	Rasva (kg)			Rasva (%)			BMI		
	Min	Keskmine	Max	Min	Keskmine	Max	Min	Keskmine	Max
18	5,2	9,51	21,3	7,9	12,36	22,8	18,93	22,93	26,34
19	3,3	8,47	16,5	4,8	11,59	22,3	18,56	22,46	27,03
20	3,0	9,57	24,9	4,5	11,86	29,2	18,61	23,51	32,11
21	2,9	10,72	24,7	4,2	13,13	25,0	16,47	23,91	30,57
22	3,6	10,67	29,2	4,8	12,45	23,6	20,01	24,26	33,28
23	4,6	14,82	33,5	6,5	16,90	28,3	23,16	27,20	35,74
24	4,2	10,05	16,3	6,5	12,94	17,0	20,55	23,15	26,73
25	4,4	10,80	18,8	15,0	28,20	21,1	17,59	23,29	28,52
26	5,4	14,98	28,2	6,6	15,13	25,5	23,28	25,95	30,31
27	4,1	11,59	17,5	6,1	14,47	19,7	21,28	24,56	27,57
28	4,7	14,05	36,1	7,9	15,63	30,5	20,15	26,07	37,45
29	3,3	12,80	19,2	5,3	15,62	21,7	22,09	24,78	28,12
30	8,2	8,95	9,7	10,4	12,00	13,2	22,16	23,53	25,39
Üldine	2,9	10,63	36,1	4,2	13,13	30,5	16,47	23,89	37,45

Tabel 4. Noorte naiste keharasvasisaldus vanuserühmiti

Vanus	Rasva (kg)			Rasva (%)			BMI		
	Min	Keskmine	Max	Min	Keskmine	Max	Min	Keskmine	Max
18	4,5	12,21	20,2	9,0	20,19	27,7	17,52	20,82	25,16
19	4,7	11,75	39,3	7,6	19,30	37,9	16,93	21,19	34,18
20	4,0	11,62	24,4	8,3	18,88	30,8	16,37	21,10	27,23
21	5,2	11,57	23,5	10,1	19,67	30,8	17,99	21,18	28,68
22	5,4	12,48	24,9	10,1	20,35	33,4	16,17	21,76	28,83
23	4,3	11,44	21,4	8,9	19,02	28,0	17,75	21,14	27,47
24	6,1	11,81	22,4	13,0	19,72	29,1	18,76	21,11	25,17
25	4,7	13,27	23,3	9,5	21,21	30,7	18,13	22,10	27,81
26	5,8	14,98	35,6	11,8	22,78	37,2	17,50	22,82	33,63
27	7,3	15,23	36,3	15,7	22,71	38,9	17,69	22,98	34,10
28	9,0	16,04	32,0	16,6	23,55	34,6	18,63	23,01	32,28
29	4,8	15,55	30,8	12,3	22,93	34,8	15,64	22,60	30,02
30	11,2	15,09	19,9	16,6	24,09	28,5	19,26	22,42	24,74
Üldine	3,0	12,652	39,3	7,6	20,37	38,9	15,64	21,56	34,18

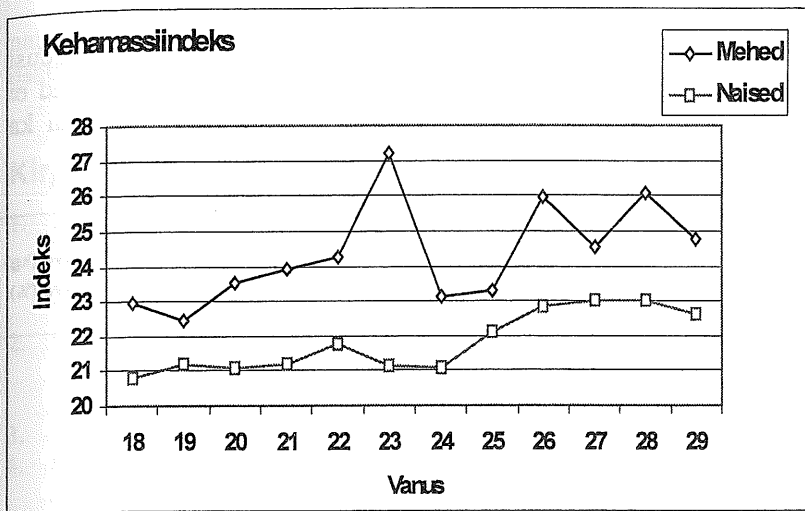


Joonis 2. Meeste ja naiste keskmised kaalud vanuserühmiti

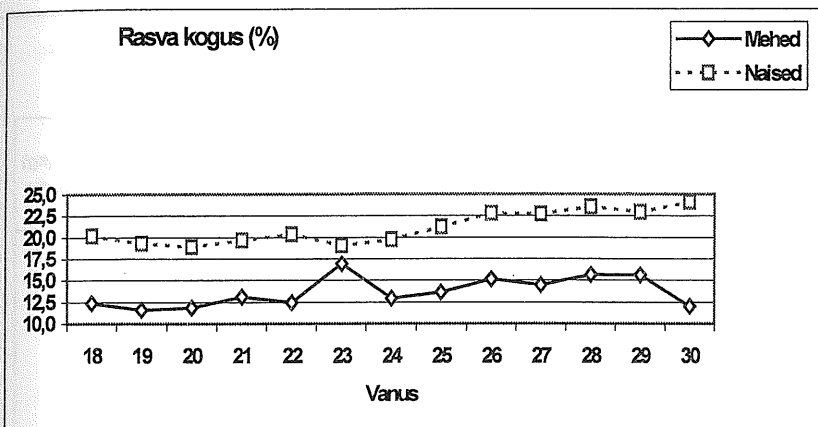


Joonis 3. Meeste ja naiste keskmised kaalud vanuserühmiti

Joonistel 6 ja 7 on esitatud nii meeste kui ka naiste kehakaalu, üldrasva hulga, rasvaprotsendi ning kehamassiindeksi näitajad.



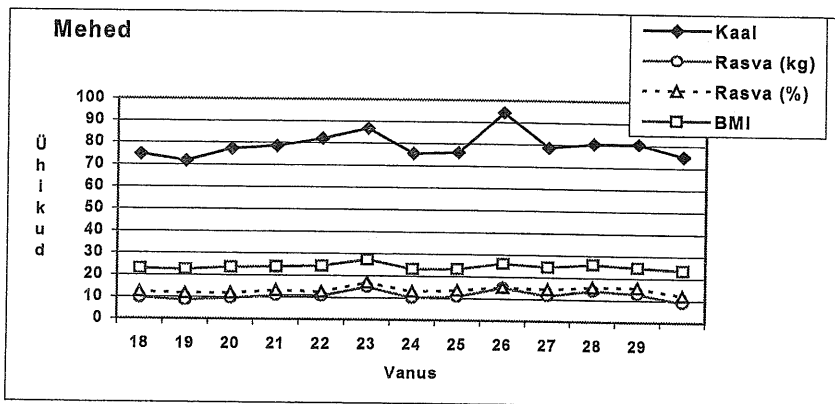
Joonis 4. Meeste ja naiste keskmised kehamassiindeksid vanuserühmiti



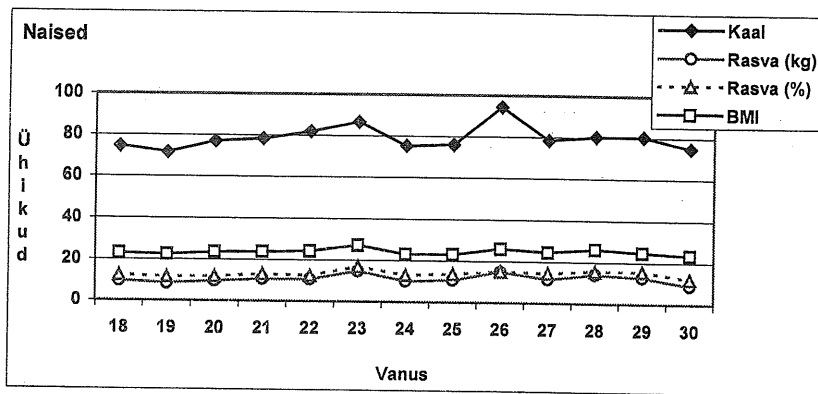
Joonis 5. Meeste ja naiste keskmised keha rasvaprotsendid vanuserühmiti

Nendelt joonistelt ilmneb, et ei noortel meestel ega ka naistel ei ole olulisi vanusest sõltuvaid erisusi ei keha rasvasisalduses ega rasvaprotsendis.

Võrreldes meeste ja naiste kehamassiindeksit ja keha rasvasisaldust ilmneb, et kehamassiindeks kirjeldab küll keha üldist täidlust, kuid ei asenda keha rasvasisalduse määramist, mis on nii meditsiinis kui ka terviseedenduses eriti suure tähtsusega.



Joonis 6. Meeste kaal, üldrasva hulk, rasvaprotsent ja kehamassiindeks sõltuvalt vanusest



Joonis 7. Naiste kaal, üldrasva hulk, rasvaprotsent ja kehamassiindeks sõltuvalt vanusest

Võrreldes noorte meeste ja noorte naiste näitajaid ilmnes, et meeste kehamassiindeks on suurem, kuid keha rasvasisaldus väiksem kui

naistel. Põhjuseks on meeste rasvavaba massi (luud, lihased) suurem osatähtsus naistega võrreldes.

Kirjandus

1. Lintsi, M., Saluste, L., Kaarma, H., Koskel, S., Aluoja, A., Liivamägi, J., Mehilane, L., Vasar, V. (1997). Characteristic traits of anthropometry in 17–18 years old schoolboys Tartu. *Papers on Anthropology* VII. Tartu, 222–231.
2. Lintsi, M., Kaarma, H., Saluste, L. (1999). First experience of application of Omron BF 300 body fat monitor to members of a Tartu sports club. *Papers on Anthropology* VIII. Tartu, 95–102.
3. Omron BF 300 Body Fat Monitor Instruction manual. 3–15.
4. Saluste, L., Koskel, S. (2002). Tartu linna täiskasvanud eesti meeste ja naiste kehaehitus ja rasvasisaldus aastatel 1998–2001. *Eesti antropomeetriaregistri aastaraamat 2002*. Tartu, 204–212.

SÜNDIMUSKÕVERATE LÄHENDAMISEST

Ene Käärik, Ene-Margit Tiit

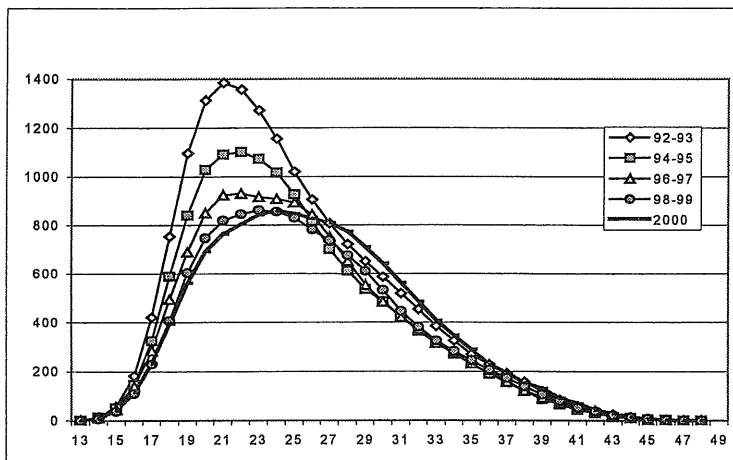
TÜ matemaatilise statistika instituut

enek@ut.ee

Empiiriliste kõverate jaoks sobivate teoreetiliste lähendite leidmine on ülesanne, mis pakub huvi erinevates biostatistika valdkondades. Näiteks prognoosimaks Eesti demograafilisi muutusi ja rahvastikuprotsesse ning hindamaks Eesti naiste reproduktiivkäitumist, on oluline modelleerida sündimuskõvera jaotust, mis üldiselt muutub ajas (vt joonis 1).

Definitsioonid ja põhimõisted

Vaatame vanuseid x_1, x_2, \dots, x_p , mida loetakse kuuluvat naise viljakasse ehk reproduktiivikka (tavaliselt 15, ..., 45). Olgu vanuses x_j elanikkonnas a_j naist ning olgu neil vaatlusaasta jooksul sündinud vastavalt b_j last ($j = 1, \dots, p$).



Joonis 1. Vanuselised sündimuskordajad, Eesti 1992–2000

Seega a_1, a_2, \dots, a_p naist on sünnitanud vastavalt b_1, b_2, \dots, b_p last. Siis on iga vanuses x_j naise kohta sündinud $f_j = b_j/a_j$ last. See on vanuseline

sündimuskordaja. Enamasti korrutatakse see 1000-ga, saades nii sündide arvu aastas 1000 vastavas vanuses naise kohta. Vanuselisi sündimuskordajaid arvutatakse kogu viljaka ea kohta. Näeme, et vaatlus-

Tabel 1. Viljakas eas naiste sündide arv Eestis 2000–2001

Aasta	Sündide arv	Viljakas eas naiste arv
2000	12983	342859
2001	12538	342461

aastail on Eestis olnud keskmiselt 37 sündi aastas 1000 naise kohta, mis teeb keskmiselt 1,3 last naise kohta kogu viljaka ea jooksul.

Summaarne sündimuskordaja B näitab, mitu last sünnitaks naine keskmiselt oma viljaka ea jooksul siis, kui sündimuskäitumist iseloomustavad vanuselised sündimuskordajad oleksid kogu tema viljaka ea jooksul samasugused kui vaatlusaastal, $B = (f_1 + f_2 + \dots + f_p)/1000$.

Sündimuskordajate lähendamise ülesanne

Empiiriliste sündimuskordajate lähendamiseks teoreetilise jaotusega (nn *sündimuskõvera*) läbitakse järgmised sammud.

- Leitakse vanuselised sündimuskordajad.
- Normeeritakse need nii, et tekiks tõenäosusfunktsioon.
- Valitakse sobiv jaotusseadus.
- Valitakse lähendamiskriteerium.
- Otsitakse vaadeldava jaotusseaduse perest sobivaim jaotus.

Sündimuskõverate lähendamisel on kõige levinum beeta- ja gamma-jaotuse kasutamine; lisaks on tuletatud mõningaid spetsiaaljaotusi, sh Hadwigeri jaotus. Lisaks standardsetele lähendusmeetoditele vaatleme ka jaotuste segu kasutamist. Olulised muutused sündimuskäitumises kajastuvad sündimuskõvera parameetrite muutumises.

Modelleerimisel on aluseks Eesti meditsiinilise sünniregistri andmed aastaist 1992–2000 [Eesti Meditsiiniline, 2001] ja ESA rahvastikuandmed [Rahvastik, 2001]. Kasutatud tarkvara on SAS 8.2 ja Excel'97/Solver.

Sündimuskõvera lähendamine beeta- ja gamma-jaotusega

Gamma- ja beetajaotus on lihtsad ning tuntud klassikalised tõenäosus-jaotused. Gammajaotust iseloomustab teravam ja järsem tipp, beetajaotus seevastu on mõnevõrra laugema kujuga. Mõlemal jaotusel on kaks parameetrit, mis on momentide meetodil võrdlemisi lihtsalt hinnatavad.

Gammajaotuse tihedusfunktsioon:

$$f(x) = \frac{x^{\gamma-1} \exp(-x\beta)}{\beta \Gamma(\gamma)},$$

kus γ ja β on jaotusparameetrid ja $\Gamma(\gamma)$ tähistab Euleri gamma-funktsiooni. Argumentide väärtused on positiivsed reaalarvud.

Beetajaotuse tihedusfunktsioon:

$$f(x) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)},$$

kus α ja β on jaotusparameetrid ning $B(x, y)$ tähistab Euleri gamma-funktsiooni. Argumentide ja parameetrite väärtused on positiivsed reaalarvud, kusjuures rahuldatud peab olema ka tingimus $x < \alpha$.

Mõlema jaotuse parameetritele on momentide meetodil võimalik lihtsalt leida algandmete põhjal hinnangud, mille nihe on valimi suure mahu tõttu (üle 12 000 vaatluse) tühiselt väike. Lihtsuse mõttes teisendame lähteandmed vahemikku $[0, 1]$ vastavalt järgmise lineaarse eeskirjaga:

$$v = (x - 15)/30,$$

kus x tähistab esialgseid, v aga teisendatud vanuseid. Nooremate kui 15-aastaste ja vanemate kui 45-aastaste sünitused jäeti kõrvale, sest neid ei olnud igal aastal ja nad moodustasid sünituste üldarvust alla 0,05 %.

Gammajaotuse parameetrid

Gammajaotuse parameetrid leitakse seostest:

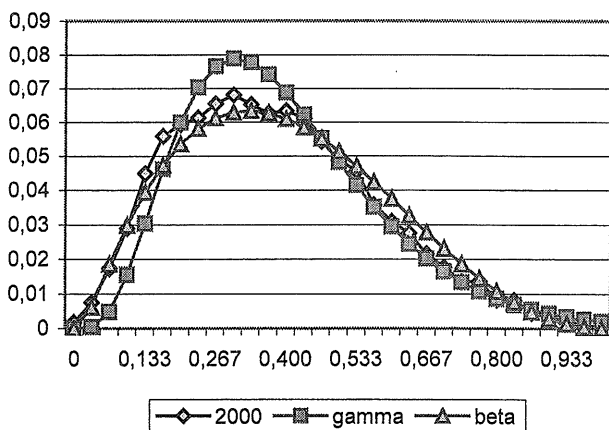
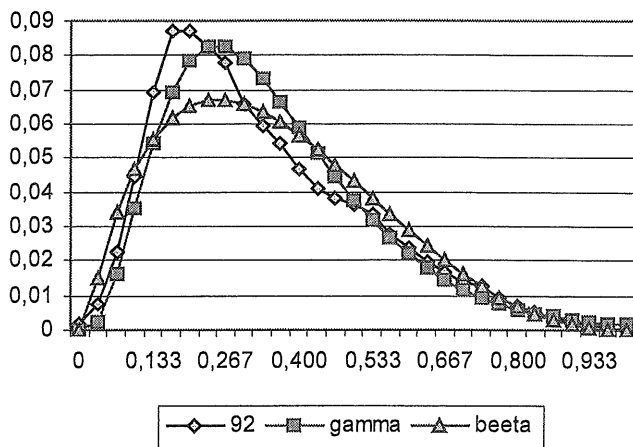
$$\beta = \frac{D}{M}, \quad \gamma = \frac{M}{\beta}.$$

Siin M ja D tähistavad vastavalt jaotuse keskvaartust ja dispersiooni.

Beeta jaotuse parameetrid leitakse seostest:

$$\alpha = \frac{M}{D}(M(1-M) - D), \quad \beta = \frac{1-M}{M}\alpha.$$

Selliselt lähendati erinevate aastate sündimuskõveraid. Leitud lähendid on üsna jämedad; pole kasutatud mingeid parameetrite korrigeerimise



Joonis 2. Sündimuskõverate lähendamine beeta- ja gammajaotusega

meetodeid. Ilmneb, et sündimuskõverad olid 1990. aastate esimesel poolel hästi lähendatavad gammajaotusega, kuid perioodi lõpul sobib lähendiks paremini beetajaotus (vt ka joonis 2, 1992 ja 2000).

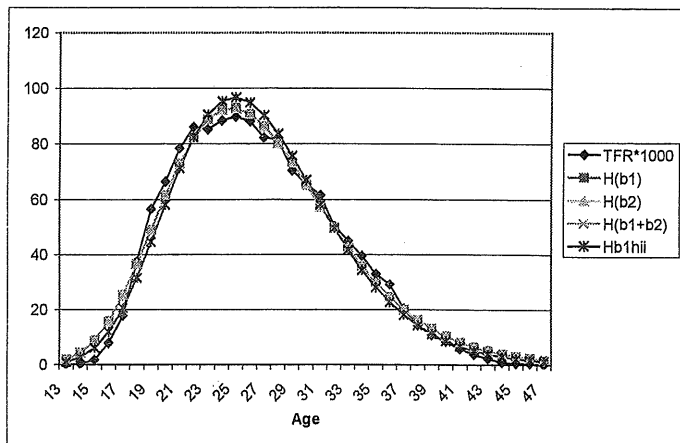
Seega võime järeldada, et 20. sajandi lõpuaastateks on eesti naiste sündimuskäitumine põhimõtteliselt muutunud, sünnitused ei ole enam koondunud lühikesse ajaperioodi naise viljaka ea alguskümnendil, vaid jagunevad viljaka ea ulatuses mõnevõrra ühtlasemalt.

Hadwigeri funktsioon ja selle kasutamine

Paljudes Euroopa maades¹ on sündimuskõverate lähendamiseks kasutatud Hadwigeri funktsiooni [Gilje, 1969; Chandola *et al*, 1999], mis avaldub järgmisel kujul [Hadwiger, 1940]:

$$h(x) = \frac{\alpha\beta}{\gamma\sqrt{\pi}} \left(\frac{\gamma}{x}\right)^{\frac{3}{2}} \exp\left\{-\beta^2\left(\frac{\gamma}{x} + \frac{x}{\gamma} - 2\right)\right\},$$

kus x tähistab sünnitaja vanust ning α , β , γ on hinnatavad parameetrid, mille hinnangud tähistame vastavalt a , b , c .



Joonis 3. Sündimuskõvera (Eesti, 2000) lähendamine Hadwigeri jaotusega

¹ Näiteks Austrias, Taanis, Prantsusmaal, Rootsis.

Tähistame vanuselise sündimuskordaja f_x (vanuses x) ja defineerime summa $R_K = \sum_x x^K f_x$. Siis saab Hadwigeri funktsiooni parameetrite hinnangutele anda sisulise interpretatsiooni. Sel korral on R_0 summaarne sündimuskordaja ja ta on parameetri α hinnanguks ($\alpha=R_0$). Parameetri γ hinnang c avaldub sel juhul suhtena $c=R_1/R_0$.

Parameetri β hindamiseks on mitu varianti: b võib omandada väärtuse b_1, b_2 või b_1+b_2 , kus

$$b_1 = \sqrt{\frac{0,5R_1^2}{R_0R_2 - R_1^2}} \text{ ja } b_2 = \frac{cf_c\sqrt{\pi}}{R_0}.$$

Modelleerimisel katsetati kõiki parameetri b valikuvõimalusi, erinevused saadud tulemustes polnud märgatavad. Kooskõla hindamiseks kasutati hälvete ruutude summat ja χ^2 -statistikut.

Üldistatud Hadwigeri funktsiooni ja jaotuste segu kasutamine

Üldistatud Hadwigeri funktsioon saadakse, kui vanust korrigeeritakse [Gilje, 1999] ja arvestatakse vanust korrigeerivat parameetrit d . Seega on üldistatud Hadwigeri jaotusel järgmine kuju:

$$H(x) = \frac{\alpha\beta}{\gamma\sqrt{\pi}} \left(\frac{\gamma}{x+d} \right)^{\frac{3}{2}} \exp \left\{ -\beta^2 \left(\frac{\gamma}{x+d} + \frac{x+d}{\gamma} - 2 \right) \right\}$$

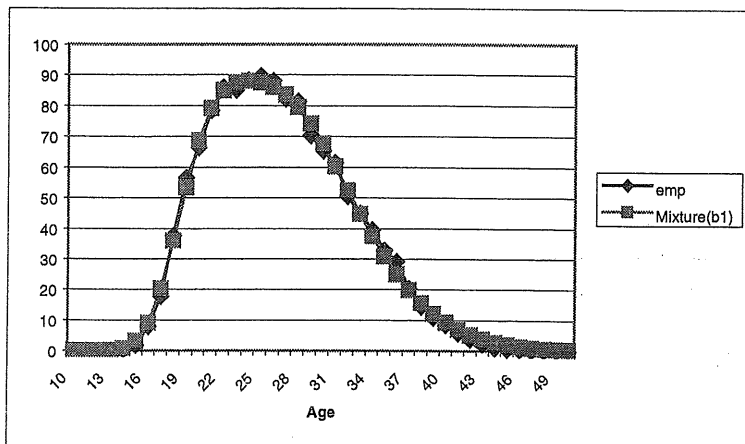
Soovitav korrigeeriva parameetri väärtus $d > -15$.

Kui uuritav populatsioon koosneb mõnevõrra erinevatest alampopulatsioonidest (nt Suurbritannias või Iirimaal), soovitatakse kasutada jaotuste segu [Chandola *et al*, 1999]

$$h(x) = mh_1(x) + (1-m)h_2(x), \quad (1)$$

kus $h_1(x)$ ja $h_2(x)$ on Hadwigeri funktsioonid alampopulatsioonides ja m ($0 < m < 1$) seguparameeter. Eestis on eestlaste ja muulaste sündimuskäitumine mõnevõrra erinev, mistõttu on mõeldav sündimuskõvera lähendamine kahekomponendilise seguga kujul (1).

Selleks hinnati Hadwigeri jaotuse parameetrid mõlemas populatsioonis eraldi ning segu kaalud määrati vastavalt eestlaste ja muulaste osakaalu-
le sünditajate seas. Siiski ei andnud niiviisi määratud parameetrid
optimaalset lähendit ning parameetrite täpsustamine (vähimruutude
mõttes) andis lõpptulemuseks jaotuse, mis empiirilist sündimuskõverat
väga hästi lähendas, vt joonis 4.



Joonis 4. Sündimuskõverate lähendamine Hadwigeri jaotuste seguga

Kahe Hadwigeri jaotuse segu on seitsmeparameetriline jaotus, mis võimaldab empiirilist sündimuskõverat väga hästi lähendada, kuid selle parameetritel ei tarvitse olla sisulist tähendust. Sündimuskõverat hakati diskreetse seguga lähendama ka Eesti Statistikaametis, kus lähendamiseks kasutati kahe beetajaotuse segu [Maamägi, 2002].

Kirjandus

1. Chandola, T, Coleman, D. A., Hiron, R. W. (1999) Recent European fertility patterns: Fitting curves to 'distorted' distributions. *Population Studies*, 53, 317–329.
2. Eesti meditsiiniline sünniregister 1992–2000. Eesti abodiregoster 1996–2000. (2001) EKMI, Tallinn.
3. Gilje, E. (1969) Fitting curves to age-specific fertility rates: Some examples. *Statistical Review of the National Central Bureau of Statistics of Sweden*, 3rd series, 7, 118–134.
4. Maamägi, A. (2002). Vanusekordajate tasandus. *Rahvastik 2001*, ESA, 94.
5. *Rahvastik 2001* (2002), Tallinn, ESA.

EESTI KOOLIJUHTIDE TÖÖSTRESS JA LÄBIPÕLEMINE

Mart Kals¹, Eda Merisalu², Krista Fischer²

¹TÜ matemaatilise statistika instituut

²TÜ tervishoiu instituut, kristal@ut.ee

Käesoleva uuringu aluseks on Eesti koolijuhtide andmestik, mis põhineb TÜ tervishoiu instituudi dotsendi Eda Merisalu tehtud ankeetküsitlusel. Kõigi maakondade koolijuhtide teabepäevadel (1999–2001) tutvustati Eesti õpetajate tööstressi ja läbipõlemisuuringu tulemusi ning paluti haridusjuhtidel täita ankeetküsimustik, mida oli kasutatud õpetajate uuringus 1997. a kevadel.

Järgnevalt tutvustame tagastatud ankeetide põhjal loodud andmestiku analüüsitulemusi. Analüüsi peaeesmärgiks on välja selgitada peamised koolijuhtide tööstressi tekitavad riskitegurid ning leida sobivad mudelid koolijuhtide läbipõlemise kirjeldamiseks.

Küsitlusuuringu tagapõhi

Õpetajaametit võib tänapäeval pidada väga stressirohkeks. Rahvusvahelised uuringud on näidanud, et kolmandik õpetajatest on väga kõrges stressis. Arvatakse, et rohkem tööpinget on reformitavates riigikoolides, eelkõige alg- ja põhikoolides.

Eesti keeruline ja kiiresti muutuv sotsiaal-majanduslik olukord ei ole jätnud puutumata ka koolikeskkonda. Õpetajatööd saadab suur suhtlemispinge probleemsete õpilaste ja nende vanematega. Stressi tekitab töö suurearvuliste klassidega ja vastutus. Suhtlemispinged juhtkonna või kolleegidega ei mõjuta üksnes kooli mikrokliimat, vaid halvendavad ka õpetajate töövõimet ja töö tulemuslikkust [Merisalu, 2000]. Seega peab koolijuht juhtima paindlikult sageli väga keerukat struktuuri ja lahendama tihti ka külaelu probleeme. Kuna tegemist on koolistressi kui suure sotsiaalse probleemiga, peavad empiirilised uuringud aitama meil mõista selle tagamaid.

On rohkelt tõendeid, et paljud õpetajad, kes puutuvad kokku kõrge ja pikaajalise tööstressiga, tunnetavad sagedamini depressiivsust, kurnatust ja vähenevat töövõimet, mis omakorda kutsuvad esile mitmesuguseid haigusi ja enneaegset töölt lahkumist. Kroonilise tööülevinge foonil võib kujuneda *läbipõlemissündroom*. Viimast iseloomustab emotsionaalne kurnatus, depersonalisatsioon ning süvenev madal enesehinnang. Emotsionaalne kurnatus viitab sellele, et ollakse pidevast kõrgeast suhtlemispingest psühholoogiliselt väsinud. Depersonalisatsioon osutab inimese kütünilisele ja tundetule suhtumisele nendesse, kellele on suunatud tema teenus – klientidesse, patsientidesse või õpilastesse. Inimese alanenud enesehinnang tuleneb vähesest professionaalsest efektiivsusest ja kroonilise väsimuse foonil langenud tööproduktiivsusest. Kuna juhtivamet nõuab suurt vastutust ja kõrget psühhootsionaalset pinget, siis tuleb eelkõige uurida koolijuhtide tööstressi ja läbipõlemise seoseid ning nende nähtuste mõju tervisele. Samuti on uurimused näidanud, et juhi läbipõlemise tagajärjed on rasked nii talle endale kui ka alluvatele.

Metoodika

Uuritavad. Kokku jaotati küsimustikke käestkätte 400 haridusjuhile, kelle hulgas olid koolidirektorid, nende asetäitjad, õppealajuhatajad, ja omavalitsusjuhile (viimaste ankeedid jäeti käesolevast uuringust välja). Küsimustiku tagastas vaid 139 haridusjuhti. Üks põhjusi, miks vastamismäär (34,8%) nõnda madalaks jäi, võis olla üsna konfidentsiaalsete küsimuste sisaldumine ankeedis, millest võis tuleneda uuritavate hirm andmete avalikustamise ees. Teine põhjus võis olla emotsionaalne kurnatus, mille foonil näib iga lisakohustus üleliigse pingutusena.

Kasutatav küsimustik. Uuringus kasutati anonüümset ankeetküsimustikku, mis koosnes viiest osast.

- Demograafilised ja ametialased näitajad.
- AB-test [Rosenman *et al*, 1964].
- Adizes' enesetest [Adizes, 1996].
- Maslachi läbipõlemisküsimustik [Maslach, Jackson, 1981].
- Töökeskkonna küsimused [Cooper, 1984; Mykletun, 1996; Merisalu, 2000].

Ankeedi üldandmete osa sisaldas 14 küsimust **demograafiliste ja ameti näitajate** kohta. Need olid sugu, vanus, rahvus, perekonnaseis, laste arv, haridustase, töökoht, amet, tööstaaz hariduses ja praegusel ametikohal, töökoormus, töötasu, personali arv koolis kokku ning see, kas inimene õpib töökoha kõrvalt.

AB-testi kasutati juhtide käitumistüübi kindlakstegemiseks. AB-testi 20 küsimuse abil sai välja selgitada, kas juht on A-, B- või C-tüüpi käitumisega. A-tüüpi inimest iseloomustab füsioloogiliselt domineeriv sümptaatiline erutus. Ta on alati äärmiselt hõivatud, kohusetundlik, töökas ja tõeline protsesside edasiviija. B-tüüpi inimene on närvisüsteemi eripäralt tasakaalustatud, alati valmis koostööks ja igati hea meeskonnaliige. C-tüüpi inimesel on kummagi eelneva tüübi iseloomujooni, ta on loova naturiga, usaldatav, kontrollib oma elu ja tööd.

Adisez' enesetesti kasutati juhtimisstiili määramiseks. Selle abil sai igale inimesele leida tema P-A-E-I profiil. Tähed P, A, E, I tulevad ingliskeelsetest sõnadest produtseerija (*producer*), administraator (*administrator*), ettevõtja (*entrepreneur*) ja integraator (*integrator*) — need on juhtimisstiili tüübid, mida enesetesti 27 küsimusega oli võimalik määrata. Loetletud juhitüüpide isiksuse omadused on järgmised: produtseerija on loomult aktiivne, domineeriv ning kergelt haavatav, administraator vajab tunnustust, on skeptiline, tugeva enesekontrolliga ning pidevas ülepingeohus, ettevõtjat iseloomustab sõltumatus, mittekiirustav stiil ning paindlikkus, integraator on hea lõõgastuja, tagasihoidlik, töös tugineb meeskonnale.

Maslahi läbipõlemisküsimustiku abil tuvastati juhtide läbipõlemist. Esitatud 22 küsimuse põhjal sai uurida kolme läbipõlemise indikaatorit — emotsionaalset kurnatust, depersonalisatsiooni (küünilisust) ja enesehinnangut. Vastajad hindasid ennast hinnangu skaalal 0–5, kus võimalikud vastusevariandid olid: 0 — mitte kunagi, 1 — paar korda aastas, 2 — kord kuus, 3 — kord nädalas, 4 — mitu korda nädalas, 5 — iga päev.

Töökeskkonna küsimuste põhjal oli eesmärgiks välja selgitada koolijuhtide tööstressi tekitajad. Kokku oli küsimusi 58, mille korral koolijuht pidi hindama esitatud väidet kahest aspektist lähtudes: kui sageli on vaadeldav probleem viimase tööaasta jooksul esile kerkinud ja kui palju on nimetud probleem vastajat häirinud. Võimalikud variandid vasta-

miseks sageduse skaalal olid 0–4, kus: 0 — mitte kunagi, 1 — harva, 2 — pool tööajast, 3 — kolmveerand tööajast, 4 — enamus tööajast. Tugevuse skaalal olid vastusevariandid vahemikus 0–5 vastavalt: 0 — ei häiri, 1 — (häirib) mõningal määral, 2 — (häirib) vähe, 3 — (häirib) mõõdukalt, 4 — (häirib) tugevalt, 5 — (häirib) väga tugevalt. Uurimuses kasutati vastuseid, mis puudutasid probleemi esinemis-sagedusi.

Kasutatud statistikameetodid. Andmete analüüsimisel kasutati järg-misi statistikameetodeid:

- faktoranalüüsi tööstressi tekitajate leidmisel;
- regressioonanalüüsi koolijuhtide läbipõlemise kirjeldamiseks.

Ankeetküsitluse andmete analüüs

Töökeskkond. Andmestiku töökeskkonda puudutava osa analüüsil kasutame faktoranalüüsi (peakomponentide meetodil varimaks-pöörami-sega [Ehasalu, Tiit, 1993]), mis võimaldab grupeerida 58 töökeskkonna küsimust 10 faktoriks.

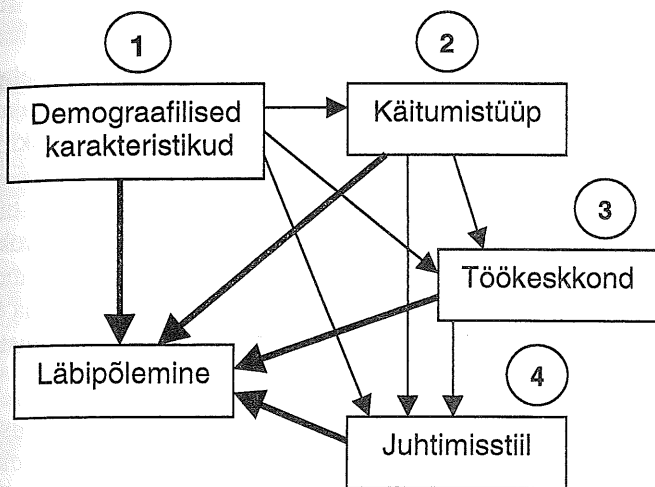
Saadud faktoreid oli sobiv nimetada järgmiselt:

1. töö ülekoormus (9 küsimust);
2. koostöö- ja suhtlemisprobleemid (11 küsimust);
3. tunnustuse ja juhtnõrde puudumine (4 küsimust);
4. puudulik väljaõpe ja tööalane ebakompetentsus (5 küsimust);
5. rusuv tööõhkkond (7 küsimust);
6. kartus koondamise ja juhtkonna ees (3 küsimust);
7. töö ja kodu konflikt (6 küsimust);
8. halb tööruumide mikrokliima (3 küsimust);
9. rahulolematuse karjääriga (2 küsimust);
10. töökoha ergonoomilised probleemid (3 küsimust).

Kõigi faktorite puhul vastas tõsisemale probleemile faktori suurem vää-rtus. Neid faktoreid kasutame edaspidi argumentidena koolijuhi läbi-põlemise prognoosimisel.

Põhjuslikud seosed andmestikus ja mudeli hierarhiline ülesehitus. Vastavalt põhjuslikele seostele andmestikus valime statistilise metoo-dika koolijuhtide läbipõlemise prognoosimiseks. Eeldame, et kehtib olu-kord (vt joonis 1), kus demograafilised karakteristikud võivad mõjutada

käitumistüüpi, juhtimisstiili, hinnangut töökeskkonnale ja läbipõlemist. Käitumistüüp võib mõjutada hinnangut töökeskkonnale ja juhtimisstiili. Töökeskkond võib mõjutada juhtimisstiili. Samuti eeldame, et kõik neli komponenti võivad mõjutada läbipõlemist. Nende mõjude selgitamine ongi käesoleva töö peaesmärgiks.



Joonis 1. Läbipõlemise mõjurite hüpoteetiline mudel koolijuhtide uuringus, mis tugineb põhjuslikele seostele andmestikus

Tuginedes põhjuslikele seostele andmestikus kasutame koolijuhtide läbipõlemise kirjeldamiseks hierarhilist mitmest regressioonanalüüsi, sisestades ülalnimetatud tunnusgruppe mudelisse järjest vastavalt joonisel 1 toodud järjekorranumbrile. Funktsioontunnusteks on kolm läbipõlemise näitajat (emotsionaalse kurnatus, küünilisus ja enesehinnang). Argumenttunnuste mõju statistiline olulisus määrati vahetult pärast vastava bloki lisamist mudelisse.

Seega sisestame esimese blokina mudelisse individuaalsed demograafilised ja tööalased näitajad. Sellega arvestame nii demograafiliste karakteristikute otsest mõju läbipõlemisele kui ka kaudseid mõjusid käitumistüüp, töökeskkond ja juhtimisstiil. Järgmisena lisame mudelisse käitumistüübi. Sellega hindame summaarselt nii käitumistüübi otsest mõju läbipõlemisele kui ka kaudseid mõjusid läbi töökeskkonna ja juhtimisstiili. Analoogiliselt lisame kolmanda blokina töökeskkonna kü-

simuste põhjal leitud 10 tööstressi tekitajat ning viimase blokina juhtimisstiili, mille otsesest mõju läbipõlemisele oli võimalik hinnata.

Tulemused

Hierarhilise mitmese regressioonanalüüsi tulemused on toodud tabelis 1. Otsuste tegemisel kasutame olulisuse nivood 0,1.

Alustame emotsionaalset kurnatust põhjustavate tegurite väljaselgitamisega. Esimese blokina arvesse võetud demograafilistest näitajatest osutus oluliseks vanus. Siit võib järeldada, et koolijuhtide keskmine emotsionaalne kurnatus väheneb vanuse suurenedes. Ka järgnevalt lisatud käitumistüübi mõju osutus oluliseks – seega, mida rohkem A-tüüpi käitumisega on juht, seda suurem on tema kurnatus. Kolmanda blokina lisatud töökeskkonna tunnustest osutus oluliseks kolme faktori mõju: töö ülekoormus, rusuv tööõhkkond, kartus koondamise ja juhtkonna ees. Viimasena arvesse võetud juhtimisstiili tunnustest osutus olulisteks stiilide P ja A mõju emotsionaalsele kurnatusele.

Pärast kõigi blokkide arvessevõtmist saame hinnatud mudelis (osaliselt) eristada tunnuseid, mis mõjutasid läbipõlemist otseselt ja mis kaudselt. Emotsionaalse kurnatuse korral jäid kõik tunnused peale käitumistüübi lõplikus mudelis olulisteks (vastavad tulemused on esitatud tabeli 1 parempoolses osas). Seega ei mõjuta käitumistüüp juhi emotsionaalse kurnatuse taset otseselt, vaid suhtumise kaudu töökeskkonda ja juhtimisstiili. Kõigil ülejäänud argumenttunnustel on tõenäoliselt kas otsene mõju läbipõlemisele või kaudne mõju läbi vaatlemata/mõõtmata tunnuste. Sarnased arvutused tehti ka juhtudel, kui sõltuvateks tunnusteks olid läbipõlemise näitajad depersonalisatsioon ja enesehinnang.

Järgnevalt on esitatud kokkuvõtte hierarhilise mitmese regressioonanalüüsi tulemustest — teguritest, mis mõjutasid koolijuhtide läbipõlemise taset.

Demograafilised näitajad. Uuringus osalenud haridusjuhtide keskmine vanus oli 47,2 aastat (min 27, max 66), 56% olid naised ja 44% mehed. Küsitletuist 45% töötas põhikoolis, 29% keskkoolis või gümnaasiumis, 13% lasteaedalgkoolis ja vaid 4% algkoolis (analüüsimisel ühendati viimased kaks rühma). Koolijuhtidest 67% töötas rohkem kui ühe kohaga ja 28% õppis töö kõrvalt.

Tabel 1. Läbipõlemise mõjurid.

	Hierarhiline mudel			F ²	Lõplik mudel		
	β	Std- viga	p		β	Std- viga	p
Emotsionaalne kurnatus							
Demograafilised näitajad				0,08			
Vanus	-0,02	0,007	0,001		-0,01	0,005	0,009
Käitumistüüp				0,22			
AB	0,55	0,111	<,0001		0,15	0,098	0,117
Töökeskkond				0,53			
Töö ülekoormus	0,43	0,077	<,0001		0,40	0,078	<,0001
Rusuv tööhkkond	0,50	0,141	0,001		0,49	0,144	0,001
Koondamis- ja juhtkonnakartus	0,21	0,089	0,021		0,23	0,089	0,012
Juhtimisstiil				0,55			
Produtseerija	2,29	1,324	0,086		2,29	1,324	0,086
Administraator	2,66	1,228	0,032		2,66	1,228	0,032
Depersonalisatsioon							
Demograafilised näitajad				0,12			
Vanus	-0,02	0,007	0,005		-0,01	0,007	0,169
Töökoht: algkool	-0,25	0,151	0,101		-0,16	0,134	0,244
Amet	0,41	0,219	0,065		0,23	0,190	0,237
Tööstaaž praegusel kohal	0,02	0,008	0,034		0,01	0,007	0,196
Personal koolis kokku	0,004	0,002	0,083		-0,001	0,002	0,797
Käitumistüüp				0,19			
AB	0,34	0,098	0,001		0,18	0,094	0,053
Töökeskkond				0,35			
Koostöö- ja suhtlemisprobleemid	0,41	0,149	0,007		0,40	0,149	0,008
Koondamis- ja juhtkonnakartus	0,27	0,097	0,006		0,30	0,096	0,002
Töö ja kodu konflikt	0,18	0,095	0,056		0,16	0,094	0,097
Rahulolematus karjääriga	-0,14	0,077	0,081		-0,17	0,076	0,026
Juhtimisstiil				0,40			
Administraator	3,33	1,301	0,012		3,33	1,301	0,012
Enesehinnang							
Demograafilised näitajad				0,07			
Amet	-0,49	0,246	0,047		-0,45	0,216	0,038
Tööstaaž hariduses	0,02	0,006	0,010		0,01	0,006	0,059
Käitumistüüp				0,09			
Töökeskkond				0,30			
Koostöö- ja suhtlemisprobleemid	-0,85	0,161	<,0001		-0,83	0,165	<,0001
Puudulik väljaõpe ja tööalane ebakompetentsus	-0,25	0,114	0,029		-0,30	0,117	0,013
Rahulolematus karjääriga	0,20	0,091	0,028		0,22	0,091	0,016
Juhtimisstiil				0,33			
Administraator	-2,87	1,545	0,066		-2,87	1,545	0,066

Emotsionaalse kurnatuse ja küünilisuse ilmingud on seda tugevamad, mida noorem on koolijuh. Enesehinnang on madalam lühikest aega

töötanud juhtidel. Seega mõjutab haridusjuhi läbipõlemist tema vanus ja tööstaž – mida noorem on koolijuht, seda suurem on läbipõlemise oht.

Käitumistüüp. Ootuspäraselt leidis oluline seos juhi käitumistüübi ja läbipõlemistaseme vahel. A-tüüpi käitumine, mida iseloomustab jõuline ja kiirustav stiil, põhjustas saadud hinnangutele tuginedes nii emotsionaalset kurnatust kui ka depersonalisatsiooni (küünilisust). Siin pole ka midagi imestada, kuna A-tüüpi inimene arvab, et kui tema suudab üleloomulikult suurt töökoormust taluda, siis peavad seda suutma ka teised. Teiste inimeste jõuvarude mittearvestamine põhjustabki küünilisust.

Töökeskkond. Hinnatud mudeli põhjal on alust väita, et emotsionaalset kurnatust kutsub esile rusuv töökeskkond, ülekoormus, kartus juhtkonna ja koondamise ees. Küünilisust süvendavad koostöö- ja suhtlemisprobleemid, hirm koondamise ja juhtkonna ees, rahulolematuse karjääritegemise võimalustega ning probleemid perekonnas. Koostöö puudumine, suhtlemisprobleemid ja tööalaste teadmiste vähesus on seotud madala enesehinnanguga. Samas olid madalama enesehinnanguga uuritavad oma karjääriga koolijuhina rohkem rahul kui kõrgema enesehinnanguga ja ambitsioonikamad vastajad.

Juhtimisstiil. Juhtimisviisidest oli kõigi läbipõlemisnäitajatega seotud juhistiil A. Seega on administratiivse juhtimisstiiliga koolijuhid kõige vastuvõtlikumad läbipõlemisele.

Arutelu ja järeldused

Käesolevas töös näitasime, et psühholoogilise küsitlusuuringu andmete korrektne ja sisuliselt tõlgendatav statistiline analüüs on võimalik vaid siis, kui on hoolikalt läbi mõeldud andmetes esineda võiv põhjuslik seosestruktuur. Koolijuhtide uuringu puhul on see skemaatilisel kujutatud joonisel 1. Põhjuslike seoste skemaatilisest kujutamisest ning nende hindamise võimalikkusest on lähemalt kirjutanud Robins [2001].

Joonisel 1 toodud skeemi tuleb vaadelda kui andmeanalüüsi eeldust — kirjeldatud hierarhilise regressioonanalüüsi tõlgendus on jõus vaid selle eelduse kehtimisel. Kahtlusi ei tekita eeldus, et inimese sugu, vanus ja mõned teised demograafilised näitajad võivad küll mõjutada teisi tunnusblokke (töökeskkond, juhistiil, käitumistüüp ja läbipõlemine), aga ei ole ise nende poolt mõjutatavad. Samas eeldame, et juhtimisstiil ei

mõjuta juhi käitumistüüpi ega suhtumist töökeskkonda ning töökeskkond (ja sellesse suhtumine) ei mõjuta käitumistüüpi. Lõpuks eeldame ka seda, et läbipõlemise näitajad ise ei mõjuta ühtegi argumentidena kasutatud tunnusblokki. Viimased eeldused on küll psühholoogilisest vaatenurgast põhjendatud, ometi puudub absoluutne kindlus nende kehtimises. Seega tuleb ka analüüsi tulemustesse suhtuda reservatsioonidega.

Arvestades, et joonisel 1 toodud struktuur on siiski üsna realistlik, saame analüüsile tuginedes järeldada, et eesti koolijuhtide läbipõlemistase sõltub üsna oluliselt demograafilistest näitajatest, käitumistüübist, töökeskkonna teguritest ning juhtimisstiilist.

Emotsionaalne kurnatus

- Kurnatuse all kannatavad rohkem nooremad juhid.
- Kurnatuse tase on kõrgem A-tüüpi käitumisega juhtidel, samuti neil, kes on produtseeriva või administratiivse juhtimisstiiliga.
- Juhtide vastuvõtlikkust emotsionaalsele kurnatusele suurendab rusuv tööõhkkond, kartus koondamise ja juhtkonna ees ning ülekoormusega töötamine.

Küünilisus

- Mida noorem on juht, seda küünilisemalt ta oma kaastöötajatesse suhtub. Põhi- ja keskkoolis töötavad juhid on depersonalisatsioonile vastuvõtlikumad, võrreldes kolleegidega algkoolist. Küünilisemalt suhtuvad ka juhid, kes peavad direktori või juhataja ametit, kes on töötanud pikka aega ühel kohal ning kellel on palju alluvaid.
- A-tüüpi inimesed, samuti administratiivse stiiliga juhid kalduvad oma kaastöötajatesse tundetult suhtuma.
- Mida suuremad on probleemid suhtlemisel ning perekonnas, mida suurem on kartus koondamise ja juhtkonna ees, seda küünilisem on juht. Küünilisus on suurem ka paremate karjäärivõimalustega juhtidel.

Enesehinnang

- Direktoriks olek alandab, kuid pikaajaline töötamine tõstab juhi hinnangut oma kompetentsile.
- Mida kergem on teha koostööd ja suhelda kaastöötajatega ning mida suurem on tööalane teadlikkus, seda kõrgema enesehinnanguga on

juht. Madalama enesehinnanguga juhid on oma karjääriga koolijuhina rohkem rahul kui kõrgema enesehinnanguga uuritavad.

- Madala enesehinnanguga on juhid, keda iseloomustab administratiivne juhtimisstiil.

Järelikult on koolijuhi läbipõlemistase seda suurem, mida

- noorem ta on,
- rohkem A-tüüpi käitumisega ta on,
- suuremad on koostöö- ja suhtlemisprobleemid,
- suurem on kartus koondamise ja juhtkonna ees,
- väiksemad on tema erialateadmised ja väljaõpe,
- rohkem administraatori tüüpi ta on.

Antud järeldustesse tuleb suhtuda ettevaatlikult, sest töös on kasutatud väikest andmestikku, võrreldes suure tunnuste hulga. Peale selle vastas koolijuhtidest suhteliselt vähe, vaid 34,8%. Tulevikus tuleb saadud tulemusi kinnitada täiendavate uuringutega, mistõttu käesolevaid tulemusi võib vaadelda kui pilootprojekti. Juhtivas-vastutavas ametis läbipõlemist soodustavate tegurite tundmaõppimine aitab teadlikult ära hoida selle sündroomi raskemaid tagajärgi.

Kirjandus

1. Merisalu, E. (2000). Eesti õpetajate tööstress ja tervis. Tartu.
2. Rosenman, R. H., Friedman, M. and Straus, R. (1964). A predictive study of CHD. *Journal of the Medical Association*, 189, 15–22.
3. Adizes, I. (1996). The Pursuit of Prime: Maximize Your Company's Success with the Adizes Program. The Knowledge Exchange.
4. Maslach, C., Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Behavior*, 2, 99–113.
5. Cooper, C. L. 1984. Executive stress: a ten country comparison. *Human Resource Management*, 23, 395–407.
6. Mykletun, R. J. (1996). Working environment and health. Stavanger.
7. Ehasalu, E., Tiit, E.-M. (1993). Komponent-, faktor- ja kanooniline analüüs. Tartu: Tartu Ülikool.
8. Robins, J. (2001). Data, Design, and Background Knowledge in Etiologic Inference. *Epidemiology*, 11, 3, 313–320.

RAHVUSVAHELISE KOOLINOORTE SUITSETAMIS- UURINGU MITMETASANDILISED MUDELID

Triin Umbleja¹, Kersti Pärna², Krista Fischer²

¹TÜ matemaatilise statistika instituut

²TÜ tervishoiu instituut, kristal@ut.ee

Sissejuhatus

Mitmed uuringud on näidanud erinevate tegurite seotust suitsetamisega nii otseselt kui ka kaudselt, nii üksikuna kui ka kombineeritult [Green & Kreuter, 1991]. Kooliõpilaste suitsetamisel on kõige tugevamad seosed leitud sõprade suitsetamisega [Urberg *et al*, 1997; Williams & Covington, 1997], samuti vanemate ja õdede-vendade suitsetamisega (Zhu *et al*, 1996; Duncan *et al*, 1995). Enamlevinud on aga nende tegurite käsitlemine indiviidi tasandil, kusjuures laiem sotsiaalne kontekst (koolikeskkond, elamispiirkond) jäetakse kõrvale [Karvonen & Rimpelä, 1996; Duncan *et al*, 1993]. Kirjanduse andmetel on koolil kui sotsiaalsel keskkonnal mõju kooliõpilaste suitsetamisele [Karvonen & Rimpelä, 1996]. Mitmetasandiline andmeanalüüs võimaldab kirjeldada suitsetamise seoseid erinevate teguritega nii indiviidi kui ka sotsiaalse keskkonna tasandil [Rice & Leyland, 1996; O'Campo *et al*, 1997].

Töö eesmärgiks on kirjeldada suitsetamise levimust Tallinna, Helsingi ja Moskva kooliõpilaste seas ning leida tegurid, mis on seotud tubaka tarbimisega kolme riigi tütar- ning poeglastel nii indiviidi kui ka sotsiaalse keskkonna tasandil.

Andmete kogumise meetodika

Uuritavad

Uuritavad olid Tallinna ja Helsingi 8., 10., 12. klassi õpilased ning Moskva 9., 10., 11. klassi õpilased (vanus 13–18 aastat). Uuring tehti Tallinna kümnes eesti ($n = 1268$) ja seitsmes vene õppekeelelega

($n = 901$) koolis, Helsingis kuues ($n = 1396$) ja Moskvas üheksas ($n = 618$) koolis. Koolid valiti uuringusse juhulikumult telefonikataloogi alusel, välja jäeti süvaõppe- ja erikoolid. Helsingis koguti andmed sügisel 1994, Moskvas kevadel 1995 ja Tallinnas sügisel 1995.

Uuringumeetodid

Andmed koguti ankeetküsitlusega. Õpilased täitsid ankeedid ühe õppetunni jooksul klassiruumis uuringugrupi liikmete juuresolekul. Ankeetide täitmine oli anonüümne. Ankeet koosnes 79 tervisekäitumise ja keskkonnatervise valdkonna küsimusest. Käesolevas töös on kasutatud küsimusi õpilase soo, vanuse, suitsetamisharjumuse, religiooni, perekonna tüübi, perekonna ülalpidaja hariduse ja vanemate töö kohta. Lisanduvad veel küsimused laste arvu kohta perekonnas ning sõprade, õdede-vendade, vanemate ja õpetaja suitsetamisest.

Statistilised meetodid: mitmetasandilised mudelid

Saadud andmestiku eripäraks on klastriline struktuur, kus klastriteks on koolid. Üksikindiiviidi suitsetamise staatus on kindlasti mõjutatav suitsetamistasemest tema klassis ja koolis. Seda arvestades on sobiv kasutada mitmetasandilisi mudeleid – segamudelite erijuhtu, mis arvestab grupi juhuslikku mõju. Järgnevalt tutvustame lähemalt mitmetasandiliste andmete olemust, selgitame vastavate mudelite ideed ning meetoodika kasutatavust antud uuritud andmestikul.

Mitmetasandilised andmed

Mis on mitmetasandilised andmed?

Ideed, kuidas statistilist modelleerimist ja analüüsi mitme tasandiga andmete korral rakendada, tekkisid 1980. aastate keskpaigas. 1990. aastate alguseks oli selles valdkonnas olemas juba rida arvestatavaid tulemusi [Goldstein, 1995]. Mitmetasandiliste andmete korral koosneb üldkogum objektidest, mis on grupeerunud või grupeeritud eri tasanditel. Neid grupe nimetatakse *klastriteks* ning öeldakse, et andmed on klastrilise ehk hierarhilise ehk mitmetasandilise struktuuriga. Üldiselt on hajuvus igal sellisel tasandil erinev.

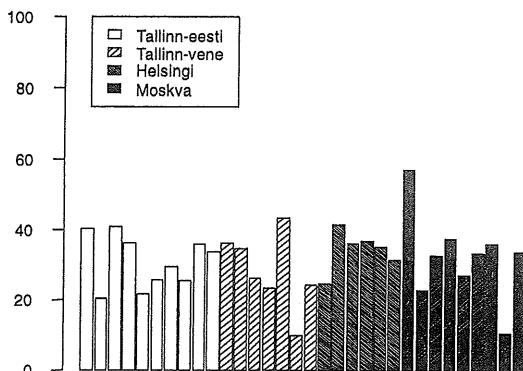
Mitmetasandilised mudelid võimaldavad kaasata ühte mudelisse eri tasandid (nt kooli ja õpilase tasand) ja neid mõjutavad tegurid ning

toovad paremini välja fikseeritud mõjud, sest arvestavad erineva hajuvusega kooli ja õpilase tasandil. Klastrilise struktuuriga andmete puhul saadakse informatsiooni vähem kui juhuslikult valitud valimi korral. Tavaliselt kalduvad grupiliikmed olema sarnased ja grupid üksikestist erineva isegi siis, kui nad on moodustatud juhuslikult. Grupp ja selle liikmeskond mõjutavad üksikest vastastikku: grupp kui tervik määrab, kes sellesse kuulub ning grupi liikmed määravad, missuguste omadustega grupp on. Selliseid mõjusid arvestav mitmetasandiliste mudelite meetodika võimaldab täpsemaid ja usaldusväärsemaid hinnanguid kui klassikalised analüüsimeetodid.

Koolinoorte andmestiku mitmetasandiline struktuur

Käesolevas uuringus on esimese tasandi ühikuteks õpilased. Et kõik loetletud tunnused on mõõdetud õpilastel, on nad esimese tasandi argumenttunnused, tunnus **maa** aga on teise, koolide tasandi argumenttunne.

Osakaal (%)



KOOL

Joonis 1. Suitsetajate osakaal eri maades koolide kaupa

Mudelites kasutatud andmete struktuur on järgmine:

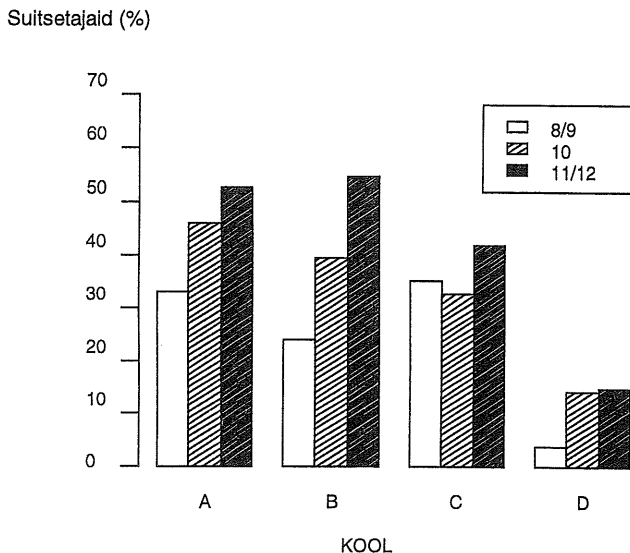
Esimese tasandi mõjud:

- klass (8. ja 9., 10., 11. ja 12.);
- laste arv peres;

- sõprade suitsetamine;
- õdede-vendade suitsetamine;
- vanemate suitsetamine.

Teise tasandi argumenttunnusel **maa** on neli väärtust – Tallinna eesti ja vene koolid, Helsingi koolid ja Moskva koolid.

Joonisel 1 on näha nii koolide- kui maadevaheline varieeruvus. Kõige suurem erinevus suitsetamise levimuses on Moskva koolides, samas on Helsingi ja Tallinna vene koolides tase küllaltki ühtlane. Vaadeldes nelja juhuslikult valitud kooli (A, B, C ja D) õpilaste suitsetamist erinevates klassides, selgub, et nendes koolides erineb nii üldine suitsetamistase kui ka suitsetamistaseme sõltuvus klassist (joonis 2). Nii on näiteks koolis B suitsetamise tase vanemates klassides oluliselt kõrgem kui nooremates, seevastu koolis C on erinevused üsna väikesed. Koolis D on erinevused küll olemas, aga üldine suitsetamistase on oluliselt madalam kui teistes koolides. Arvestades, et ka küsitletud õpilaste arv on kooliti ja klassiti erinev, ei tohi analüüsis arvestamata jätta koolide erinevusi.



Joonis 2. Suitsetamistase (%) neljas juhuslikult valitud koolis klasside kaupa

Kahetasandiline logistiline mudel

Antud näites on tegemist kahetasandiliste andmetega: esimese tasandi ühikud on õpilased ning teise tasandi ühikud koolid. Olgu p_{ij} suitsetamise tõenäosus j -nda kooli i -ndal õpilasel. Kahetasandiline suitsetamise tõenäosuse logistilise regressiooni mudel avaldub siis kujul

$$\text{logit}(p_{ij}) = \alpha_j + \beta_j x_{ij} + \varepsilon_{ij}.$$

Funktsioon logit [Gail, Benichou, 2000] on defineeritud seosega

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right).$$

Argumendiks x_{ij} on klass, milles käib j -nda kooli i -s õpilane (kodeeritud arvudega 1, 2 või 3). Mudeli vabaliige α_j ja tõus β_j avalduvad fikseeritud (α , β) ning juhuslike osade (u_j , v_j) kaudu järgmiselt

$$\alpha_j = \alpha + u_j \quad \text{ning} \quad \beta_j = \beta + v_j,$$

kus

$$u_j \sim N(0, \sigma_u) \quad \text{ja} \quad v_j \sim N(0, \sigma_v).$$

Parameeter α iseloomustab siin keskmist suitsetamise taset üle kõigi koolide ning juhuslik osa u_j j -nda kooli taseme erinevust üldisest tasemest. Analoogiliselt on tõlgendatavad ka tõusu β_j fikseeritud ja juhuslik osa. Seega tuleb kahetasandilise mudeli hindamisel lisaks fikseeritud parameetritele α , β ja juhuslikule veale ε_{ij} hinnata veel kolme juhuslikku parameetrit

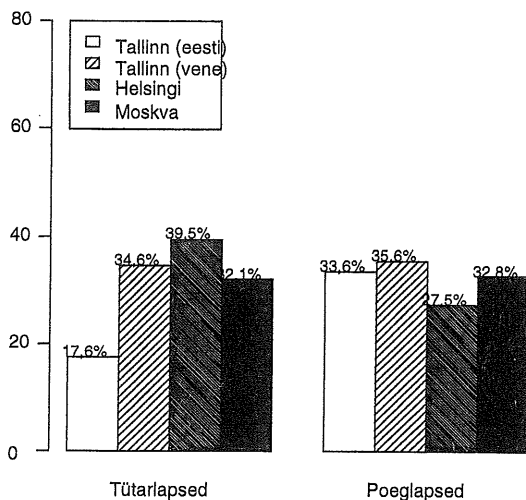
$$\text{Var}(u_j) = \sigma_u^2, \quad \text{Var}(v_j) = \sigma_v^2 \quad \text{ja} \quad \text{Cov}(u_j, v_j).$$

Tulemused

Üldine suitsetamise tase eri maades

Vaadeldes suitsetajate osakaalu sugude kaupa, selgub, et poeglaste hulgas on suitsetamise tase tunduvalt ühtlasem, tütarlaste puhul on erinevused maade vahel suuremad (joonis 3). Samuti on suitsetamise levimusmäär tütarlaste hulgas enamasti madalam kui poeglaste hulgas. Vaid Helsingis on tütarlaste suitsetamise osakaal (39,5%) palju suurem poeglaste vastavast näitajast (27,5%). Need erinevused annavad põhjust

arvata, et poeg- ja tütarlapsed pole suitsetamist mõjutavate tegurite suhtes ühtmoodi vastuvõtlikud, mistõttu tuleb neile eraldi mudelid hinnata.



Joonis 3. Suitsetajate levimus (%) Tallinna, Helsingi ja Moskva kooliõpilaste seas

Koolinoorte andmetel hinnatud suitsetamise mudelid

Et uuritav tunnus SUITSETAMINE on binaarne, siis on kasutatud logistilise regressiooni kahetasandilist versiooni. Tütar- ja poeglastele vastavaid mudeleid on hinnatud eraldi. Selleks on kasutatud tarkvarapaketi SAS/STAT makrot GLIMMIX.

Esmalt on vaadeldud mudeleid tunnustega, mis ei saa ise olla vastaja suitsetamise poolt mõjutatud, mistõttu saab rääkida nende põhjuslikust mõjust suitsetamisele (tabel 1, mudel I). Need tunnused on:

- **maa** (Helsingi, Tallinna vene, Moskva ning Tallinna eesti koolid);
- **klass** (8. ja 9. klass, 10. klass, 11. ja 12. klass);
- **vanemate suitsetamine**;
- **laste arv peres**.

Mudeli hindamisel saadud tulemused on esitatud šansside suhetena ja nende 95%-liste usalduspiiridena. Et mudelis ei ole koosmõjusid, siis kerkitab hästi esile tunnuse **maa** mõju. Nii Tallinna vene, Helsingi kui ka Moskva koolides õppivate tütarlaste šansid suitsetamiseks on oluliselt suuremad kui Tallinna eesti koolides käivate tütarlaste šansid, Helsingis koguni 3,4 korda. Poeglaste korral ei erine suitsetamise tõenäosus Tallinna eesti koolides teistest koolidest oluliselt. Samuti suitsetavad rohkem vanemate klasside õpilased ja need, kelle vanemad suitsetavad. Õdede-vendadega poeglaste hulgas on šansid suitsetada 1,4 korda suuremad kui üksiklastel. Tütarlaste hulgas vastav seos oluline ei ole, kuid selleks, et poiste ja tüdrukute mudelid paremini võrreldavad oleksid, on ka seal jäetud mudelisse laste arv peres.

Seejärel on hinnatud mudelid, millesse on kaasatud kõik olulisteks osutunud tunnused, ka need, mis samal ajal võivad nii mõjutada vastaja suitsetamist, kui ka ise sellest mõjutatud olla. Seetõttu ei pruugi leitud kordajad peegeldada nende näitajate põhjuslikku mõju vastaja suitsetamisele. Seega on lisaks tunnustele, mis ka esimestesse mudelitesse olid kaasatud, lisatud veel sõprade ja õdede-vendade suitsetamine (tabel 1, mudel II).

Suitsetavate sõpradega poeglaste hulgas on šansid suitsetamiseks 4,1 korda suuremad kui teistel. Tütarlaste hulgas on vastav šansside suhe koguni 9,1. Samuti selgub, et tütarlapsed on suitsetamist mõjutavate tegurite suhtes vastuvõtlikumad, st mõjutatavamad kui poeglapsed. Nende korral on mudelis olulised ka koosmõjud maa ja sõprade suitsetamine ning maa ja õdede-vendade suitsetamine. Sõprade mõju on suurim Helsingis, perekonna (õed-vennad) mõju aga Moskvas. Kui õdede-vendade suitsetamise mõju poolest Tallinna eesti ja vene koolid oluliselt ei erine, siis sõprade suitsetamise mõju on Tallinna vene koolides oluliselt väiksem (šansside suhe on 0,3). Sõprade mõju poolest on Tallinna vene koolid sarnasemad Moskva koolidega.

Tabel 1. Seosed suitsetamise ja mitmesuguste faktorite vahel indiviidi ja kooli tasandil, esitatuna šansside suhete abil

Faktor	Mudel I*		Mudel II*	
	Tütarlapsed	Poeglapsed	Tütarlapsed	Poeglapsed
	Šansside suhe (95% UP)	Šansside suhe (95% UP)	Šansside suhe (95% UP)	Šansside suhe (95% UP)
Indiviidi tasand				
<i>Klass</i>				
8. või 9. klass	1	1	1	1
10. klass	1,6 (1,3–2,0)	1,4 (1,1–1,8)	1,3 (1,0–1,7)	1,2 (0,9–1,5)
11. või 12. klass	1,9 (1,5–2,4)	2,1 (1,6–2,7)	1,5 (1,2–2,0)	1,6 (1,2–2,1)
<i>Vanemate suitsetamine</i>				
Ei	1	1	1	1
Jah	1,8 (1,5–2,1)	1,3 (1,0–1,6)	1,7 (1,4–2,1)	1,3 (1,0–1,6)
<i>Õdede, vendade suitsetamine (õde-vend)</i>				
Ei			1	1
Jah			1,5 (0,9–2,5)	1,6 (1,2–2,2)
<i>Sõprade suitsetamine</i>				
Ei			1	1
Jah			9,1 (4,8–17,2)	4,1 (3,2–5,3)
<i>Laste arv peres</i>				
üksiklaps	1	1	1	1
kaks või rohkem	1,0 (0,8–1,2)	1,4 (1,1–1,8)	0,9 (0,7–1,1)	1,3 (1,0–1,7)
Koosmõjud				
<i>Õde-vendxmaa</i>				
Tallinna eesti koolid			1	
Tallinna vene koolid			0,9 (0,5–1,7)	
Helsingi koolid			1,4 (0,7–2,8)	
Moskva koolid			2,0 (1,0–4,2)	
<i>Sõprade suitsetaminexmaa</i>				
Tallinna eesti koolid			1	
Tallinna vene koolid			0,3 (0,1–0,6)	
Helsingi koolid			1,4 (0,6–3,3)	
Moskva koolid			0,4 (0,2–1,0)	
Kooli tasand				
Tallinna eesti koolid	1	1	1	1
Tallinna vene koolid	2,5 (1,7–3,6)	1,0 (0,6–1,8)	7,6 (3,6–15,7)	1,0 (0,6–1,6)
Helsingi koolid	3,4 (2,4–5,0)	0,7 (0,4–1,2)	2,1 (0,9–5,1)	0,7 (0,4–1,1)
Moskva koolid	2,4 (1,6–3,6)	1,0 (0,6–1,7)	4,1 (1,8–9,3)	1,2 (0,7–1,9)
Vabaliikme dispersioon	0,075 (<i>p</i> =0,07)	0,209 (<i>p</i> =0,01)	0,073 (<i>p</i> =0,08)	0,131 (<i>p</i> =0,03)

* Sõprade ja õdede-vendade suitsetamist arvestamata;

** Sõprade ja õdede-vendade suitsetamist arvestades.

Järeldused ja arutelu

Käesolevas töös aitab mitmetasandiline andmeanalüüs eristada indiviidi ja sotsiaalse keskkonna tasandi tegurite mõjusid suitsetamisele. Nii saame tulemustele tuginedes toetust teadmisele, et suitsetamist ennetavad ning vähendavad strateegiad peaksid peale indiviidi pöörama tähelepanu ka koolikeskkonnale.

Sarnase uuringu planeerimisel tuleks edaspidi sotsiaalse keskkonna tegurina vaadelda kooliklassi, mis oleks üheks täpsemaks tervisekäitumise kujunemise keskkonda iseloomustavaks teguriks kooliõpilaste seas. Kahjuks ei olnud käesolevas andmestikus säilinud andmed klassi kuuluvuse kohta.

Mitmetasandiliste mudelite hindamine ei ole nüüdisaegset tarkvara (nt SAS GLIMMIX) kasutades tehniliselt keerukas, samas on selle meetodi kasutegur suur. Nii näiteks oleks antud juhul üheks alternatiiviks kooli mõju arvestamata jätmine. See aga tooks kaasa mudeli jääkide varieeruvuse olulise suurenemise, mille tõttu langeks tunduvalt analüüsi võimsus — ülejäänud mõjude avastamise tõenäosus. Kui ülejäänud tegurite jaotus kooliti erineb, siis kaasnevad sellega nihked nende mõjude hindamisel. Teiseks alternatiiviks oleks kooli mõju arvestamine fikseeritud mõjuna — nii tekiks mudelisse lisaparameeter iga kooli kohta. Selline lähenemine tooks kaasa oluliselt suurema hinnatavate parameetrite arvu, seega väheneks jäägi vabadusastmete arv ja ülejäänud parameetrite hindamise võimsus.

Seega võib nii käesoleva kui ka kõigi teiste sarnaste uuringute korral pidada sobivaimaks andmeanalüüsi meetodiks just mitmetasandilisi mudeleid. Arvestades taoliste uuringute suhtelist rohkust, tuleks mõelda ka selle meetodika laiemale propageerimisele eluteadlaste hulgas.

Kirjandus

1. Duncan, C, Jones, K, Moon, G. Do places matter? (1993). *Social Science & Medicine*, 37, 725–733.
2. Duncan, T. E., Tildesley, E., Duncan, S. C., Hops, H. (1995). The consistency of family and peer influences on the development of substance use in adolescence. *Addiction*, 90, 1647–1660.

3. Gail, M. H., Benichou, J. (2000). Encyclopedia of Epidemiologic Methods, Wiley, 481-491.
4. Goldstein, H. (1995). Multilevel Statistical Models. 2. trükk. New York: Halstead Press.
5. (<http://www.arnoldpublishers.com/support/goldstein.htm>).
6. Green, L.W. & Kreuter, M. W (1991). Health promotion planning: An educational and environmental approach. Mountain View: Mayfield Publishing Company, 44-187.
7. Karvonen, S., Rimpela, A. (1996). Socio-regional context as a determinant of adolescents' health behaviour in Finland. *Social Science & Medicine*, 43, 1467-1474.
8. O'Campo, P., Xue, X., Wang, M. C., Caughy, M. (1997). Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *American Journal of Public Health*, 87, 1113-1118.
9. Rice, N., Leyland, A. (1996). Multilevel models: applications to health data [review]. *Journal of Health Services & Research Policy*, 1, 154-164.
10. Zhu, B. P., Liu, M., Shelton, D., Liu, S., Giovino, G. A. (1996). Cigarette smoking and its risk factors among elementary school students in Beijing. *American Journal of Public Health*, 86, 368-375.
11. Urberg, K. A., Degirmencioglu, S. M., Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology*, 33, 834-844.
12. Williams, J. G., Covington, C. J. (1997). Predictors of cigarette smoking among adolescents. *Psychological Reports*, 80, 481-482.

REGRESSIOONIVÖRRANDITE LINEARISEERIMINE

Artur Nilson

EPMÜ metsakorralduse instituut, nilson@eau.ee

Sissejuhatus

Eluteadustes (sh metsanduses) pole meie praeguste ettekujutuste kohaselt parimaks peetavate võrrandite parameetrid enamasti vahetult hinnatavad kasutajasõbralike, töökindlate ja kiirete lineaarregressiooni programmidega. Põhjamaades on viimastel kümnenditel metsa kasvu ja ehitust kirjeldavate mudelitena eelistatud multiplikatiivseid valemeid, kus astmenäitajatena esinevate parameetrite hindamiseks lineaarregressiooni programmidega kasutatakse sõltuva muutuja väärtuse asemel selle logaritmi. Paljudes metsa kasvu funktsioonides on hinnatavad parameetrid murru nimetajas ja parameetrite hindamiseks teisendatakse võrrand pöördväärtuseks. Kasutatakse ka keerukamaid teisendusi. Taoliselt teisendatud andmetest leitud regressioonivõrrandite pöördteisendusega saadud võrrandid on küll teisendatud andmete tingliku keskmise parimad lähendid, kuid tunnuse algväärtute tinglike keskmiste jaoks on hinnangud nihutatud. Taoliste teisenduste kasutaja võib teisendusnihke probleemi mitteteadvustamise korral langeda ise teisendusnihke lõksu ja lisaks osutada "heauskseks statistilise vale" levitajaks.

Probleemi ajaloost ja uuritusest

Teisendusnihke küsimusi on sporaadiliselt empiirilisel uuritud varemgi, kuid küsimuse süstemaatilist käsitlemist kirjandusest ei õnnestunud leida. Põgus ülevaade sellistest töödest on esitatud kirjutises [Nilson, 2002]. 1980. a paiku tuletasime ja võtsime regressioonanalüüsis teisendusnihke vähendamiseks kasutusele nn teisenduskaalud. Pisut hiljem leidsime koos tõestusega soovitusel [Петрович, 1982] kasutada kaaluna teisenduse tuletise pöördväärtuse ruutu. Sellest peale loobusime enese (osaliselt Petrovitši soovitusel kattuvate ja osaliselt erinevate) teisenduskaaludega tegelemisest ning Eesti metsanduses hakati kasutama M. Petrovitši soovitatud teisenduskaalusid. Eesti metsade

statistilise inventeerimise andmete töötlemistulemusi analüüsidest selgus, et hoolimata M. Petrovitši teisenduskaalu kasutamisest on regressioonivõrrandite tulemused märgatava ja oodatule vastasmärgilise nihkega. Sama ilmnes ka muudel juhtudel ja nii tuli teisenduskaalude küsimus uuesti üle vaadata.

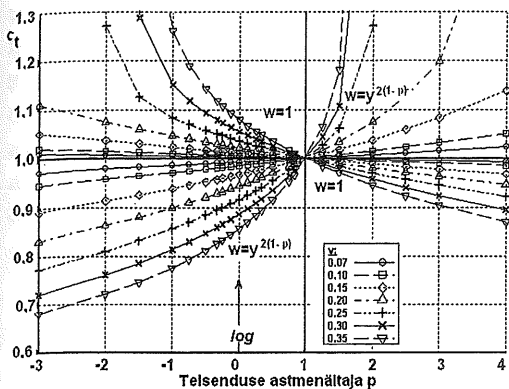
Regressioonivõrrandite lineariseerimine mittelineaarsete teisenduste abil on bioloogias ja metsanduses jätkuvalt levinud tegevus, kuid enamasti jääb märkimisväärse teisendusnihke tekkimine seejuures tähelepanuta. Kui teisendusnihet siiski arvestatakse, siis kasutatakse nihke kompenseerimiseks vaid erijuhtudeks sobivaid meetmeid. Taoliste näiteks on poole jääkdispersiooni lisamine logaritmilise teisendusega saadud võrrandi vabaliikmele [Baskerville, 1972], mis on korrektne vaid tingliku variatsioonikordaja konstantsuse juhul. M. Petrovitši teisenduskaal osutus sobivaks vaid väikese suhtelise jääkhälbega juhtudele.

Värsked teaduslikke publikatsioone lehitsedes [Holmström *et al*, 2001; Jonsson *et al*, 2002; Thorsen *et al*, 2001 jne] veendume, et teisendusnihke probleem on jätkuvalt aktuaalne. Viidatud allikad on vaid väike juhuvalim suurest hulgast taolistest.

Teisendusnihke suurus

Teisendusnihke suurus sõltub teisenduse mittelineaarsuse tasemest ja suhtelise jääkhälbe ehk tingliku jaotuse variatsioonikordaja v väärtusest. Teisendusnihke tekkimist uuriti rangelt sümmeetrilistel normaaljaotuse simulatsioonidel esmalt ühemõõtmelise jaotuse astmete teisenduse $g(y, p) = y^p$ kaudu hinnatavate teisenduskeskmiste $\hat{y} = (\overline{y^p})^{1/p}$ nihke ja seejärel kahemõõtmelise tingliku jaotuse regressioonanalüüsi nihke hindamise ning nihet vähendavate teisenduskaalude katsetamise abil. Teisendusnihke asemel hinnatakse edasises suhet $c_t = \bar{y}/\hat{y}$, mis on ühtlasi teisenduse vahendusel leitud hinnangu \hat{y} korrigeerimiskordajaks.

Teisendussuhte c_t väärtuse sõltuvust teisenduse astmenäitajast p , kasutatud teisenduskaalust w ja jaotuse variatsioonikordajast v ilmestab joonis 1. Piirkonnas, kus suhe $c_t > 1,2$, on selle hindamine variatsioonikordaja v väärtustele tuginevate valemite järgi väga ebakindel. See on teisenduskaalu kasutamisetä või Petrovitši kaalu kasutamisel astmete teisendusteks selgelt sobimatu piirkond. Joonisel 1 on kujutatud teisen-

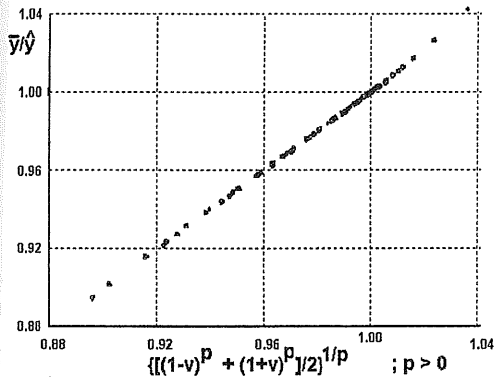


Joonis 1.

Jooniselt 1 selgub, et eluteaduse andmetes tavapäraselt suure varieeruvuse korral on astmeteisenduste abil saadud tulemused suure moonutusega nii Petrovitši kaalu kasutamise korral kui ka teisenduskaaluta arvutustes. Edasises vaatleme teisendussuhte suurust peamiselt teisenduskaaluta teisenduste juhu.

Katsetes arvupaaride ja normaaljaotusega selgus, et astendaja p positiivsete väärtuste ja kaalu $w = 1$ korral lähendab suhet c_t hästi valem

$$c_t = \{[(1+v)^p + (1-v)^p]/2\}^{1/p}, \quad (1)$$



Joonis 2

dussuhte $c_t = \bar{y}/\hat{y}$ sõltuvus teisenduse astmenäitajast p , kasutatud teisenduskaalust w ja jaotuse variatsioonikordajast v . Esimeses ja neljandas veerandis on näidatud suhte c_t väärtused teisenduskaalu kasutamisetä $(w = 1)$, teises ja kolmandas veerandis Petrovitši teisenduskaalu $w = y^{2(1-p)}$ korral.

mida demonstreerib joonis 2, kus on näidatud aritmeetilise keskmise ja "astmekeskmise" suhte

$$c_t = \bar{y}/\hat{y}$$

sõltuvus teisenduse astmest p ja normaaljaotuse variatsioonikordajast v .

Valem on täpne astendaja $p = 2$ puhul suvalistel jaotustel.

Kasutasime seda kujul $c_t = (1 + v^2)^{-1/2}$ arvutisimulatsioonide täpsuse kontrolliks. Arvutustes kasutatud 6 tükikohta olid õiged. Jooniselt näeme, et astendaja p ühest suurema väärtuse korral on ühest väiksem suhe c_t üsna täpselt lähendatav valemiga (1), kuid joonise ülemises parempoolses osas ($p < 1$) muutub seos tajutavalt nõgusaks kõveraks. Tabelis 1 on toodud rida empiirilisi valemeid suhte c_t arvutamiseks, neist esimesed on teisenduskaalu kasutamata ja viimased kaks M. Petrovitši teisenduskaaluga $y^{2(1-p)}$ saadud astmeteisenduste aritmeetilise keskmise pöördteisendustele.

Tabel 1. Paranduskordaja arvutamine, kus p — teisenduse y^p astmenäitaja; e — jääkstandardhälve, v — variatsioonikordaja

p	Paranduskordaja c_t valem	e
<i>Teisenduskaalu kasutamata ($w = 1$) saadud tulemustele</i>		
> 1	$\{[(1 + v)^p + (1 - v)^p]/2\}^{-1/p}$	0,0005
0,5	$\{[(1 + v)^p + (1 - v)^p]/2\}^{-1/p} + 0,628v^{2,122}$	0,0002
0	$1 + \exp[-0,715 + 1,993\ln(v) + 2,096v^2 + 4,484v^4]$	0,0001
-1	$1 + \exp[-0,0214 + 1,994\ln(v) + 2,659v^2 + 36,3v^4]$	0,0007
< 1	$1 + 0,04694\{[(1 - p)v^2 + 1]^{8,61558} - 1\}$	0,0027*
<i>M. Petrovitši teisenduskaaluga $w = y^{2(1-p)}$ saadud tulemustele</i>		
0	$1 - 0,9727v^{1,782}$	0,0009
-1	$1 - 1,2185v^{1,565}$	0,0030

* Veahinnang kehtib paranduskordaja väärtustel alla 1,2.

Teisendusnihke vähendamine teisenduskaaludega

Teisendus pöördväärtuseks on üks sagedasemaid lineariseerimisvõtteid, sest väga paljud populaarsed asümptootilised valemid on üldistatavad kujule

$$y = f(x)/[f(x) + c]$$

ning on lineariseeritavad pöördväärtusena. Esmalt tõestame, et sümmeetrilise jaotusega tunnuse Y harmooniline keskmine on teisenduskaalu y kasutamise korral aritmeetilise keskmise nihketa hinnang.

Alustame tõestust väärtuste paarist $y_1 = y_k - c$ ja $y_2 = y_k + c$. Kasutades kaaluna neid väärtusi endid, saame harmoonilise keskmise avaldiseks $y_h = [(y_1/y_1 + y_2/y_2)/(y_1 + y_2)]^{-1}$, millest pärast lihtsustamist saame $y_h = (y_1 + y_2)/2$. Tulemus kehtib ka ühe ja sama tsentriga väärtusepaaride

suvalise hulga ehk suvalise sümmeetrilise jaotuse kohta, sh ka tinglike jaotuste puhul. Väite kehtivust regressioonanalüüsis kinnitasid nii algul heteroskedastiliste kui ka homoskedastiliste jaotuste pöördväärtusteks teisendatud regressiooni tulemused. Regressioonivõrrandite pöördteisendused osutusid kõigis katsevariantides tegeliku sõltuvuse täpseteks nihketa koopiatega. Katsevariandid sarnanesid edasises kirjeldatavate logaritmilist teisendust vajavate variantidega (vt joonis 3).

Lineariseerides näiteks võrrandi $y = 1/(a + bx)$, saame võrrandi $1/y = a + bx$, millest teisenduskaalu y kasutamisel saame süsteemi

$$\begin{cases} a \sum y + b \sum xy = \sum 1 \\ a \sum xy + b \sum x^2 y = \sum x \end{cases}$$

Arvutuslikult on saadu samaväärne püüdega leida seose $ay + bxy = 1$ kordajad. Eelnevast selgus, et sellise süsteemi lahendamisel saame parameetrite a ja b nihketa hinnangud. Kui taolisel viisil on võimalik regressioonivõrrandeid lineariseerida, siis oletasime, et nihketa lähendeid peaks saama leida muudelgi juhtudel, kus sõltuv muutuja Y esineb võrrandis ilmutamata kujul. Mõned näited (neist esimest kasutame sageli):

$$\begin{array}{ll} a + bxy + \dots = y, & a + bx/y + \dots = y, \\ ax + bxy + \dots = y, & a/y + bx + \dots = y. \end{array}$$

Seni on sel viisil saadud tulemused olnud head. Võrrandisüsteemi teine ja järgnevad võrrandid saadakse, korrutades algvõrrandi teguritega, mis sisaldavad tavapäraselt regressoreid, kuid mitte sõltuvat muutujat y ega tema teisendusi. Regressioonanalüüsi programmid on selleks kohaldatavad, kui sisestada muutujat Y sisaldavad tegurid kahel eraldi kujul (eraldi regressor ning regressor koos tunnuse Y avaldistega).

Paljud metsanduses populaarsed valemid on lineariseeritavad pöördväärtuseks teisendamise teel ja need seni suure teisendusnihke tõttu riskantseks peetud valemid on sõltuva muutuja väärtuse kasutamisel teisenduskaaluna ja tingliku jaotuse sümmeetrilisuse korral rehabiliteeritud täielikult nihkevabadena.

Eksitav signaal korrelatsioonikordaja või -suhte väärtusest on katsetuste üheks tähelepanu väärivaks tulemuseks. Kasutades pöördväärtuseks teisendamisel variandis $0,3x$ (vt joonis 3) teisenduskaaluna suurust y , saime korrelatsioonikordaja hinnanguks $0,06707$ ja täiesti nihketa regres-

sioonivõrrandi (kordajate hinnangutes olid kõik 7 tüvikohta õiged ka suu- rima variatsioonikordaja $v = 0,3$ korral). Korrelatsioon puudus, kuid reg- ressoonanalüüsi tulemus oli ideaalne.

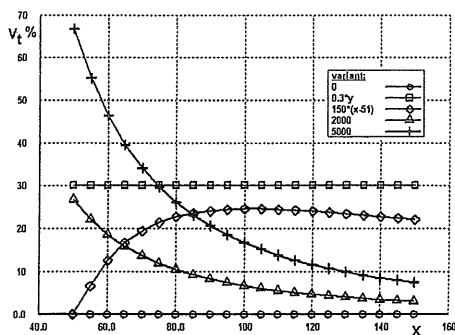
Kasutades kaaluna seni soovitatud (Петрович, 1982) väärtust y^4 , saime korrelatsioonikordaja väärtuseks 0,62669, kuid täiesti moonutatud kordaja- te hinnangud ning lähendite minimaalse suhtelise süstemaatilise vea 20%, maksimaalse 24% ja keskmise suhtelise vea 22%. Korrelatsioon oli küll arvestatav, kuid saadud regressioonivõrrand oli praktiliselt kõlbmatu. Prak- tiliselt samasugune oli tulemus ka teises heteroskedastilises katsevariandis.

Sellise paradoksi põhjustas asjaolu, et kaalu y kasutamiseega muudeti sõltuva muutuja tinglik keskmine e regressioonijoon arvutuste esimeses faasis praktiliselt konstandiks, kuid hälbed selle ümber säilisid, mistõttu jääkhälve osutus ligikaudu võrdseks üldise standardhälbega ja korrelat- sioonikordaja väärtus kujunes nullilähedaseks. Toodud näide manitseb äärmisele ettevaatusele mittelineaarselt teisendatud tunnuste regressioon- analüüsiga kaasnevate hinnangute tõlgendamisel.

Teisenduskaalu y kasutamine osutus pöördväärtuseks teisendamisel muu- dest katsetatud kaalu variantidest nii analüütilise kui empiirilise uuringu tulemusena sedavõrd selgelt eelistatavaks, et nende katsetuste detailide edasine esitamine pole mõttekas.

Logaritmilisele teisendusele vastab astmeteisenduste $g(y) = y^p$ hulgas teisendusastendaja $p = 0$, analoogia põhjal tuletise astmenäitajaga. Jaotustele iteratiivselt leitud teisendusnihet kõrvaldava teisenduskaalu $w = y^q$ astmenäitaja q paikneb mõõdukate variatsioonikordaja v väärtuste korral sirgel $q = (1 - p)/2$ samuti kohal $p = 0$. Seega võis eeldada, et parima teisenduskaalu väärtus logaritmilisele teisendusele on ka regressioonanalüüsis $y^{0,5}$. Katsetati erinevate astendaja q väärtustega vahemikus 0...2. Logaritmiliselt lineariseeritavate katsevariantide tingli- ku variatsioonikordaja väärtusi teise katse algandmetes iseloomustab joonis 3. Variandis 0 oli sõltuvus funktsionaalne, $y = 3x^2$, teiste varian- tide tähistus näitab lisatud rangelt sümmeetrilise normaalse jääkstan- dardhälbe väärtust. Tabelis 2 on toodud huvitavamad tulemused käes- olevas töös soovitatava kaaluga $y^{(1-p)/2}$, M. Petrovitši (Петрович, 1982) soovitatud kaaluga $y^{2(1-p)}$ ja teisenduskaaluta variandiga (tähistatud 0, $y^0 = 1$) saadud tulemuste kohta. Logaritmilise teisenduse korral ($p = 0$) on kaks esimest kaalu vastavalt $y^{0,5}$ ja y^2 . Mõlema viimatinimetatud

kaalu korral saadi funktsionaalse sõltuvusega andmetest (variant 0) täpsed kordajate a ja b hinnangud ja neid seetõttu tabelis 2 ei esitata, teisenduskaaluta hinnangud osutusid aga pisut nihutatuteks ka funktsionaalse sõltuvuse korral. Joonisel 3 on kujutatud tinglik variatsioonikordaja v_i protsentides regressori X funktsioonina logaritmiliselt lineariseeritud regressioonivõrrandi leidmiseks kasutatud algandmete variantides. Korrelatsioonikordaja eksitavaid väärtusi näeme ka tabelis 2, kuid siin on see efekt märksa nõrgem kui pöördväärtuseks teisenduse korral.



Joonis 3

Teisendusnihed on kaalu $y^{0,5}$ kasutamisel üldiselt suurusjärgu kuni paari võrra väiksemad teisenduskaaluta ja M. Petrovitši kaalu y^2 kasutamisel tekkinutest. Korrelatsioonikordaja väärtused on aga homoskedastilistel jaotustel (variandid 2000 ja 5000) kaalu $y^{0,5}$ puhul pisut väiksemad kui kaalu y^2 kasutamisel.

Astmeteisenduse $g(y, p) = y^p$ korral vähendab teisenduskaalu $y^{(1-p)/2}$ kasutamine teisendusnihet ligikaudu suurusjärgu võrra ka juhul kui teisenduse aste $p \neq -1$ ja kuulub vahemikku $-2,5 \dots 3$. Saadud tulemused võib üldistada järgmiselt.

Arvestatava tingliku variatsioonikordaja väärtuse ($v_i > 0,02$) korral algandmetes ja astmeteisendusetaoliste siledate lineariseerimisteisenduste rakendamisel sobib teisendusnihet vähendavaks teisenduskaaluks ruutjuure teisenduse tuletise pöördväärtusest. Väikese suhtelise jääkhälbega ($v_i < 0,02$) andmete korral sobib M. Petrovitši [Петрович, 1982] soovitusel kaaluks teisenduse tuletise pöördväärtuse ruutjuure asemel selle ruut. Variatsioonikordaja v_i väikeste väärtuste korral muutub teisendusnihe vähetundlikuks valitud teisenduskaalule ja seetõttu pole kaalu valik selles piirkonnas kriitiline. Sümmetrilise tingliku jaotuse ja teisenduse korral pöördväärtuseks ($p = -1$) tagab kaalu y^{-1} kasutamine parameetritele nihketa hinnangu.

Tabelis 2 on esitatud seose $y = ax^b$ parameetrite $a = 3$ ja $b = 2$ hindamis-
tulemused logaritmilise teisenduse abil. Korrelatsioonikordaja R on hin-
natud logaritmilises teisenduses ja kasutatud kaalukordajate y^q korral.
Suhtelised nihked on hinnatud protsentides lähendatava suuruse y
tegelike tinglike keskmiste suhtes.

Tabel 2.

Variant	Korrelats- kordaja R	Kaaluk- aste q	Kordajate hinnangud		Suhteline nihe (%)		
			a	b	min	max	keskm.
<i>Regressor X on ühtlase jaotusega, $X \sim U(51, 150)$</i>							
0,3y	0,87741	0,5	2,986092	2,001401	0,09	0,24	0,18
150(x-51)	0,92769	0,5	2,997030	2,000297	0,02	0,05	0,04
2000	0,98710	0,5	3,003965	1,999724	-0,01	0,02	0,01
5000	0,92424	0,5	3,373808	1,975465	-0,55	2,12	0,55
0	1,00000	0	3,001052	1,999924	0,00	0,01	0,00
0,3y	0,84263	0	3,044056	1,985069	-5,85	-4,32	-5,22
150(x-51)	0,93427	0	3,215671	1,979287	-3,38	-1,19	-2,48
2000	0,98326	0	2,654704	2,025404	-2,21	0,50	-0,63
5000	0,73396	0	0,534130	2,361344	-26,30	8,85	-6,76
0,3y	0,82727	2	3,013677	2,024102	10,44	13,35	12,14
150(x-51)	0,87366	2	3,346078	1,992268	7,30	8,20	7,67
2000	0,98979	2	3,523830	1,967495	-0,19	3,37	1,27
5000	0,94080	2	7,215759	1,823250	-0,79	20,05	7,49
<i>Regressor X on tsenseeritud normaaljaotusega, $X \sim N(100, 30)$, $50 < x < 151$</i>							
0,3y	0,73356	0,5	2,985605	2,001349	0,05	0,20	0,14
150(x-51)	0,80145	0,5	2,997268	2,000300	0,03	0,06	0,05
2000	0,97550	0,5	3,001217	1,999914	0,00	0,01	0,00
5000	0,85894	0,5	3,109378	1,992368	-0,24	0,58	0,10
0	1,00000	0	2,999979	2,000001	0,00	0,00	0,00
0,3y	0,70190	0	3,010500	1,988429	-5,30	-4,11	-4,81
150(x-51)	0,79835	0	3,080973	1,987536	-3,52	-2,21	-2,98
2000	0,97406	0	2,786393	2,015405	-1,32	0,33	-0,35
5000	0,78925	0	1,385440	2,162706	-12,40	4,36	-2,81
0,3y	0,76190	2	2,954151	2,027880	9,88	13,23	11,84
150(x-51)	0,80177	2	3,508751	1,983452	7,65	9,59	8,45
2000	0,97811	2	3,423548	1,973061	-0,29	2,65	0,92
5000	0,87209	2	6,711340	1,836249	-1,52	17,51	6,06

Teisenduse diferentsi kasutamine tuletise asemel

Tuletise analüütilise avaldise leidmine võib keerukamate teisenduste korral osutada mittematemaatikutele raskeks ja eksimisriskiga tegevuseks, näiteks metsa kasvufunktsioonide teatmikus [Кивисте, 1988] leiduvad lineariseerimisteisendused $\ln(1 - y/c)$, $1/\ln(y - 1)$, $\ln(c/y - 1)$ jt. Kuna teisendust $g(y)$ realiseeriv valem või avaldis on teisenduse rakendamisel nagunii vaja programmeerida, siis on otstarbekas seda kasutada ka tuletise aset täitva diferentsi leidmiseks ja kasutada teisenduskaaluna avaldist

$$w = \{\Delta_y / |g(y + \Delta_y) - g(y)|\}^{1/2}.$$

Sellise teisenduskaalu arvutamine on täielikult automatiseeritav ja vabastab tellija teisenduskaalu leidmise probleemidest. Meie katsetused selleks koostatud programmiga *wReg.prg* kinnitasid öeldut ja näitasid, et sammu Δ_y suuruse valik arvutustäpsuse piiri ja jääkstandardhälbe vahel ei ole eriti kriitiline. Võib siiski oletada, et sammu Δ_y suuruse valikus võib leiduda mõningaid tulemuste parandamise võimalusi.

Turvanõuetest lineariseerimisteisenduste kasutamisel

Esmane ja obligatoorne turvanõue on lineariseerimisega saadud lähen-dite korrelatsioonisuhte ja prognoosijääkide hindamine otse teisen-damata algandmetel. Teaduslikes kirjutistes tuleks lineariseerimisega saadud regressioonivõrrandite ja nende karakteristikate vahetu esitamise ilma järelkontrollita algandmetel pidada eksitava teabe levitamiseks ja taolised kirjutised märgendada kui teisendusnihkekahtlased. Niisugune kontroll on soovitatav isegi käesolevas kirjutises ettepanud teisenduskaalu kasutamise korral, sest kui teisenduse aste $p \neq -1$, võib tulemusse ikkagi jääda väike teisendusnihe (vt ridu $q = 0,5$ tabelis 2). Kui järelkontrolliga selgub kriitilise suurusega teisendusnihe, siis tuleb võrrand parandada sõltuvalt nihke või teisendussuhte iseloomust, kor-rutades seda korrigeerimiskordajaga, liites sellele korrigeerimiskonstan-di või koguni kasutada sõltuva muutuja teistkordset või regressiooni jääkide regressiooni, kus regressoritena võivad esineda nii prognoo-siväärtused kui esialgsed regressorid. Enamasti piisab korrigeerimiseks ühest regressorist.

Kokkuvõte

1. Regressioonivõrrandite lineariseerimine tekitab eluteadustes tavaliste suure varieeruvusega andmete regressioonanalüüsi tulemustes suuri moonutusi.
2. Moonutuste suurus sõltub tingliku variatsioonikordaja väärtusest ja lineariseerimisteisenduse mittelineaarsuse tasemest.
3. Lineariseerimisel saadud võrrandite prognoosijääke tuleb kontrollida ja korrelatsioonisuhet hinnata teisendamata algandmetel.
4. Aruannetes ja publikatsioonides tuleks kasutada vaid p 3 kohaselt kontrollitud ja vajadusel korrigeeritud hinnanguid.
5. Astmeteisenduse $g(y, p) = y^p$ korral vähendab teisenduskaalu $y^{(1-p)/2}$ kasutamine teisendusnihet vähemalt suurusjärgu võrra ja teisenduse korral pöördarvuks ($p = -1$) muudab nihke olematuks.
6. Astmeteisenduste ja nende sarnaste siledate teisenduste korral sobib teisenduskaaluks ruutjuur teisenduse tuletise pöördväärtusest.
7. Teisenduskaalu arutamist võib automatiseerida, asendades teisenduse tuletise teisenduse diferentsiga.

Kirjandus

1. Baskerville, G. I. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research*, 2, 1, 49–53.
2. Holmström, H., Nilsson, M., Ståhl, G. (2001). Simultaneous Estimations of Forest Parameters using Aerial Photograph Interpreted Data and the k Nearest Neighbour Method. *Scandinavian Journal of Forest Research.*, 16, 1, 67–78.
3. Jonsson, A., Ericsson, G. (2002). Nitrogen Productivity as a Predictor of Growth in *Pinus sylvestris*. *Baltic Forestry*, 8, 1, 8–13.
4. Nilson, A. (2002) Teisenduskaalust regressioonivõrrandite lineariseerimisest. *Metsanduslikud uurimused XXXVII*. Tartu: EPMÜ (avaldamisel).
5. Thorsen, Å., Mattson S., Weslien, J. (2001). Influence of Stem Diameter on the Survival and Growth of Containerized Norway Spruce Seedling Attacked by Pine Weevils (*Hylobius spp.*). *Scandinavian Journal of Forest Research.*, 16, 1, 54–66.
6. Кивисте, А. К. (1988). Функции роста леса. Тарту: Эстонская сельскохозяйственная академия.
7. Петрович, М. Л. (1982). Регрессионный анализ и его математическое обеспечение на ЕС ЭВМ (практическое руководство). Москва: Финансы и статистика.

UUS ASÜMPTOOTILINE FUNKTSIOON

Artur Nilson

EPMÜ metsakorralduse instituut, nilson@eau.ee

Sissejuhatus

Eluteadustes kasutatavate naturaalsete tunnuste väärtused on sageli piiratud altpoolt nulliga ja ülalt antud tingimustes suurima võimaliku väärtusega — piirväärtusega või asümptoodiga. Kasvu analüüsid nimetame asümptootilist funktsiooni *kasvufunktsiooniks*; kui vanuse või aja asemel on argumendiks mingi muu tunnus, siis lihtsalt *seoseks* ja jaotusi analüüsid *jaotusfunktsiooniks*. Jaotusfunktsioonid sobivad kasvufunktsioonideks ja vastupidi. Globaalse modelleerimise korral tunnuste ruumis pürgib eluteaduste funktsioonide asümptootilisus peaaegu looduseaduse tasandile ja seda saab arvestamata jätta vaid lokaalsetes mudelites. Neiski annab protsessi olemust kajastava asümptootilise mudeli kasutamine katse- või vaatlusandmete piiril hinnangutele täiendava stabiilsuse. Loomulikust nullpunktist algavate protsesside korral saab nullpunkti kasutada kindla punktina ja sobiva funktsiooni laiendamist vaatlusandmete piirkonnast nullpunkti suunas võib vaadelda interpoleerimisena. Teabe olemasolul asümptoodi väärtusest on interpoleerimise sugemetega ekstrapoleerimine võimalik ka edasisuunas. Asümptootilistel funktsioonidel on eluteadustes oluline ja asendamatu roll.

Metsanduspraktikas ja metsabioloogias on nt üsna tavaline otsida paljude tunnuste (puu kõrgus, tüve maht, võra laius, puu kogu biomass, lehestiku pindala jpt) sõltuvust puu tüve diameetrist. Nn *kõrgusekõver* $h = f(d)$ on näide asümptootilisest seosest, mida tuleb alata kasutada nii metsanduspraktikas kui metsabioloogilistes uuringutes. Töö asümptootiliste funktsioonidega on keerukavõitu, alates parima funktsiooni otsingu probleemidest ja lõpetades parameetrite hindamise ning tõlgendamise, sobivuse testimisega ning funktsiooni kasutamise suuremates mudelites.

Asümptootiliste sõltuvuste kuju suure varieeruvuse katmiseks on aegade jooksul loodud sadu funktsioone. Orienteerumine funktsioonide rägastikus pole lihtne ülesanne. Seda hõlbustavas metsa kasvu-funktsioonide teatmeteoses [Кивисте, 1988] toodud 75 funktsiooni hulgas on 47 olemuselt asümptootilised ja enamuse ülejäänutestki käitub realselt huvipakkuvas vanusevahemikus nendega sarnaselt. Käesolevas kirjutises esitatakse paindlik funktsioon, millega seda kirevust vähendada ja osaliselt korrastada. Funktsiooni erijuhtudena saab käsitleda paarikümnet kasvu- ja jaotusfunktsiooni. Viie võimaliku kuju- ja kahe mastaabiparameetri kasutamise võimalus annab funktsioonile suure paindlikkuse. Funktsioon ise ja kolm selle kujuparameetrit on autori originaaltulemused.

Funktsiooni saamisloost ja omadustest

Koostades hüperboolse tangensi $th(x)$ väärtuste arvutamiseks kasutaja-funktsiooni, tuletasime arvutamisevalemi

$$th(x) = 1 - 2/(e^{2x} + 1).$$

Seda pisut modifitseerides saime kuju

$$rth(x, r) = 1 - r/(e^x - 1 + r), \quad (1)$$

mille abil saame erijuhtudena arvutada kasvufunktsioonides sagedase avaldise $1 - e^{-x} = rth(x, 1)$ ja hüperboolse tangensi $th(x) = rth(2x, 2)$ väärtusi. Funktsioon (1) on parameetri r mittenegatiivsete reaalarvuliste väärtuste korral argumenti x järgi monotoonselt kasvav lõigul $[0, 1]$ ja jääb selliseks ka siis, kui selle komponente astendada positiivsete astendajatega s, t, u, v alljärgnevalt:

$$rth(z; r, s, t, u, v) = \{1 - [r/(r + (\exp(z^s) - 1)^t)]^u\}^v. \quad (2)$$

Valem (2), kus argument z on asendatud avaldisega x/c , on traditsioonilisel publitseerimiskujul järgmine:

$$y = \left\{ 1 - \left[\frac{r}{\{\exp[(x/c)^s] - 1\}^t + r} \right]^u \right\}^v. \quad (3)$$

Kujuparameetrite r, s, t, u ja v roll on pisut selgemini läbinähtav esituses

$$y = \left\{ 1 - \left[\frac{1}{1 + \frac{\{\exp[(x/c)^s] - 1\}^t}{r}} \right]^u \right\}^v \quad (4)$$

Kujuparameetrite ja mastaabiparameetri c väärtused on omavahel tugevas korrelatsioonis ning nende vastastikune mõju on osaliselt kompenseeriv, kuid parameetrite vahel puudub täpne funktsionaalne sõltuvus. Astendaja u mõjub kujuparameetrina ja tema kasutamisel on mõtet vaid siis, kui $r \neq 1$ või $t \neq 1$. Kui parameetrid r ja t võrduvad mõlemad ühega, siis astendaja u mõjub täiendava mastaabiparameetrina, sest funktsioon (4) lihtsustub avaldiseks

$$y = \{1 - 1/\exp[u(x/c)^s]\}^v.$$

Arvestades funktsiooni (4) suurt võimalike parameetrite hulka, on tema kasutamiseks soovitatav leida esmalt vaadeldava andmestiku (juhusliku) alamhulga põhjal paari-kolme parameetri parimad fikseeritud väärtused. See on põhimõtteliselt samaväärne kirjeldatud valikuga parameetri r väärtuste 1 ja 2 vahel, st sellega määratakse funktsiooni tüüp. Otsitavate parimate fikseeritud väärtuste leidmiseks saab kasutada kas ekspert-hinnanguid või töökindlat mittelineaarse regressiooni programmi, mis ei tõrgu ka parameetrite tugeva korrelatsiooni korral. Praeguste teadmiste (õigemini funktsiooni mittetundmise) foonil ei ole välistatud ka *katse ja eksituse* meetodi kasutamise vajadus. Selle esimese raske faasi läbimise järel leitakse iga konkreetse juhu tarvis parimad hinnangud muutuvaks jäetud parameetritele. Niisugust etapiviisilist lähenemist kasutades on lootust jõuda rahuldava kompromissini ühest küljest nähtuse ja mudeli adekvaatsuse ning teisest küljest mudeli lihtsuse ja käepärasuse vahel.

Tuletatud rth-funktsiooni kuju

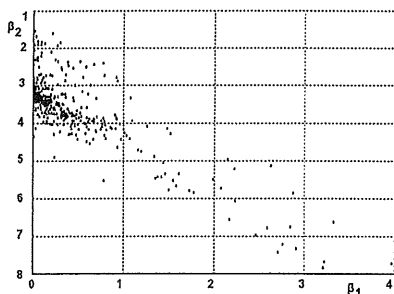
Vaatleme esmalt funktsiooni tavalise jaotusfunktsioonina, iseloomustades teda jaotustüübi valikuks sageli kasutatavate parameetri-funktsioonide β_1 (asümmeetriakordaja ruut) ja β_2 (ekstsess + 3) tasandil [Никитин, Швиденко, 1978]. Andsime funktsiooni parameetritele

juhuslikud väärtused $r \sim U(0,03, 25)$ ja ülejäänutele $s, t, u, v \sim U(0,03,)$ ning genereerisime 500 jaotust.

Tabel 1. Juhuslike parameetritega genereeritud rth -jaotuste jaotusparameetrite piirid

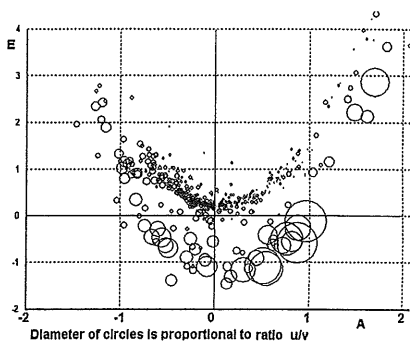
Karakteristik	Miinumum	Maksimum
Aritmeetiline keskmine	0,02	9×10^8
Variatsioonikordaja	2,14	219,6
Asümmeetriakordaja A	-2,70	5,98
Ekstsess E	-1,45	48,9

Joonistel esitame vaid tsentraalse osa tulemustest. Punktid osutusid paiknevateks kõikjal enamkasutatavate unimodaalsete jaotuste piirkonnas ja ulatusid ka bimodaalsete jaotuste piirkonda (joonis 1). Seega võib eeldada, et rth -funktsiooniga peaks saama lähendada väga paljusid empiirilisi jaotusfunktsioone. Joonisel 1 on punktid koondunud kahte vasemalt paremale langevasse ritta, kus ülemises hõredas reas esinevad ka bimodaalsed jaotused.

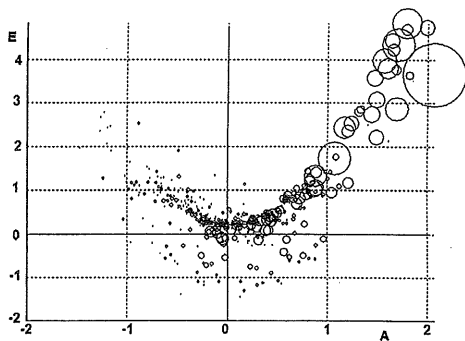


Joonis 1. rth -funktsiooni jaotuskarakteristikad β_1 ja β_2 parameetrite juhuslike väärtuste korral: $r \sim U(0,03; 25)$ ja $s, t, u, v \sim U(0,03; 5)$

Joonisel 2 näeme parameetrite u ja v suhte u/v mõju jaotuse asümmeetriakordajale ja ekstsessile ülejäänud parameetrite juhuslike väärtuste foonil. Suhte u/v suurenedes domineerib liikumine piirkonnast $A \approx 0,5$ ja $E \approx 0,5$ ülemist paraboolikujulist kogumit pidi vasemale negatiivse asümmeetria suunas ja seejärel juba alumist, peamiselt negatiivse ekstsessiga paraboolset kogumit pidi tagasi paremale positiivse asümmeetria suunas.



Joonis 2. Parameetrite u ja v suhte mõju jaotuse asümmeetriakordajale A ja ekstsessile E . Ringide diameeter joonisel on võrdeline suhtega u/v



Joonis 3. Parameetri s mõju jaotuse asümmeetriakordajale A ja ekstsessile E . Ringide diameeter joonisel on võrdeline parameetri s pöördväärtuse ruuduga $1/s^2$

Teiste parameetrite juhuslike väärtuste foonil eristub üsna hästi ka parameetri s mõju punktide asendile karakteristikute A ja E tasandil. Astendaja s suurenedes väheneb asümmeetriakordaja A (joonis 3).

Kasvufunktsioonide esitus rth-funktsiooni kujul

Varem analüüsitud kasvufunktsioonides [Prodan, 1968; Борачев и Свалов, 1978; Кивисте, 1988] leidsime analoogiaid vaid rth-funktsiooni parameetri r ning astendajate s ja v kasutamisele. Parameetritele r on seni antud vaid väärtusi 1 või 2 valikuna avaldiste $th(x)$ või $1 - e^{-x}$ vahel, argumendiks x on enamasti olnud tunnuse tegeliku väärtuse korrutus mastaabikordajaga (vahel astendatuna). Kasutame edaspidi mastaabikordajana tegurit $1/c_x$, kus jagaja c_x (või lihtsalt c) on uue suuruse z , $z = x/c_x$, mõõtühiku rollis. Allpool esitame põgusa ülevaate kasvufunktsioonidest, mida saab esitada rth-funktsiooni abil.

Tabel 2. Parameetrite r , s ja v kasutamine kasvufunktsioonides. Funktsioonide tähistus vastab A. Kiviste teatnikus [Kueucme, 1988] toodule.

Funktsiooni		Argument	Parameetrid		
number	nimetus		r	s	v
5.1	Weber	z	1	1	1
5.3	Mitscherlich	z	1	1	v
5.6	Weibull II	z	1	s	1
5.7	Todorovič III	z	1	s	v
7.6	Sevastjanov I	z	2	1	1
7.7	Sevastjanov II	z	2	1	3
5.2	Schubert	$z - z_0$	1	1	1
5.4	Bertalanffy	$z - z_0$	1	1	v
5.5	Weibull	$z - z_0$	1	s	1

Weibulli jaotusfunktsioon on esitatav rth-jaotuse variandina, kus argument esineb kujul $z - z_0$ ning parameetrid t ja u võrduvad ühega (tabel 2, number 5.5). Samamoodi võib jaotuse nullpunkti nihutada ka rth-funkt-

siooni muude variantide kasutamise korral (tabel 2, numbrid 5.2 ja 5.4). Paljudel juhtudel on rth-funktsioon kasvufunktsioonides vaid üks komponent või kasvufunktsioonid on rth-funktsioonide vahed, suhted või koguni rth-funktsiooni rth-funktsioonid. Allpool toome mõned näited funktsioonidest viidatud teatmikus [Кивисте, 1988].

- 5.8. van der Vliet: $y = c_y [1 - P_2 + P_2 \text{rth}(2x/c_x, 1, 1, 1, 1, 1)];$
 $P_2 = 1 + 2x/c_x + 2(x/c_x)^2$
- 5.9. Kövessy: $y = c_y [c_1 \text{rth}(x/c_1, 1, 1, 1, 1, 1) - c_2 \text{rth}(x/c_2, 1, 1, 1, 1, 1)]$
- 5.10. Thomasius: $y = c_y \text{rth}[xc_1, \text{rth}(x/c_2, 1, 1, 1, 1, 1), 1, 1, 1, 1, 1]$
- 5.11. Thomasius: $y = c_y \text{rth}[xc_1, \text{rth}(x/c_2, 1, 1, 1, 1, 1), 1, 1, 1, 1, v]$
- 5.12. Bass: $y = c_y \text{rth}(x/c_x, 1, 1, 1, 1, 1) [b \text{rth}(x/c_x, 1, 1, 1, 1, 1) - b + 1]$
- 7.4. Scharf IV: $y = b + c_y \text{rth}((x - x_0)/c_x, 2, 1, 1, 1, 1)$
- 7.8. Grigaliunas: $y = c_y f_x \text{rth}(x, c_x, 2, 1, 1, 1, 1); f_x = 1 - 1/(1 + x)^c$
- 4.7. Gompertz: $y = \exp\{b_0 + b_1 [1 - \text{rth}(x/c_x, 1, 1, 1, 1, 1)]\}$
- 4.9. Sloboda: $y = \exp\{b_0 + b_1 [1 - \text{rth}(x/c_x, 1, s, 1, 1, 1)]\}$

Kasutajafunktsioonid ja tähistus

Sageli kasutatavaid avaldiseid on soovitatav nende lülitamiseks rakendus- või uuringusüsteemidesse vormistada nn kasutajafunktsioonidena. EPMÜ metsakorralduse instituudi mudelite baasis (www.eau.ee/~mbaas/) on rth-kasutajafunktsioonide komplekt olemas nii statistikasüsteemi R (<http://cran.r-project.org>) kui ka andmebaasi ohjesüsteemi MS Visual FoxPro keskkonna tarvis. Komplektid sisaldavad funktsiooni enese, tuletise ja pöördfunktsiooni kasutajafunktsioone nii vastavalt valemile (2), mis sisaldab kõiki valitavaid parameetreid kui ka viieteistkümnele lihtsustatud variandile, kus lisaks parameetritele r on vaid mõni kujuparameeter vabalt valitav. Kolmele lihtsamale variandile on olemas ka käänupunkti (jaotuse moodi) arvutusvalemid. Tähistuse esimesed kolm sümbolit on alati rth , millele lihtsustatud variantide korral järgneb punktiga eraldatuna variandis kasutatud kujuparameetrite r , s , t , u ja v loend (nt $rth.rsv$). Tuletise korral lisatakse veel laiend dif (nt $rth.rsv.dif$) ja pöördfunktsiooni korral inv (nt $rth.rsv.inv$). Käänupunkti valemile lisatakse viimase laiendina $nflz$ või $nfly$ sõltuvalt sellest, kumba koordinaati valem arvutab. Funktsiooni nimele järgneb kasutajafunktsiooni poole pöördumisel parameetrite väärtuste või nimede loend sulgudes. Kasutajafunkt-

sioonide põhikujul on argumendiks $z = x/c$. Samu funktsioone saab kasutada ka tuletise leidmiseks argumendi x järgi, kasutades teisendust $dy/dx = (dy/dz)/c$, ja pöördfunktsiooni leidmiseks valemi $x = c f^{-1}(y, \dots)$ abil. Esitame näitena statistikasüsteemi R tähistuses täisfunktsiooni kasutajafunktsioonid. Neid on otstarbekas katsetada sobivate tüüpvariantide väljaselgitamisel uuringu esmafaasis.

```
rth<-function(z,r,s,t,u,v){return((1-r/((exp(z^s)-1)^t+r))^u)^v)}
rth.inv<-function(y,r,s,t,u,v){return((log((((1-y^(1/v))^(1-u))*r-r)^(1/t)+1))^(1/s))}
rth.dif<-function(z,r,s,t,u,v){ezs<-exp(z^s)-1;ret<-r/(ezs^t+r);
return(r*s*t*u*v*(1-ret^u)^(v-1)*ret^(u-1)*ezs^(t-1)*(ezs+1)*z^(s-1)/(ezs^t+r)^2)}
```

Kokkuvõte

Tutvustatud rth-funktsiooni variantide pere võimaldab senisest paremini süstematiseerida kasvufunktsioone ja on koos kasutajafunktsioonidega paindlikuks töövahendiks asümptootiliste protsesside modelleerimisel. Selle funktsiooni uurimine on alles algusjärgus. Autor loodab, et seda hakkavad uurima pädevad huvilised ja peatselt selguvad rth-funktsiooni olulisemad head ja vead nii matemaatilises kui ka rakenduslikus mõttes.

Kirjandus

1. Prodan, M. (1968). Forest Biometrics. Pergamon Press. Oxford.
2. <http://cran.r-project.org>
3. Богачев А. В., Свалов С. Н. (1978). Методы таксации лесного и лесосеченого фонда. *Лесоведение и лесоводство*. Том 2. Москва: ВИНТИ. 7–109.
4. Кивисте А. К. (1988). Функции роста леса. (Учебно-справочное пособие). Тарту: Эстонская сельскохозяйственная академия.
5. Никитин К. Е., Швиденко А. З. (1978). Методы и техника обработки лесоводственной информации. Москва: Лесная промышленность.

NÄIDISTE ABIL KLASSIFITSEERIMINE JA PROGNOOS

Kalle Remm

TÜ geograafia instituut, kalle_r@ut.ee

Sissejuhatus

Nähtuste ja objektide klassifitseerimist kasutatakse pea kõigis elu ja eluta looduse uurimise valdkondades. Maailma tohtus mitmekesisuses orienteerumiseks on kogemuste liigitamine mingitesse kategooriatesse juba tunnetuslikult vajalik. Seoses satelliit- ja aerofotoinfo rohkusega on viimastel aastakümnetel klassifitseerimise probleemidega eriti palju tegeldud kujutise töötluse ja tõlgitsuse valdkonnas. Pilditöötlust peetakse kõige efektiivsemaks taimkatte ning maakasutuse keskmise- ja väikesemõõtkavalise kaardistamise viisiks. Nii pildandiandmete kõrval kui ka nende lisana saab kasutada mitmesuguseid topograafilisi andmes- tikke [Remm, 2002]. Eesti varieeruva koosseisuga segametsade puhul on metsa koosseisu ning ka mitmesuguses metsastumise järgus olevate põllumaade pildandiandmete järgi eristamine siiski keerukas [Aaviksoo *et al*, 2000]. Üksikute liikide leviku ja ohtruse kaardistamisel distants- kujutiste järgi on võrdlemisi häid tulemusi saadud vaid monokultuuride ja liigivaeste looduslike koosluste puhul. Eesti segametsade koosseisu hindamine aero- ja kosmosefotode järgi on tunduvalt keerukam ülesanne.

Üks suuremaid tehnilisi probleeme kujutise järgi prognoosimisel ja klassifitseerimisel on otsitavate klasside heterogeensus detailsemate pil- diandiandmete korral. Näiteks metsa kujutis koosneb nii puuvõrade valguse- kui ka varjupoolsetest pikslitest, niidetud ja niitmata rohumaa võivad kuuluda samasse klassi, kuid on pildil erinevate kiirgusomadustega. Peale selle sõltuvad mingi koha kiirgusomadused mitmetest lokaalsetest tingimustest, näiteks on kevadisel kujutisel lõunanõlvadel olevad puud rohkem lehes kui põhjanõlvadel. Heterogeensete lähteandiandmete puhul ei sobi kujutise parameetrite (kiirgusvahemike intensiivsus, värvitoon, tumeduse lokaalne varieeruvus, lokaalne autokorrelatsioon) keskmised

ja standardhälbed otsitavate klasside esindajateks. Kujutiste klassifitseerimise klassikalised vahendid kasutavad aga just kiirgusvahemike keskmisi ja kovariatsioonimaatrikseid.

Alternatiivsete prognoosimeetodite hulgas (mõistes siin ja järgnevalt prognoosi all ka klassikuuluvuse prognoosi) on kasutatud näiteks kunstlikke närvivõrke, mitmeastmelist klassifitseerimist ja hierarhilist segmenteerimist. Kasutatud on ka interpoleerimist tunnuseruumis, so hinnangut k -lähima naabri järgi ehk prognoosi ja klassifitseerimist näidiste abil. Sarnaste analoogide järgi prognoosimist on kasutanud näiteks karude elupaigaelistuste modelleerimiseks [Clark *et al*, 1993], taimkatte kaardistusel [Remm, 2002; Wilds *et al*, 2000], puidutagavara hindamisel [Muinonen *et al*, 2001], paleoökoloogilistes rekonstruktsioonides [Birks, 1993; Flower *et al*, 1997], psühholoogias [Settle *et al*, 1994; Ward *et al*, 2001], meditsiinilises diagnostikas (Seitz *et al*, 1999; Frize ja Walker, 2000), tehnoloogias (Suh *et al*, 1998; Verdenius ja Broeze, 1999; Xia ja Rao, 1999), majanduse planeerimisel ja mitmesuguste riskide hindamisel (Johnson, 2002; Knight *et al*, 1999; Yeh ja Shi, 2001). Tehisintellekti uurimise valdkonnas tuntakse meetodit eelkõige näidistele tuginevate järelduste (*case-based reasoning*) ja "laisa probleemikäsitluse" (*lazy problem solving*) nime all (Aha, 1998).

Vähemalt teoreetiliselt võiks prognoos näidiste abil olla üks paindlikumaid meetodeid heterogeensete klasside ja eritüübiliste lähteandmete korral. Suure paindlikkusega kaasneb aga vajadus leida sobiv sarnasuse määramise viis, iga prognoosi puhul kasutatavate näidiste või sarnasuste hulk (k), minimaalne lubatav sarnasus näidise kasutamiseks, kasutatavate tunnuste hulk ning iga tunnuse ja iga näidise kaal.

Järgnevalt tutvustatakse näidiste abil prognoosi mõningaid tulemusi Otepää looduspargi kaardistamisel maakatte ja metsa koosseisu aero- ja kosmosefotode ning topograafiliste andmete järgi, samuti prognoosiva süsteemi ehitust ja andmetega sobitamist. Näidiste abil prognoosi eesmärgiks oli võimalikult hea prognoosisüsteemi väljatöötamine ja erinevate lähteandmete indikaatorväärtuste hindamine.

Andmed

Välivaatluse teel kirjeldati Otepää looduspargis asuvas 1000 vaatluspunktis taimkatte kasvukohatüüp J. Paali [Paal, 1997] järgi, samuti puude katvus ja puistu valem. Euroopa EUNIS klassifikatsioonüsteemi [Davies ja Moss, 2002] teisele tasemele vastavad elupaigatüübid tuletati kasvukohatüübist ja puistu andmetest. Vaatluspunktid genereeriti kihitatud juhuvalimina dünaamiliste kvootidega, mis tõkestasid klasside üleesindatust [Remm, 2000]. Etteantud teemakihtidena kasutati varasema kasvukohatüüpide kaardi üksusi [Remm, 2002]. Iga genereeritava juhupunkti aktsepteerimise täiendavaks tingimuseks oli $10 \times 10 \text{ m}^2$ suuruste naaberpikslite sama klassikuuluvus mainitud varasemal kasvukohatüüpide kaardil. Naabruse ühetaolisust nõuti selleks, et vältida asukoha määramise väikesest ebatäpsusest tulenevaid suuri vigu kohtade kirjeldustes. Asukoha määramiseks kasutati välitöödel 1:10 000 põhikaarti, ortofotosid ja GPS-vastuvõtjat.

Elupaigatüübi ja puistu koosseisu hindamiseks kogu looduspargi ulatuses kasutati prognoosivahenditena Landsat ETM+ kaardrit 16. maist 2000. a (8 kiirgusvahemikku + halltoon), halltoonis ja värvilisi ortofotosid, nõlva kaldenurka ja suhtelist kõrgust 100 m raadiuses oleva ümbruse keskmise suhtes, mulla lõimist ja mulla tüüpi 1:10 000 mullakaardilt (vt ka joonis 4 tulemuste osas). Halltoonis ortofotod pärinevad aastast 1995 ja värvilised ortofotod aastast 1999. Värvilised ortofotod katavad vaid kolmandiku uuritavast alast. 1999. a halltoonis ortofotod on saadud värvifotode teisendamisel. Kõik tunnused arutati eraldi ja omistati $10 \times 10 \text{ m}^2$ Eesti põhikaardi koordinaatsüsteemis võrgustiku iga piksli kohta. Ortofotodelt, mille pikslid on suurusjärgus 1 m^2 , loeti iga piksli puhul 10 m raadiuses ning 20...30 m ümbruses tumeduse keskväärtus, tumeduse standardhälve, tumeduse ruumiline autokorrelatsioon Morani I järgi, värvifotodelt ka värvitoon ning eraldi punase, roheline ja sinise tooni intensiivsus. Kokku saadi seega 27 argumenttunnust iga $10 \times 10 \text{ m}^2$ pinnaosa kohta looduspargi lääne- ja lõunaosas, kus värvilised ortofotod on olemas, ja 19 argumenttunnust looduspargi ülejäänud osades. Mullakaardi andmebaasis kasutatud tuhanded erinevad klassifikatsiooniüksused üldistati 20 lõimise- ja 50 mullaklassiks. 1:10 000 põhikaardi alad üldistati 11 klassiks, metsa ja

rohumaal puhul eristati turbapinnasel olev maa ja nii saadi 13 maakatteklassi, mida kasutati prognoosimisel eelklassifikaatoritena.

Näidiste abil prognoosimise mudeli sobitamisel jäeti kasutamata nende vaatluskohtade andmed, kus maastik oli satelliidipildi tegemise järel ilmselgelt muutunud (värsked raiesmikud, uued hoonestusalad, muutunud künnimaa). Need kohad, mille muutumist fotode tegemise järel ei saa kindlalt väita, jäeti näidiste hulka. Kokku kasutati mudeli hindamisel 934 väliuuringutes kirjeldatud näidist.

Valdava osa välitööst tegid magistrandid Kaupo Kohv ja Jaanus Remm, kellele autor avaldab tänu ja tunnustust hoolsa töö eest.

Meetodid

Näidiste abil prognoosimiseks tuleb moodustada prognoosiv süsteem – arvutiprogramm ja andmete kogu, mis koosneb näidiste komplektist ja prognoosi arvutamise reeglitest ning parameetritest. Efektiveks prognoosiks on vaja leida parim või vähemalt piisavalt usaldusväärne tunnuste komplekt ja sobivaimad algparameetrite väärtused. Kui kasutada vaid täisarvulisi kaale vahemikus 0–9 ja lubada lähimate naabrite arvul varieeruda vahemikus 1–10, on 27 tunnuse puhul võimalike variantide arv väga suur (10^{28}). Kuna eritüübiliste omavahel korreleerunud tunnuste puhul võib küllalt hea prognoosi saada argumenttunnuste ja mudeli parameetrite erinevate kombinatsioonide korral, on nii suurest hulgast võimalikest lahenditest parima leidmine keerukas. Enamasti otsitakse selliste probleemide lahendit mingi iteratiivse meetodiga põhimõttel, et pime kana leiab tera, kui ta piisavalt kaua nokib ja suudab kogemusest õppida.

Näidiste abil prognoosivale süsteemile kaalude leidmine on sarnane kunstliku närvivõrgu õpetamisega. Erinevalt kunstlikust närvivõrgust ei sisalda näidiste abil prognoosiv süsteem suurt hulka lineaarseid võrrandeid, vaid näidiste komplekti ja sarnasuse määramise reegleid. Kui võrrelda programmeeritud klassifitseerimise ja prognoosi vahendeid automatiseerimata meetoditega, siis on kunstlik närvivõrk analoogiline eksperdi käest küsimisega, otsuste puud määramistabeli kasutamise, signatuuride kasutamine taksonite kirjeldustega võrdlemisega ning näidiste abil prognoosiv süsteem muuseumieksemplaridega võrdlemisega.

Otepää looduspargis tehtud vaatluste puhul määrati klassifitseeritava objekti ja näidise sarnasus sarnasuse kaalutud keskmisena üle kõigi tunnuste. Pideva tunnuse puhul arvati seejuures nullsarnasuseks erinemine vähemalt ühe standardhälbe võrra, nominaaltunnuse puhul näidise ja objekti erinev klassikuuluvus. Sarnasuse mõõduks loeti 1 väärtuste kokkulangevuse puhul (kuuluvus samasse nominaalklassi). Kui objektil ei olnud näidiste hulgas ühtegi sarnast, suurendati sarnasuse piirväärtust järk-järgult seni, kuni leiti kõige sarnasem näidis.

Näidistega prognoosi mudeli sobivuse kontrollimiseks kasutati nominaalse tunnuse puhul klassifikatsioonide vastavuse indeksit κ (*kapa*) ja puistu koosseisu prognoosi puhul puistu üldkatvusega kaalutud puistu valemite ühisosa. Üksiku liigi ohtruse prognoosi headust mõõdeti prognoosi ja vaatluse kahekordse ühisosa ja nende summa suhte abil, mida tuntakse Sørenseni indeksina. Kontrollandmestik saadi, prognoosides igat näidist ülejäänud näidiste järgi.

Näidiste abil prognoosiva süsteemi jaoks parima tunnuste komplekti otsimisel on võimalik 1) võrrelda prognoosi vastavust kontrollandmetele argumenttunnuste ühekaupa kasutamise korral; 2) kasutada sammuviisilist tunnuste lisamist või väljajätmist; 3) otsida sobivamaid kaalude kombinatsioone juhuslike kaaludega saadud lahenditest parimate hulgas; 4) kasutada sama meetodit mitte juhuslike kaaludega, vaid kaale korrapärase sammuga muutes; 5) liikuda sobivusmaastikul kaale nihutades lahendi maksimaalse paranemise suunas või (lokaalse maksimumini jõudmisel) minimaalse halvenemise suunas (suunatud triiv). Peale selle on võimalik kombineerida erinevaid meetodeid.

Mitme sobitamismeetodi katsetamise tulemusel jõuti elupaigaklasside prognoosimisel parima tulemuseni (80%-line vastavus), valides suure hulga juhuslike lahendite seast parimate puhul kõige sagedamini esinenud kaalukombinatsioonid ja täpsustades neid pikaajalise suunatud triiviga. Kuna selline parimate kaalude leidmise viis on üsna aeganõudev ja sisaldab ekspertotsust, siis ei ole see parim siis, kui lahend on tarvis leida kiiresti — näiteks, kui parim sõltumatute tunnuste komplekt ja tunnuste kaalud on vaja leida tunni aja jooksul. Kiirematest ja ekspertotsuseid vältivatest meetoditest olen efektiivseimaks pidanud argumenttunnuste komplekti sammuviisilist valimist (võrdse kaaluga) ja

parima tunnuskomplekti kaalude täpsustamist piiratud kestusega suunatud triivi abil.

Näidiste abil prognoosi arvutamise aeg sõltub oluliselt kasutatavate näidiste hulgast. Paarisajast suurema näidiste arvu korral on enamasti mõistlik kasutada näidiste juhuvalimit ja/või eelklassifitseerimist. Selles uurimuses alustati piiratud kestusega suunatud triive 20% valimi mahuga ja iga seeria järel suurendati valimi mahtu kuni kõigi andmete kasutamiseni. Koos valimi mahu suurendamisega vähendati kaalude muutmise ulatust. Seega parima argumenttunnuste komplekti leidmise järel esmalt destabiliseeriti otsimisparameetrid, et leida kiiremini paremaid variante, ning seejärel stabiliseeriti need järk-järgult, et mitte sobivusmaastiku "künka tipust üle astuda". Kirjeldatud meetod ei pruugi leida parimaid kaale, aga on suuteline mõistliku ajakuluga leidma üsna häid kaale. Juhuvallimite kasutamisel ei tarvitse samade lähtetingimustega sobitamisprotsessid anda sama tulemust. Juhuvallimitega otsimisprotsess on riskantsem, aga tunduvalt kiirem. Korduv automatiseeritud kaalude otsimine nõuab eelkõige arvuti-, mitte inimtööaega. Juhuvallimeid kasutades võib juhuslikult leida sobivusmaastiku selliseid tippe, milleni jõudmine kogu andmestiku kasutamisel poleks suure ajakulu tõttu reaalne.

Kõik selle uurimuse arvutusprogrammid on koostatud programmeerimiskeskonnas Visual Basic.

Tulemused

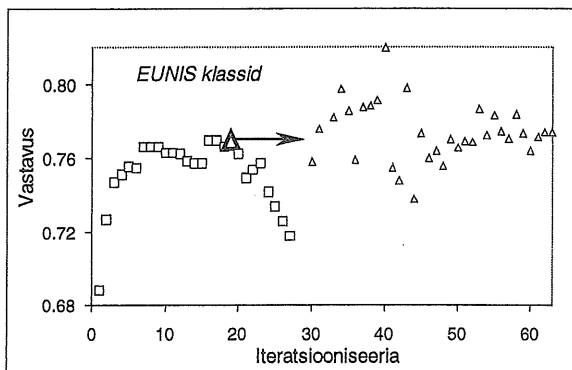
Metoodika osas esitatud kokkulangevuse ulatuse arvutamise eeskirjade kohaselt leiti järgmised vastavused välihinna ja prognoosi vahel: 1) puistu kogukatvus 91%; 2) kuuse katvus 64%; 3) kase katvus 56%; 4) männi katvus 37%; 5) halli lepa katvus 34%; 6) sarapuu katvus 30%; 7) põõsaspajude katvus 20%; 8) haava katvus 17%; 9) tamme katvus 16%; 10) puistu üldkatvusega kaalutud koosseis 40%.

Elupaigaklasside prognoosis 934 näidise ja 27 tunnuse järgi jõuti 80%-lise vastavuseni, mis on palju kõrgem kui samade etalonide ja vaid Landsat-pildi kasutamisel saadud 33%-line vastavus, mis saadi, kasutades suurima tõepära klassifikatsioonimeetodeid tarkvarapaketi IDRISI. 19%-lise mittevastavuste peamised põhjused on ilmselt maastiku muutused (välitöö ja distantsseire andmete ajaline vahe on 2–7 aastat), sujuvad

üleminekul põllu, rohumaa ja noore metsa seoses põllumajanduse hääbumise tõttu uuritaval alal, sujuvad üleminekul leht-, sega- ja okasmetsa ning mitmesuguste raiejärgsete koosluste vahel. Ülemineku nähtuste klassifitseerimine on teatud määral tinglik ja juhuslikku viga sisaldav nii välitöö eksperdi jaoks kui ka formaalsete reeglite kasutamise puhul. 80%-lise klassifikatsioonitäpsuse andnud kaalude vahekorra leidmine sisaldas ekspertotsust, korduv automaatne kaalude otsimine andis tulemusi 78%-lise täpsusega.

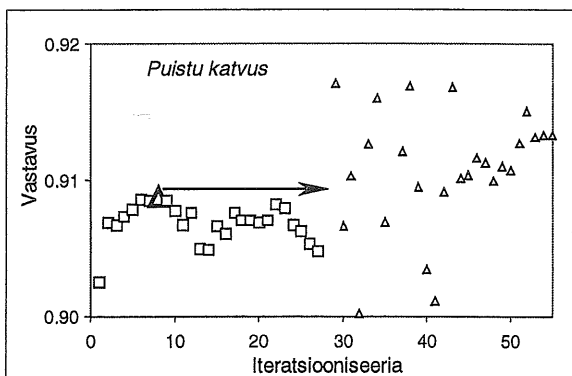
Joonistelt 1–3 saab hinnata prognoosi ja välivaatluste vastavust tunnuste ükshaaval lisamisel (\square) näidiste abil prognoosivas (NAP) süsteemis ning sellele järgneval kaalude iteratiivsel sobitamisel järjest suurenevates juhuslikes valimites (Δ). Nool tähistab iteratsioonide jätkumist esimesel etapil saadud parima komplektiga.

Parimate kaalude otsimise tulemuslikkus optimaalse tunnuste komplekti leidmise järel on erinev. Metsa katvuse hinnangu vastavus 90%-le on tagatud ka vaid ühe tunnuse kasutamisel, edasine mudeli sobitamine prognoosi märkimisväärselt ei paranda (joonis 2).



Joonis 1. NAP-süsteemi sobitumine EUNIS elupaigaklasside prognoosil

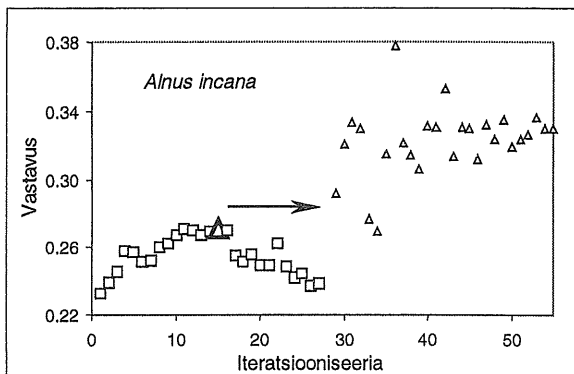
Prognoosi olulist paranemist ei saagi loota, sest kirjeldamata jäänud kümnekond protsenti vastab selle uurimuse väliandmete keskmisele täpsusele. Lisaks mõõtmisvigadele sisaldavad andmed ajalisest nihkest tulenevaid hälbeid. Ka on raske saada head prognoosi nähtuste puhul, mille suurus on lähedane mõõtmisveale.



Joonis 2. NAP-süsteemi sobitumine puistu katvuse hindamisel

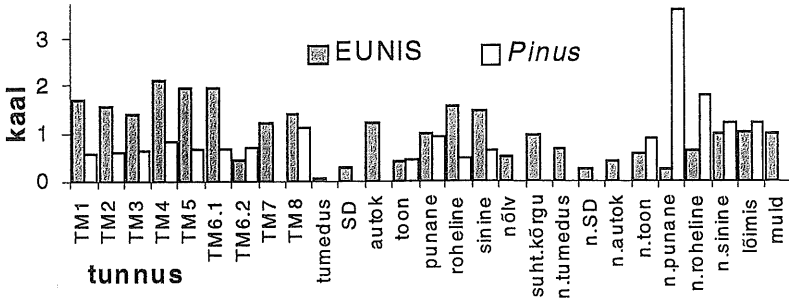
Lepa katvuse hindamisel võimaldab aga tunnuste valimine ja kaalumine parandada prognoosi 23%-lt 33%-ni (joonis 3). Kahe viimase näite puhul kasutati vaid neid vaatlusi, mis põhikaardi järgi on metsamaal.

Argumenttunnuste indikaatorväärtused on erinevate tunnuste prognoosimisel erinevad (joonis 4). Ilmselt sõltuvad parimad indikaatorkaalud näidiste komplektist ning teistest muutujatest.



Joonis 3. NAP-süsteemi sobitumine halli lepa katvuse hindamisel

Iga andmestiku ja prognoositava nähtuse jaoks tuleb nii argument-tunnuste komplekt kui ka kaalud uuesti sobitada. Mõningaid tendentse võib leitud paremate kaalude järgi siiski esile tuua.



Joonis 4. Parimad senileitud kaalud EUNIS-süsteemi elupaigaklasside ja männi (Pinus) katvuse hindamiseks Otepää looduspargist kirjeldatud näidiste järgi.

Kõigi 27 tunnuse kasutamine ei anna võrdsete kaalude korral parimat prognoosi. Heaks prognoosimiseks piisab 2–14 tunnusest. Minimaalse vajaliku tunnustekomplekti juurde peaks jääma ka jälgides säästvusreeglit (*parsimony rule*).

Halltoonis ortofotode parameetrid tulevad automaatselt valitud primaarsetesse kompleksidesse suhteliselt harvem kui värvilistelt ortofotodelt loetud tunnused ja Landsat ETM kanalite kiirusväärtused. Kasutada olnud 1995. a halltoonis aerofotode kvaliteet on üsna ebahütlane, tihti on sama pilt servades teise tooniga kui keskel ja vanemad ortofotod kahest või enamast erineva tumeduse, kontrastsuse ja pildistamissuunaga fotost kokku monteeritud. Seetõttu on satelliidiandmed vähemalt sama kaadri piires palju paremini võrreldavad kui seni kasutada olevad aerofotod. Kaalude summade järgi automaatselt leitud prognoosikomplektides üle tunnuste on metsa koosseisu ja puude katvuse prognoosil sellest andmestikust olulisemad värvifoto punase ja sinise tooni intensiivsus ning Landsat ETM kaheksanda (halltoon), neljanda ja viienda (lähi- ja keskmine infrapunane kiirus) kanali andmed. Halltoonis ortofotode infost on tumedus ja tumeduse lokaalne varieeruvus enam-vähem võrdse

väärtusega, tumeduse lokaalsest autokorrelatsioonist on eelkõige abi puuvõrade suuruse hindamisel (korrapäraselt istutatud noorendike eristamisel), mitte aga metsa koosseisu või üksikute puuliikide eristamisel. Landsat ETM halltooni andmed on kõigi analüüsitud tunnuste puhul halltoonis ortofoto tumeduse andmetest eelistatumad.

Kokkuvõte

Näidiste abil prognoosimise eelis kõigi teiste meetodite ees on universaalsus ja paindlikkus. Näidistega võrdlemise teel saab lahendada väga erinevaid prognoosi- ja klassifikatsiooniülesandeid, seda võib nimetada universaalseks prognoosimeetodiks. Eriliselt väärib esiletõstmist metoodika sobivus heterogeensete nähtuste jaoks, mille puhul tunnuste kesk-väärtused on halvad klassi esindajad. Näidiste puhul võib ühte heterogeenset klassi esindada mitu täiesti erinevat näidist (nt klassifikatsiooniüksuse laanemetsa puhul: raieküpse laanemetsa jaoks üks, noorendiku jaoks teine ja raiesmiku jaoks kolmas).

Näidiste abil prognoosisüsteemi empiirilistele andmetele sobitamine sarnaneb kunstliku närvivõrgu väljaõpetamisega, see on arvutusmahukas ja seda on vaja teha iga andmestiku puhul eraldi. Tänapäeva arvutivõimsuste juures on näidiste abil siiski võimalik lahendada küllaltki keerukaid prognoosiülesandeid suurte andmemahtude korral. Näidiste abil prognoosi eelis kunstliku närvivõrgu ees on selle avatus. Kunstlik närvivõrk on musta kasti taoline süsteem, millest on raske saada teavet kasutatud tunnuste indikaatorväärtuse kohta. Näidistega võrdlemise korral näitavad tunnuste kaalud tunnuste indikaatorväärtust, samuti võib teada saada igal üksikul juhul kasutatud näidised, näidiste kasutamissageduse ning millisel sarnasuse tasemel üks või teine otsus langetati.

Eelkõige on näidistega prognoosi puhul oluline sõltumatu tunnuse prognoositavus kasutatavate argumenttunnuste järgi ja näidiste komplekti esindavus. Prognoosi tõepärasus sõltub kasutatavate tunnuste optimaalsest komplektist ja prognoosivate tunnuste kaaludest.

Näidistega võrdlemise peamised puudused on: 1) arvutuste mahu eksponentsiaalne kasv koos näidiste arvu kasvuga; 2) interpoleeriv iseloom (vähemalt juhul, kui hinnang arvutatakse näidiste kaalutud keskmisena). Ekstrapoleerimiseks väljapoole empiirilise andmestiku haaret tuleks näi-

diste komplekti täiendada teoreetiliste näidistega või arvutada prognoos mitte sarnasustega kaalutud keskmisena, vaid statistilise mudeli abil.

Uurimust finantseeriti Euroopa Liidu V raamprojekti projekti SPIN (*Spatial Indicators for Nature Conservation*) ja Eesti Teadusfondi grandid nr 5261 kaudu. Mullakaarti ja põhikaarti kasutati Riigi Maaameti litsentsilepingu nr 107, ortofotosid litsentsilepingute nr 46 ja 334 alusel.

Kirjandus

1. Aaviksoo, K., Paal, J., Dišlis, T. (2000). Mapping of wetland habitat diversity using satellite data and GIS: an example from the Alam-Pedja Nature Reserve, Estonia. *Proc. Estonian Acad. Sci Biol. Ecol.* 49 2, 177–193.
2. Aha, D.W. (1998). The omnipresence of case-based reasoning in science and application. *Knowledge-Based Systems*, 11, 261–273.
3. Birks, H. J. B. (1993). Quaternary palaeoecology and vegetation science — Current contributions and possible future developments. *Review of Paleobotany and Palynology*, 79, 1–2, 153–177.
4. Clark, J. D., Dunn, J. E., Smith, K. G. (1993). A multivariate model of female Black Bear habitat use for a Geographic Information System. *Journal of Wildlife Management*, 57, 3, 519–526.
5. Davies, C. E., Moss, D. (2002). *EUNIS habitat classification 2001 work programme final report*. European Environmental Agency. European Topic Centre on Nature Protection and Biodiversity.
6. Flower, R. J., Juggins, S., Battarbee, R. W. (1997). Matching diatom assemblages in lake sediment cores and modern surface sediment samples: The implications for lake conservation and restoration with special reference to acidified systems. *Hydrobiologia*, 344, 27–40.
7. Frize, M., Walker, R. (2000). Clinical decision-support systems for intensive care units using case-based reasoning. *Medical Engineering & Physics*, 22, 671–677.
8. Johnson, C. (2002). Software tools to support incidence reporting in safety-critical systems. *Safety Science*, 40, 765–780.
9. Knight, B., Taylor, S., Petridis, M., Ewer, J., Galea, E. R. (1999). A knowledge-based system to represent spatial reasoning for fire modelling. *Engineering Applications of Artificial Intelligence*, 12, 213–219.
10. Muinonen, E., Maltamo, M., Hyppänen, H., Vainikainen, V. (2001). Forest stand characteristics estimation using a most similar neighbor approach

and image spatial structure information. *Remote Sensing of Environment*, 78, 223–228.

11. Paal, J. (1997). *Eesti taimkatte kasvukohatüüpide klassifikatsioon*. Tallinn, 297.
12. Remm, K. (2000). Eesti ruutkilomeetrite andmebaas. In: T. Frey (ed.). *Problems of Contemporary Ecology*. Tartu, 241–247.
13. Remm, K. (2002). Otepää looduspargi taimkatte kasvukohatüüpide kaart. In: T. Frey (toim.), *Eesti süsinikubilansi ökoloogiast ja ökonoomikast. Eesti XIII ökoloogiapäev*. Tartu, 62–76.
14. Seitz, A., Uhrmacher, A. M., Damm, D. (1999). Case-based prediction in experimental medical studies. *Artificial Intelligence in Medicine*, 15, 255–273.
15. Settle, R. H., Sommerville, B. A., McCormick, J., Broom, D. M. (1994). Human scent matching using specially trained dogs. *Animal Behaviour*, 48, 1443–1448.
16. Suh, M. S., Jhee, W. C., Ko, Y. K., Lee, A. (1998). A case-based expert system approach for quality design. *Expert Systems With Applications*, 15, 181–190.
17. Verdenius, F., Broeze, J. (1999). Generalised and instance-specific modelling for biological systems. *Environmental Modelling & Software*, 14, 339–348.
18. Ward, B. O., Billinton, A., Wilkinson, L. S. (2001). Learning, remembering and applying an arbitrary non-matching to position rule in mice. *Behavioural Brain Research*, 125, 1–2, 229–236.
19. Wilds, S., Boetsch, J., Manen, F. T. van., Clark, J. D., White, P. S. (2000). Modeling the distributions of species and communities in Great Smoky Mountains National Park. *Computers and Electronics in Agriculture*, 27: 389–392.
20. Xia, Q., Rao, M. (1999). Dynamic case-based reasoning for process operation support systems. *Engineering Applications of Artificial Intelligence*, 12, 343–361.
21. Yeh, A. G. O., Shi, X. (2001). Case-based reasoning (CBR) in development control. *JAG*, 3, 3, 238–251.

ÜHE ELUTEADUSE LAINETEL: KONVERENTSIJÄRGSEID MÖTTEMÕLGUTUSI

Mati Rahu

Eksperimentaalse ja Kliinilise Meditsiini Instituut,
epidemioloogia ja biostatistika osakond, rahum@ekmi.ee

Võimatu on vältida lainetamist, kui sa tead, et koosned
95% ulatuses veest, ja maailm on alatasa tuuline ...

Arvo Valton (2001)

Sissejuhatus

Omanikuvahetuse tõmbetuulde sattunud Marguse spordibaasi aknast paistab künklik maastik sügisrüüts metsa, räsitud Väike-Munamäe ja jääkattes paisjärvega. Äsja lõppesid kahe päeva jooksul kõöksaali ja kaminatuba täitnud arutelud elu- ja statistikateaduste seoste üle. Rääkijatekuulajate-küsimiste kogemustepagas ulatus tudengist professorini, teemad geenist rahvastikuni, õhkkond innukast ... innukani. Meeldivas olustikus viibimine ja tavategemistest eemalolek andis mullegi tõuke heietada mõtisklusi ühe eluteaduse (epidemioloogia) tegemistest ning kokku-
saamistest biostatistikaga. Alljärgnevad uitmõtted võiksid pakkuda osale lugejaist eeskätt nostalgilist, teistele tunnetuslikku huvi.

Epidemioloogide ja biostatistikute rongisõit

Kord sõitsid biostatistikud ja epidemioloogid samas rongis ühisele nõupidamisele. Iga epidemioloog oli ostnud sõidupileti, biostatistikutel oli kamba peale vaid üks pilet. Loomult teadmishimulised epidemioloogid küsisid biostatistikutelt, kuidas nad saavad konduktoriga hakkama, kui nende käsutuses on nii väike valim piletitest. Biostatistikud vastasid: "See on lihtne. Me kasutame asjakohast metoodikat."

Pärastpoole, kui konduktor tuli pileteid augustama, lipsasid biostatistikud märkamata tualettruumi. Kui konduktor uksele koputas, pistis biostatistikute boss pileti ukse alt läbi ja tõmbas kogenematu konduktori haneks.

Pärast nõupidamist sattusid biostatistikud ja epidemioloogid jällegi samasse rongi. Alati kiire taibu poolest tuntud epidemioloogid olid seekord ostnud üheainsa pileti. Biostatistikud (alati teaduse eesliinil rühkijatena) ei olnud ostnud ühtegi. Hämmeldunud epidemioloogid pärisid: "Me mõistame, kuidas teie metoodika töötas ühe pileti korral, aga mil moel te saate hakkama piletitä?" "Lihtsalt," vastasid biostatistikud, "Meil on niisuguse olukorra jaoks erisugune metoodika."

Hiljem, kui konduktor jõudis kõrvalvagnisse, ruttasid epidemioloogid oma ainsa piletiga ühte, biostatistikud teise tualetti. Varsti hiilis biostatistikute boss epidemioloogide peidukoha juurde ja koputas nõudlikult uksele. Nagu õpetatud, torkasid epidemioloogid pileti ukse alt läbi. Boss võttis pileti ja naasis võidukalt teiste biostatistikute juurde. Loomulikult jäid epidemioloogid vahele ja neid häbistati avalikult.

LOO MORAAL: ärge kasutage statistilisi meetodeid, kui te ei mõista, mis põhimõtetel nad rajanevad [Bernier, Pierre, 2000].

Miks ma selle ammuse, paljude jaoks vahest kulunudki naljaloo eestindasin? Ikka ja ainult eesmärgil rõhutada *biostatistika ülisuurt tähtsust* epidemioloogia ja teiste eluteaduste jaoks.

Mõned biostatistikud

1960. aastate lõpul oli Eestis minule teadaolevalt kolm biostatistikut: Leo Võhandu, Artur Nilson ja Tõnu Möls. Oma silmaga nägin neist esimesena Võhandut, kunagi hiljem külastasime koos Nilsonit (seoses siseprefokaartide töötusega), ning Mölsi kohtasin alles 1980. aastate lõpupoole. Nii et kui siin-seal tuli juttu Eesti biostatistikute arvust, sain loendamisel hakkama ühe käe sõrmedega.

Püsisidemed tekkisid Leo Võhanduga, kellega esimene kokkusaamine leidis aset 1966. aasta kevadsemestri alguses. Kolmanda kursuse tudengina astusin tema tööruumi, tutvustasin ennast, teatasin, et mul on eriprogrammi kohaselt ette nähtud üld- ja meditsiinistatistika kursus, ja küsisin, kas võiksin tema käe alla jääda. Võhandu soovitas mulle kohe raamatu (Austin Bradford Hill'i (1955) 1958. aastal vene keeles ilmunud meditsiinistatistika õpiku), mille alusel pidin talle arvestuse ja eksami tegema. Võhandu käe alla ma ei jäänud, aga tema kaelal olen istunud juba aastakümneid. Kaugeltki mitte pidevalt, kuid ikka siis, kui mingi

meetodi või termini kallal pusides kimpu jään või kui arvutusvõimsusest puudu tuleb. Alati on Võhandul aega, hea tuju ja lõpmatu hulk ideid. Et viimased on minu jaoks osutunud pahatihti üleliia keeruliseks, olen olnud sunnitud kõrvad longu laskma ja näole arusaava ilme manama.

Ehkki üksnes mõne korra olen palgest palgesse näinud Uno Merestet, kes ennast ilmselt biostatistikuks ei pea, on tema kirjutatu mind püsivalt abistanud ja innustanud. Nimetaksin ideid teaduste klassifitseerimisest ning raamatuid statistika üldteooria, arvjooniste, kompleksanalüüsi, süsteemse mõtlemisviisi, majandusanalüüsi teooria, rahvaloenduse ja oskuskeeleõpetuse kohta. Oli põnev avastada, et majanduslikus võrdlevanalüüsis tegelesid saksa autorid 1950.–1960. aastatel samaga, mis minuni jõudis epidemioloogiakirjandusest hulga hiljem — so segavate tegurite eristamise ja segamise mõju kõrvaldamisega. Paraku oleme mõlemad jäänud võimetuks parasiitsõna "populatsioon" (minule lühidalt "popu") väljarookimisel meditsiinikirjutistest. Ilmselt ei tea paljud, et meil Eestis elutsevad "laste-", "Tartu maakonna laste", "Tallinna mees-", "HI-viirusega nakatunud meeste", "vanemate inimeste", "seksuaalselt aktiivsed" ja muudki populatsioonid.

Edaspidi ilmus minu vaatevälja Ene-Margit Tiit — esmalt tema ja tema tiimi kirjutatud ülivajalike statistikaõpikute kaudu, pärastpoole ühise istumisega tänaseks põhja lastud järjekorras teise meditsiiniregistrite nõukogu koosolekutel. Ja nii me püüame teineteist ergutades — senini veel õhkõrn lootusekübe südames — arendada ametnikkonna ja tava-kodaniku registriteadvust.

Hetkel tundub, et biostatistikute kokkulugemisel jääb kahe käe sõrmedest väheks. Krista Fischeri hoogsa tuleku ja entusiasmi kiiluvees ning teisteski hoovustes on biostatistikute pere täienenud ja täienemas uute nägudega. Kui tavaolud uusikuid ära ei muserda ega ravimifirmad neist valdavat enamikku erinišile ei maanda, peaksid paljude eluteadlaste tööd omandama korraliku biostatistilise aluse.

Imesõltuvus

Tom Clancy on öelnud, et mingil ajal olid naelad kõrgtehnoloogia [Freedman, Schafer, 1995]. Nii üksikisikust eluteadlasele ja biostatistikule kui ka nende põlvkondadele tervikuna on iga arvutustehnoloogiline

uuendus järjekordne tõeline ime, mis on võimaldanud senisest kiiremini ja tõhusamalt rehkendusi teha.

Millistest imedest olen ise osa saanud? 1964. aasta suvel asendas välitöödel varasemaid tuttavaid — paberit-pliiatsit ja arvutuslükakit — üürikeseks *arvelaud*, millega pidin korrutamise ära õppima. Umbes aasta pärast ilmus käeulatusse *Felix*. Oli teine selline vanda ja kangikestega instrument, mille sain kauplemise peale ülikooli majandusteaduskonnast, et teha arvutusi aegridade tasandamiseks. Minu kätte usaldatud *Felxi* eksemplar osutus pöduraks, ta läks õnneks üliruttu täitsa kihva ja andis seega õigustatud võimaluse minna sujuvalt üle sama teaduskonna *elektromehaaniliste klahvarvutite* kasutamisele. Need meenusid kujult kandilisemaid kirjutusmasinaid, tegid arvutamisel lärmi ja muutusid tundetuks, kui juhtus mõni neile vastukarva vajutus klahvile; eluvaim tärkas neis uuesti alles mehaaniku abiga.

Järgmine ime ilmutas ennast *suurarvuti* kujul. Kuigi olin teinud eksami programmeerimiskeeles ALGOL-60, polnud ma oma ihusilmaga tõelist elektronarvutit kaenud. Viimase kursuse tudengina 1968. aasta kevadsemestril nägin TPI arvutuskeskuses hiidu nimega *Minsk-22*. Klassivend Mati Tombak, kes oli mulle χ^2 arvutamiseks programmijupi kirjutanud, lõi mind hilisõhtuses masinasaalis täiesti pahviks — elektronarvutist voogas välja Mozarti meloodia. Et mul oli kange tahtmine usaldada masinale haiguste leviku kaartide valmistamine, tegutsesin Leo Võhandu ettepanekul mõnda aega Eesti Raadio arvutuskeskuses. Sealsele *Razdan-3*-le tehti 1969–1972 parajasti andmetöötlussüsteemi SODI. Õppisin ära *Razdani* masinakeele ja mulle vajalike moodulite programmeerimise. Olin rahul, et minu programmid võimaldasid vähkihaigestumuskaarte välja trükkida. Paraku ei jätkunud elevust kauaks. Programmeerimisguru (nii vist praegu öeldakse) Tarmo Aus sõnas lakooniliselt, et ma tegelen mitte programmeerimise, vaid ainult käskude rittapanekuga. Sedapsi, mõningase kibedusega hinges, potsatasin maa peale tagasi, ent ... aju sai treenitud ja kaardid valmisid.

1970. aastate esimesel poolel pruukisin TPI keldris asunud *WANG-2200 A-d*. Et vajasin seda riistapuud paar korda aastas, tundsin teda pinnapealselt ja jäin töötamisel alailma jänni. 1970. aastate keskel andis kolleeg Malle Lill mulle statistikaalase nõuandmise eest välismaise *taskuarvuti*, millega sai isegi ... ruutjuurt võtta; kingitus oli kuninglik ja tegi

meele härdaks. Täiesti ootamatu pauk käis 5. novembril 1979, kui külaskäigu ajal Harvardi ülikooli kinkis epidemioloogia *grand old man* Brian MacMahon mulle programmeeritava taskuarvuti *HP-67*. Arvutiga olid kaasas magnetkaardikesed epidemioloogilisteks arvutusteks vajalike programmidega ja andmeanalüüsi käsiraamat [Rothman, Boice, 1979]. Õnneliku hetkega sain oma kätte parima, mida maailm suutis taskuarvuti näol tol ajal epidemioloogile pakkuda. Kingitu innustas süvenema epidemioloogia teoriasse ning pakkus kergendust efektiivajate ja nende hajuvuse arvutamisel. Kui õigesti mäletan, rehkendasin esimesed šansside suhted ja nende usalduspiirid 2×2-tabelite seeria jaoks käsiloleva rinnavähi *case-control*-uuringu raames. (*Case-control study* praeguse emakeelse vaste *juhtkontrolluuring* tekkeni jäi 14 aastat.)

Järgmine tähelepanuväärne päev — 11. juuni 1982. Soome Vähiregister ostis Eero Pukkala vahendusel mulle õhukese taskuarvuti *Canon Card F-57T*. Viimase üldotstarbelistest võimalustest olen tihti kasutanud äratuskella, mistõttu lähetustel on see Jaapani ime mind saatnud juba 20 aastat.

1980. aastate keskpaiku, tänu Arvi Sinisalu vastutulelikkusele, pääsesin tervishoiuministeeriumi arvutuskeskuses õhtuti ja puhkepäeviti ligi *personaalarvutile Olivetti*. Erilist rõõmu valmistasid *Olivetti* tehtud arvjoonised elulemuskõveratega [Aareleid, Rahu, 1987]. Jooniste tehniline tase jättis küll soovida, ent lõpuks ometi sain jätta hüvasti aegadega, mil pidin mööda linna otsima joonestajaid, rublad näpu vahel või piiritusepudel diplomaadikohvril ...

Andmetöötamise okkalisel rajal tekitas parasjagu närvikõdi uuring, milles mõõdeti rinnavähiriski kopsutuberkuloosi põdenud naiste hulgas. Et uuringu iseloomust tulenevalt oli vaja kasutada Coxi võrdelise riski regressioonimudelit, pöörasime lootusrikkad pilgud statistikapaketi BMDP uue versiooni poole, mis sisaldas vastavat programmi. Kuulduste kohaselt hankis Nõukogude Liit BMDP pakette Indiast, kuhu need saabusid tehnoloogilise abi korras USAst. Tavategevusena tõlgiti BMDP Nõukogude Liidus osaliselt vene keelde ja pikka aega oli ta ainuke laialdaselt kasutatav üldotstarbeline tarkvarapakett. Üks BMDP loojatest, Will Dixon, kes mitmeid kordi külastas Nõukogudemaad, veendus oma silmaga, kui võrd populaarne oli seal ebaseaduslikult imporditud pakett [Flournoy, 1993]. Meid huvitav versioon pidi kohe-

kohe jõudma Moskvasse ja edasi Tartu ülikooli, kust tahtsimegi asjatundlikku abi saada. Kuid Indira Gandhi mõrvamine 1984. aastal tekitas poliitikataevasse pilved ja pakett jäi tulemata. Nüüd aitas hädast välja biostatistikust doktorant Leena Tenkanen, kes Soome Vähiregistris analüüsi ette võttis. Sama kümnendi lõpul tegin näpuharjutuse GW-BASICus, mille tulemusena *Olivetti M24* personaalarvutist väljusid Chernoffi näod kirjeldamaks vähkihaigestumuse ja vähiregistreerimise taset riigiti [Rahu, 1989].

Aprillis 1991 oli 150 töötajaga Eksperimentaalse ja Kliinilise Meditsiini Instituudis üks personaalarvuti. Kõnekas tõik, mis iseloomustab teadusetegemise tõhusust. Seejärel pais vallandus: tänu kaalukale rahvusvahelisele projektile hakkas minu osakonda 1990. aastate keskel personaalarvuteid juurde siginema ja peatselt lebas igaihe laual arvuti. Me ei kujutagi ette, et elu võiks olla korraldatud kuidagi teisiti – ilma sõbraliku kuvarita meie vastas.

Olla või mitte olla

Biostatistiku elukutseks valmistuja peaks olema kindlasti kursis, millised ahvatlused teda ees ootavad (tabel 1). Eriti kosutab hinge teadmine, et leitud töökohta on raske kaotada, sest teiste erialade inimesed konkureerima ei kipu. Epidemioloogi elukutse valinut nii helge tulevik ei oota. Raskused johtuvad pidevast vajadusest tunda keeruka naaberteaduse uusimat sõna, ning märkimisväärsest painest, mis kaasneb epidemioloogiks olemisega (tabel 2). Teised epidemioloogi kohta eriti ei himusta. Kuid epidemioloogi tööd matkivad paljud, sest arvatakse tõemeeli, et epidemioloogia tegeleb üksnes küsitlusuuringutega, et *küsida oskab igaiüks* ja et *küsida võib kõike*. Biostatistikuid ja eluteadlasi (epidemiolooge) seob palju ühist, kuid mitte ainult meetodika valdkonnas. Nende huvide ühtsust väljendab kaudsel moel X linna teletõrjekomando aastalõpul antud töötus: “Kui me käesoleval aastal hakkasime tuld kustutama pool tundi pärast tulekahju algust, siis järgmisel aastal oleme kohal vähemalt viis minutit enne süttimist.” Ehk teisisõnu, mõlemad pooled on eluliselt huvitatud, et juba uuringu kavandamise algjärgus saaks sõlmitud tutvus ja üheskoos tehtud uuringuprotokoll. Nii vajuksid möödanikku ajad, mil eluteadlane kougib biostatistikuga esmakohtumisel oma paunast hunnikute viisi välja täiskritseldatud paberilehti ja

Tabel 1. Kümme põhjust, miks hakata (bio)statistikuks
[Ramseyer, 2002]

-
- 1 (Bio)statistikutena peame hälvet normaalseks.
 - 2 Me tunneme ennast täielike ja piisavatena.
 - 3 Me oleme armastaja etalon.
 - 4 Me teeme seda diskreetselt ja pidevalt.
 - 5 Meil on õigus 95%-l juhtudest.
 - 6 Me ei pea kunagi ütleva, et oleme selles kindlad.
 - 7 Me võime mitte olla normaalsed, kuid oleme teisendatavad.
 - 8 Me võime legaalselt kommenteerida kellegi posterioorset jaotust.
 - 9 Me oleme tõemeeli oluliselt erinevad.
 - 10 Mitte keegi ei himusta meie töökohta.
-

Tabel 2. Kümme põhjust, miks loobuda epidemioloogi elukutsesest
[The top ..., 2000]

-
- 1 Me ei tarvitse enam õppida statistikat rohkem, kui soovime.
 - 2 Me ei pea tavavestluses kasutama sõna "oluline".
 - 3 Meie järeleandmatus nullhüpoteesi jahil muudab me elu testimiseks ja kummutamiseks.
 - 4 Meil ei ole vaja selgitada, et hoolimata meie sobitatud paaridest ei tööta me kosjakontoris.
 - 5 Me ei pea ealeski käsitama juhust, nihet ega segamist oma elu tähtsama jõuna.
 - 6 Me tahame tunnustada individuaalsust ega mitte mõelda rühmitamisele soo, vanuse ja rahvuse järgi.
 - 7 Me ei pea enam leidma ennast kolmandat korda meelsasti täitmas rahvaloenduse küsitluslehte.
 - 8 Me ei pea panema auto kaitserauale kleepekat tekstiga: "Teen kõik, et vastasmäär suureneks!"
 - 9 Meil ei tule enam kuulda: "Oo, Te olete üks *nende* hulgast! Täna Te ütlete mulle, et ma ei tohi kohvi juua, aga mõni kuu hiljem kirjutate terviseõpetuses, et kohvi joomine on okei."
 - 10 Meil ei tule enam ette kriisiolukorda, kui keegi röögib: "Kas majas on mõni epidemioloog?!"
-

avaldata soovi andmete "statistiliseks analüüsiks". Või saavad nad kokku alles hädaolukorras, kui eluteadlase näpu vahel on ajakirja toimetusest tulnud retsensioonid, milles tehakse suuri etteheiteid uuringukavandi, valimialuse, uuritavate värbamise taktika, analüüsimeetodite jms suhtes.

Biostatistikat ja epidemioloogiat lähendab veelgi teadmine, et neid õppeaineid peavad näiteks Suurbritannia ja USA meditsiinitudengid igavaks ja kasutuks [Hebel, Morton, 1979; Clarke jt, 1980]. Olukorra parandamise nimel otsivad biostatistikud ja epidemioloogid pedagoogilisi nippe, kuidas kõita "vastupunnivat õppurit" [Floreay, 1992; Simpson, 1995].

Biostatistika ja epidemioloogia on ajaloo jooksul teineteist täiendanud ja igati sõbralikult läbi saanud. Tean üht suurt musta kassi, kes alustas nende vahelt läbijooksmist 1950. aastate teisel poolel. Nimelt mitmed statistikud, sealhulgas tubakatööstuse palgalise konsultandina töötanud kuulus Ronald Fischer [Stolley, 1991], eitasid põhjusliku seose olemasolu suitsetamise ja kopsuvähi vahel.

Mida edasi teha?

Eesti juurde tagasi pöördudes loetleksin mõned lühikesed sammud, mille astumine võiks edendada (bio)statistika ja epidemioloogia (eluteaduste) (koos)arengut.

- 1. Eesti Statistikaameti (ESA) juriidilise staatuse muutmine.** ESA muutmine avalik-õiguslikuks juriidiliseks isikuks aitaks suurendada tema autonoomsust ja autoriteeti. Lisaks vajab rõhutamist — niisugune samm võimaldaks ESA-l jõulisemalt kaitsta oma au ja vääriskust. Niiviisi avaneks lõpuks võimalus luua ESAs *riiklik suremuse register*, mille täisvereline tegevus huvitab Eestis paljusid institutsioone, eriti neid, kus tehakse terviseuuringuid. Praeguse suremuse andmekogu tegutsemine juriidilises vaakumis [Rahu, 2002] ja surve isikut tuvastavate andmete eemaldamiseks sellest andmekogust teevad lihtsalt häbi kogu Eesti riigile.
- 2. Eluteadlase ja biostatistiku ameti väärtustamine ühiskonnas.** Senisest võrratult suurema tähelepanu pööramine avalikkuse ja teiste erialade inimeste teavitamisele meie uuringutulemustest ka juhul, kui nad on ilmunud mujal kui maailma tippajakirjades. Kui me sel-

le töö unarusse jätame, riskime oma igapäevategevuse põlu alla satumise ja rämpsteaduse tegija maine omandamisega. Meie vaikiva oleku tõttu on teised — *loodus tühja kohta ei salli* — oma terve mõistuse najal hakanud meid innukalt paika panema ja õpetussõnu jagama, mis meetodikaga peab tegema epidemioloogilisi uuringuid ja andmetöötlust.

- 3. Loodud andmekogude säilitamine korduvkasutamiseks.** Teadustöö käigus luuakse palju andmekogusid, mis võiksid (tulevikus) pakkuda olulist huvi teistele teadusrühmadele. Uuringusse võetud isikute edasise jälgimise (nt terviseseisundi suhtes) ja andmete linkimise seisukohalt on tähtis *isikut tuvastavate tunnuste olemasolu, andmekogu üksikasjaline kirjeldamine ja selle pikaajaline säilitamine*. Eeltoodu on ühtlasi maksev administratiivtöö käigus rajatud andmekogude kohta, mida kahjuks tavatsetakse käsitada hetkeülesannete täitmise instrumendina, mitte laiemat eesmärki teenivana. Siia sobib näitena *ravikindlustuse andmekogu*, mille säilitamisaeg on seitse aastat (<http://www.riik.ee/arr/register/ravikindl.html>).

Kokkuvõtte asemel

Epidemioloogia ja biostatistika on köitvad teadusharud, mille ühistegevus arvestab *epidemioloogia kesket paradigmat* — haiguste rahvastikus esinemise süstemaatiline analüüs aitab mõista haiguste põhjusi ja ennetada haigusi [Bhopal, 2002]. Epidemioloogide ja biostatistikute edukus Eestiks nimetatud maalapikesel sõltub mitmete tuulte suunast ja tugevusest: riigi üldisest arengutasemest, ametnikkonna pädevusest, teadus- ja arendustegevuse sihtidest, avalikust arvamusest, seadusloome suhtumisest terviseuuringutesse, ning kindlasti — põhitegijate visadusest ja tarmukusest. Julgustagu meid igapäevatoos kunagi kellegi poolt öeldu: "Kui nalja ei saa, ei ole tegemist epidemioloogiaga ..."

Naljalugude hulgas pärineb üks seik 1970. aastatest, kui Tallinna onkoloogiadispanserit külastas Nõukogude Liidu epidemioloogiakorüfee professor Aleksandr Tšaklin. Keset vestlust haaras professor äkki kahe käega peast kinni ja kurtis, et peab olema järgmisel päeval Kiievis väitekirja kaitsmisel, kuid ei suuda meenutada, kas ta on juhendaja või oponent [Rahu, 1982].

Teine, samuti tõsisündmus, juhtus 1987. aastal Lyonis. Sealsesse vähiuurimiskeskusesse saabus täiendusõppele Vilniuse epidemioloog Romas G. Ta valis välja tüürikorteri, mille peremees hakkas talle näitama tube. Ühes ruumis osutas peremees mingile valgele esemele ja küsis: "Kas Te teate, mis see on?" "Bidee", vastas Romas. Peremees küsitles kohe edasi: "Kas Te borši armastate süüa?" Epidemioloog, näitamata välja hämmingut kahe seosetu küsimuse üle, teatas naeratades, et talle borš maitseb. Seepeale tõmbus peremehe nägu tõsiseks ja ta lausus murelikul häälel: "Jah, paar aastat tagasi elasid minu juures kaks selli Teie suurest riigist (st Nõukogude Liidust). Nad lõhkusid ära bidee, sest keetsid spiraalidega selles borši ..."

Kolmas lugu puudutab Eesti epidemioloogidele lähedal seisvat isikut, kes 1995. aastal töötas eksperdina mõne nädala Brüsselis. Et ta pidi oma töötasust maksma ühtlasi hotelli eest, otsustas ta mõne päeva möödudes otsida odavama elamise. Varsti leidiski ta hulga sobivama hotelli ja pikemalt ringi vaatamata möllis ennast sinna. Hiljem, tagantjärele tarkuse-na, arvas ta, et sisse kirjutamisel tekitas päevade arvu nimetamine hoteliametniku pilgus hetkelise elevuse-imetlusehelgi, kuid toona ei osanud ta sellele mitte mingit tähendust omistada. Tuba ise ei andnud veel alust midagi arvata, kuigi peegelklaasist lagi voodi kohal tundus võõrastavana. Tõde selgus öösel, kui läbi seinte kostsid naabertubadest tõeliste proffide megaorgasme väljendavad hääliksused, ja nii ... kaks nädalat jutti. Ehkki taolises mitte just päris tavalises hotellis tuli magada, pea kahe padja vahel, tasus raskuste trotsimine ennast ära, sest rahaline kokuhoid oli märkimisväärne.

Endiselt, epidemioloogia ja biostatistika on köitvad teadusharud ...

Kirjandus

1. Aareleid, T., Rahu, M. (1987). Pahaloomuliste kasvajatega haigete eluloomus Eestis 1968–1981. *Eesti Arst*, 66, 3, 171–173.
2. Bernier, R. H., Pierre, J. L. St., eds. (2000) *Epidemiology wit & wisdom: the best of The Epidemiology Monitor*. Roswell: The Epidemiology Monitor, 245.
3. Bhopal, R. S. (2002). *Concepts of epidemiology: an integrated introduction to the ideas, theories, principles and methods of epidemiology*. New York: Oxford University Press, 3.

4. Clarke, M., Clayton, D. G., Donaldson, L. J. (1980). Teaching epidemiology and statistics to medical students – the Leicester experience. *Int J Epidemiol*, 9, 2, 179–185.
5. Florey, C. V. (1992). Teaching the reluctant student. In: Olsen, J., Trichopoulos, D., eds. *Teaching epidemiology: what you should know and what you could do*. Oxford: Oxford University Press, 23–32.
6. Flournoy, N. (1993). A conversation with Wilfrid J. Dixon. *Stat Science*, 8, 4, 458–477.
7. Freedman, D. H., Schafer, S. (1995). Vonnegut and Clancy on technology. <http://www.inc.com/magazine/19951215/2653.html>
8. Hebel, J. R., Morton, R. F. (1979). Teaching experience with an independent study programme in epidemiology and biostatistics. *Int J Epidemiol*, 8, 2, 177–179.
9. Hill, A. B. (1955). *Principles of medical statistics*. 6th ed. London: The Lancet.
10. Rahu, M. (1982). Aleksandr Tšaklin. *Eesti Loodus*, 25, 6, 402–403.
11. Rahu, M. (1989). Graphical representation of cancer incidence data: Chernoff faces. *Int J Epidemiol*, 18, 4, 763–767.
12. Rahu, M. (2000). Kas Tšernobõl rikkus meeste tervise? *Luup*, 4, 40–41.
13. Ramseyer, G. (2002) First Internet gallery of statistical jokes: 44. <http://www.ilstu.edu/~gcr Ramseyer/Gallery.html>
14. Rothman, K. J., Boice, J. D. (1979). *Epidemiologic analysis with a programmable calculator*. Washington, NIH.
15. Simpson, J. M. (1995). Teaching statistics to non-specialists. *Stat Med*, 14, 2, 199–208.
16. Stolley, P. D. (1991). When genius errs: R. A. Fischer and the lung cancer controversy. *Am J Epidemiol*, 133, 5, 416–425.
17. The top ten reasons not to be an epidemiologist (2000). *The Epidemiology Monitor*, 21, 11, 13, 17.
18. Valton, A. (2001). *Seltskondlik üksindus: aforismid*. Tallinn: Angestel, 17.

EESTI STATISTIKASELTSI KÜMMETEGEVUSAASTAT

Ene-Margit Tiit

1992. aasta sügisel, kui taasiseseisvunud Eestis intensiivselt loodi ja taasloodi seltse ning ühendusi, pandi alus ka Eesti Statistikaltsile. Siin on sobiv koht õiendada ka eesti entsüklopeedia 11. köites sisalduv väike ebatäpsus: seal on Eesti Statistikaltsi meeles peetud leheküljel 514, kus on kirjas, et selts taasloodi 1994. aastal — tegelikult rajati Eesti Statistikalts siiski n-ö tühjale kohale.

Loomisel seadis Eesti Statistikalts eesmärgiks Eesti statistikute ühendamise, samuti akadeemilise statistika lõimumise praktikaga. Nende eesmärkide teostamiseks otsustati korraldada konverents, mis peaks toimuma Eesti erinevates paikades ja anda välja häälekandjat ESS Teabevihik. Nende põhiliste punktide täitmine on seltsil kulgenud üsna edukalt, nagu näitab lisatud tabel. Kuigi konverentside tihedus on viimastel aastatel olnud kavandatust mõnevõrra väiksem, on ettekannete tase, ettekannete ja konverentsil osalejate (sh ka noorte) arv pigem tõusnud. Sama tuleb öelda teabevihikute kohta. Isegi geograafilise hajutatuse osas on suudetud plaane täita — Tartu ja Tallinna kõrval on koos käidud Viljandis, Rakveres ja Otepääl.

Peale kohalike konverentside on Statistikalts olnud mitmel rahvusvahelisel konverentsil kaaskorraldajaks, samuti on lisaks Teabevihikutele välja antud mõningaid teisi väljaandeid. Seltsil on oma kodulehekülg, kuid selle kasutamise aktiivsus on suhteliselt väike. Ka pole õnnestunud korraldada teiselaadilisi üritusi peale konverentside, kuigi mõningaid ettepanekuid on tehtud.

Statistiline kokkuvõte Eesti Statistikaltsi 10 aasta tööst hõlmab 15 konverentsi kokku 26 päeval, 259 ettekande ja 985 osavõtjaga. Avaldatud (ja avaldamisel) on 13 teabevihikut kokku 1286 leheküljel. Liikmeskond on püsinud 100 inimese ringis, kuigi on toimunud arvestatavad muutused — on lahkunud neid, kes ei tunne end enam olevat statistikaga seotud, ja juurde on tulnud uusi noori tegijaid.

ESS KONVERENTSID JA NENDE KAJASTAMINE ESS TEABEVIHIKUTES

Konverents							Teabevihik		
Konverents	Koht	Aeg	Ette- kandeid	Osa- võtjaid	Korraldajad*	Partner*	Nr	Aasta	Lk
Asutamis- konverents	Tartu	30.09.92	8	90	E.-M. Tiit	TÜ**	1	1993	35
Eesti rahvuslik arvepidamine	Tallinn	15.12.92	4	25	A.Tihemets	ESA			
Registrid Eesti sta- tistika- ja informaa- tikasüsteemis	Viljandi	2.-3. 03.93	12	40	M. Servinski	Viljandi SB	2	1993	64
Statistika tarkvara Eestis	Tartu	20.- 21.05.93	22	50	K. Ääremaa, Ü. Randaru	Tartu SB			
Majandus- ja inseneristatistika	Tartu	2.-3.11.93	16	50	V. Tamm, T. Kollo	TÜ MTK	3	1994	52
Matemaatiline statistika tänapäeva koolis	Tallinn	8.-9.05.94	14	60	E. Saks	TPedI	4	1994	69
Statistika eluteadustes	Tartu	24.03.94	15	50	E. Käärik	TÜ, EPMÜ	5	1995	85
Statistika ühiskonna- uuringutes	Tartu	22.- 23.03.95	24	60	L.-M. Too- ding	TÜ STK, TPedI	6	1995	190
7. Konverents. Riigistatistika	Tallinn	19.- 20.09.95	16	100	H. Vigla	ESA	7	1996	108
Kvaliteedisuundu- mused Eestis	Tallinn	14.11.96	13	90	K. Pärna	Eesti KÜ	8	1997	116
9. Konverents Meditsiinistatistika ja -registrid	Rak- vere	17.- 18.04.97	22	80	R. Malbe	Med. statistika Büroo	9	1997	93
10. Konverents Majandus- ja kindlustusstatistika	Tartu	29.- 30.10.97	28	70	V. Tamm, T. Kollo	TÜ MTK	10	1998	167
11. Konverents Statistika keskkonnakaitse ja ökoloogias	Tartu	16.- 17.04.98	9	30	T. Möls	ELUS, EPMÜ	11	2001	73
13. Konverents Rahvastikustatis- tika ja registrid	Tallinn	8.-9.11.00	24	120	Ü. Randaru, E.Tamm	ESA	12	2001	109
14. Konverents Statistika eluteadustes	Otepää	1.-2.11.02	32	70	E.-M. Tiit, K. Fischer	TÜ ATk	13	2003	125

*Kõigi konverentside korraldamisel osales ESSi president ja üks organi-
seerivaid asutusi oli TÜ matemaatilise statistika instituut.

KROONIKA

KAITSTUD BAKALAUREUSETÖÖD 2001–2002

TÜ matemaatilise statistika instituut

2001

1. Jevgeni Budalov. *Loglineaarne analüüs*. Juhendaja dotsent A.-M. Parring.
2. Olga Gorjajeva. *Erindid statistilises andmestikus*. Juhendaja lektor S. Koskel.
3. Jane Jostov. *Logistiline regressioon statistikapaketis SPSS*. Juhendaja dotsent A.-M. Parring.
4. Sirje Küber. *Dispersioonanalüüs paketi SPSS*. Juhendaja lektor M. Vähi.
5. Valeri Koort. *Simulatsiooni- ja puumeetoditest Euroopa tüüpi optioonide hindade arvutamisel*. Juhendaja lektor R. Kangro.
6. Elsa Leiten. *Statistilisest metsainventeerimisest Eestis*. Juhendaja dotsent I. Traat.
7. Agneta Matthiessen. *Mitmene võrdlemine*. Juhendaja dotsent A.-M. Parring.
8. Jaana Raudla. *Eestlaste suremusnäitajad kõrges vanuses*. Juhendaja professor T. Kollo.
9. Jevgeni Rjabinin. *Bootstrap valikuteoorias*. Juhendaja dotsent I. Traat.
10. Anne Selart. *Mitmemõõtmeline asümmeetriline normaaljaotus*. Juhendaja professor T. Kollo.

2002

1. Ivika Allekand. *Asümmeetriline normaaljaotus andmemudelina*. Juhendaja professor T. Kollo.
2. Tiia Arro. *Valimimahu määramine*. Juhendaja lektor E. Käärrik.
3. Piret Bild. *Mõõtmisvead valikuteoorias*. Juhendaja dotsent I. Traat.

4. Anna Bileva. *Hüpoteeside kontrollimine ja võimsuse arvutamine Analyst Application abil*. Juhendaja dotsent A.-M. Parring.
5. Marina Faruškina. *Programmi Matrixer kasutamise juhend*. Juhendaja lektor E. Käärik.
6. Natalja Gluškova. *Kindlustuskahjude jaotuse modelleerimine*. Juhendaja professor T. Kollo.
7. Kati Hoop. *Portfelliteooriad ja SAS tarkvara*. Juhendaja professor K. Pärna.
8. Mihhail Juhkam. *Valimi värvilisus ja katvus multinomiaalse ning Poissoni skeemi korral*. Juhendaja professor K. Pärna.
9. Pille Kool. *Faktoranalüüsi mudel ja selle identifitseerimine*. Juhendaja dotsent A.-M. Parring.
10. Tanel Lumiste. *Internetiankeedi koostamine ja tulemuste statistiline analüüs*. Juhendaja emeriitprofessor E.-M. Tiit.
11. Piret Paomees. *Puu vanuse määramisest korba paksuse järgi*. Juhendaja professor T. Kollo.
12. Ivika Puusepp. *Kalibreerimismeetodite võrdlus Eesti töajõu-uuringus*. Juhendaja doktorant K. Sõstra.
13. Kadri Ruusmaa. *Kvalitatiivsete argumenttunnustega lineaarsed mudelid*. Juhendaja dotsent A.-M. Parring.
14. Ardo Säks. *Investori kasulikkusfunktsioonile vastavate optioonidest ja aktsiatest moodustatud optimaalsete portfelli koostamine Markowitzi portfelli teooria kohaselt*. Juhendaja vanemteadur R. Kangro (RMI).
15. Risto Tamme. *Peipsi mudeli realiseerimine programmeerimiskeeles Java*. Juhendaja dotsent T. Möls.
16. Taavi Tasuja. *Prioriteetidega järjekorrasüsteemid: analüüs ja optimeerimine SAS-protseduur QSIM abil*. Juhendaja professor K. Pärna.
17. Anu Teeäär. *Mitmemõõtmeline asümmeetriline normaaljaotus*. Juhendaja professor T. Kollo.
18. Randel Veerits. *Meditšiiniuuringutes kasutatavad katseplaaniid ja seosekordajate hindamismeetodid erinevate katseplaaniid korral*. Juhendaja teadur M. Möls.

KAITSTUD MAGISTRITÖÖD 2001–2002

TÜ matemaatilise statistika instituut

2001

Finants- ja kindlustusmatemaatika eriala

1. Julia Abramova. *Suremustabel ja riskipreemia Hansapanga kindlustuse näitel*. Juhendaja dotsent O. Karma.
2. Irina Keller. *Delta-hedgingu perioodi optimeerimine valuutaoptsoonide korral*. Juhendaja professor K. Pärna.

Matemaatilise statistika eriala

1. Tatjana Obidina. *Description of sampling designs by eigenvalues*. Juhendaja dotsent I. Traat.
2. Kaari Orav. *Riski väärtus ehk VaR Eesti Panga investeerimisportfelli näitel*. Juhendaja professor T. Kollo.
3. Margus Pihlak. *Mitmemõõtmeliste jaotusfunktsioonide lähendamine*. Juhendaja professor T. Kollo.
4. Heti Pisarev. *Structural nested mean models in a repeated measures setting*. Juhendaja teadur K. Fischer.

Välismaal kaitstud magistritöö biostatistika erialal

1. Inge Ringmets. *Resistance to (re)infection with Schistosoma mansoni in Senegalese epidemic foci*. Juhendaja professor M. Aerts (Limburgi Ülikool, Belgia), kaasjuhendaja J. Scott.

2002

Finants- ja kindlustusmatemaatika eriala

1. Anne Reitsak. *Laostumistõenäosuste hinnangud lõpliku ajahorisondi juhul*. Juhendaja professor K. Pärna.

Matemaatilise statistika eriala

1. Svetlana Bizjajeva. *Järjestustunnuse mudelid*. Juhendaja lektor E. Käärik.

2. Kristiina Rajaleid. *Dispersiooni hindamine Pareto valiku korral*. Juhendaja dotsent I. Traat.
3. Jelena Sõrmus. *Mittetäielike järjestuste korrespondentsanalüüs*. Juhendaja professor K. Pärna.

Rakendusstatistika eriala

1. Kadri Ohvril. *Raseda kehamõõtude muutumise statistiline mudel ja lapse sünnikaalu prognoosimine*. Juhendajad emeriitprofessorid E.- M. Tiit, H. Kaarma.

TÜ STATISTIKAÜLIÕPILASTE AUHINNAD KONKURSSIDEL

2001

- Valeri Koort (magistrant). Eesti Teaduste Akadeemia 2001. a üliõpilastööde II auhind uurimustöö *Simulatsiooni- ja puu-meetoditest Euroopa tüüpi optsioonide hindade arvutamisel* eest.
- Kristiina Rajaleid (magistrant). *International Association of Survey Statisticians* 2001. a. Cochran-Hanseni preemia noorele statistikule uurimustöö *On the order sampling design* eest.
- Kristiina Rajaleid (magistrant). *Videomat Casino Group*'i preemia edukuse eest õpingutes ja teadustöös.

2002

- Mihhail Juhkam (magistrant). Üliõpilaste teadustööde 2002. a riiklikul konkursil II preemia teadustöö *Valimi värvilisus ja katvus multinomiaalse ning Poissoni skeemi korral* eest.
- Anu Roos (magistrant). Üliõpilaste teadustööde 2002. a riiklikul konkursil III preemia teadustöö *Mitmemõõtmeline asümmeetriline normaaljaotus* eest.

Kuidas toetavad statistikameetodid laboratoorsete andmete kasutamist diagnoosimisel?

Kuidas on Eestis seni õnnestunud neerusiirdamised, missugused on eluväljavaated siiratud neeruga patsientidel?

Mis põhjustab koolijuhtide stressi ja "läbipõlemist" Eestis ja kas samad probleemid kimbutavad ka Eesti koolijuhtide ametivendi välismaal?

Kuidas kirjeldada mudeli abil naiste sündimuskäitumist?

Kuidas õnnestub edukalt modelleerida metsakasvu ja klassifitseerida ja tüpiseerida maastikku aero- ja satelliitfotode järgi?

Kõik need küsimused kuuluvad eluteaduste valdkonda ja nendele vastamisel kasutatakse statistikameetodeid. Statistikameetodite kasutamist eluteadustes arutati Otepää külje all Margusel toimunud Eesti Statistikaseltsi konverentsil 1. ja 2.novembril 2002 ning käesolev raamatuke sisaldab valikut esitatud ettekannetest.