



KÄSKKIRI

Tallinn

14.01.2013 nr 1-4/88

**Tervise Arengu Instituudi Tervisestatistika ja terviseuuringute andmebaasi kontseptsioon**

Tervise Arengu Instituudi põhimääruse § 6 punkti 7 alusel, kinnitan käesoleva käskkirjaga Tervise Arengu Instituudi Tervisestatistika ja terviseuuringute andmebaasi kontseptsiooni.

Lisa 1: Tervise Arengu Instituudi Tervisestatistika ja terviseuuringute andmebaasi kontseptsioon

Maris Jesse  
direktor

# **Tervise Arengu Instituudi Tervisestatistika ja terviseuuringute andmebaasi kontseptsioon**

## **1. Andmebaas**

Tervisestatistika ja terviseuuringute andmebaas (edaspidi: andmebaas) on tervisega seotud statistika ja uuringute tulemuste elektroonilise publitseerimise koht.

Andmebaasis avaldatakse andmeid, mis annavad ülevaate Eesti rahvastiku tervise seisundist ja seda mõjutavatest teguritest ning tervishoiu valdkonna ressurssidest ja nende kasutamisest. Andmebaasis on esitatud olulised tervisenäitajad, riiklike strateegiate ja rahvusvahelised indikaatorid ning terviseprofiilide koostamiseks vajalikud andmed, mis vastavad andmebaasi kriteeriumidele. Andmebaasi haldamise eest vastutab Tervise Arengu Instituut (TAI). Avaldatavad andmed on pärit erinevatelt asutustelt nagu Ravimiamet, Haigekassa, Terviseamet, Sotsiaalministeerium, Maanteeamet, Päästeamet jne ning Tervise Arengu Instituudi registritest, regulaarsetest uuringutest ja statistikast.

Andmebaasi eesmärgiks on teha tervisestatistika Internetis kättesaadavaks ja koondada ühte kohta, kust andmekasutajatel oleks seda lihtne leida ja kasutada. Andmebaas loodi Euroopa Sotsiaalfondi programmide „Tervislikke valikuid toetavad meetmed 2008-2009“ ja „Tervislikke valikuid toetavad meetmed 2010-2011“ raames. Andmebaas on kohalikele omavalitsustele ja maakondadele kohaks, kus nad saavad enda paikkonna kohta andmeid terviseprofiili koostamiseks. Andmebaas on avalik ja selle kasutamine on tasuta. See on mõeldud kasutamiseks poliitikakujundajatele, analüütikutele ning kõigile neile, kes vajavad ja kasutavad tervisestatistikat.

Andmebaas põhineb Rootsi Statistikaameti poolt välja töötatud tarkvaral PX-Web, mis võimaldab kasutajal luua olemasolevatest tabelitest uusi tabeleid, konverteerida tabeleid teistesse vormingutesse, esitada valitud andmetabelid graafiliselt, teha lihtsaid statistilisi tehteid jms. Andmebaasis esitatakse tervisestatistikat valdkondade kaupa. Valdkonnad jagunevad alamvaldkondadeks, moodustades puustruktuuri. Lisaks andmetele on iga alamvaldkonna juures toodud ära metaandmed. Andmed on andmebaasis esitatud tabelitena, millel on unikaalne kood. Valdkonnatabelite loetelu on sortitud tabeli koodi järgi. Loetelus on tabeli nimetuse all esitatud tabeli muutujate nimetused koos muutuja mõne väärtusega ja sulgudes muutuja väärtuste arv. Kui tabel mingil põhjusel katkeb, s.t uute perioodide andmeid ei lisandu, märgitakse tabeli pealkirjas sulgudesse esitatavate andmete algus- ja lõpuaasta.

## **2. Kvaliteedinõuded andmebaasis avaldatavatele andmetele**

### **2.1 Andmete kvaliteedi tagamine**

Kõik andmebaasis esitatud andmed peavad olema kvaliteetsed ja representatiivsed. Avaldatud statistika meetodika ning definitsioonid on põhjalikult kirjeldatud, metaandmed kasutajale kättesaadavaks tehtud ning vastavad EL kvaliteedikriteeriumidele.<sup>1</sup> Andmete õigsuse ning meetodika ja definitsioonide kirjeldamise eest vastutab andmeallikas (andmete kontaktisik) ning metaandmete kasutajale kättesaadavaks tegemise eest andmebaasihaldur.

### **2.2 Andmete kogumise jätkusuutlikkus ja regulaarsus**

Andmebaasis avaldatakse regulaarselt kogutavat statistikat ning regulaarselt läbiviidavate uuringute andmeid. Pidevalt kogutav statistika (registrite andmed) avaldatakse vähemalt kord aastas vastavalt avaldamiskalendrile ning küsitlusuuringute statistika vastavalt uuringu läbiviimise intervallile. Erandina võib andmebaasis avaldada ka ühekordselt läbiviidavate uuringute andmeid, kui uuringu valim on kogurahvastiku suhtes esinduslik või teema rahvatervise seisukohalt oluline, kuid teiste uuringute andmetega katmata. Kõik andmed avaldatakse avaldamiskalendris toodud kuupäeval. Andmete õigeaegse esitamise eest vastutab andmeallikas (andmete kontaktisik) ning määratud kuupäeval avaldamise eest andmebaasihaldur.

### **2.3 Andmete dubleerimine**

Andmebaasis võib dubleerida mujal (teistes andmebaasides, terviseinfo kogumisega tegelevate asutuste veebilehtedel) avaldatud rahvatervise ja/või tervishoiu seisukohalt olulisi andmeid. Sellisel juhul juhatas tervisestatistika andmebaas andmeallika juurde.

Lisaks võib andmebaasihaldurit nõustava töörühma otsuse alusel dubleerida andmeid, mis on vajalikud andmebaasi kasutajale.

## **3. Andmete esitamine andmebaasis**

### **3.1 Üldpõhimõtted**

Statistika avaldamisel andmebaasis jälgitakse, et tabelid oleksid koostatud tarbijale arusaadavalt. Kus võimalik, ei ületa tabelisse valitavate tunnuste arv nelja ning näitajate arv sadat. Tunnuste nimetuste maksimaalne pikkus on 99 tähemärki. Osakaalude avaldamisel näidatakse alati, millest osakaal on arvutatud. Võimalusel näidatakse ka üldarvu (nt „Eesti kokku“, „vanusrühmad kokku“ jms). Küsitlusuuringute andmete puhul tuuakse negatiivse (nt „suitsetajate osakaal“) kõrval välja ka positiivne („mittesuitsetajate osakaal“).

---

<sup>1</sup> EUROOPA PARLAMENDI JA NÕUKOGU MÄÄRUS (EÜ) nr 223/2009, 11. märts 2009, Euroopa statistika kohta ning Euroopa Parlamendi ja nõukogu määruse (EÜ, Euratom) nr 1101/2008 (konfidentsiaalsete statistiliste andmete Euroopa Ühenduste Statistikaametile edastamise kohta), nõukogu määruse (EÜ) nr 322/97 (ühenduse statistika kohta) ja nõukogu otsuse 89/382/EMÜ, Euratom (millega luuakse Euroopa ühenduste statistikaprogrammi komitee) kehtetuks tunnistamise kohta - <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:et:PDF>

Avaldatav statistika on varustatud metaandmetega („mõisted ja metoodika“) ning klassifikaatorite kasutamisel on andmed klassifikaatori koodide alusel ühtselt tähistatud.

Statistika avaldamisel tagatakse andmekaitse, st andmeid avaldatakse või edastatakse andmesubjekti<sup>2</sup> identifitseerimist võimaldavate tunnusteta, vähemalt kolme andmesubjekti koondandmetena (kusjuures ühegi isiku andmete osatähtsus koondandmetes ei ole suurem kui 90%). Andmesubjekti kirjalikul nõusolekul avaldatakse andmebaasis ka asutuste andmeid.

### 3.2 Kordajate arvutamine

Võimalusel esitatakse andmed kordajana 1000 või 100 000 elaniku kohta. Kordajate arvutamiseks kasutatud rahvaarv on ära toodud metaandmetes.

### 3.3 Puuduvate väärtuste tähistamine

Puuduvaid väärtusi tähistatakse lähtuvalt üldisest puuduvate väärtuste tähistamise tavast (Tabel 1).

**Tabel1.** Andmebaasis kasutatavad tähised ja nende tähendused

| Tähis | Selgitus                                                                                                                                                  | Näide                                                                                                              |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| .     | Mõiste pole rakendatav.                                                                                                                                   | Abortide arv meestel.<br>Emakakaela vähk meestel.                                                                  |
| ..    | Andmeid pole saadud või on avaldamiseks ebakindlad.<br><u>Uuringute puhul</u> : vastajate arv alla 20 isiku.<br><u>Teksti sees</u> „sealhulgas“ tähendus. | Konkreetsel aastal pole näitaja kohta andmeid kogutud.<br><br>Harju maakond<br>.. Tallinn<br>st sealhulgas Tallinn |
| 0     | Nähtuste mitte esinemine või ümardamisel saadud 0.                                                                                                        |                                                                                                                    |
| -     | Ravimistatistikas: ravimit ei tarvitatud.                                                                                                                 |                                                                                                                    |

Tähised märgib andmeallikas enne tabelite andmebaasihaldurile edastamist.

<sup>2</sup> Andmesubjekt on nii füüsiline kui juriidiline isik, kelle andmeid töödeldakse.

## 4. Metaandmed

### 4.1 Metaandmete koostamise ja avaldamise põhimõtted

Metaandmed on andmed, mis aitavad kasutajal aru saada, milliste andmetega on tegu ning kuidas neid tõlgendada. Andmebaasis on iga konkreetse valdkonna kohta käivad metaandmed koondatud dokumendi „mõisted ja metoodika.“ Metaandmetes on lahti seletatud kõik mõisted. Indikaatorite ja kordajate puhul välja toodud arvutuskäik ning kordajate puhul ka millist rahvaarvu kasutati. Samuti on kirjeldatud metoodika ning lahti kirjutatud või viidatud klassifikaatorid. Lisaks on toodud välja loetelu andmetega seotud väljaannetest ja võimalusel lingid väljaannete juurde; samuti viited rahvusvaheliste andmeteni jõudmiseks. Välja on toodud kontaktisiku kontaktandmed. Juhul kui andmete avaldamine nõuab spetsiifilisi meetmeid nende anonümiseerimiseks, on metaandmetes välja toodud, milliseid tunnuseid on muudetud. Metaandmete uuendamine toimub samaaegselt iga-aastase andmete uuendamisega ning selle käigus kontrollib andmebaasihaldur ka metaandmetes kasutatavate terminite vastavust teiste andmebaasis esitatud valdkondade terminitele.

### 4.2 Metaandmete struktuur

4.2.1 Iga andmetabeli juurde kuuluvad järgmised metaandmed:

4.2.1.1 Tabeli pealkiri, mis sõltuvalt teemast sisaldab enamasti uuritud nähtust ja esitatud lõikeid (nt vanusrühmad, geograafilised määratlused, vaatlusühik jm);

4.2.1.2 Ridade ja veergude pealkirjad tabelites;

4.2.1.3 Mõõtühikud;

4.2.1.4 Aja/perioodi määratlus;

4.2.1.5 Piirkondlikud allüksused;

4.2.1.6 Vörreldavus ajas (katkestused aegridades, puuduvad andmed);

4.2.1.7 Allmärkused, mis viitavad spetsiaalsetele soovitustele;

4.2.1.8 Informatsioon andmeallika ja statistika kasutamispääsude kohta.

4.2.1.9 Viimase uuenduse kuupäev

4.2.2 Iga valdkonna juurde kuuluvas „Mõisted ja metoodika“ failis on minimaalne andmebaasis avaldatav komplekt metaandmeid järgmine:

4.2.2.1 Mõistete definitsioonid. Kõik andmetabelites esinevad mõisted on defineeritud.

4.2.2.2 Metoodika kirjeldus, kus on välja toodud andmete saamise protseduur ning selgitatud/põhjendatud erinevate klassifikaatorite kasutamine.

4.2.2.3 Klassifikaatorid. Lisatud on viited kasutatud klassifikaatoritele (RHK, NOMESCO, EHAK jne).

4.2.2.4 Väljaanded. Lisatud on loetelu teemaga seonduvatest raportitest, metodoloogia ülevaadetest, standardtabelite kogumikest, aastaraamatutest jt. väljaannetest.

4.2.2.5 Viited. Lingid rahvusvaheliste andmeteni jõudmiseks, et saaks Eesti andmeid võrrelda rahvusvaheliste andmetega.

4.2.2.6 Kontaktisik. Kontaktisik on inimene, kelle poole saab pöörduda, kui on soov lisapäringuid teha; harilikult andmete levitamise eest vastutava osakonna juhataja.

Kontaktisiku kohta avaldatakse nimi, osakond ja asutus, telefoninumber ja e-posti aadress.

## 5. Avaldamiskalender

Tervisestatistika avaldamisaegadest teavitab avaldamiskalender, mis on kasutajatele kättesaadav veebilehel

<http://pxweb.tai.ee/esf/pxweb2008/dialog/Info/Avaldamiskalender.html>

Alates 01.01.2012 on andmed andmebaasis tarbijatele kättesaadavad avaldamiskalendris teatatud päeval hiljemalt kell 12.00. Olulistest andmete uuendustest teatakse TAI veebilehel, andmebaasi veebilehel ning sotsiaalmeedias (TAI Facebooki-leht, andmebaasi Twitteri-leht).

Oktoobri alguses (tavaliselt 01.10) avaldatakse avaldamiskalendris järgmise aasta avaldamisajad. Avaldamiskalendri koostamise eest vastutab andmebaasihaldur.

Andmete uuendamisel lisab andmebaasihaldur tabeli juurde uuendamise kuupäeva. Andmete korrigeerimisel lisab andmebaasihaldur tabelialuse märkuse, kus on ära toodud muutuse sisu ja põhjused.

Avalikkusele enim huvipakkuva statistika valmimisega seoses avaldab TAI pressiteateid. Pressiteadetes edastatav info avaldatakse samaaegselt ka statistika andmebaasis. Analüütilisi ülevaateid publitseeritakse trüki- ja e-väljaannetena.

## 6. Andmete arhiveerimine

Iga kord, kui andmebaasi uuendatakse, salvestatakse andmebaasi üks varukoopia TAI serverisse ja teine välisele andmekandjale, mis asub Tervise Arengu Instituudi ruumides. Andmebaasi regulaarse varundamise eest vastutab andmebaasihaldur.

## 7. Andmebaasis andmete uuendamise kord

### 7.1 Tervise Arengu Instituudi kogutavate andmete avaldamine

a) Andmetabelite uuendamine. Üks nädal enne avaldamiskalendris märgitud kuupäeva (avaldamistähtaega) peavad olema andmed kontrollitud ja tabelid koostatud. Andmed esitatakse nõutavas Exceli tabeli vormis<sup>3</sup> kolm tööpäeva enne avaldamistähtaega andmebaasihaldurile – vastutab andmeid töötlev TAI analüütik.

b) Metoodika ja definitsioonide uuendamine. Metoodika ja definitsioonid kirjeldatakse andmete esmakordsel avaldamisel ning vaadatakse üle iga uuendamise eel. TAI

<sup>3</sup> Andmetabelivormid on andmetega täitmata Exceli tabelid, kus on veergude ja rea pealkirjade sisu.

analüütik vastutab, et koos andmete uuendustega kolm tööpäeva enne avaldamistähtaega edastatakse andmebaasihaldurile ka metoodika ja definitsioonide uuendused nii eesti kui inglise keeles.

c) Pressiteade või TAI poolne lühiteade sotsiaalmeedia jaoks<sup>4</sup> (kui pressiteadet ei tule). TAI analüütiku koostatud pressiteade (tähtsamate uuenduste korral) peab olema valmis viis tööpäeva enne avaldamistähtaega, mille järel saadetakse see kooskõlastamiseks vastavale osakonnajuhatajale ja TAI avalike suhete nõunikule. TAI analüütiku koostatud lühiteade sotsiaalmeedia jaoks peab olema valmis kolm tööpäeva enne avaldamistähtaega, mille järel selle kooskõlastavad vastav osakonnajuhataja ja TAI avalike suhete nõunik. Andmebaasihaldurile saadab TAI analüütik kooskõlastatud pressiteate või lühiteate sotsiaalmeedia jaoks hiljemalt avaldamistähtajale eelneva tööpäeva lõpuks.

d) Andmete avaldamine veebis. Andmebaasihaldur paneb andmed ning pressiteate/lühiteate veebi üles avaldamistähtaja päeval kell 12.00.

## 7.2 Teiste asutuste kogutavate andmete avaldamine

a) Andmetabelite uuendamine. Vähemalt üks nädal enne avaldamistähtaega edastab andmete kontaktisik andmed nõutavas Exceli tabeli vormis<sup>5</sup> andmebaasihaldurile.

b) Metoodika ja definitsioonide uuendamine. Metoodika ja definitsioonid kirjeldatakse andmete esmakordsel avaldamisel ning vaadatakse üle iga uuendamise eel. Andmete kontaktisik vastutab, et koos andmete uuendustega üks nädal enne avaldamistähtaega edastatakse andmebaasihaldurile ka metoodika ja definitsioonide uuendused nii eesti kui inglise keeles.

c) Pressiteade või lühiteade sotsiaalmeedia jaoks (kui pressiteadet ei tule) edastab andmete kontaktisik koos andmete uuendustega üks nädal enne avaldamistähtaega. Lühiteate/pressiteate vaatab kolme tööpäeva jooksul üle TAI avalike suhete nõunik ning võimalikud parandused kooskõlastatakse andmete kontaktisikuga.

d) Andmete avaldamine veebis. Andmebaasihaldur paneb andmed ning pressiteate/lühiteate veebi üles avaldamistähtaja päeval kell 12.00.

## Tervisestatistika ja terviseuuringute andmebaasi loomist toetas Euroopa Liit.



<sup>4</sup> Lühiteade võtab paari lausega kokku olulisemad tulemused – nt koondnäitajate kasv/kahanemine võrreldes eelneva aastaga.

<sup>5</sup> Andmetabelivormid on andmetega täitmata Exceli tabelid, kus on veergude ja rea pealkirjade sisu.