

Koolkonnad statistikas

Biostatistika probleemast

Krista Fischer

TÜ Matemaatilise statistika professor

Frekventistlik lähenemine: tõenäosus

- Tõenäosused on vaadeldavad suhteliste sagedustena (teoreetiliselt lõpmatus üldkogumis). Kui meil on vaja leida sündmuse A (nt täringuviskel “6” saamise) tõenäosust, siis seda saab hinnata kui teeme n katset (n viiset) ja seal on k -l korral sündmus A toimunud:

$$P(A) \approx \frac{k}{n}$$

ja kui n suurendada, siis

$$P(A) = \lim_{n \rightarrow \infty} \frac{k}{n}$$

Bayesi lähenemine: tõenäosus

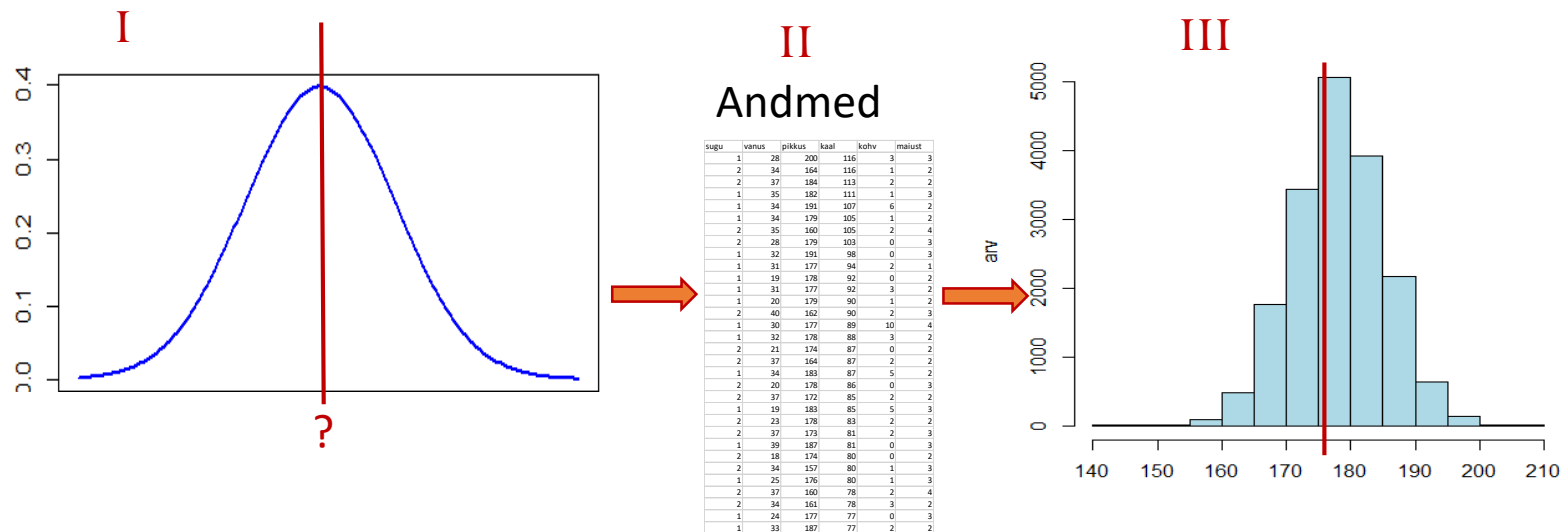
- Nn Bayesi koolkonna esindajad defineerivad tõenäosuse kui arvulise hinnangu sündmuse tõenäolisusele – realistlikule ootusele, vastavalt hetketeadmistele
- Bayesi vaatepunktist on igal hüpoteesil teatud tõenäosus – frekventistlik lähenemine seisneb aga hüpoteeside testimises ilma eelteadmisteta tõenäosuse kohta
- Bayesi andmeanalüüsi idee seisneb selles, et andmed täpsustavad eelteadmisi – nt eelnevalt paika pandud hinnangut/olemasolevaid teadmisi hüpoteesi tõenäosuse kohta

Frekventistlik lähenemine: mis on juhuslik ja mis on fikseeritud?

- Eeldame et andmed pärinevad üldkogumist jaotusega F ja parameetri(te)ga θ :

$$X \sim F(\theta)$$

- Enne andmete analüüsimist tehakse küll eeldus F kohta, kuid kogu info θ kohta saadakse kogutud andmetest – st hinnatakse parameeter θ



Frekventistlik lähenemine: mis on juhuslik ja mis on fikseeritud?

- Eeldame et andmed pärinevad üldkogumist jaotusega F ja parameetri(te)ga θ :

$$X \sim F(\theta)$$

- Enne andmete analüüsimist tehakse küll eeldus F kohta, kuid kogu info θ kohta saadakse kogutud andmetest – st hinnatakse parameeter θ
- Eeldame, et üldkogumi parameetrid on fikseeritud, kuid valimi põhjal leitud hinnang $\hat{\theta}$ on varieeruv (teoreetiliselt saaks samast üldkogumist lõputult valimeid võtta ja iga kord saaks erineva $\hat{\theta}$)

Bayesi lähenemine

- Idee: meil on uuritava probleemi kohta olemas teatud eelteadmised. Andmed täpsustavad (või parandavad) meie eelteadmisi.
- Andmed on fikseeritud, parameetrid on juhuslikud
- Enne analüüsi tehakse eeldus parameetrite jaotuse kohta – nn eeljaotus („prior“)
- Analüüsi käigus leitakse nn järeljaotus („posterior“), kombineerides infot eeljaotuse ja andmete kohta:



Thomas Bayes (1701-1761)

Bayesi valem ja Bayesi statistika

Bayesi valem:

Sündmuse A tõenäosus
tingimusel, et on
toimunud sündmus B

Sündmuse B tõenäosus
tingimusel, et on
toimunud sündmus A

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Kasutame seda, et leida parameetri eeldatav jaotus meie nähtud andmete põhjal:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} = \boxed{P(\theta)} \cdot \frac{P(X|\theta)}{P(X)}$$

Posterior = **Prior** x Likelihood

- Kui andmeid saadakse juurde, leitakse uus järeljaotus, kasutades nüüd varasemalt saadud järeljaotust eeljaotusena

Bayesi lähenemine: mis on juhuslik ja mis on fikseeritud?

- Eeldame et andmed pärinevad üldkogumist jaotusega F ja parameetri(te)ga θ :

$$X \sim F(\theta)$$

- Samal ajal eeldatakse, et ka θ on juhuslik – nt jaotusega

$$\theta \sim G(\alpha),$$

kusjuures θ jaotuse parameetreid α nimetatakse hüperparameetriteks.

- Enne andmeanalüüsi on paika pandud parameetrite **eeljaotus** (*prior distribution*) $G(\theta|\alpha)$, andmeanalüüsi käigus leitakse **järeljaotus** (*posterior distribution*) $G(\theta|\alpha; \mathbf{X})$, kus $\mathbf{X}$ tähistab analüüsitavaid andmeid.

Üks vana näide (R. Fisher)

- Tädi väitis, et ta suudab teed juues maitse järgi eristada, kumb pandi tassi enne, kas suhkur või tee
- Tehti katse, kus ta pidi otsustama seda 20 tassist maitsemise põhjal
- Tädi arvas 15 tassi puhul õigesti
- Frekventistlik lähenemine:
 - $H_0: p=0,5$
 - Leiame p-väärtuse: $P(X \geq 15 \mid p=0,5) = 0,02$
(Kui tädi otsus oleks n-ö samaväärne mündiviskega, st täiesti juhuslik, siis sellise tulemuse saamise tõenäosus on ca 2%)
 - Seega saaks otsustada, et tädi tõepoolest suudab eristada seda, kas suhkur pandi tassi enne või pärast tee valamist

Üks vana näide - järg

Bayesi meetodil lahendamine, kasutades šansisuhet:

(šansid = $p/(1-p)$)

Ehk siis:

$$\frac{P(T|X)}{P(-T|X)} = \frac{P(T)}{P(-T)} \frac{P(X|T)}{P(X|-T)},$$

kus T ja $-T$ tähistavad kahte alternatiivset hüpoteesi. Näiteks võiksime eeldada, et $T : p = 0,5$ ja $-T : p = 0,75$. Edasi saame arvutada:

$$\frac{P(p = 0,5|X = 15)}{P(p = 0,75|X = 15)} = \frac{P(p = 0,5)}{P(p = 0,75)} \frac{P(X = 15|p = 0,5)}{P(X = 15|p = 0,75)} \approx \frac{P(p = 0,5)}{P(p = 0,75)} \cdot \frac{1}{13,7}$$

Seega sõltub vastus valitud eeljaotusest: kui palju on meie arvates $p=0,5$ tõenäolisem kui $p=0,75$. Näiteks kui me eeldame, et 99 korda tõenäolisem (st tõenäosus 1%) on, et tädil sellist võimet pole, siis ülaltoodud võrduse vastus on $99/13,7$ ehk ligikaudu 7,2 (mille põhjal võib hinnata ka, et tõenäosus, et tädil on see võime, on ca 12%).

Bayesi statistika keerulisemates olukordades

Aposterioorse tõepära (järeljaotuse) leidmine matemaatiliselt:

$$P(\theta|\mathbf{X}) = \frac{P(\theta)P(\mathbf{X}|\theta)}{\int P(\theta)P(\mathbf{X}|\theta)d\theta}$$

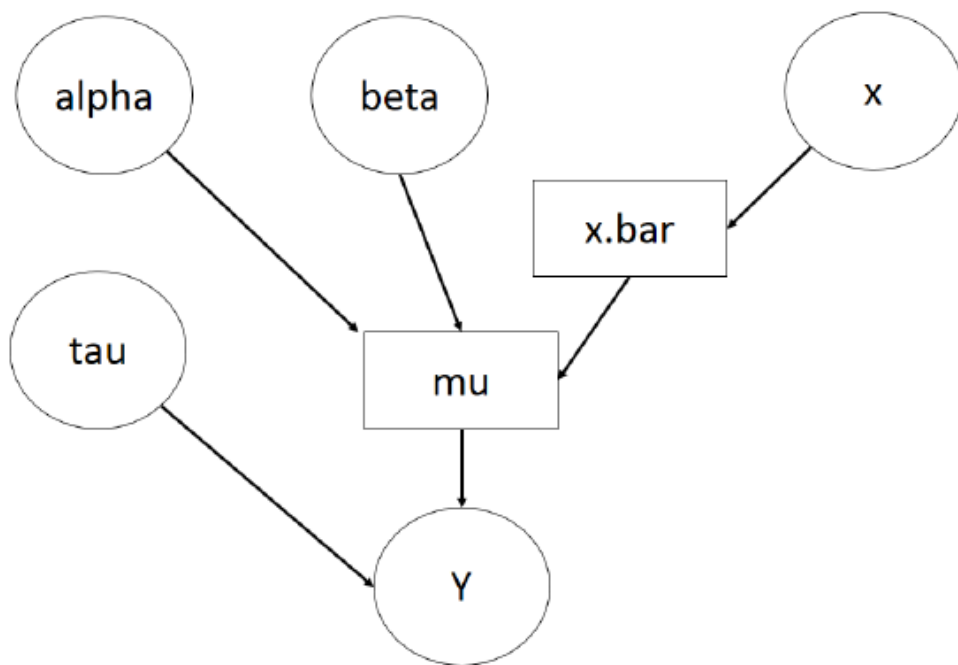
... võib osutuda väga keerukaks.

Seetõttu on läbi ajaloo Bayesi statistikat siiski üsna vähe kasutatud ...
kuni algas arvutite võidukäik

Bayesi statistika arvutite ajastul

- Bayesi statistika on tänapäeval siiski üsna populaarne, seda tänu nn MCMC-meetoditele (Markov Chain Monte Carlo)
- Idee seisneb selles, et järeldaotust ei tuletata analüütiliselt, vaid tema leidmiseks kasutatakse juhuslike arvude generaatorit
- Väga jämedalt:
 - Esimesel sammul võetakse üks parameetri algväärtus x_0
 - Teisel sammul võetakse järgmine juhuslik väärtus x_1 (tihti x_0 lähedalt) ja hinnatakse, kas selle väärtuse korral on olemasolevate andmete tõepära suurem kui algväärtuse korral
 - Tõepärasuhte kaudu määratakse tõenäosused uue väärtuse vastuvõtmiseks või tagasilükkamiseks ja järgmiseks algväärtuseks saab kas x_1 või x_0 , olenevalt sellest kas x_1 võeti vastu või mitte
 - Ja nii jätkatakse kuni algoritmi koondumiseni

Tihti kasutatakse mudeli graafilist esitust:

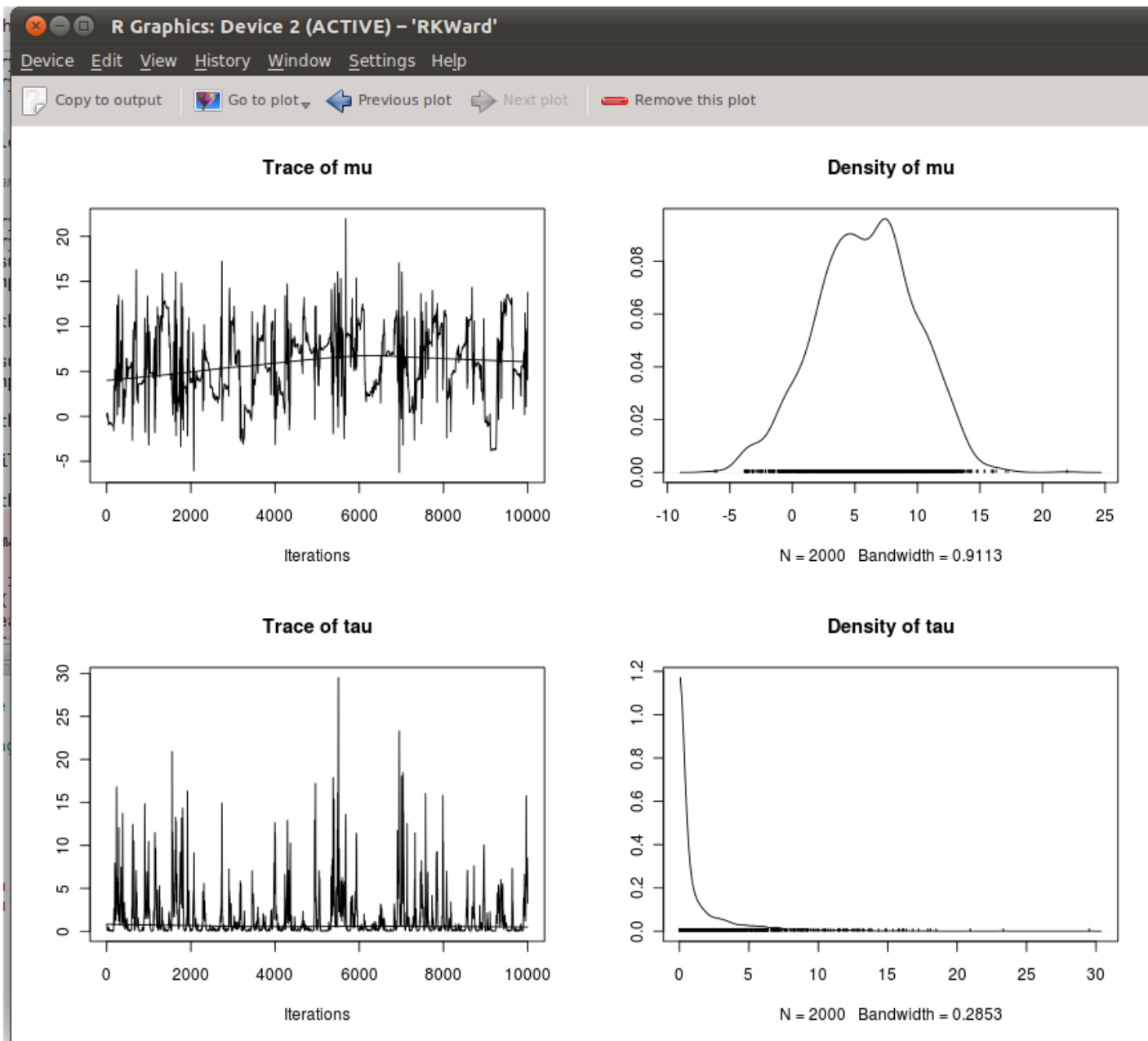


$$Y_i = \alpha + X_i\beta + \epsilon_i,$$

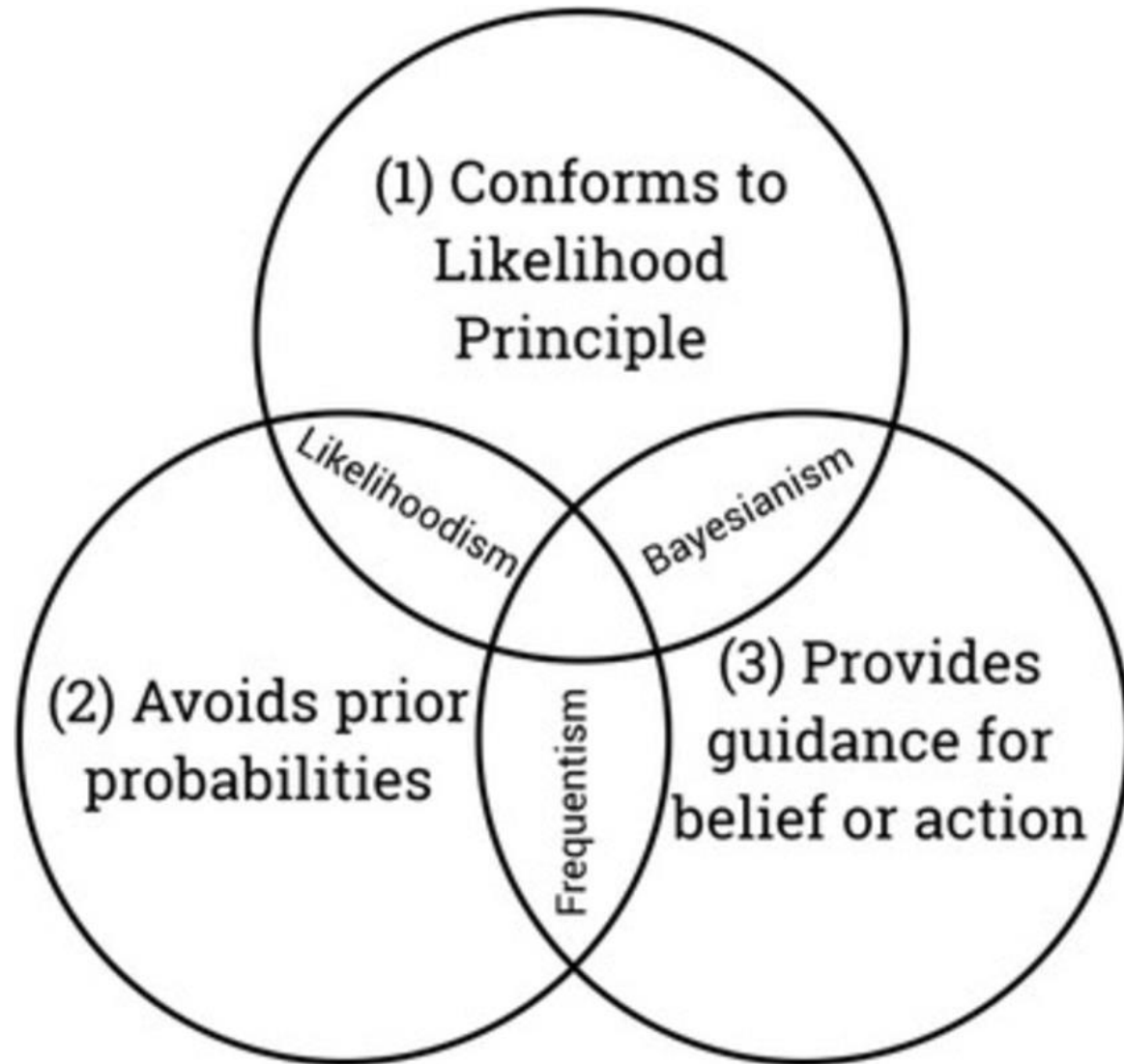
$$\text{kus } 1/\text{Var}(\epsilon_i) = \tau$$

$$\text{ja } E(Y_i|X_i) = \mu$$

Jälgitakse protsessi....



Veel koolkondadest (ka „Likelihoodism“ on olemas)



Koolkondi on tegelikult veelgi rohkem (nt „Fiducial inference“), kuid teised on siiski üsna marginaalse tähtsusega – põhiline jagunemine on siiski Bayesi ja frekventistliku suuna vahel

Koolkonnad ja tulevik – kas jäädaksegi vaidlema?

- Arvutite ajastul on üha enam näha ka mitme koolkonna segunemist
- Frekventistlikke ja Bayesi lähenemisi ühendatakse ja kombineeritakse vastavalt sellele, mis on mingi projekti puhul optimaalseim
- Lisaks on oluliseks muutunud ka hüpoteesi-, mudeli- ja jaotusevabad meetodid – nt andmekaeve. Siin võib rääkida mitte ainult erinevate statistika koolkondade kombineerimisest, vaid ja statistika ja informaatika piirimaile olevatest meetoditest

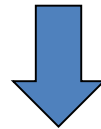
Statistika ja eluteadused ehk **biostatistika**

Mis asjad on eluteadused?

- Bioloogia
- Meditsiin
- Põllumajandusteadused
- Metsandus
- Kalandus
- jne

Tüüpiline teadusprojekt eluteadustes

1. Uurimisküsimuse püstitamine (hüpoteesid)
2. Uuringuplaani koostamine
3. Andmete kogumine
4. Andmeanalüüs
5. Analüüsitulemuste tõlgendamine, järeldused uurimisküsimuse kohta, publitseerimine



Iga etapi juures on vaja teadmisi statistikast!

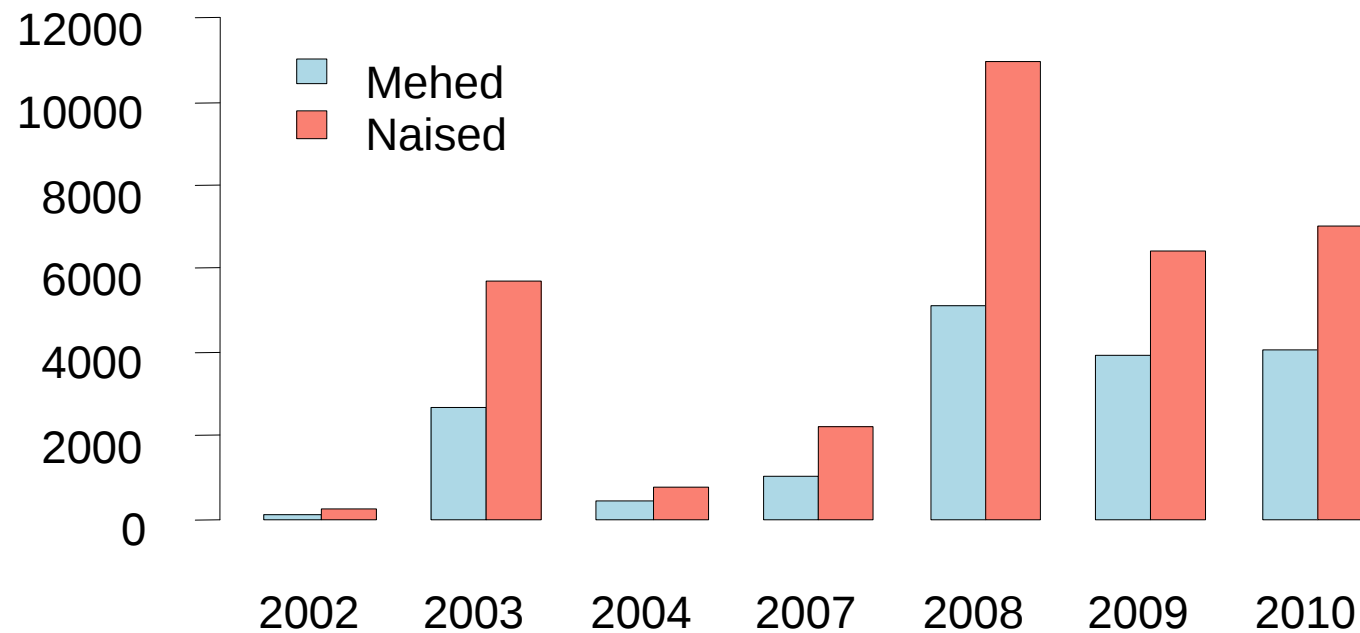
...aga kas ka statistikuid?

Biostatistika ja karjäärivõimalused

- Rahvatervis – epidemioloogia (TAI, TÜ Rahvatervise instituut, ...)
- Uuringud kliinilises meditsiinis (kliinikud)
- Ravimuuringud (StatFinn, Pasi Korhonen)
- Geeniteadus, biotehnoloogia (TÜ Genoomika instituut, TÜ Eesti Geenivaramu)
- Põllumajanduse temaatika, sh ka loomade aretus/geneetika (EMÜ, Tanel Kaart)
- Metsandus
- Zooloogia, botaanika – teadusasutused

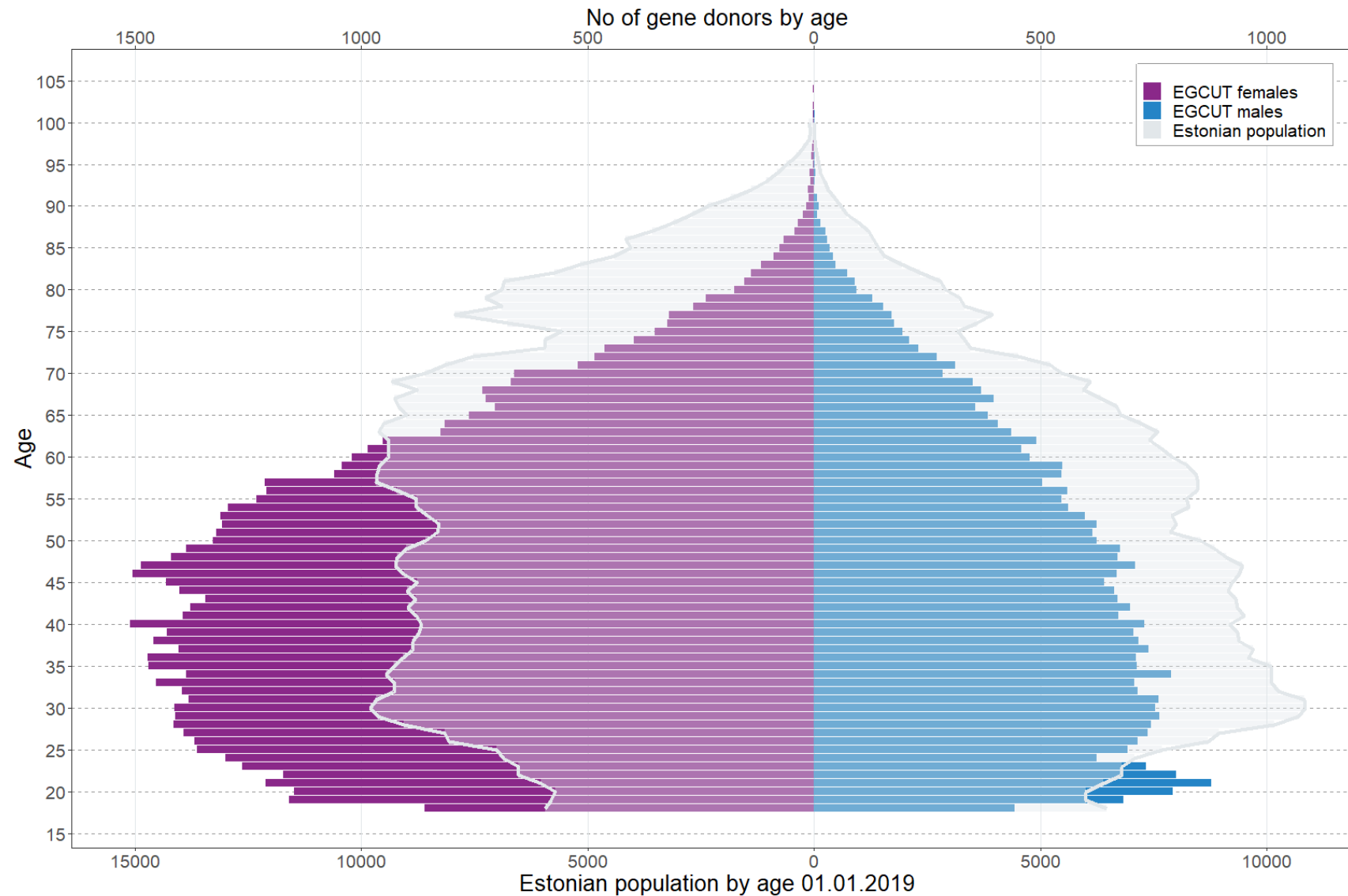
Tartu Ülikooli Eesti Geenivaramu

- Enam kui 51000 geenidoonorit, liitunud 2002-2010
- Vanus liitumisel: 18-103
- Liitunud aasta kaupa:



Geenidoonorid vs Eesti rahvastik

Representative sample of Estonians



Total adults: 1 070 375
Total adult gene donors: 104 317
9.75 %
Age at the date of the agreement
Date: 2019-06-17

Geenidoonorite kohordi jälgimine ajas

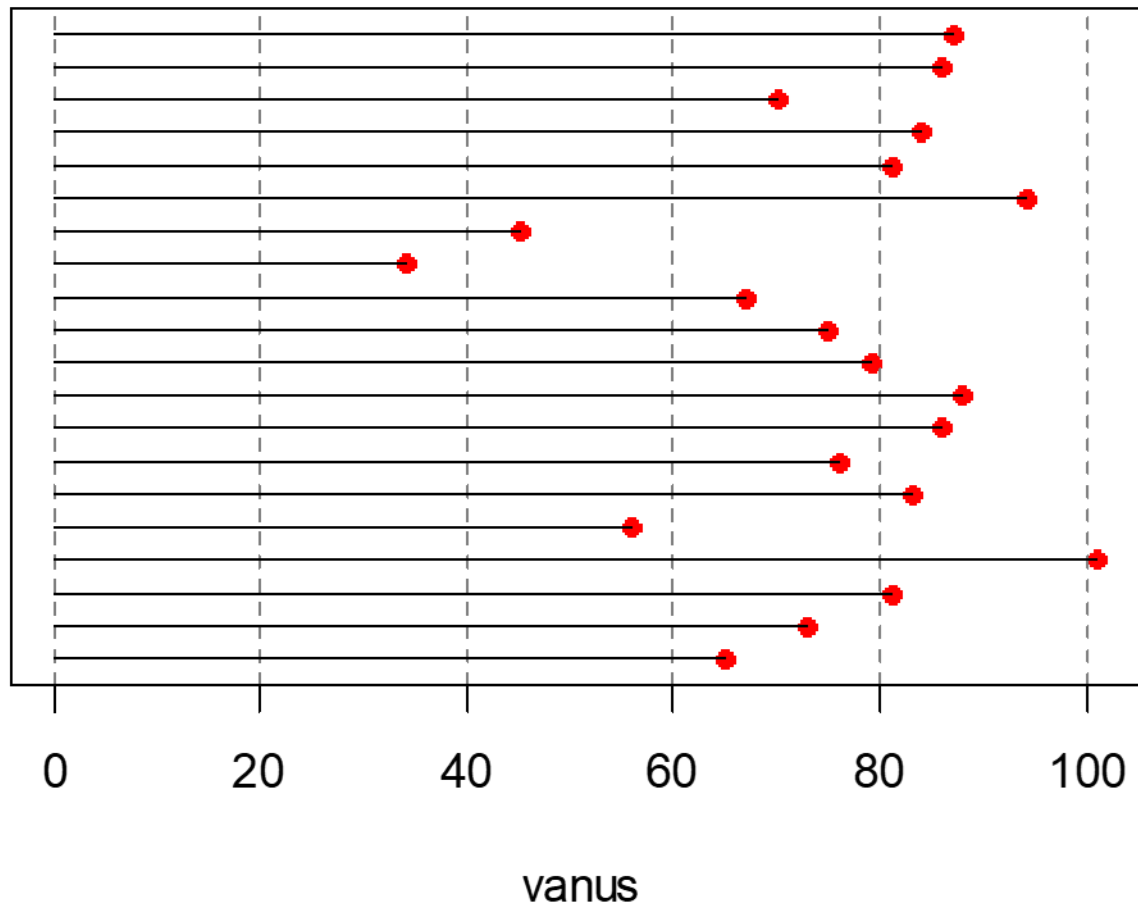
Regulaarsed linkimised riiklike registritega jm andmebaasidega:

- Rahvastikuregister
- **Surma põhjuste register**
- Vähiregister
- + haigekassa andmebaas, e-tervis, jne

Ligi 5000 geenidoonorit (veidi alla 10% neist, kes liitusid enne aastat 2014) on surnud aprilliks 2019

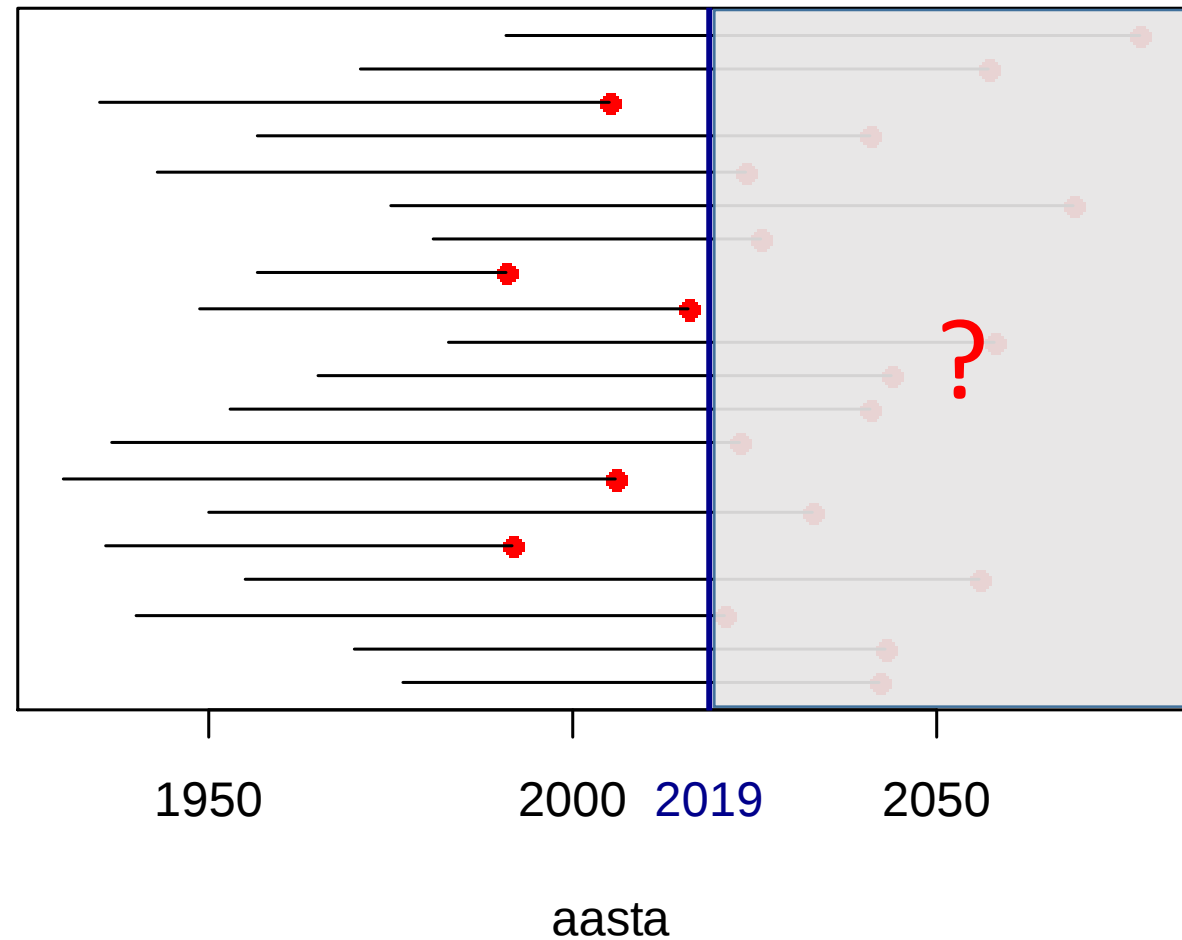
Kuidas analüüsida inimese elu pikkust?

20 inimese eluead sünnist surmani

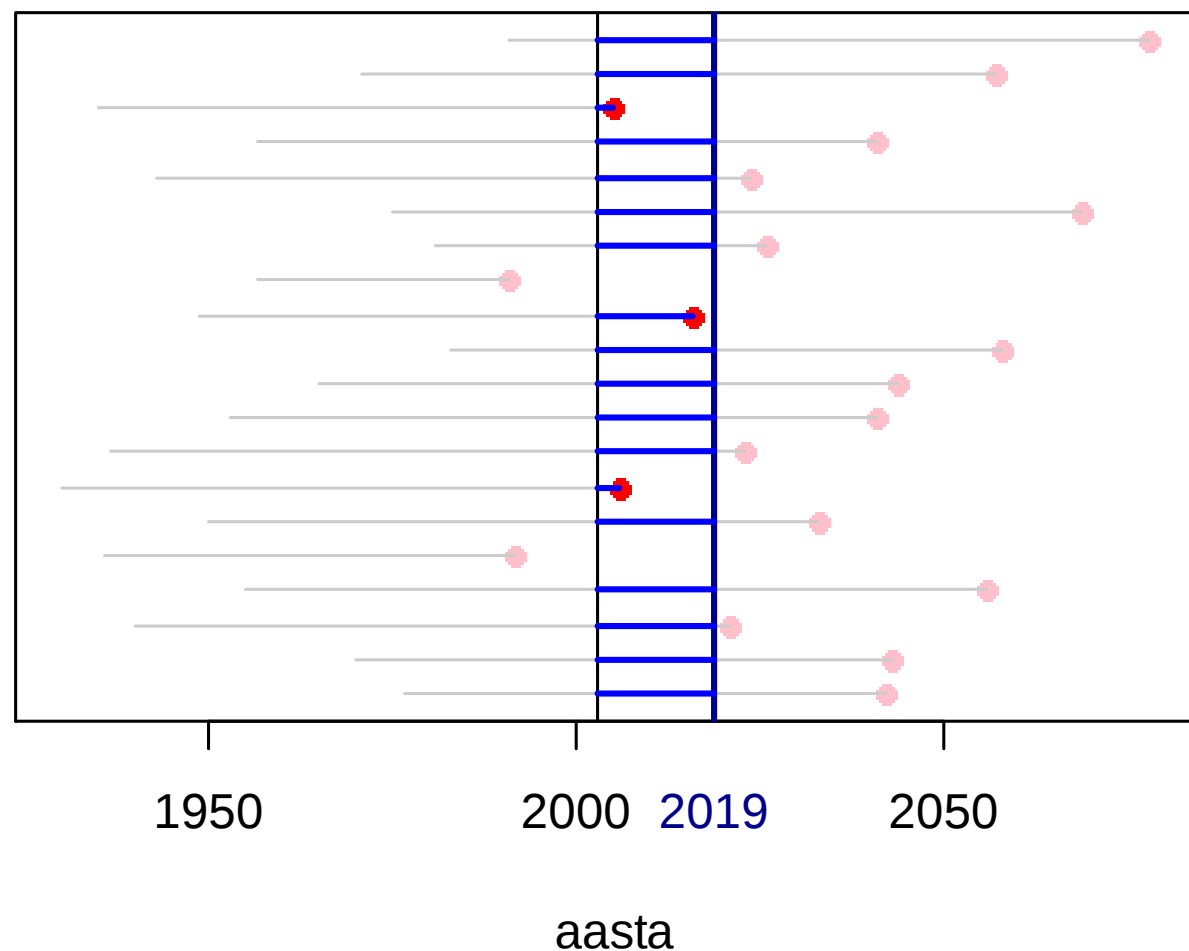


Mis siin keerulist on? (arvud on arvud, keskmist saame ju arvutada?)

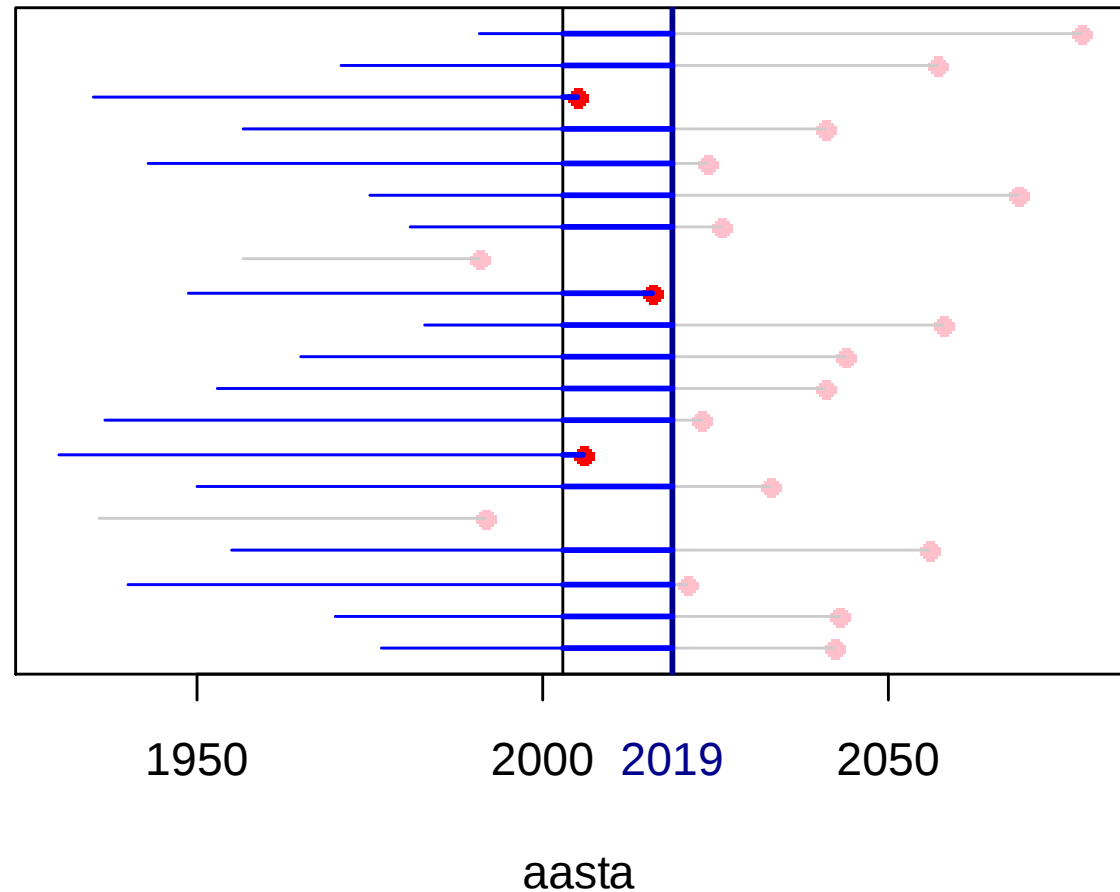
Probleem 1: täna elavate inimeste eluea saame me teada alles tulevikus



Uurides eluiga mõnes uuringukohordis (nt TÜ Eesti Geenivaramu andmebaasis), näeme vaid neid, kes on andmekogumise hetkeks elus

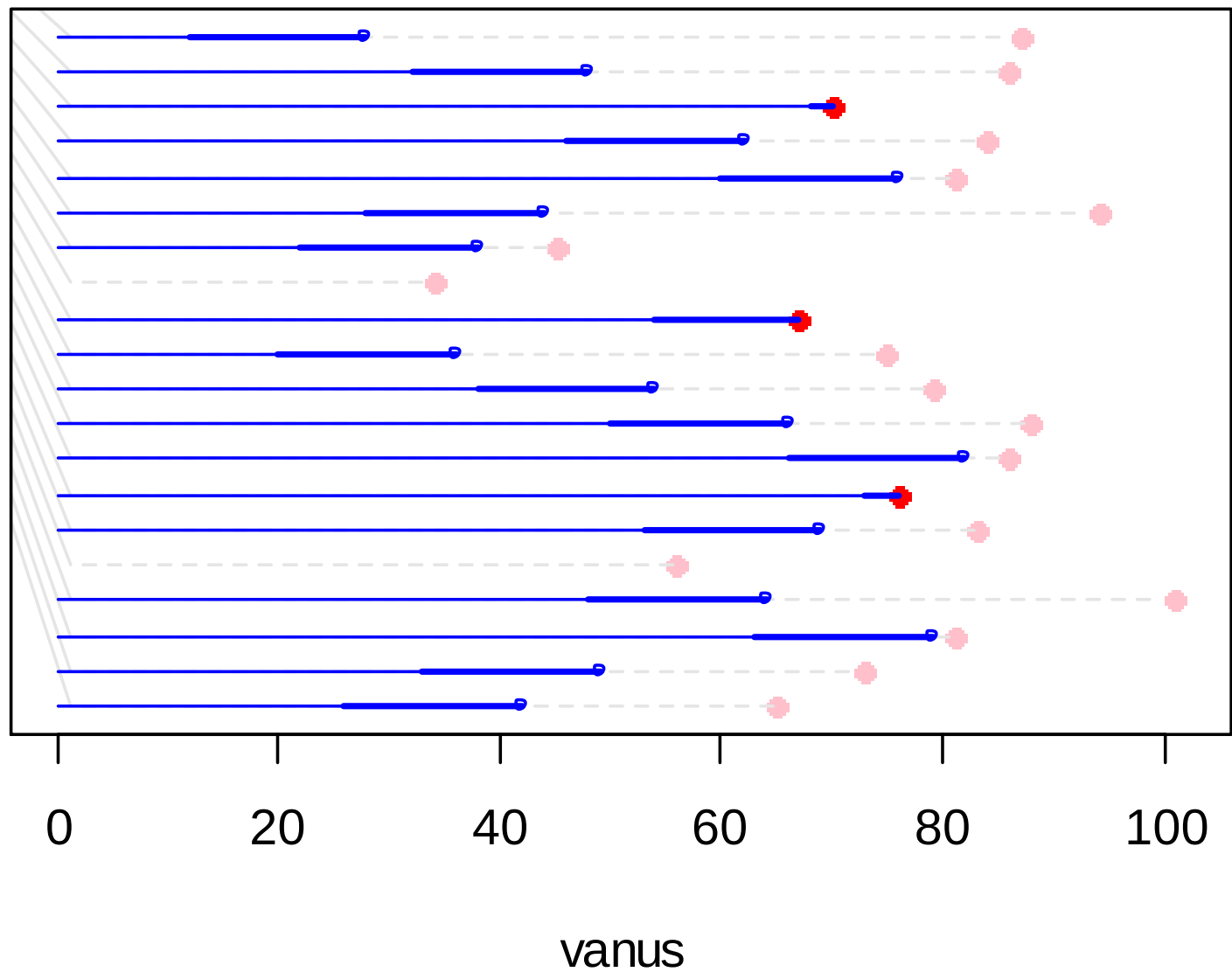


... nende kohta on meil andmeid ka tagasiulatuvalt

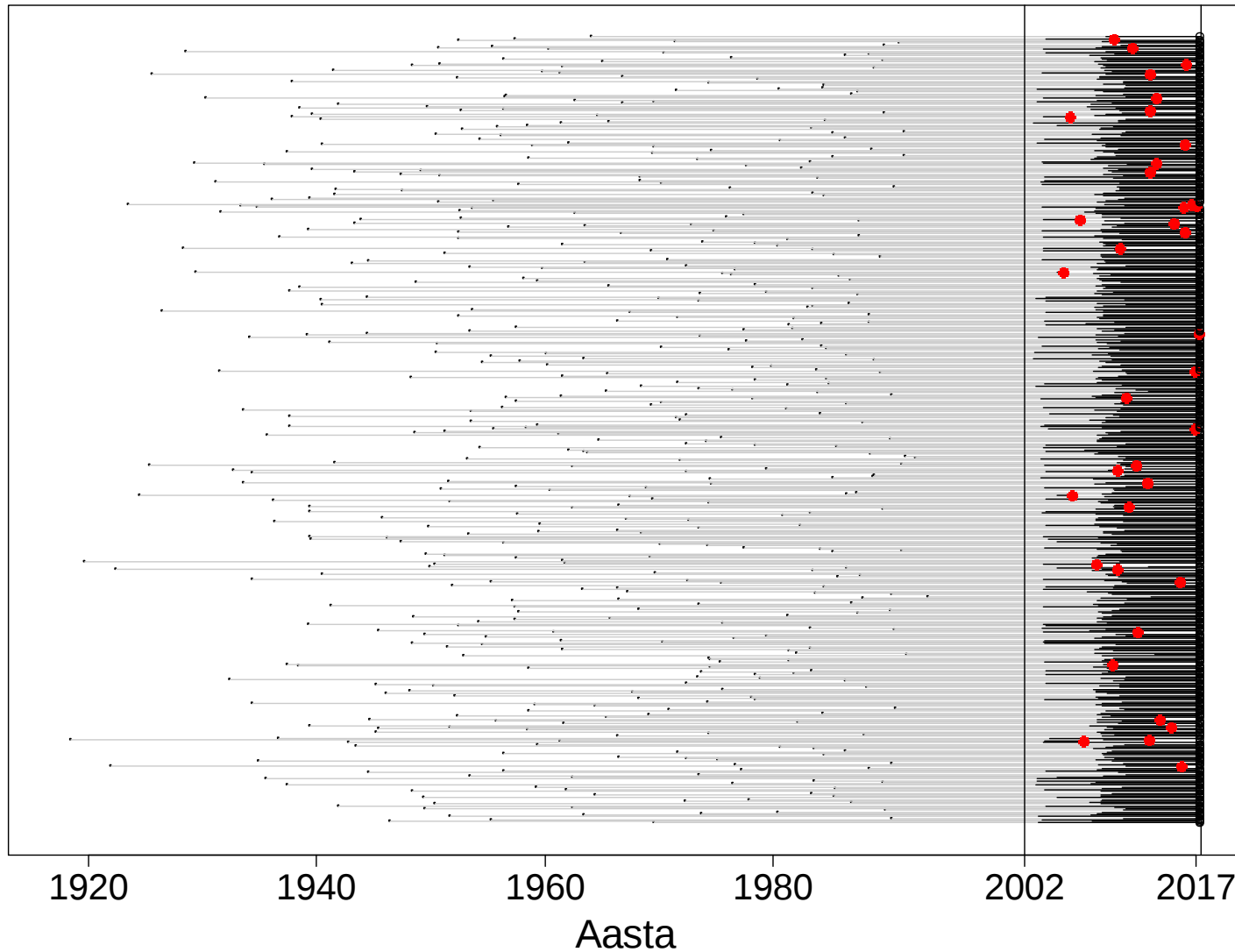


Selles näites oleks teada 20 inimesest vaid 3 täielik eluiga, ning 2 inimese kohta puuduksid andmed üldse.

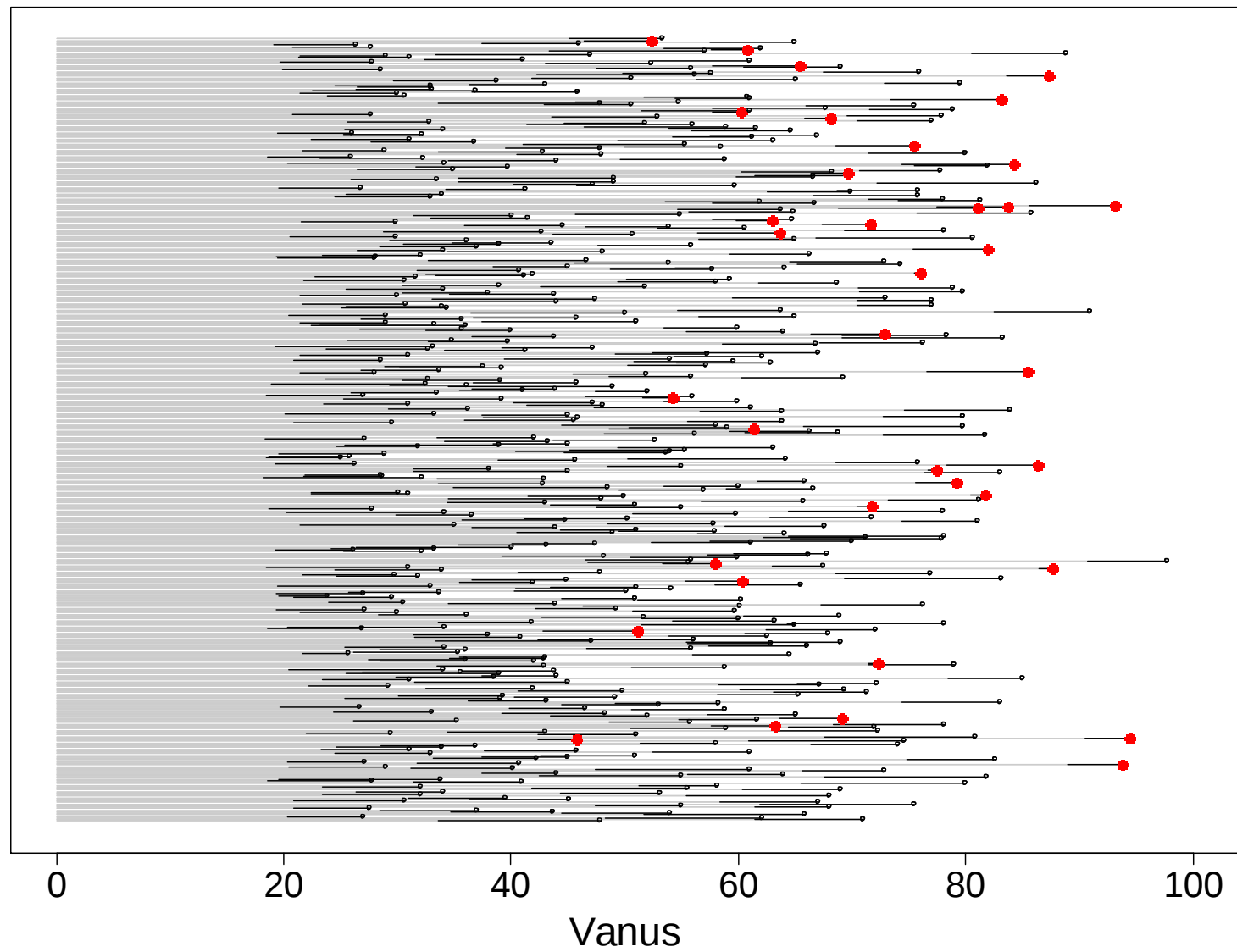
Vanuse skaalal:



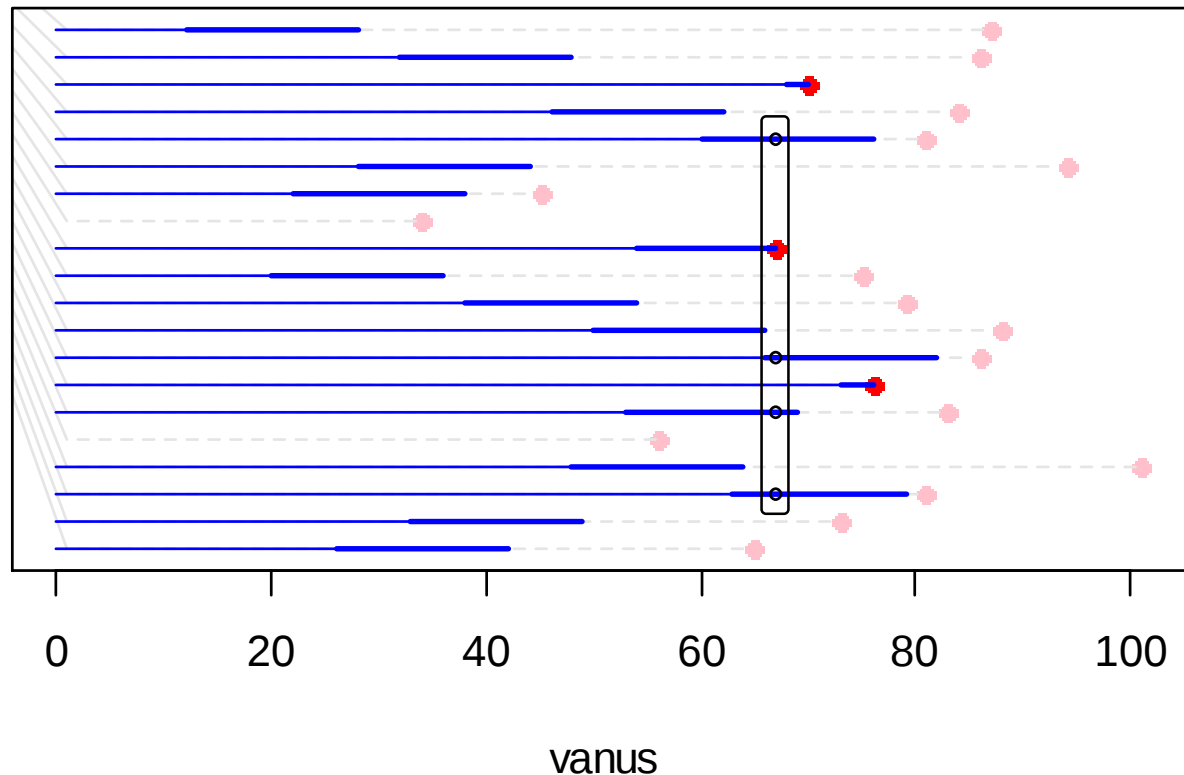
Sarnane olukord ka TÜ Eesti Geenivaramu andmestikus (joonisel vaid <1% kohordist)



Vanuse skaalal

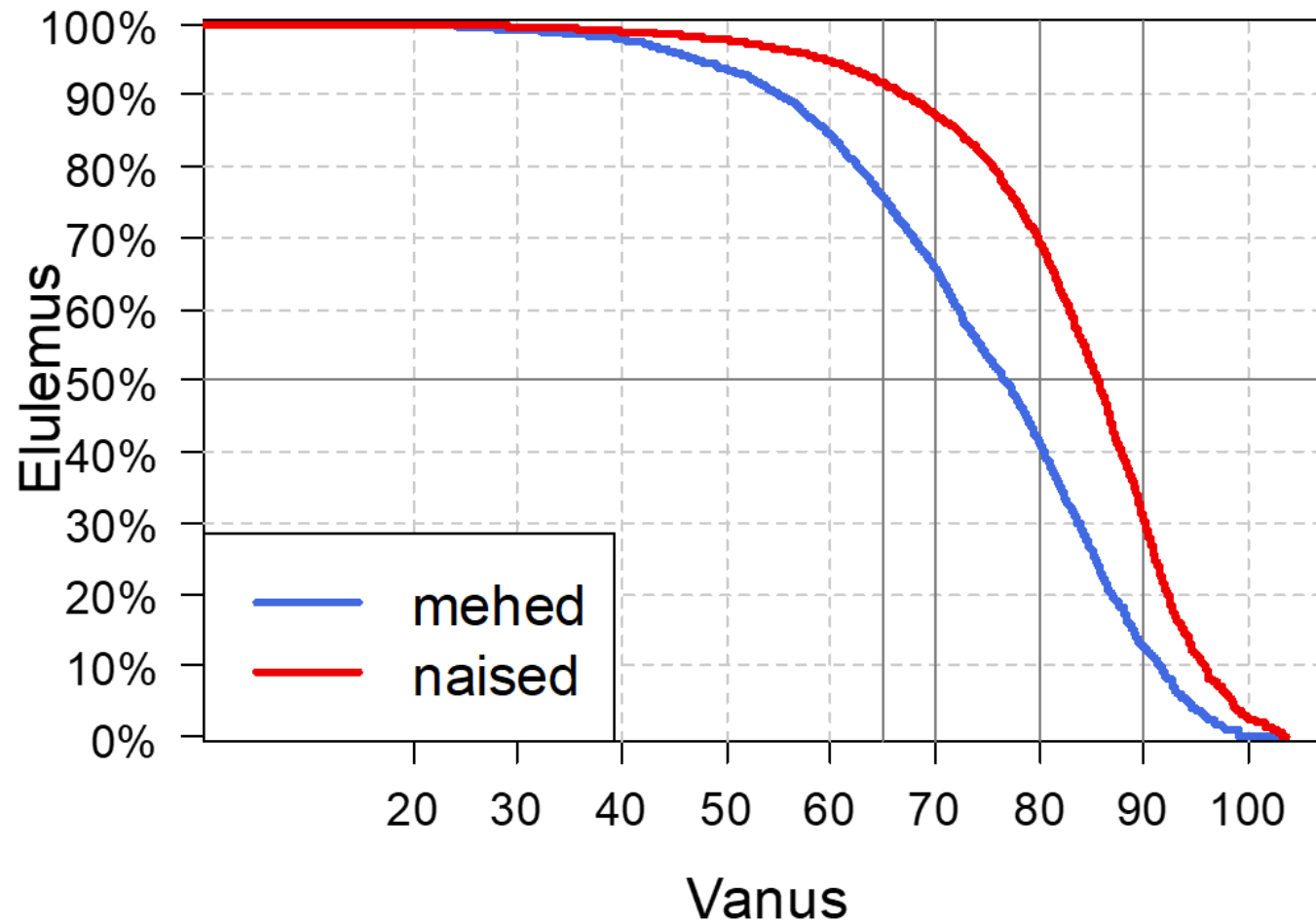


Kuidas selliseid andmeid analüüsida?



Läbiv idee: iga surnu andmeid võrreldakse nendega, kes samas vanuses veel **vaatluse all** olid (matemaatiliselt on selle põhimõttega leitud hinnangute tuletuskäigud väga huvitavad!)

Geenidoonorite elulemus (soo-spetsiifiline tõenäosus teatud vanuses elus olla)

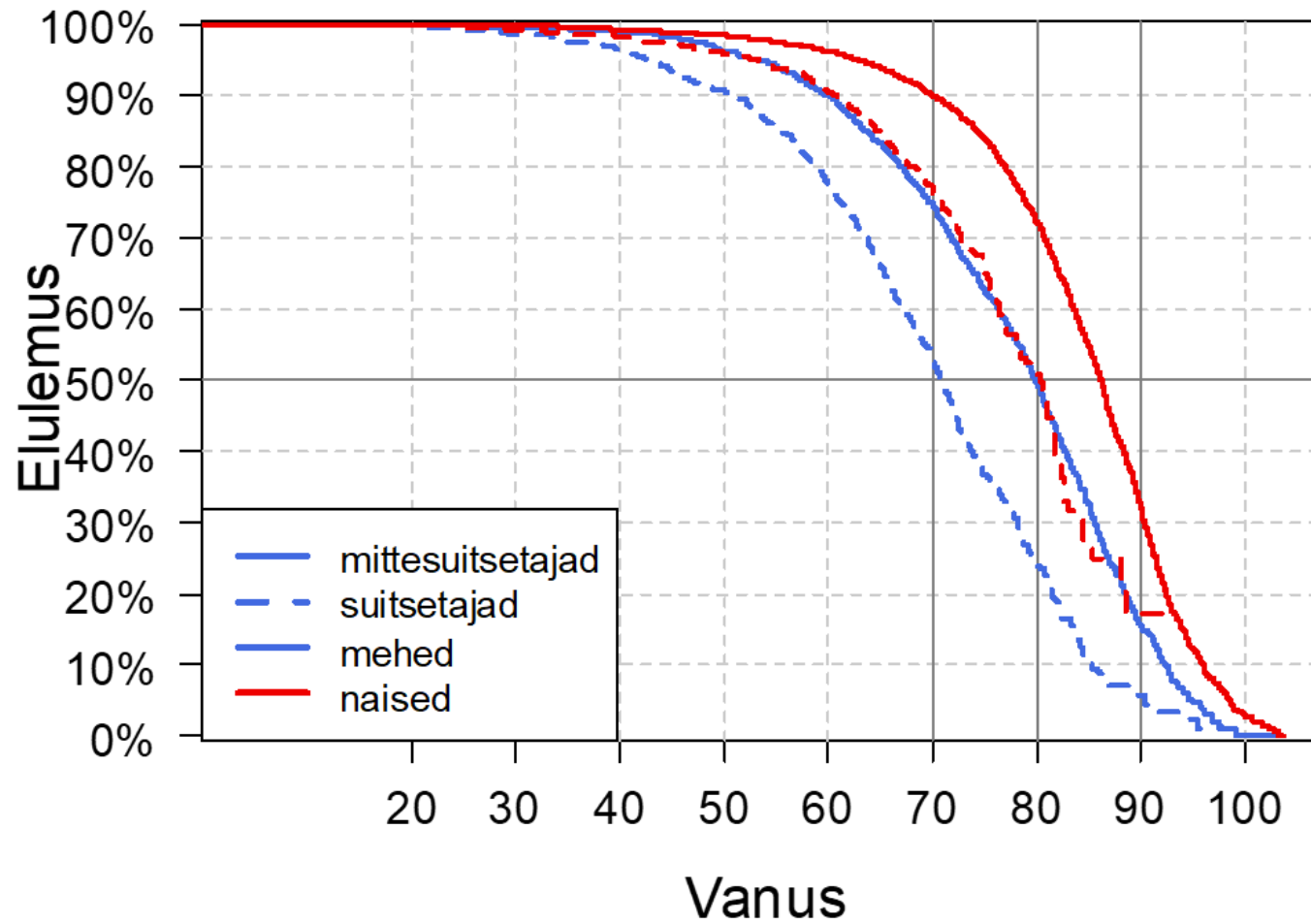


See ei ole päris sama, mis Eesti rahvastiku elulemus!

(kuid üllataval kombel on praktiliselt identne, kui arvestada vanus-spetsiifilist suremust aastal 2017)

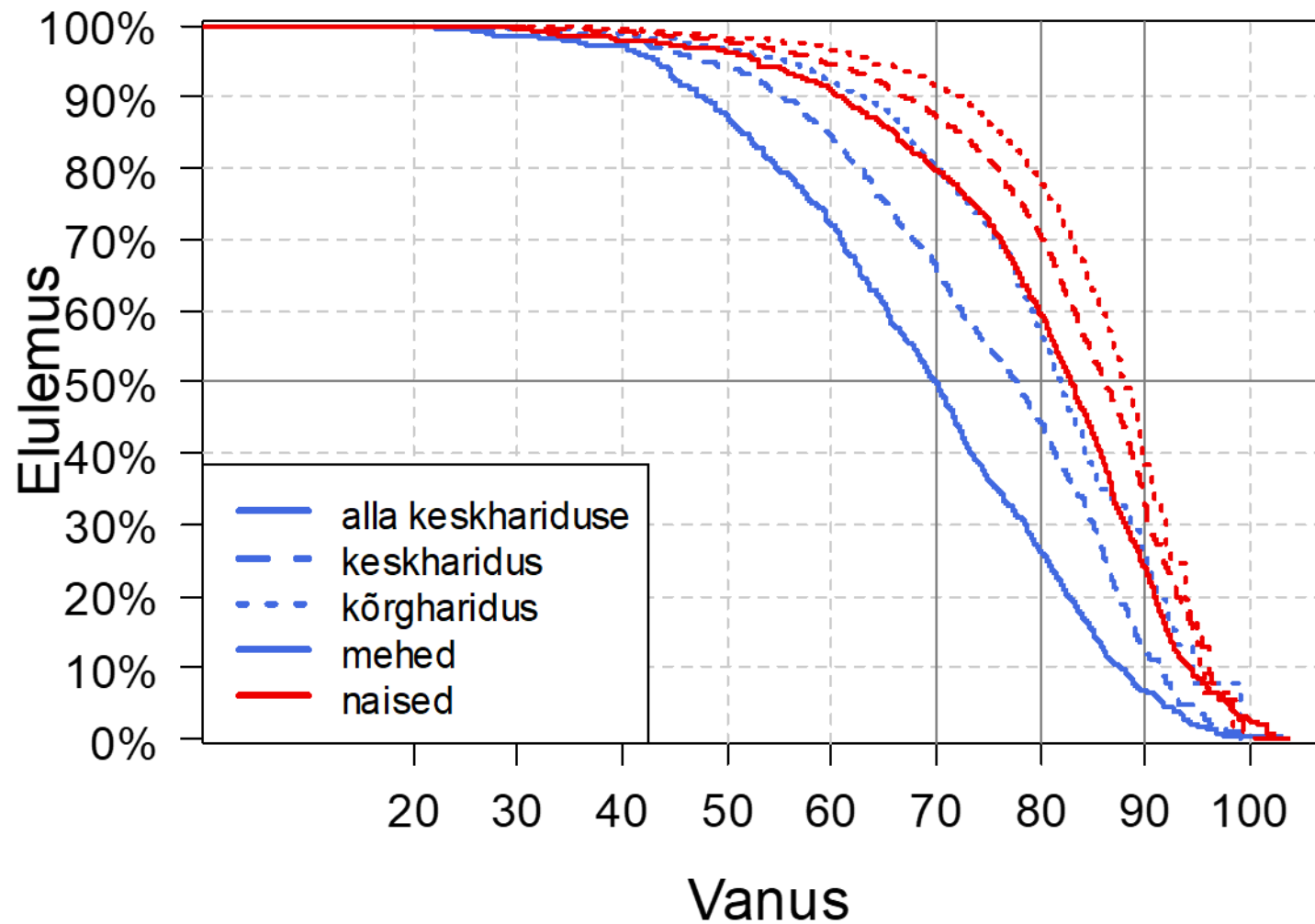
Mis mõjutab elukestust?

Suitsetamine?



Mis mõjutab elukestust?

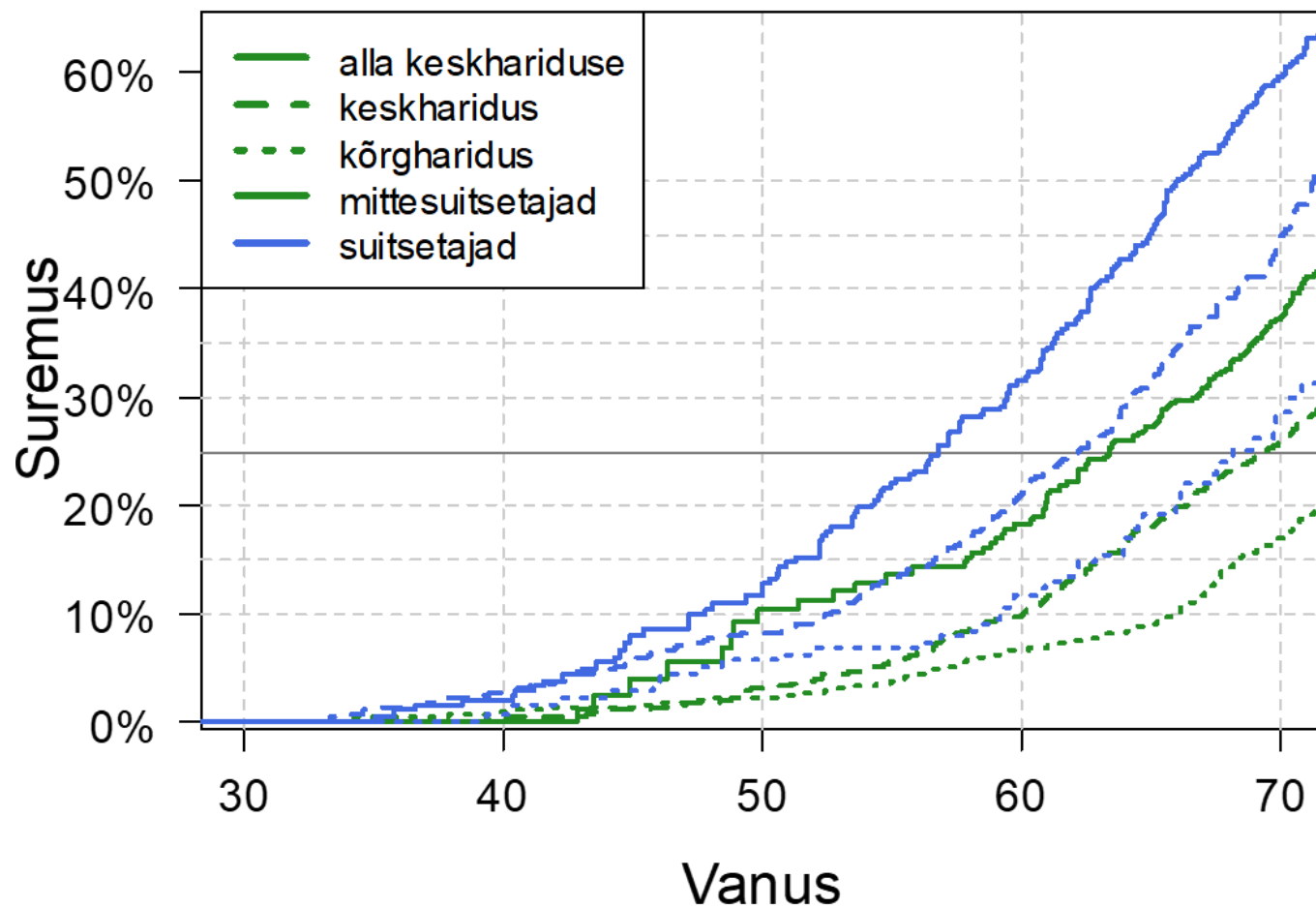
Haridus?



Mis mõjutab suremust alla 70-aastastel?

Haridus ja suitsetamine koos?

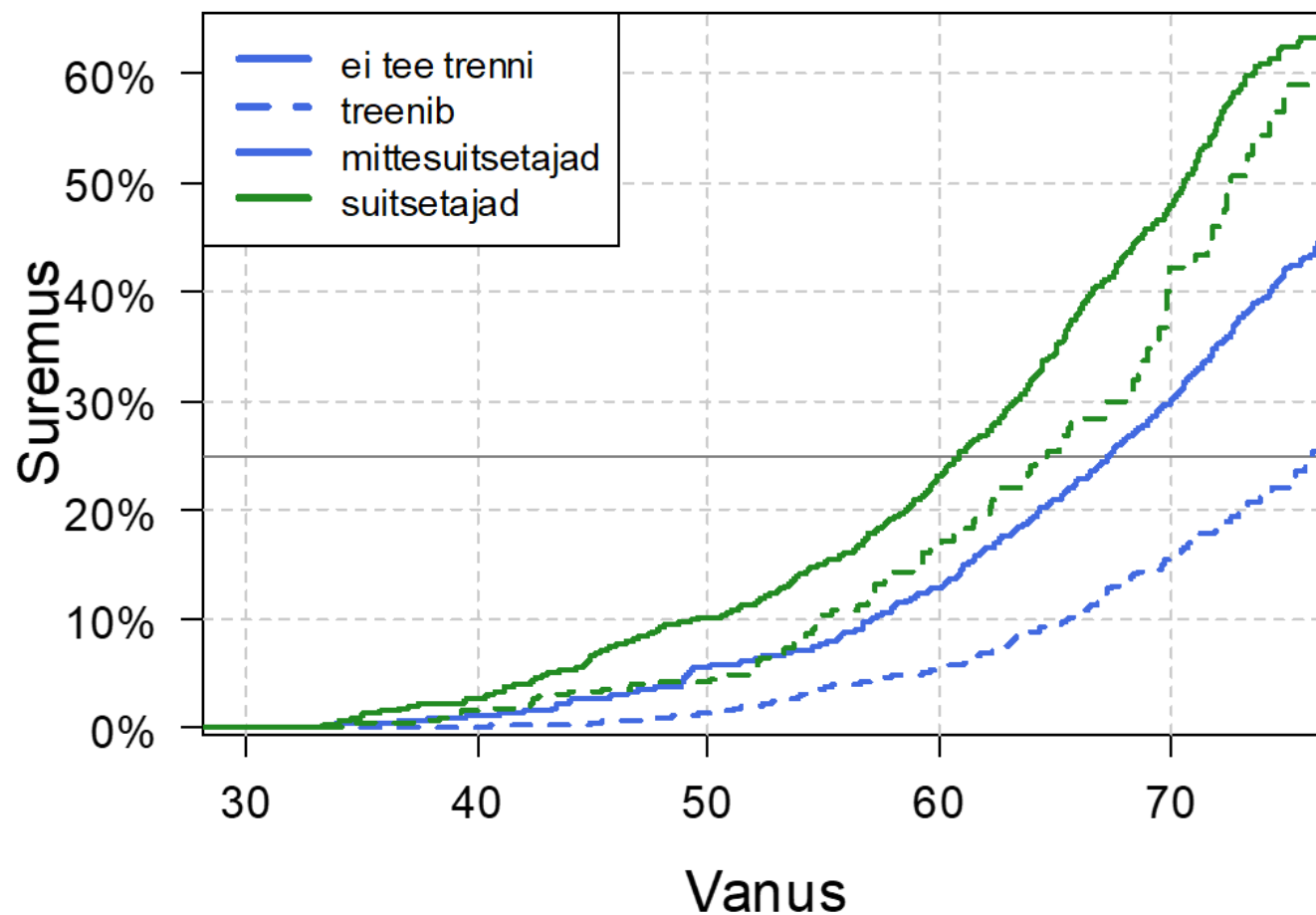
Mehed 30-70



Mis mõjutab suremust alla 70-aastastel meestel?

Sportimine vs suitsetamine?

Mehed 30-70



Mis mõjutab riski – kui kõik kokku panna? Kuidas seda teha?

- Vastus – mudelid
- Kõige levinum meetod: Coxi võrdeliste riskide mudel (Cox proportional hazard's model)



Sir David Cox
(Oxford, UK)

$$h_i(t) = h_0(t) \times \psi_{1i} \times \psi_{2i} \times \dots \times \psi_{ki}$$

Coxi mudeli hindamismetoodika – nn osalise tõepära meetod (ei kasutata tegelikke eluigasid, vaid tuletatakse hinnangud, võrreldes teatud vanuses surnute andmeid nende indiviidide andmetega, kes samas vanuses elus ja jälgimise all olid. Vastav artikkel avaldati ajakirjas „Journal of the Royal Statistical Society. Series B“ aastal 1972)

Mis mõjutab riski – kui kõik kokku panna (Coxi mudelisse)?

Hinnatud riskisuhted koos 95% usaldusvahemikuga

	Kõik naised	Kõik mehed	Noored mehed (30-70)
praegune suitsetamine	2.5 (2.2, 2.7)	2.4 (2.1, 2.8)	1.9 (1.7, 2.3)
endine suitsetamine	1.5 (1.4, 1.7)	1.4 (1.2, 1.6)	1.2 (1.0, 1.4)
ei treeni	1.5 (1.3, 1.7)	1.7 (1.4, 2.0)	1.4 (1.2, 1.7)
keskhariduse puudumine	1.7 (1.5, 2.0)	1.7 (1.5, 2.1)	2.1 (1.7, 2.6)
kõrghariduse puudumine	1.2 (1.1, 1.4)	1.3 (1.1, 1.5)	1.4 (1.2, 1.7)
ei söö värskaid puu- ega juurvilju	1.7 (1.4, 2.0)	1.4 (1.1, 1.9)	1.7 (1.3, 2.2)
alkohol ≥ 0.5 ühikut päevas	1.3 (1.2, 1.5)	1.6 (1.4, 1.8)	1.2 (1.1, 1.4)
rasvumine: KMI >35	1.3 (1.1, 1.4)	1.2 (1.0, 1.5)	1.5 (1.2, 1.8)
<u>Diagnoositud haigused:</u>			
vähk	1.8 (1.6, 2.0)	1.6 (1.4, 2.0)	2.1 (1.7, 2.6)
müokardiinfarkt	2.2 (2.0, 2.4)	2.0 (1.7, 2.2)	2.2 (1.9, 2.5)
I tüübi diabeet	2.1 (1.6, 2.7)	2.4 (1.7, 3.4)	2.4 (1.7, 3.4)

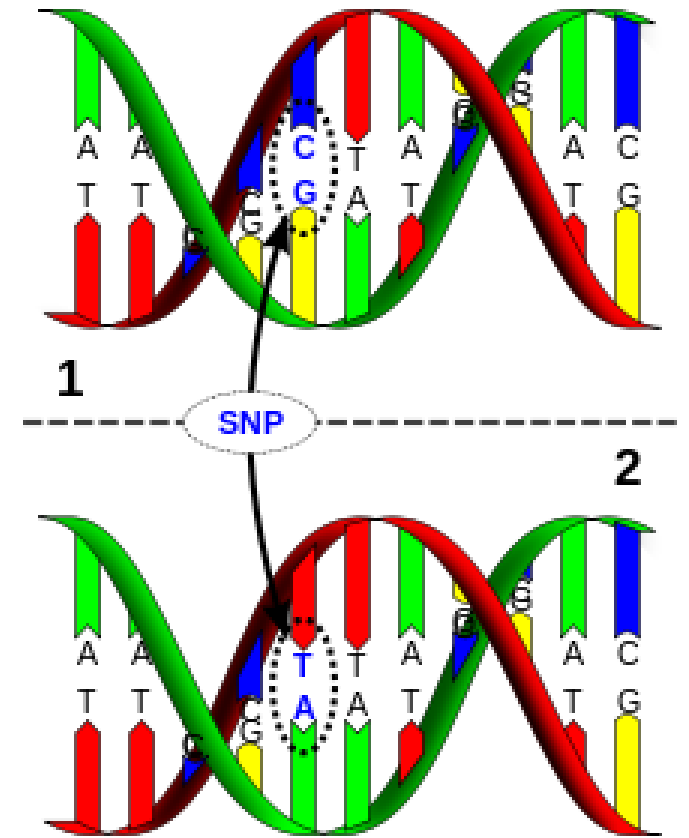
Siiski ei tea kas kõik seosed on põhjuslikud!

Geeniandmed: mil moel saavad 2 DNA-d erineda?

>99% osas on kõigi inimeste DNA-d identsed – inimestevahelised erinevused on määratud väikese varieeruva osa poolt.

Enim uuritud (sagedaseimad) geneetilised variandid:

- SNP-d (ühenukleotiidsed polümorfismid, *single nucleotide polymorphism*)
- SNP-de määramiseks kasutatakse **genotüpiseerimist** – see on oluliselt kiirem ja odavam kui täisgenoomi sekveneerimine



Tänaseks on genotüpiseeritud 150000 geenidoonori DNA

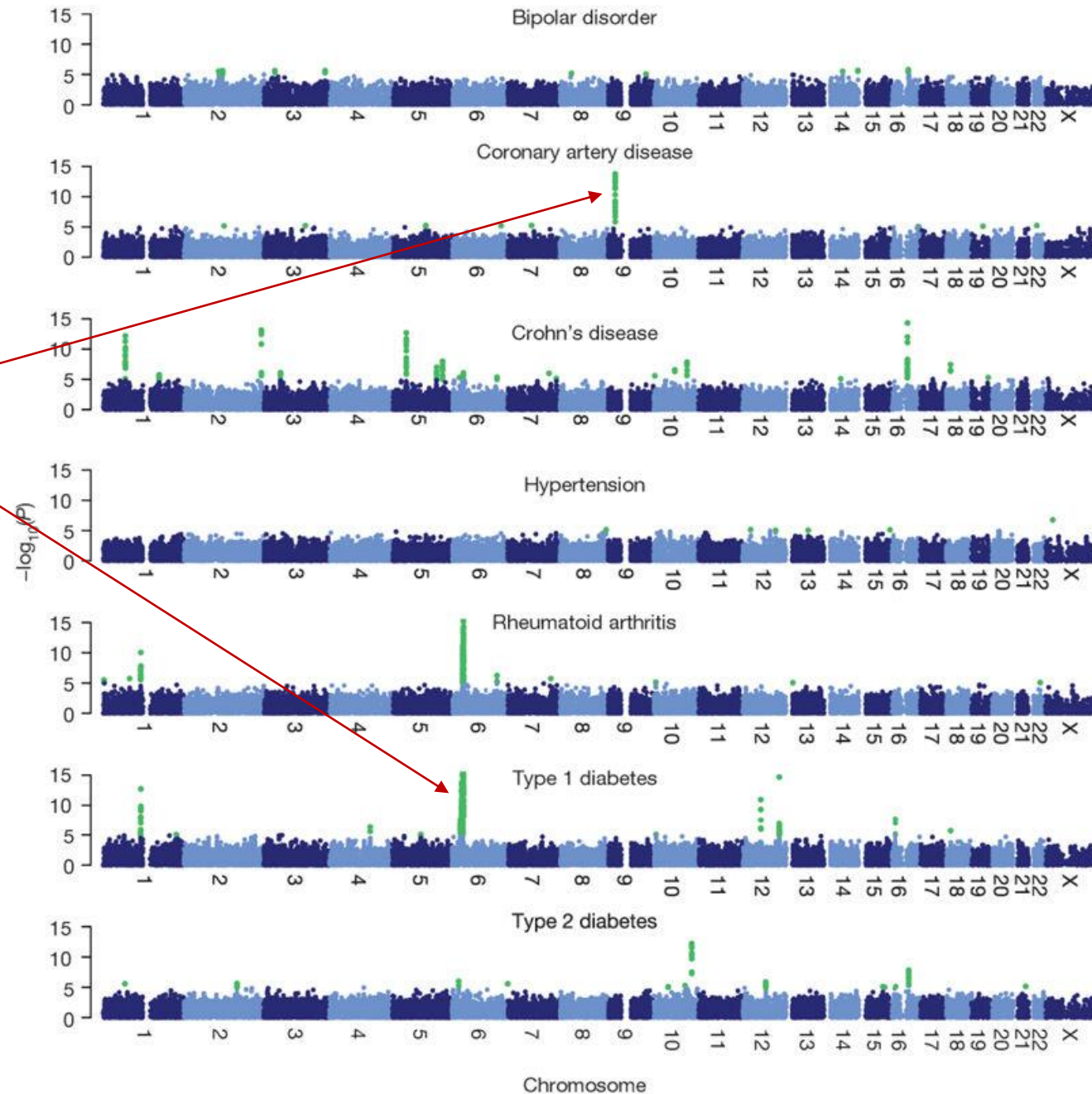
Ülegenoomsed uuringud (*Genome-Wide Association Study, GWAS*)

- Idee: testitakse iga määratud SNP seost haiguse vm tunnusega (>10 miljoni testi) (statistilised mudelid/kiired algoritmid)
- Ajalugu: esimene oluline artikkel aastal 2007 (WTCCC - Wellcome Trust Case Control Consortium), 14000 juhtu (7 haigust), 3000 kontrolli
- 2019: maailmas on läbi viidud tuhandeid GWAS uuringuid, mille abil on uuritud enam kui 1800 haigust vm tunnust

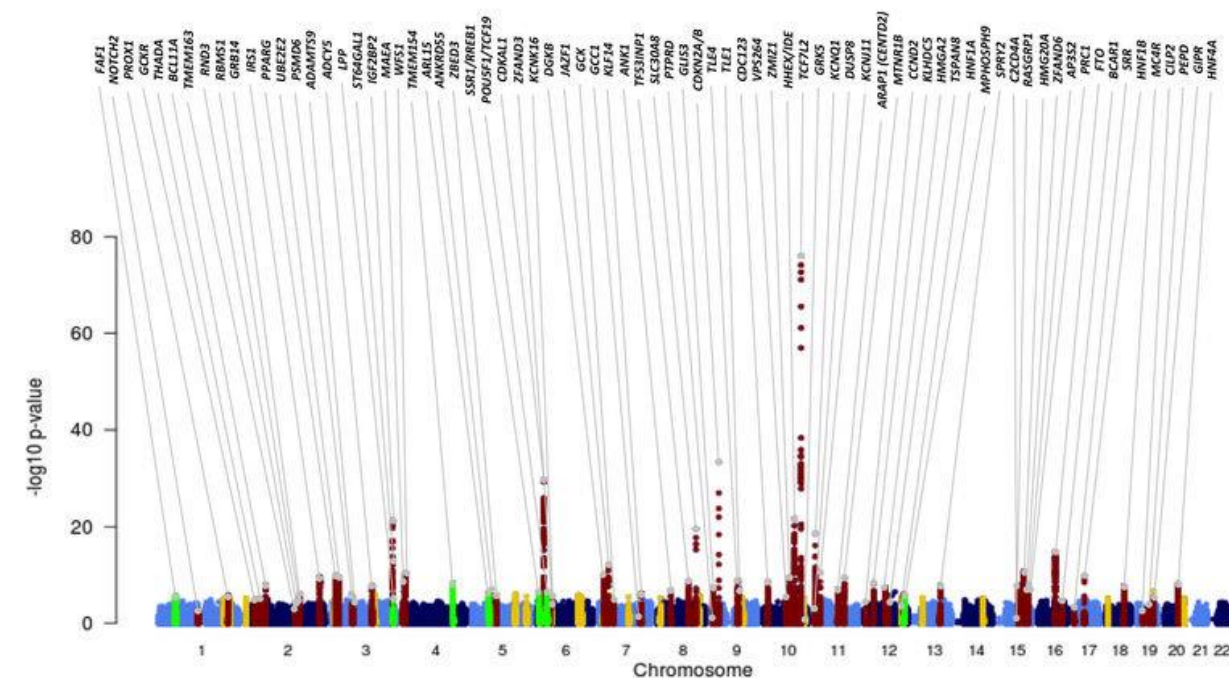
WTCC 2007

„Manhattan plot“ –
näidatakse olulisi seoseid
igal kromosoomil

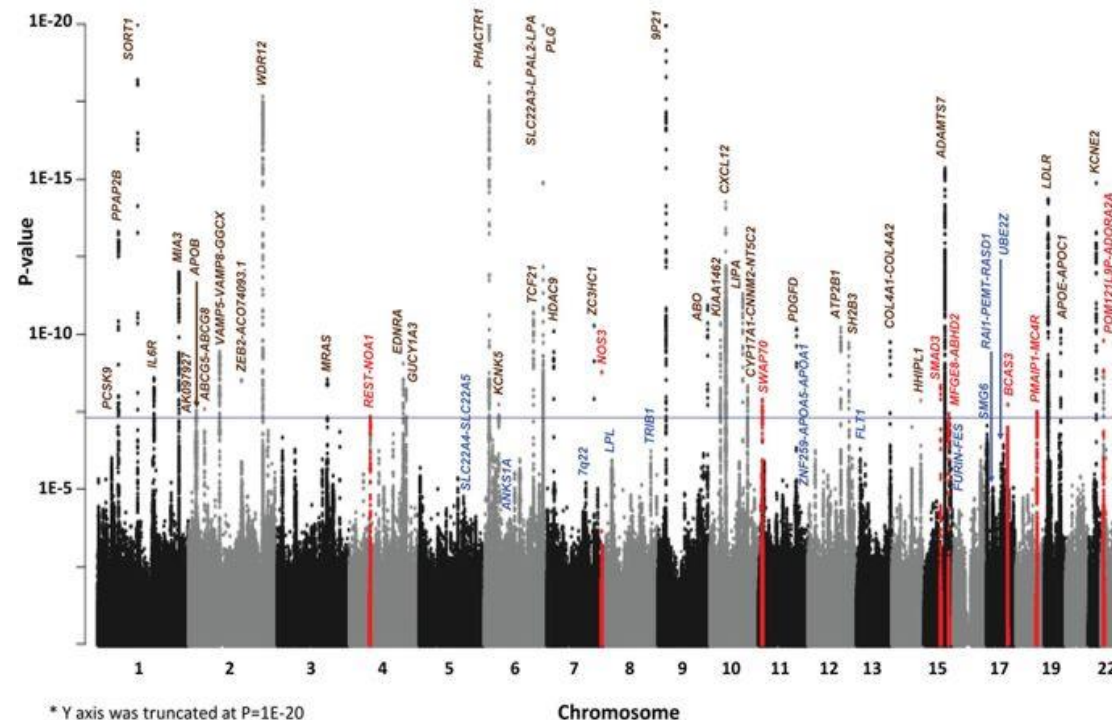
Leitud olulised
seosed



7 aastat hiljem....



Teist tüüpi diabeet:
Mahajan jt, 2014
(48000 diabeedijuhtu ja 140000
kontrolli)



Südame isheemiatõbi
(Coronary Artery Disease),
Nikpay jt, 2015,
N>185000

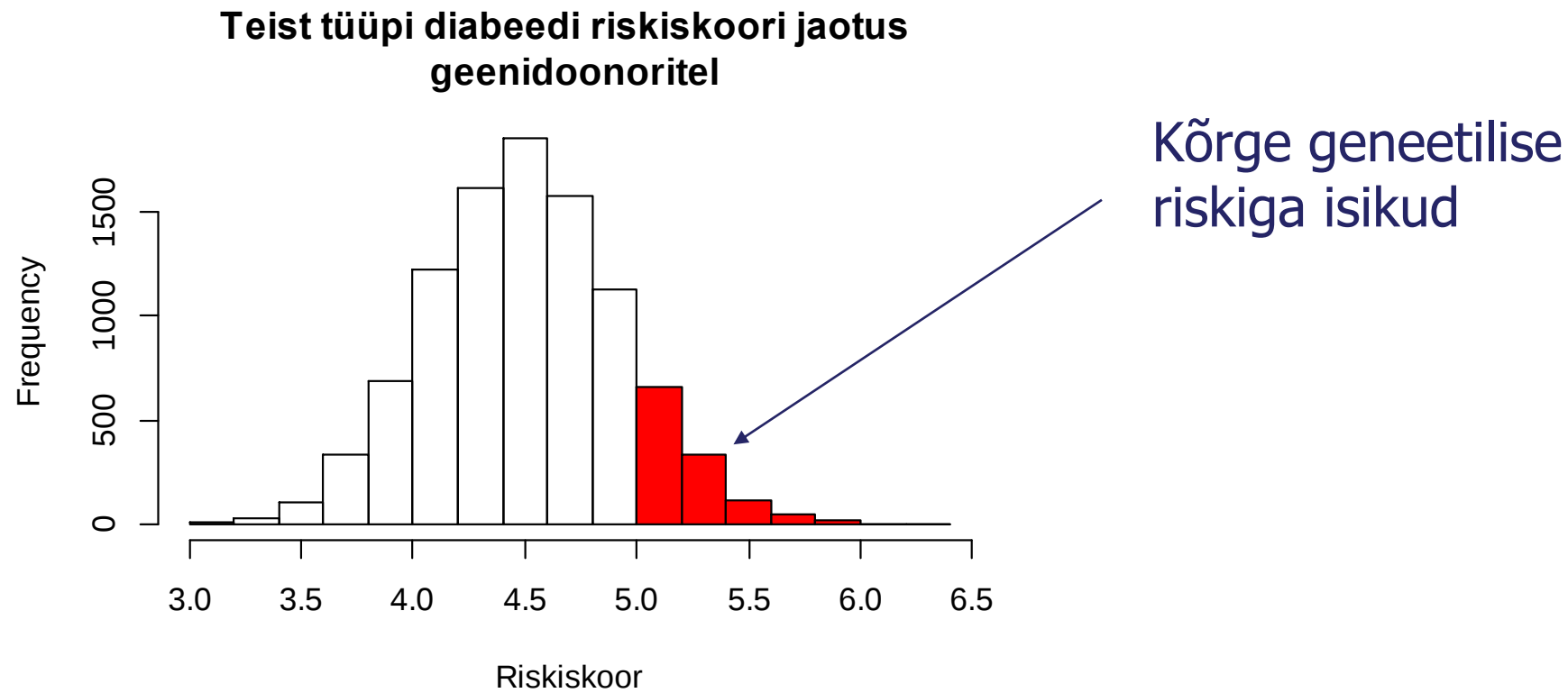
Geneetiline riskiskoor kombineerib tuhandete SNP-de andmed

Genetic (polygenic) risk scores (GRS, PGRS, PRS)

Arvutatud kui: $S = w_1X_1 + w_2X_2 + \dots + w_kX_k$,

$X_1, \dots, X_k$ – k sõltumatu SNP alleelidoosid (väärtus 0-2)

$w_1, \dots, w_k$ – kaalud



Geneetiline (polügeenne) riskiskoor

Arvutatud kui: $S = w_1X_1 + w_2X_2 + \dots + w_kX_k$,

$X_1, \dots, X_k$ – k sõltumatu SNP alleelidoosid (väärtus 0-2)

$w_1, \dots, w_k$ – kaalud

Küsimused statistikutele:

- Mitu SNP-i valida riskiskoori?
- Millised markerid valida riskiskoori? (suurima efektiga või olulisema efektiga?)
- Kuidas valida SNP-ide kaalud w_i ?

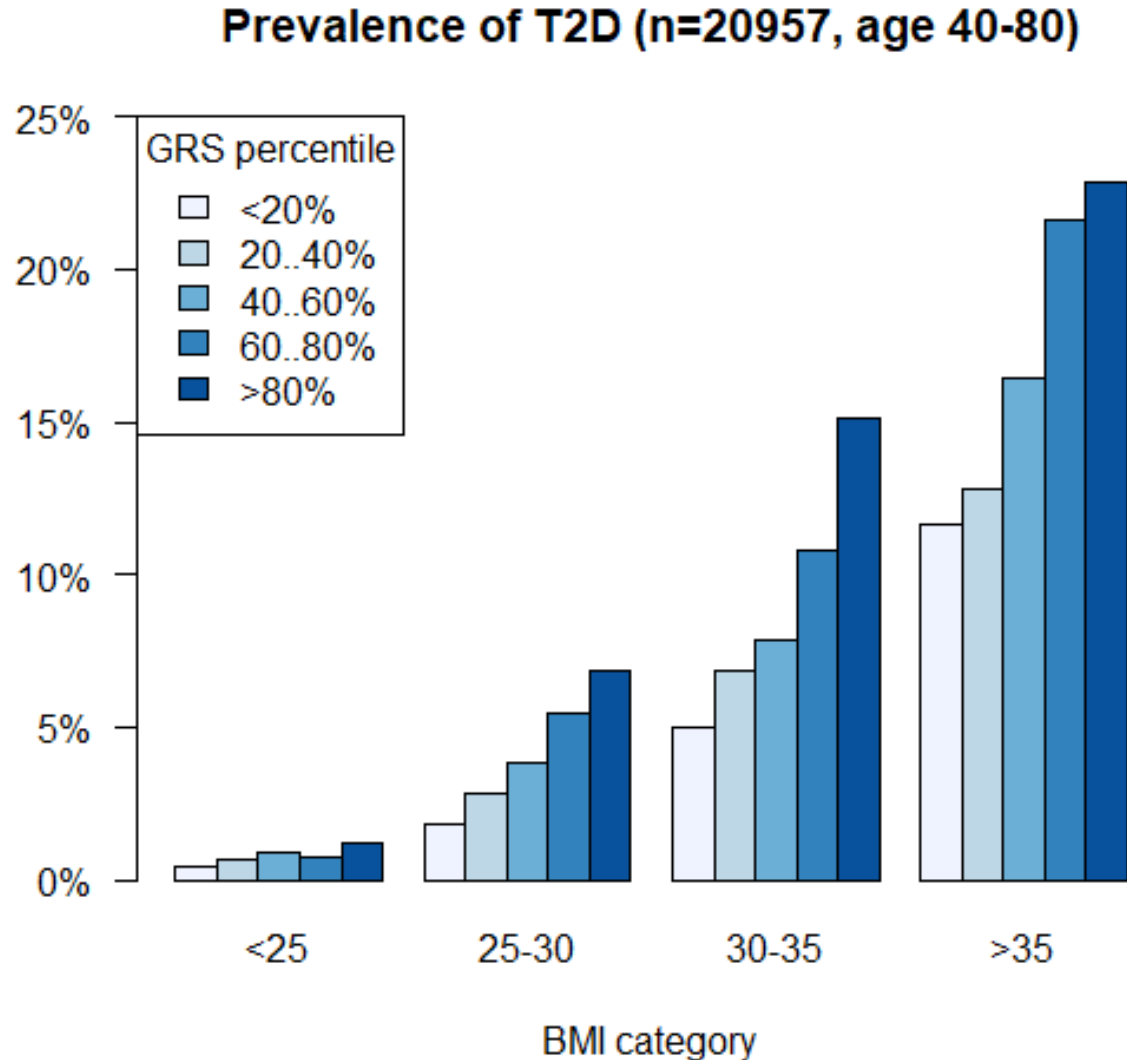
Geneetiliste riskiskooride teemal kaitses juunis 2019 doktoritöö Kristi Läll (TÜ Matemaatika ja statistika instituut)



NN topeltkaalutud skoor:

$$\text{dGRS}_k = \sum_{i=1}^K \hat{\pi}_i(k) \hat{\beta}_i X_i,$$

Teist tüüpi diabeedi esinemissagedus vs geneetiline riskiskoor (>5000 SNP kombinatsioon) ja kehamassiindeks



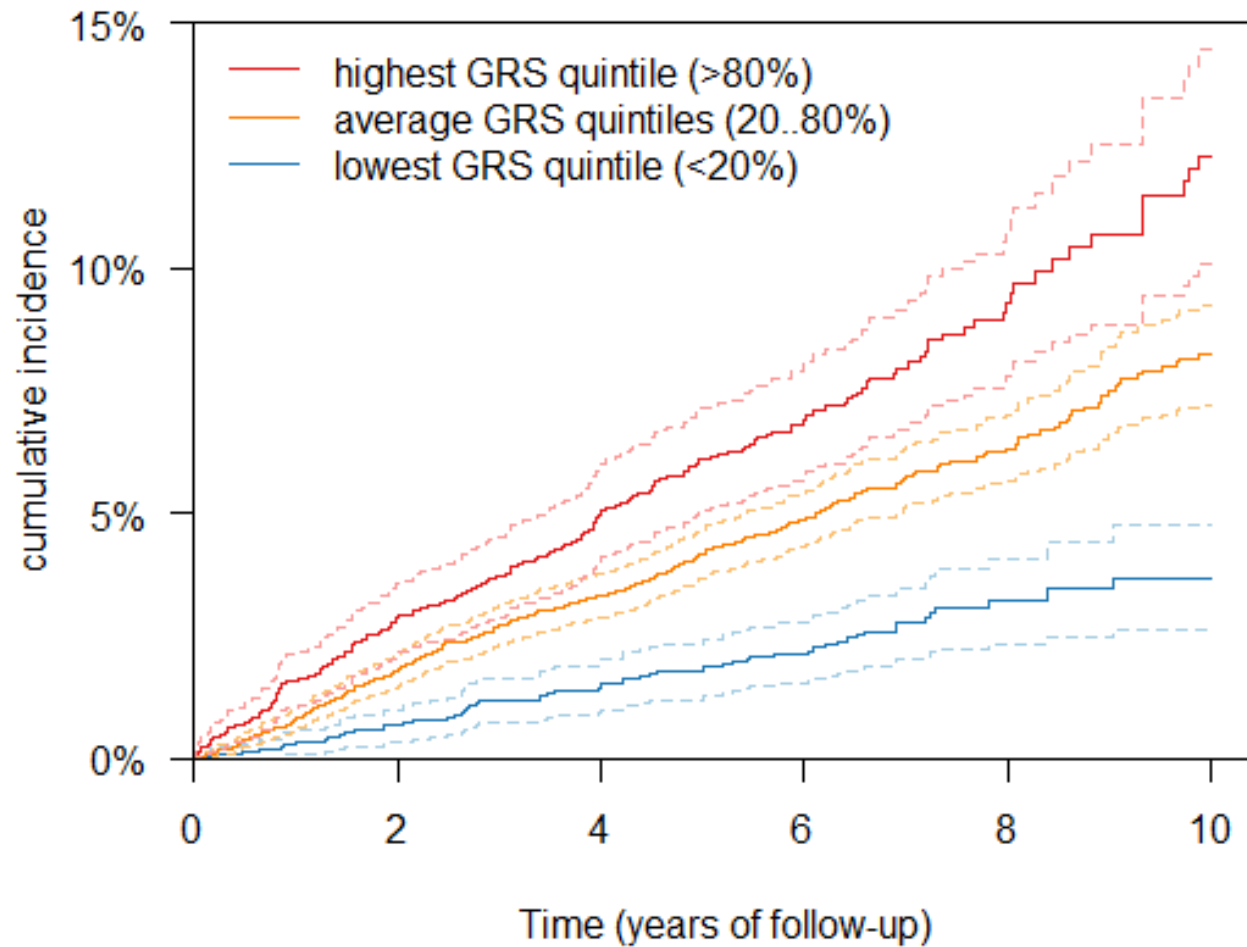
Kuigi rasvumine ja ülekaal on kõige olulisemad riskitegurid, on just keskmise ja mõõduka ülekaalu korral väga suur mõju ka geneetilisel riskil

Läll K., et al. 2017, Genetics in Medicine

Tuleviku prognoosimine...

10402 isikut vanuses 35-79, kellel on ülekaal, aga mitte rasvumine (KMI 25-30) ning liitumishetkeks diabeedi diagnoosi pole – neist 638-l tekkis diabeet jälgimisaja jooksul

**T2D incidence by GRS category
(n=10402, age 35-79, BMI 25-30)**



HR=3.22, $p=1.2 \times 10^{-15}$
184 T2D juhtu

HR=2.19, $p=7.9 \times 10^{-9}$
391 T2D juhtu

63 T2D juhtu

GRS = $dGRS_{1000}$

Tagasiside
geenidonorile:
kasutades
geneetilisi
riskiskoore,
suudame
ennustada ja
tagasisidet anda
teist tüüpi
diabeedi ja
südame
isheemiatõve
päriliku riski kohta

Teise tüüpi diabeet

Teise tüüpi diabeet on komplekshaigus, mille tunnuseks on kõrge veresuhkru tase ning vähemalt osaline insuliiniresistentsus. Haigestumise riski mõjutavad nii elustiilitegurid kui ka geneetiline eelsoodumus.

Geneetiline eelsoodumus

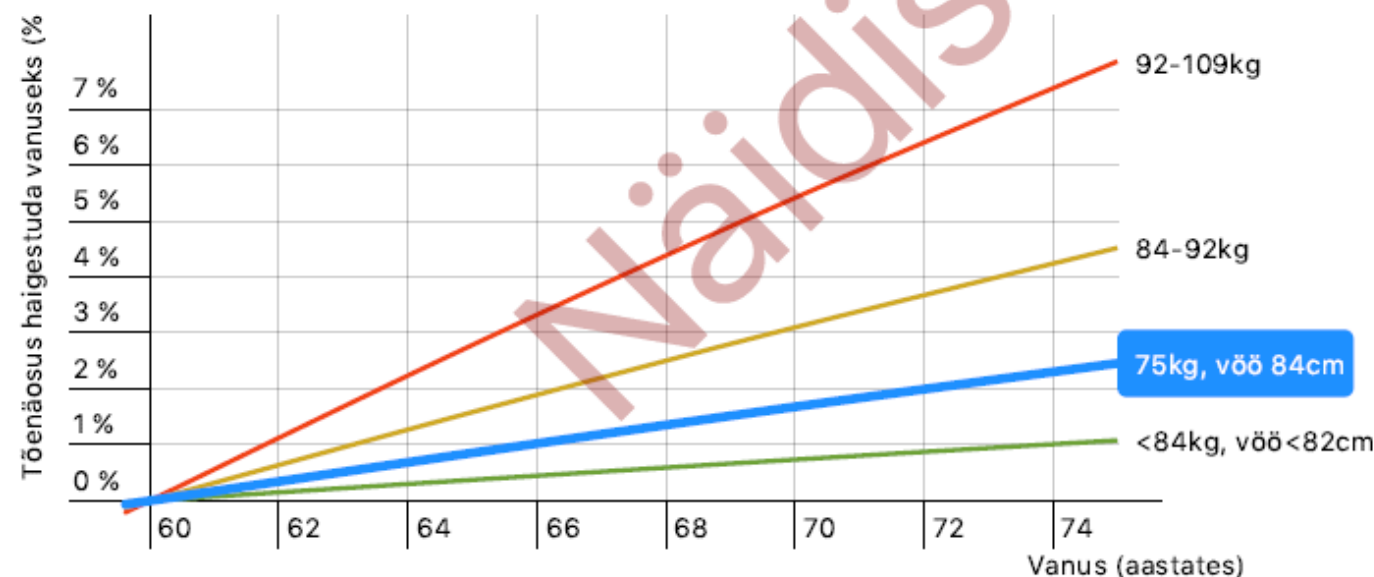
Teie geenide poolt määratud diabeedirisk on **madal**.



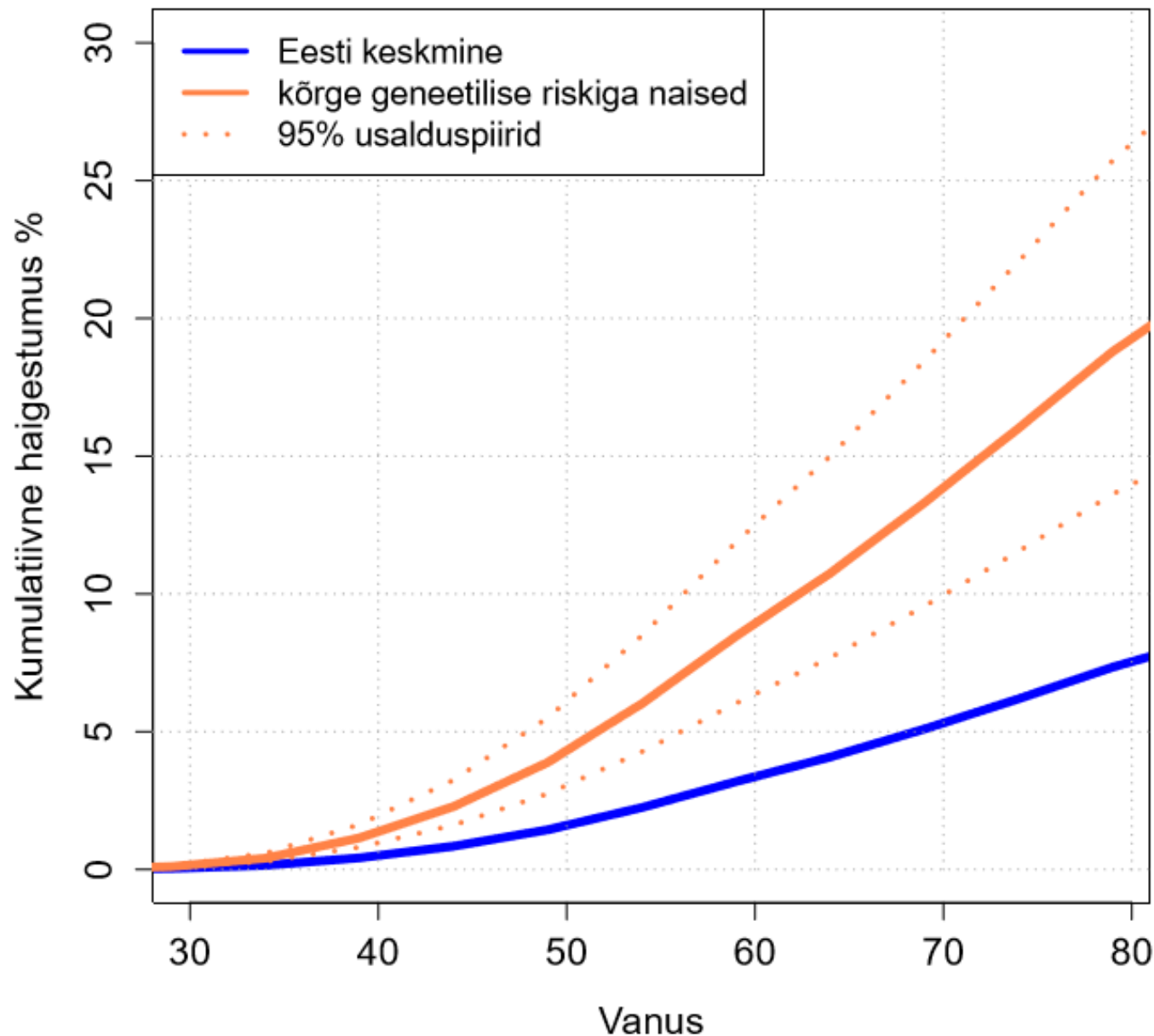
Üheksal inimesel kümnest
on Teist suurem pärilik risk

Kehakaalu mõju diabeedi koguriskile

Haigestumise tõenäosus sõltuvalt kehakaalust ja vanusest Teie geneetilise eelsoodumuse korral. Teie elustiili poolt määratud risk on **madal**.



**Rinnavähi kumulatiivne haigestumus vastavalt vanusele –
Eesti keskmine ja kõrge geneetilise riskiga naised (top 5%)**



...käimasolev pilootuuring (RITA projekt koostöös Sotsiaalministeeriumi ja Eesti suuremate kliinikutega – kõrge geneetilise riskiga naiste nõustamine ja varasem sõeluuringule kutsumine

Geneetilistest mõjudest elukestusele: 1. ülegenoomne analüüs UK Biopanga andmete põhjal, kasutades vanemate eluea andmeid

Received 28 Sep 2015 | Accepted 29 Feb 2016 | Published 31 Mar 2016

DOI: 10.1038/ncomms11174

OPEN

Variants near *CHRNA3/5* and *APOE* have age- and sex-related effects on human lifespan

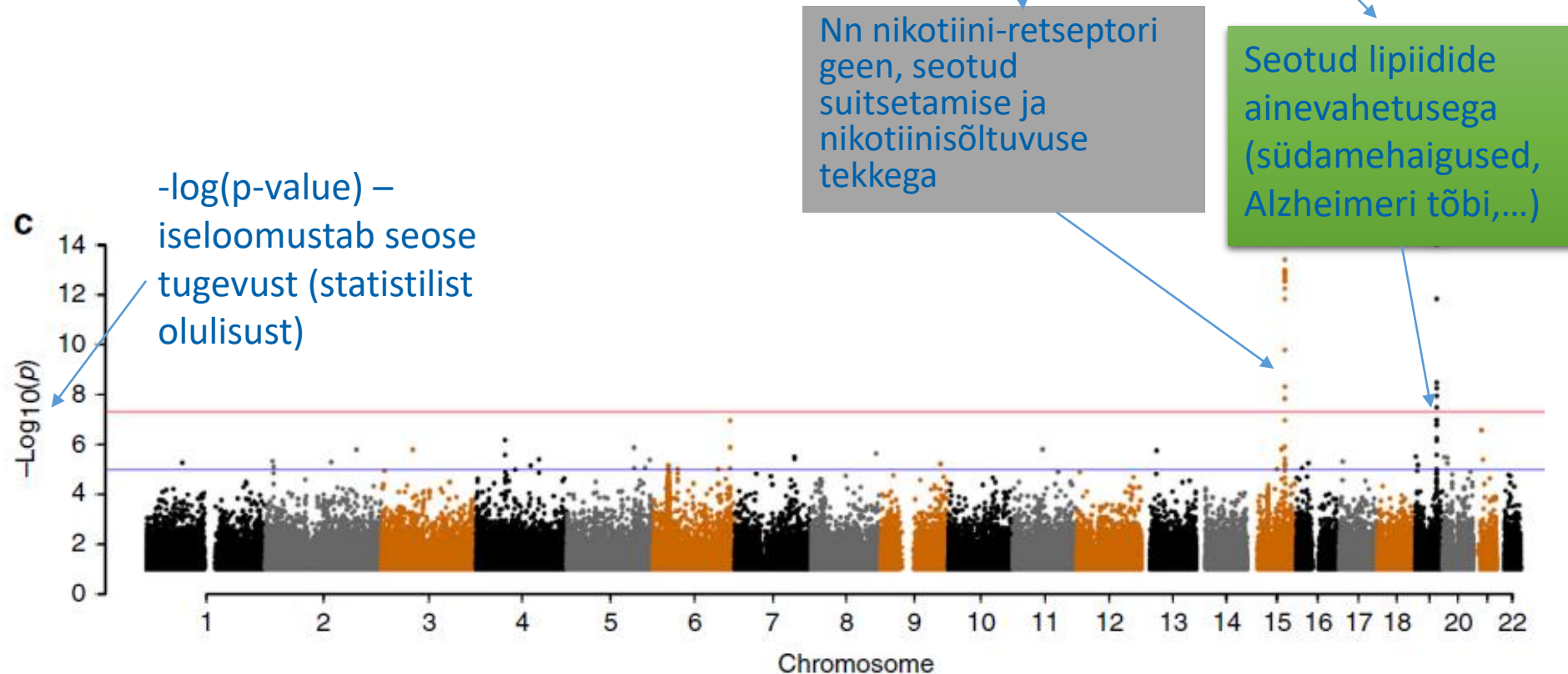
Peter K. Joshi¹, Krista Fischer², Katharina E. Schraut^{1,3}, Harry Campbell¹, Tõnu Esko^{2,4,5,6} & James F. Wilson^{1,7}

- Analüüsiti 140000 isiku geenandmete seost nende vanemate elukestusega (kokku ca 270000 vanemat)

Mida me leidsime?

Variants near *CHRNA3/5* and *APOE* have age- and sex-related effects on human lifespan

Peter K. Joshi¹, Krista Fischer², Katharina E. Schraut^{1,3}, Harry Campbell¹, Tõnu Esko^{2,4,5,6} & James F. Wilson^{1,7}



Kõige tugevamat seost näidanud geenivariandid on seotud teada-tuntud riskiteguritega!

Tänaseks on seda uuringut jätkatud märksa suuremate
valimimahtudega ja tuvastatud 14 elueaga seotud geenipiirkonda



ARTICLE

DOI: 10.1038/s41467-017-00934-5

OPEN

Genome-wide meta-analysis associates *HLA-DQA1/DRB1* and *LPA* and lifestyle factors with human longevity

Peter K. Joshi et al.[#]

[Nat Commun.](https://doi.org/10.1038/s41467-017-00934-5) 2017 Oct 13;8(1):910. doi:
10.1038/s41467-017-00934-5.

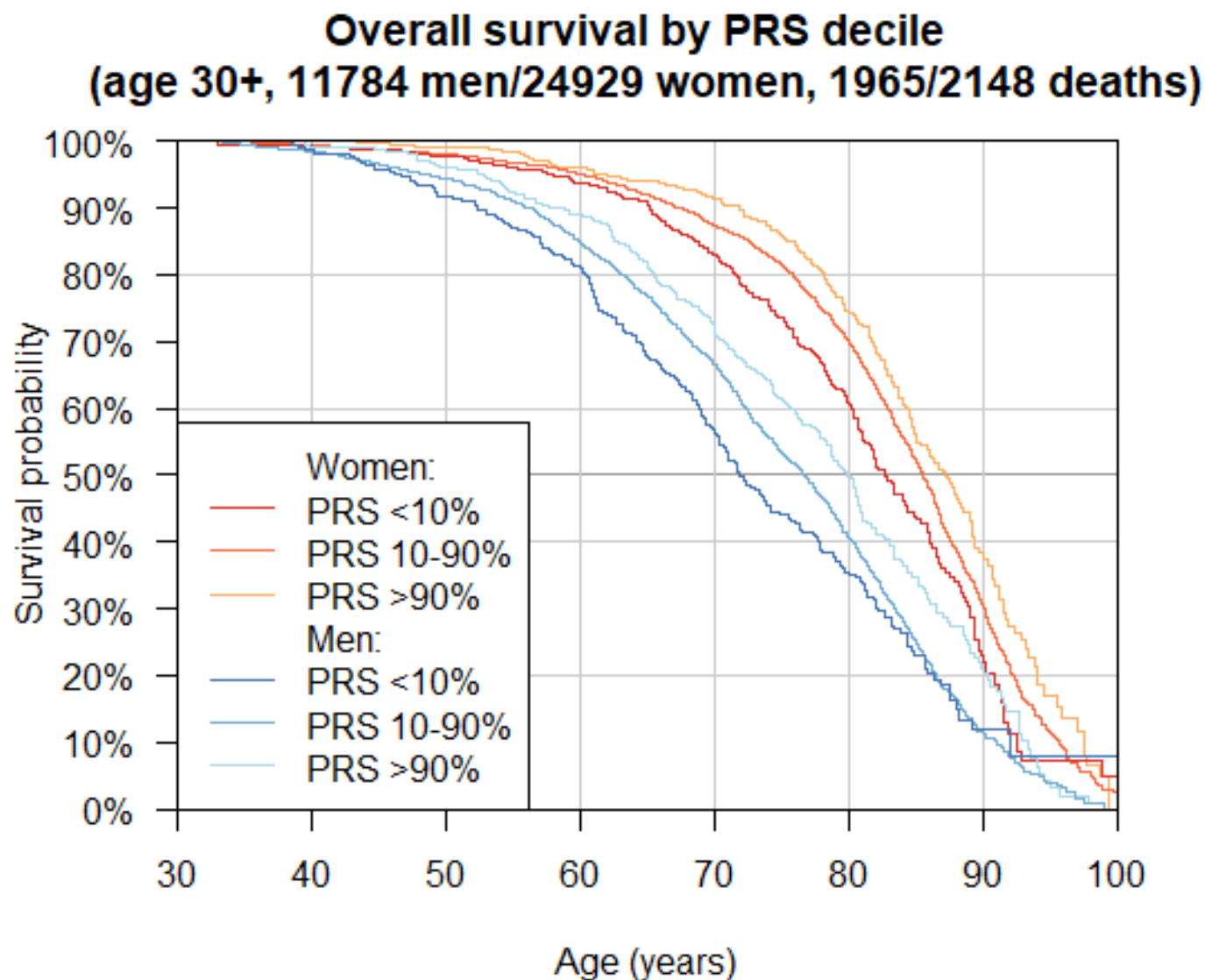
Eesti kaasautorid: Krista Fischer, Tõnu Esko, Lili Milani,
Reedik Mägi



Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances

Paul RHJ Timmers¹, Ninon Mounier^{2,3}, Kristi Lall^{4,5}, Krista Fischer^{4,5}, Zheng Ning⁶,
Xiao Feng⁷, Andrew D Bretherick⁸, David W Clark¹, eQTLGen Consortium,
Xia Shen^{1,6,7}, Tõnu Esko^{4,9}, Zoltán Kutalik^{2,3}, James F Wilson^{1,8}, Peter K Joshi^{1,2*}

Elukestuse geneetiline skoor



Skoor sisaldab sadade geenimarkerite mõjusid. Paljud neist geenidest on eelkõige seotud riskiteguritega (suitsetamine, ülekaalulisus), aga ka otsest südame-veresoonkonnahaiguste riskiga

Biostatistika – kellele ja milleks?

- Valdkonnad: bioloogia, meditsiin, metsandus, loomakasvatus, kalandus, ...
- Seoste leidmine keerulistes andmetes
- Meeskonnatöö
- Õpetamine, selgitamine, konsulteerimine
- Panus meditsiini/bioloogia vm eluteaduste arengusse

Põnevaid teemasid biostatistikast (ainult mõned näited)

- Ravimuuringute planeerimine ja analüüs
- Epidemioloogiliste uuringute (haiguste riskitegurid) planeerimine ja analüüs
- Ökoloogilised mudelid (sh kliimamuutused, taimed-loomad)
- Geenandmete analüüs
- Põhjuslikud mudelid
- Elukestusanalüüs

Kus vajatakse biostatistikuid?

- Ülikoolid
- Teadusasutused
- Suuremad haiglad/kliinikud
- Ravimifirmad
- Kliiniliste uuringute firmad
- Biotehnoloogia firmad

Biostatistika on põnev ja kiirelt arenev valdkond

Biostatistikuid on väga vaja nii Eestis kui mujal

Biostatistikuna töötades tuleb kasuks tugev matemaatiline taust

Krista.Fischer@ut.ee