

TARTU ÜLIKOOL
Loodus- ja täppisteaduste valdkond
Arvutiteaduse instituut
Andmeteaduse õppekava

Kristiina Miller

***Tervise infosüsteemi ja Tervisekassa andmete võrdlus
perioodil 2012–2019 patsientide, diagnooside, ravitüübi ja
teenuseosutajate lõikes***

Magistritöö (15 EAP)

Juhendajad: Marek Oja, PhD
Sirli Tamm, MSc

Tartu 2024

Comparison of data from the Health Information System and the Health Insurance Fund in the period of 2012-2019 by patients, diagnoses, types of treatment and health service providers

Abstract: Examining insurance claims and electronic health records is critical to improving the efficiency, quality, and effectiveness of health care delivery, while helping audits to simplify the work of health care professionals and prevent fraud. It is known that health care providers send fewer electronic health records to the Health Information System than they send insurance claims to the Health Insurance Fund, but the current researchers have looked into this issue so far based on aggregated health data. In this work, health data between 2012-2019 from the Health Insurance Fund's insurance claims and electronic health records from Health Information System are used to compare two data sources by patients, by diagnoses, types of treatment and service providers. Since those are two separate data sets, personalized pseudonymized data were used and transferred to the OMOP CDM data model. The purpose of the work, apart from comparing the data, is to find how much of the data overlaps between two data sources, whether the goal of one-time data entry and reuse is fulfilled, how big is the problem of data continuity in Estonian health data and how much has the obligation of health care providers to send data in the health information system improved. As a result of the work, it was found that even if the healthcare service provider has an obligation to send health documents to the Health Information System, this obligation was not fully fulfilled by the end of 2019. A large part of the data was available in the data of the Health Insurance Fund, but not in the Health Information System. In addition, it was found that for inpatient health records, the amount of health data is almost the same between the two data sources, and the difference comes in with outpatient health records. It was also found that most of the data related to diagnoses is duplicated between two data sources, which does not meet the goal of single data entry.

Keywords:

Health Information System, Health Insurance Fund, health records, insurance claims, OMOP CDM

CERCS: B110 Bioinformatics, medical informatics, biomathematics, biometrics

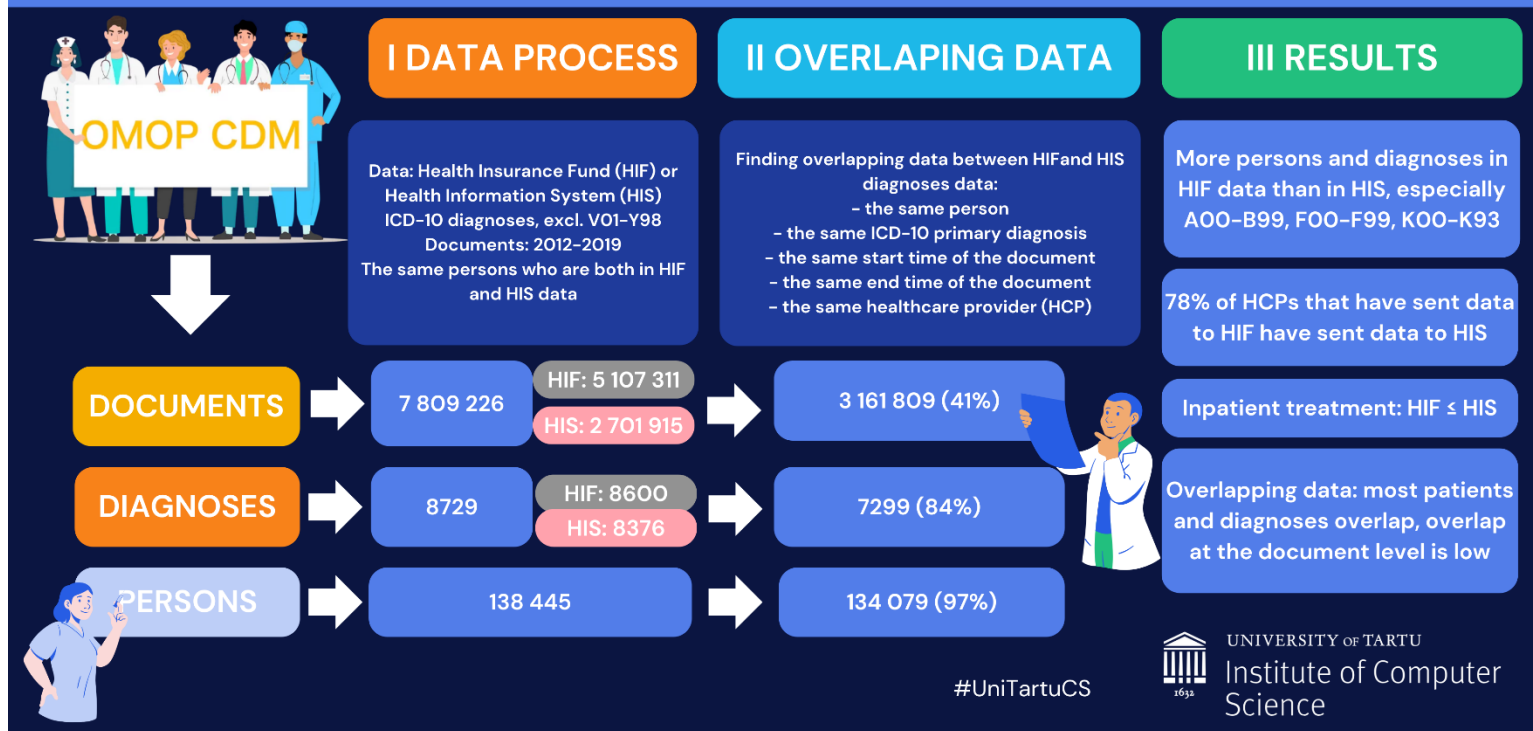
Graphical abstract:

Comparison of data from the Health Information System and the Health Insurance Fund in the period of 2012-2019 by patients, diagnoses, types of treatment and health service providers

Kristiina Miller

Data Science Curriculum (15 ECTS) graphical abstract

Supervisors: Marek Oja (PhD), Sirli Tamm (MSc)



Tervise infosüsteemi ja Tervisekassa andmete võrdlus perioodil 2012–2019 patsientide, diagnooside, ravitüübi ja teenuseosutajate lõikes

Lühikokkuvõte: Raviarvete ja elektrooniliste terviseandmete uurimine on oluline tervishoiuteenuste osutamise tõhususe, kvaliteedi ja tulemuslikkuse parandamiseks, aidates samal ajal kontrollida kulusid, lihtsustada tervishoiutöötaja tööd ja ära hoida pettusi. On teada, et üldjuhul saadavad tervishoiuteenuse osutajad tervise infosüsteemi vähem tervisedokumente kui edastavad nad raviarveid Tervisekassale, aga selliseid uurimistöid on tehtud siamaani agregeeritud andmetega. Antud töös kasutatakse Tervisekassa raviarvete ja tervise infosüsteemi edastatud tervisedokumentide 2012-2019 andmeid, et võrrelda kahe andmeallika andmeid patsientide, diagnooside, ravitüüpide ja teenuseosutajate lõikes. Kuna tegemist on kahe eraldiseisva andmekoguga, siis kasutati isikustatud pseudonüümituid andmeid, mis viidi OMOP CDM andmekujule. Töö eesmärk peale andmete võrdlemise on leida kui suur osa andmeid kattuvad kahe andmeallika vahel, kas täidetakse andmete ühekordse sisestamise ja taaskasutamise eesmärki, kui suur on andmepidevuse probleem Eesti terviseandmetes ja kui palju on paranenud tervishoiu osutajate andmete saatmise kohustus tervise infosüsteemi. Töö tulemusena leiti, et isegi kui tervishoiuteenuse osutajal on kohustus saata tervisedokumente tervise infosüsteemi, siis ei täidetud seda kohustust täielikult 2019. aasta lõpuks. Suur osa andmeid oli olemas Tervisekassa andmetes, aga mitte tervise infosüsteemis. Lisaks leiti, et statsionaarsete tervisedokumentide puhul on terviseandmete hulk peaaegu sama kahe andmeallika vahel ja erinevus tuleb sisse ambulatoorsete tervisedokumentidega. Peale selle leiti, et enamus diagnoosidega seotud andmeid on dubleerivad, mille puhul ei täideta andmete ühekordse sisestamise eesmärki.

Märksõnad:

Tervise infosüsteem, Tervisekassa, tervisedokumendid, raviarved, OMOP CMD

CERS: B110 Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika

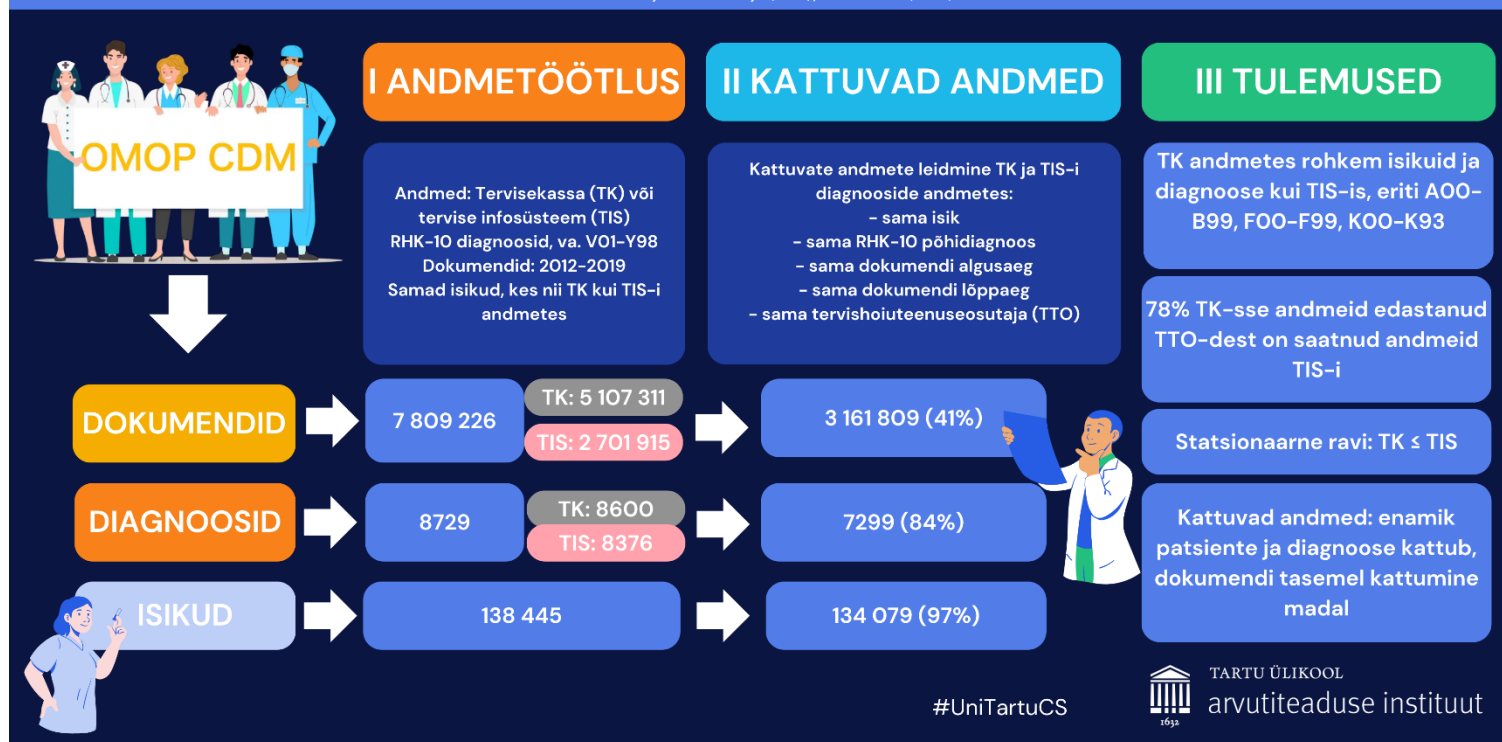
Graafiline kokkuvõte:

Tervise infosüsteemi ja Tervisekassa andmete võrdlus perioodil 2012–2019 patsientide, diagnooside, ravitüübi ja teenuseosutajate lõikes

Kristiina Miller

Andmeteaduse õppekava magistritöö (15 EAP) graafiline kokkuvõte

Juhendajad: Marek Oja (PhD), Sirli Tamm (MSc)



Sisukord

| | |
|--|----|
| Sisukord..... | 6 |
| 1. Sissejuhatus..... | 7 |
| 2. Mõisted ja lühendid..... | 9 |
| 3. Kirjanduse ülevaade | 11 |
| 3.1. Raviarvete ja elektrooniliste terviseandmete võrdlus | 11 |
| 3.2. Tervisekassa ja tervise infosüsteemi andmete võrdlus perearstiabis | 13 |
| 3.3. Eesti E-tervise arengukava ja terviseinfosüsteemid | 14 |
| 3.4. Elektrooniliste terviseandmete andmepidevuse probleem | 17 |
| 4. Kasutatud andmed ja metoodika | 19 |
| 4.1. Kasutatud andmeallikad | 19 |
| 4.2. Andmete töötlus | 21 |
| 5. Tulemused ja arutelu | 26 |
| 5.1. Patsientide võrdlus..... | 26 |
| 5.2. Diagnooside võrdlus | 29 |
| 5.3. Statsionaarne vs. ambulatoorne võrdlus | 33 |
| 5.4. TTO tüüpide ja haigla liigi võrdlus | 39 |
| 5.5. Kattuvate patsientide ja diagnooside võrdlus | 43 |
| 6. Järeldused | 45 |
| 6.1 Töö tugevused, nõrkused, edasiarendusvõimalused | 46 |
| 7. Kokkuvõte | 48 |
| Viidatud kirjandus | 50 |
| Lisad..... | 54 |
| Lisa 1. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 A00-B99 peatüki alampeatükkide lõikes | 54 |
| Lisa 2. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 F00-F99 peatüki alampeatükkide lõikes | 55 |
| Lisa 3. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 K00-K93 peatüki alampeatükkide lõikes | 56 |
| Lisa 4. Haigla liikidesse kuuluvad haiglad | 57 |
| Lisa 5. Litsents | 58 |

1. Sissejuhatus

Eesti e-tervise projektid said alguse juba 2000. aastate algul, mille peamine eesmärk oli luua infosüsteeme, mis lihtsustaksid arstide igapäeva tööd, tõstaksid tervishoiuteenuste kvaliteeti ja annaksid riigile sisendit tervishoiupoliitika jaoks. Selle tulemusel loodi tervise infosüsteem (TIS), mis võimaldab tervishoiuteenuste osutajatel pääseda turvaliselt juurde ja jagada patsientide terviseandmeid erinevate tervishoiuasutuste vahel. Lisaks on patsiendil võimalus näha endaga seotud terviseandmeid. Tervishoiuteenuste osutajatel on kohustus edastada TIS-i tervisedokumente, aga siamaani on uurimistöodes välja tulnud, et seda tehakse osaliselt. Tervisekassale saadetakse küll raviarvete dokumente, mis osaliselt sisaldavad sama andmekoosseisu nagu TIS-i saadetud dokumendid, aga TIS-i saadetakse ka saatekirju, analüüside tulemusi jm. Antud töös on vaatluse all Tervisekassa raviarved ja TIS-i edastatud epikriiside andmed. Analüüsitakse ja võrreldakse ühiseid andmetunnuseid nagu patsiendid, diagnoosid, ravitüübid ja tervishoiuteenuse osutajad ning uuritakse andmete kattuvust ja puuduvust.

Siamaani on uuritud antud probleemi Eestis agregeeritud andmete kujul, mille puhul ei ole ühendatud isikustatud Tervisekassa ja TIS-i andmed. Käesolev töö võimaldab vaadata andmeid kahe andmeallika vahel patsiendipõhiselt ja leida vastused järgnevatele uurimisküsimustele:

- **Kui palju erinevad Tervisekassa raviarvete ja TIS-i epikriiside andmed? Mis on erinevused andmetunnuste osas?**
- **Kui palju ja mis osas on andmed kattuvad Tervisekassa raviarvete ja TIS-i epikriiside andmetes?**
- **Milline on andmepidevuse probleem TIS-i andmetes?**
- **Kas ja kuidas täidetakse Eesti terviseandmetes andmete ühekordse esitamise ja taaskasutamise põhimõtet?**

Sarnaseid uurimistöid on enim tehtud Ameerika terviseandmete põhjal, mille puhul on leitud, et Ameerika tervise infosüsteemides olevad andmed ei sisalda sama hulk andmeid nagu ravikindlustuse andmetes on [1, 2]. Eestis on uuritud ja analüüsitud Tervisekassa ja TIS-i andmeid Tervise Arengu Instituudi poolt, kus nad leidsid, et perearstiabi puhul on Tervisekassas andmetes ligi 30% rohkem patsiente kui TIS-i andmetes [3]. Lisaks on Riigikontroll

analüüsinud e-tervise projekti eesmärkide saavutamist ja toonud välja põhjuseid, miks ei ole saavutatud seatud e-tervise eesmärgi [3, 4, 5].

Käesolev töö on jagatud kolme peatükki: „Kirjanduse ülevaade“, „Kasutatud andmed ja meetodika“ ja „Tulemused ja arutelu“. „Kirjanduse ülevaade“ peatükis seletatakse lahti, mis on raviarved ja elektroonilised terviseandmed. Lisaks tuuakse välja seni tehtud sarnased uurimistööd, mille puhul vaadatakse, kui palju on erinenud siamaani raviarved ja tervise infosüsteemi andmed. Peale selle antakse lühike ülevaade Riigikontrolli poolt tehtud e-tervise eesmärkidest ja nende mitte saavutamise põhjustest. „Kasutatud andmed ja meetodika“ peatükk annab ülevaate, mis andmed antud analüüsis sees on ja samuti kirjeldatakse andmeanalüüsi protsessi. „Tulemused ja arutelu“ sisaldab endas antud töös tehtud analüüsi tulemusi ja võimalikke põhjuseid andmetunnuste kaupa. Lisaks tuuakse välja analüüsi järeldused, töö tugevused ja nõrkused koos võimalike edasiarendustega. Lisades on leitavad mõned detailsemad joonised, et näidata erinevusi Tervisekassa raviarvete ja TIS-i epikriisi diagnooside andmete vahel.

2. Mõisted ja lühendid

| | |
|---------------------------------------|---|
| RHK-10 | Rahvusvaheline haiguste ja nendega seotud terviseprobleemide, epidemioloogia, sümptomite ja ebanormaalsete leidude ning kaebuste statistilise klassifikatsiooni kümnes versioon |
| TEHIK | Tervise ja Heaolu Infosüsteemide Keskuse |
| TIS | Tervise infosüsteem |
| TK | Tervisekassa |
| Tervisekassa raviarved | Patsiendi tervishoiuteenuste hüvitamise arved, mille TTO-d edastavad Tervisekassale |
| Tervise Infosüsteem | Keskne riiklik andmekogu, kuhu saavad erinevad tervishoiuteenuse osutajad isikute terviseandmeid saata ja andmeid vahetada |
| Põhidiagnoos | Peamine tervises seisund, st. seisund, miks patsienti raviti või uuriti |
| Kaasuv diagnoos | Haigus, mis esineb koos põhidiagnoosiga, aga pole peamine tervises seisund |
| Statsionaarne tervishoiuteenus | Ravi, mille osutamiseks on vajalik patsiendi ööpäevaringne viibimine haiglas |
| Ambulatoorne tervishoiuteenus | Ravi, mille osutamiseks ei ole ööpäevaringne haiglas viibimine vajalik |
| Tervishoiuteenus | Tervishoiutöötaja tegevus haiguse, vigastuse või mürgistuse ennetamiseks, diagnoosimiseks ja ravimiseks. Eesmärgiks on |

leevendada vaevusi, hoida ära tervises seisundi halvenemist ning taastada tervis.

Tervishoiuteenuse osutaja (TTO) Tervishoiutöötaja või tervishoiuteenuseid osutatav juriidiline isik, kes on taotlenud tervishoiuteenuste osutamiseks Terviseametilt tegevusloa [6]

OMOP CDM *Observational Medical Outcomes Partnership Common Data Model* ehk standardne andmemudel terviseandmete analüüsimiseks ning rahvusvaheliste uuringute tegemiseks

SNOMED SNOMED CT terminoloogia on ühine keel, mis võimaldab ühtlast ja järjepidevat kliiniliste andmete registreerimist, hoiustamist, pärimist ja agregeerimist erinevate erialade ning tervishoiuüksuste vahel [7]

LOINC *Logical Observation Identifiers Names and Codes* on Eestis kasutatav laborianalüüside kodeering

HL7 HL7 on rahvusvaheline standard, mis on kasutusel Eesti tervise infosüsteemis terviseandmete andmevahetuses. Standardi eesmärk on defineerida andmevahetuses kasutatava infohulga struktuur ja semantika nii, et see oleks võimalikult hästi taasesitatav, üheselt mõistetav ja taaskasutatav [8]

FHIR FHIR (*Fast Healthcare Interoperability Resources*) on tervishoiualase teabe elektroonilise vahetamise standard ning võimaldab erinevatel tervishoiuasutustel, rakendustel ja seadmetel suhelda ning jagada terviseandmeid tõhusalt ja turvaliselt. FHIR on loodud HL7 põhjal [8]

3. Kirjanduse ülevaade

3.1. Raviarvete ja elektrooniliste terviseandmete võrdlus

Elektroonilised terviseandmed ja ravikindlustuse raviarved (edaspidi raviarved) on mõlemad olulised komponendid terviseandmetes, et hallata tervishoiuteenuseid. Elektroonilised terviseandmed on nõ. patsientide paberkaartide digitaalsed versioonid. Need sisaldavad põhjalikku teavet patsiendi haigusloo, diagnooside, ravimite, raviplaanide, immuniseerimisandmete, allergiate, radioloogiliste jm piltide, saatekirjade kohta ja laboratoorsete analüüside tulemusi [9]. Elektroonilisi terviseandmeid kasutavad üldjuhul tervishoiuteenuse osutajad (edaspidi TTO), kes edastavad patsientide terviseandmeid nii terviseandmete infosüsteemi kui ka saavad riskasutamise põhimõttel ligi patsiendi kõikidele terviseandmetele, kui see on võimalik. See on vajalik, sest TTO vajab ligipääsu reaalajas patsiendi terviseandmetele, et toetada kliiniliste otsuste tegemist.

Raviarved on aga dokumendid, mille TTO-d esitavad kindlustusasutustele või riiklikule ravikindlustuse osutajale, et taotleda patsientidele osutatud tervishoiuteenuste hüvitamist. Kuna tervishoiuteenuse hüvitamiseks ei ole vaja nii palju andmeid saata nagu elektrooniliste terviseandmetega saadetakse, siis raviarvetes on ainult olulised andmeid tervishoiuteenuse kulude hüvitamise jaoks. Näiteks raviarvete puhul on oluline teada, mis protseduur ja mille alusel tehti, aga ei ole oluline, mis olid protseduuri detailsed tulemused. Samas kui elektrooniliste terviseandmete suhtes on olulised ka protseduuri detailsed tulemused, mida arstid saavad kasutada patsiendi haigusloos ja võimaliku raviplaan koostamisel. Küll on aga oluline raviarvete puhul, et protseduurid jm tervishoiuteenused oleksid põhjendatud patsiendi haigusloos, sest muidu on õigus kindlustusandjal teenuse eest makstud raha tagasi nõuda [10]. Raviarved üldjuhul sisaldavad patsiendi demograafilisi andmeid, teenuste kuupäevasisid, osutatavate tervishoiuteenuste kirjeldusi ja nende maksumusi, diagnoosikoode ja arvelduskoode [11]. Raviarvete esitamine ja menetlemine on vajalik, et TTO-d saaksid patsiendile osutatud tervishoiuteenuste eest tasu vastavalt lepingule.

Raviarvete ja elektrooniliste terviseandmete koosseisuline erinevus võib tulla andmete kogumise eesmärkidest, meetoditest ja kasutusjuhtudest. Elektroonilisi terviseandmeid kogutakse üldjuhul kliinilistel eesmärkidel, samas kui raviarved on seotud tervishoiuteenuste osutamisega ja vastava tervishoiuteenuse osutamise eest hüvitise saamisega. TTO-d ei pruugi võimalikult täpselt kirjeldada patsiendi haigusloo või raviga seotud terviseandmeid

elektroonilistes terviseandmetes [12: 9], sest osade dokumentide andmekoosseisud saadetakse puudulikult isegi kui on kohustus dokumendis saata kõik andmetunnused. Lisaks on leitud, et arstid kulutavad palju aega elektroonilisi terviseandmeid täites [13] ja- võivad kirja panna ainult hetkel olulise informatsiooni patsiendi haigusloo või ravi kohta.

Raviarved on üldjuhul struktureeritud dokumendid, mis sisaldavad kindlaid ja kohustuslikke andmetunnuseid raviarvete arveldamiseks ja hüvitise saamiseks [14]. Elektroonilised terviseandmed on andmekoosseisult erinevad nagu on dokumenteeritud ka näiteks Tervise ja Heaolu Infosüsteemide Keskuse (TEHIK) publitseerimiskeskuse lehel [15]. Isegi kui on etteantud täpsed juhised, mis juhtudel, kuidas ja milliseid andmeid esitada elektroonilistes terviseandmetes, võivad andmed erineda TTO-de lõikes, sest TTO-d kasutavad erinevaid tervisetarkvarasid [16], nende infosüsteemid ei pruugi kasutada kõige viimasemaid kokkulepitud klassifikaatoreid ja standardeid ning nende tervisesüsteemide koostalitlusvõime (integratsioonid teiste tervisetarkvaradega) võib olla erinev [17: 34].

Terviseandmed on patsiendi kohta käivad tundlikud andmed, millele kehtib seaduses ette nähtud kaitse. Eestis reguleerivad patsiendi kohta käivaid terviseandmeid mitmed seadused nagu näiteks õigust privaatsusele ja konfidentsiaalsusele võlaõigusseadus [18], tervishoiuteenuste korraldamise seadus [19] reguleerib tervishoiuteenuste osutamise korraldust ja nõudeid ning isikuandmete kaitse seadus [20] reguleerib isikuandmete töötlemist. TTO-d peavad kinni pidama nendest seadustest, kui nad koguvad, talletavad ja edastavad patsiendi kohta käivaid andmeid. Sellest tulenevalt võivad patsiendi terviseandmed olla piiratud vastavalt privaatsusenõuetele juhul, kui andmed ei ole kliinilises mõttes haigusloo või ravi jaoks tähtsad.

Terviseandme edastamisel peab üldjuhul kasutama erinevaid kindlaks määratud klassifikaatoreid (RHK-10, SNOMED, LOINC, kohalikud klassifikaatorid jne) või standardeid (HL7, FHIR jne) [8], mis võivad erineda TTO või kasutatavate tarkvarade lõikes. Tervishoiu eest vastutavad asutused lahendavad siimaani väljakutseid, et oleksid ühised terminoloogiad tervisevaldkonnas. Kuna ühtseid klassifikaatoreid või juhendeid ei pruugi TTO-d kasutada, siis võib juhtuda, et elektroonilised terviseandmed ei ole koondatavad või analüüsitavad tervishoiuasutuses ning infosüsteemide koostalitlusvõime ei ole tagatud. Seda ideed kinnitab ka e-tervise strateegia 2020 [17], kus raporti autorid kirjeldavad tänast terviseandmete andmehõivet ja andmekvaliteeti. Raporti autorite sõnul puudub Eestis “üheselt mõistetav terviseandmete kvaliteedi kontseptsioon, mis võimaldaks tõhusalt juhtida tervishoiu kasutatavate kvaliteetsete andmete tekkimist.” [17: 15] Isegi kui on olemas kokkulepitud

andmekoosseis, on nende kasutusele võtmine TTO-de poolt kulukas nii ajaliselt kui rahaliselt. Teine probleem on andmekoosseisu ja juhendite mitte täitmine, mida on täheldatud ka Eesti terviseandmetes [17: 15], kus andmed on ebakorrektselt täidetud, mitteõigeaegselt esitatud, mitte üldse esitatud või esitatud vabatekstina, kus vabateksti tükeldamiseks puuduvad standardid.

Edaspidises töös kasutame elektrooniliste terviseandmetena tervise infosüsteemi andmeid (edaspidi TIS) ja ravikindlustuse raviarvetena Tervisekassa raviarvete andmeid (edaspidi raviarved). TIS-i haldab „Tervise infosüsteemi põhimäärus“, mille §2 kohaselt TIS on „riigi infosüsteemi kuuluv andmekogu (edaspidi tervise infosüsteem), milles töödeldakse tervishoiuvaldkonnaga seotud andmeid tervishoiuteenuse osutamise lepingu sõlmimiseks ja täitmiseks, tervishoiuteenuste kvaliteedi ja patsiendi õiguste tagamiseks ning rahva tervise kaitseks, sealhulgas tervislikku seisundit kajastavate registrite pidamiseks, tervisestatistika tegemiseks ja tervishoiu juhtimiseks.“ [21] Tänapäevaks on enamused TTO-d kohustatud edastama andmeid TIS-i. Riiklikku ravikindlustust pakub Eestis Tervisekassa, mida reguleerib ravikindlustuse seadus. Ravikindlustus tähendab Ravikindluse seaduse §2 lõige 1 alusel „tervishoiukulude katmise süsteem kindlustatud isiku haiguste ennetamise ja ravi, ravimite ja meditsiiniseadmete ostmise rahastamiseks ning ajutise töövõimetuse hüvitiste ja muude hüvitiste maksmiseks käesolevas seaduses sätestatud tingimustel ja korras.“ [22] Hüvitamise alla kuuluvad tervishoiuteenused, nende piirmäära hind ja patsiendi omaosalus % kehtestatakse seaduses. TTO-l on õigus nõuda patsiendile tehtud tervishoiuteenuste eest hüvitist, kui TTO-l on Tervisekassaga ravi rahastamise leping. Kõik täpsemad detailid on üldjuhul määratletud lepingus ja õigusaktides.

3.2. Tervisekassa ja tervise infosüsteemi andmete võrdlus perearstiabis

Tervise Arengu Instituut on riigi teadus- ja arendusasutus, mis tegeleb rahvatervishoiu teadustöö, haiguste ennetamise programmide ja tegevuste ning terviseedendusega. 2019. aastal valmis neil analüüs „Perearstiabi patsientide pöördumiskordade arvu ja haigusjuhu kestuse prognoosid Eesti Haigekassa ning tervise infosüsteemi andmete põhjal“, kus kasutati erinevaid

statistilisi mudeleid (loendusandmete ja logistilist mudelit), et prognoosida patsientide pöördumiskordade arvu ja haigusjuhu pikkust. Autorid kasutasid 2017. aasta Eesti Haigekassa (alates 2023. aasta kannab nime Tervisekassa) ravikindlustuse perearstiabi raviarveid ja tervise infosüsteemi ambulatoorsete epikriiside andmeid. Andmete eeltötluses selgus, et nii Tervisekassa kui TIS-i andmetes oli osa andmeid puudu nagu patsiendi sugu või vanus ja põhidiagnoos. Autorid täheldasid, et TIS-is oli puuduvaid andmeid kuus korda rohkem kui Tervisekassa andmetes [3: 6]. Kokkuvõttes oli Tervisekassa andmetes ligi 30% patsiente rohkem kui TIS-i andmetes [3: 10]. Tervisekassa andmed jagati kaheks: treeningandmed ja testandmed. Mudelid treeniti Tervisekassa treeningandmete peal ja mudeleid testiti kas Tervisekassa testandmetega või TIS-i andmetega.

Kui vaadeldi ja võrreldi reaalseid perearstiabi ühe visiidi pöördumiskordasid prognoositud pöördumiskordadega, siis Tervisekassa prognoos erines 62% võrra ja TIS-i prognoos 73,5% võrra reaalistest pöördumiskordadest [3: 17, 3: 23]. Kui prognoositi, kas patsient pöördub perearstiabisse ühe korra või kaks korda aastas, siis Tervisekassa andmete põhjal oli täpsus 64% ja TIS-i andmete puhul 52% [3: 21, 3: 25]. Mõlemad prognoositud tulemused näitavad, et Tervisekassa raviarvete andmed on paremad, et prognoosida perearstiabi pöördumiskordasid ja nende arvu.

3.3. Eesti E-tervise arengukava ja terviseinfosüsteemid

E-tervise lahendusi on loodud Eestis alates 2000. aastast. Selle töö käigus on oluline vaadata, millised on olnud probleemid e-tervise arengukavas ja tervisesüsteemides, et aidata mõista analüüsi käigus saadud tulemusi ja tähelepanekuid. Lisaks saab antud töös analüüsida, kas on saavutatud osa seotud e-tervise valdkonna eesmärke. Eesti tervishoiusüsteemis on mitmeid viise, kuidas saadakse isikupõhiseid terviseandmeid [3: 18]:

- Tervisekassa isikupõhised raviarvete andmebaas, kuhu TTO-d saadavad raviarvete andmed
- Andmeid tõstetakse ümber paberkandjatele, et saata neid erinevatesse registritesse nagu sünniregister, surmaregister, vähiregister jne

- Andmed on umbisikulises ja agregeeritud kujul, et täita Tervise Arengu Instituudi tervishoiustatistika aruandeid
- Andmed saadetakse isikupõhiselt TIS-i TTO-de poolt

Isegi kui edastatakse andmeid digitaalselt andmebaasidesse, ei pruugi need olla masinloetavalt kättesaadavad kõikidele tervishoiuvaldkonna pooltele.

Tänaseks on riik teinud mitmeid analüüse e-tervise raamistiku ja visiooni osas. Peamisteks e-tervise osapoolteks on Sotsiaalministeerium, Tervisekassa, TEHIK ja tervishoiuteenuse osutajad. 2021. aastal valminud analüüsis toodi välja [23], et puudub formaalne terviklik e-tervise valitsemise mudel ja rollid ning vastutused ei ole selgelt jagatud. 2014.- 2015. aastal loodi e-tervise arengukava, mille alusel loodi strateegia ja rakendusplaan, aga mille täitmine oli ebaühtlane.

2020. aastaks seati mitmeid eesmärgi, et tervishoiu e-teenuseid parandada. Üheksast seotud eesmärgist olid neli seotud terviseandmetega [17: 12]:

- Ühekordne andmete sisestus ja taaskasutus
- Andmekvaliteet ja koostalitlusvõime
- Andmed on organiseeritud tervisesündmustena
- Andmete masinloetav taaskasutus ja koostalitlusvõime

Ühekordse andmete sisestuse ja taaskasutuse all mõeldakse põhimõtet, et terviseandmeid sisestatakse e-tervise süsteemis üks kord ja kus järgmistel kordadel kasutatakse olemasolevaid andmeid. Teise eesmärgi puhul on oluline tagada kogutud terviseandmete kvaliteet, mille puhul kasutatakse kokku lepitud standardeid, klassifikaatoreid ja loendeid, et erinevate infosüsteemide puhul oleks võimalik võrrelda terviseandmeid ja vajadusel ühendada andmeid erinevate andmeallikate vahel. Tervishoiuteenuse osutamise ajal ja sellest tulenev andmekoosseis, peab omama sisulist tähendust, mitte tervishoiuteenuste eest hüvitamise eesmärki, et andmed oleksid organiseeritud tervisesündmustena. Vastasel juhul võivad patsiendi kohta käivad terviseandmed raskendada kliiniliste otsuste tegemist. Üheks eesmärgiks oli ka terviseandmed teha automaatselt töödeldavaks infosüsteemide poolt, et andmeid oleks võimalik taaskasutada koostalitlusvõime alusel. Kui aga puudub ühtne andmestandard, klassifikaatorid jm, siis ei ole võimalik andmeid masinloetavaks teha.

Esimeseks fookusvaldkonnaks e-tervise parandamisel on kvaliteetsed terviseandmed ja andmetaristu, mille eesmärk on „saavutada standardiseeritud digitaalsete andmete kättesaadavus infosüsteemi vahendusel kõikidele osapooltele“[17: 14]. „E-tervise strateegia 2020“ analüüsis toodi välja, et kõik terviseandmete kogud ja infosüsteemid ei ole võimelised vahetama andmeid õigeaegselt ja samas vormis. Tüüpilisteks probleemideks on andmete topelt sisestamine, infosüsteemide kasutajamugavuse puudumine, andmekogumisinõuete erinevus, lisakoormus TTO-dele, kokkulepimata tervisevaldkonna mõisted ja loendid.

E-tervise valdkonna andmete suureks sihiks on tervise infosüsteemi arendamine dokumendipõhiselt info jagamiselt andmepõhise jagamise arendamisele liikuda, mis eeldab andmeaitade, registrite jm analüütiliste andmekogu väljaarendamist [17: 18]. Andmetega seotud eesmärkide elluviimiseks viiakse ellu mitmeid meetmeid [17: 16]:

- **Ühtsete andmehõive ja andmekvaliteedinõuete väljatöötamine ja rakendamine, andmekvaliteedi kontrolli võimekuse tõstmine**, mille puhul analüüsitakse kehtestatud nõudeid andmehõivele, andmekvaliteedile ja süsteemi toimimisele; analüüsitakse andmete töötlemisega tegelevate asutuste andmevajadust ja viise, kuidas vähendada dubleerivate andmete sisestamist ja kogumist; toetatakse andmekvaliteedi kontrolli tõhusust; toetatakse protsesside loomist, mis aitavad esitada vajalikke andmeid vastavalt kokkulepitud standarditele ja klassifikaatoritele.
- **Tervisevaldkonda mõistete ja loendite arendamine ning standardiarendus**, mille puhul arendatakse ja seatakse kohustus järgida kokkulepitud standardeid, protsessinõudeid, andmekoosseisu jm; luuakse standardiarenduse protsess, et kiirendada TTO-de uute infosüsteeminõuete rakendamist; arvestatakse andmete masinloetavuse nõuet; luuakse tervishoiu mõistete, loendite jm arendamise arendusprotsessid ja koolitussüsteemid, et tõsta kasutajate nagu TTO-d teadlikkust standardite muutustest; võetakse kasutusele standardihalduse platvorm, et lihtsustada standardite loomist ja selle kohaste muutuste rakendamist.
- **Teenuseosutajate poolt kasutatavate e-teenuste ja rakenduste valideerimine**, mille puhul TTO-dele rakendatakse funktsionaalseid e-tervise nõudeid, et saavutada teenuste kiire integratsioon ja andmehõive suurenemine; arvestatakse võimalust kolmandate osapoolte nagu erinevate Euroopa Liidu süsteemide liidestust Eesti e-tervise valdkonna infosüsteemidega.

Peale selle tegeletakse veel TTO-de poolt kasutatavate tarkvarade parandamise ja lihtsustamisega, et andmehõive, andmeõigsus ja andmete sisestamise aeg toetaks andmetega seotud eesmäärke.

3.4. Elektrooniliste terviseandmete andmepidevuse probleem

Antud töös on oluline uurida, kas ja kui palju erinevad TIS-i ja raviarvete andmed. Sarnaseid uurimistöid on tehtud üldjuhul Ameerika terviseinfosüsteemide põhjal, mis on erinev Eestis olevast terviseinfosüsteemist. Näiteks Eestis on universaalne tervishoiusüsteem, mille puhul selle finantseerimine tuleb töötajate maksudest. Samas kui Ameerikas puudub universaalne tervishoiusüsteem, mille puhul tervishoiusüsteem põhineb eratervisekindlustusel, tööandjate rahastatud terviseplaanidest või valitsusprogrammidel (Medicare).

Uurimistöö *“External Validation of an Algorithm to Identify Patients with High Data-Completeness in Electronic Health Records for Comparative Effectiveness Research”* [1] autorid Lin, Rosenthal, Murphy jt uurisid elektrooniliste terviseandmete andmepidevuse probleemi, mis tähendab seda, et osa patsiendi terviseabi andmeid eksisteerib väljaspool elektrooniliste terviseandmete süsteemi. Elektroonilised terviseandmed üldjuhul sisaldavad rohkem ja erinevaid andmeid patsiendi haigusloo kohta kui näiteks raviarvete andmed. Autorid toovad näiteid USA terviseandmete alusel, kus erinevate terviseandmete allikate ja tervishoiuteenuste pakkuja vahel võib andmetes olla erinevusi. Andmed, mis on väljaspool elektroonilisi terviseandmeid, võivad põhjustada andmete väärkasutamist ja valedele järelduste jõudmiseni. Nad toovad näite kahe suure elektroonilise terviseandmete süsteemi najal Massachusettsi ja Põhja-Carolina osariigis, mis on ühendatud Medicare'i raviarvetega. Nad lisavad, et andmete väärkasutamist ja valedele järeldustele jõudmist vähendaks lisada juurde terviseandmeid nagu raviarvete kasutamine, aga alati ei ole see võimalik. Autorid on loonud algoritmi, mis suudab tuvastada elektrooniliste terviseandmete andmepidevust tüüpiliste, st. teatud struktuuri ja andmetunnustega, elektrooniliste terviseandmete süsteemi puhul. Nad kasutavad seda algoritmi, et võrrelda kahe erineva terviseandmekogu andmete järjepidevust: üks sisaldab ainult elektroonilisi terviseandmeid ja teine sisaldab nii elektroonilisi

terviseandmeid kui ka raviarvete andmeid. Autorid jõudsid järelduseni, et elektrooniliste terviseandmete kogud hõlmavad ainult 21-27% andmeid võrreldes raviarvete andmetega.

Kasutades elektroonilisi terviseandmeid ja raviarveid erinevate väljundite ennustamiseks, peab meeles pidama, et nende kahe andmeallika vahel on ainult osa andmeid, mis on mõlemas andmeallikas olemas. Zeltzer, Balicer jt. autorid uurisid elektrooniliste terviseandmete ja raviarvete kasutamist ennustamise täpsuse võrdluseks. Oma töös „*Prediction Accuracy With Electronic Medical Records Versus Administrative Claims*“ [2] kasutasid nad elektroonilisi terviseandmeid ja raviarveid, et ennustada nelja erinevat väljundit:

- haiglas viibimise pikkust (≥ 7 päeva)
- surmajuhtu vastuvõtu ajal
- 30-päeva jooksul haiglasse tagasitulemist
- 1-aasta jooksul suremust

Nad jõudsid järeldusele, et nii elektroonilised terviseandmed kui raviarved annavad enam-vähem samad tulemused, aga autorid täheldasid, et elektroonilisi terviseandmeid peaks kasutama lühiajaliste tulemuste ennustamiseks, sest nende täpsus lühiajaliste tulemuste ennustamiseks oli parem kui raviarvete andmete alusel. Samas peaks pikaajaliste ennustuste jaoks kasutama just raviarveid, sest need annavad täpsemad tulemused kui elektroonilised terviseandmed. Üheks põhjuseks, miks sobivad raviarved paremini pikaajaliste ennustuste jaoks on, et elektroonilistes terviseandmetes on andmepidevuse probleem, st. patsient saab terviseabi väljaspool elektroonilise terviseandmete süsteemi.

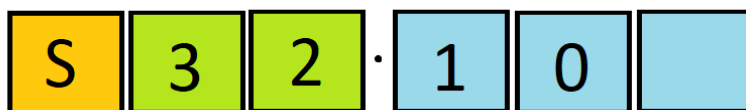
4. Kasutatud andmed ja metoodika

4.1. Kasutatud andmeallikad

Käesoleva uurimistöö andmed on pärit elektroonilisest terviseandmete andmebaasist, mis oli loodud projekti „Valdkondliku teadus- ja arendustegevuse tugevdamine“ (RITA) tegevus 1 „Strateegilise TA tegevuse toetamine – „Masinõppe ja AI toega teenused““ raames ja mille läbi viimiseks on Tartu Ülikooli inimuuringute eetikakomitee luba 300/T-23 ja Eesti Bioetika ja inimuuringute nõukogu luba (1.1-12/653). Antud töö on läbiviidud projekti PRG1844 (Kliiniliste oluliste radade tuvastamine terviseandmetest) raames.

Andmebaasis olevad andmeid pärinevad kolmest allikast: Tervisekassa andmekogu (raviarved), retseptikeskus (välja kirjutatud ja välja ostetud ravimid) ning tervise infosüsteem (haigusloomd, saatekirjad ja saatekirjade vastused). Andmestik ei sisalda kõiki Eesti elanikkonna andmeid, vaid juhuvalimina 10% kõigist Eesti isikukoodiga isikutest (n=150824 patsienti) ja nende terviseandmeid aastatel 2012-2019. Antud töö kontekstis on olulised raviarved ja tervise infosüsteemi andmed, retseptikeskuse andmeid käesolevas töös ei kasutata.

Kõik andmed andmebaasis olid viidud OMOP CDM kujule [24]. Igale patsiendile oli määratud unikaalne ID, teistest isikuandmetest oli olemas ka sugu, sünniaasta ja surmakuupäev. Teistest andmetest olid veel olemas haigusjuhtude dokumentide andmed, diagnooside andmed, terviseteenuseosutajate andmed.



I peatükk S00-T98 "Vigastused, mürgistused ja teatavad muud välispõhjuste toime tagajärjed"
II alampeatükk S30-S39 "Kõhu, selja alaosa, lülisamba nimmeosa ja vaagna vigastused"
III jaotis S32 "Lülisamba nimmeosa ja vaagnamurd"
IV alamjaotis S32.1 "Ristluumurd"
V 5.koha alajaotis S32.10 "kinnine"

Joonis 1. RHK-10 jaotiste näide S32 diagnoosikoodi näitel, kus viimane number on puudu, sest S32 saab märkida kuni viiekohalise täpsusega

Enne OMOP CDM kujule andmete viimist puhastati ja standardiseeriti diagnoosid ning võrreldi diagnoose vastu Eesti RHK-10 loetelu [25] nii Tervisekassa kui TIS andmete korral. RHK-10 on ICD-10 (*The International Classification of Diseases*, versioon 10) eestistatud versioon. RHK-10 klassifikatsiooni alusel saab haigusloo kirja panna kuni seitsmekohalise kooditäpsusega, kus viimane koht võimaldab diagnoosi detailsemalt kirja panna, aga Eestis on tavaks panna diagnoos kirja kuni 5.koha alajaotise täpsusega nagu joonisel 1 näites ja väga kuuendat ja seitsmendat kooditäpsust ei kasutata.

Kõikidest diagnoosidest, mis olid RHK-10 põhidiagnoosina või kaasuva diagnoosina TIS-i või Tervisekassa andmetes, olid enamuse alamjaotise, st. neljakohalise RHK-10 kooditäpsusega kirjas, aga oli ka diagnoose, mis olid RHK-10 peatüki kooditäpsusega haiguslool kirjas. RHK-10 diagnooside kooditäpsused jagunesid järgnevalt:

- Kolmekohaline ehk jaotise RHK-10 kooditäpsus: 15%
- Neljakohaline ehk RHK-10 kooditäpsus: 78%
- Viiekohaline RHK-10 kooditäpsus: 7 %
- RHK-10 peatüki kooditäpsus: 0.02 %

Igal RHK-10 diagnoosikoodil on oma põhimõtted kui täpselt peavad TTO-d diagnoosi kirja panema [25]. Üldiselt peaks iga diagnoosikood olema vähemalt jaotise tasemel kirjas nagu joonisel 2 näites, mis näitab, mis terviseprobleemiga on tegemist. Peale jaotist neljas, viies ja kuues koht näitavad näiteks terviseprobleemi kohta, põhjust, tõsidust. Viimased kolm kohta ei pruugi olla igal diagnoosikoodil ja nende spetsiifilised näitajad sõltuvad diagnoosikoodist endast.

Tervishoiuteenuse osutajate andmetes olid olemas TTO nimi, tegevuskoht, osutatavad tervishoiuteenused ja mis tüüpi TTO-ga on tegemist. TTO tüüpidest olid olemas: keskhaigla, üldhaigla, kohalik haigla, piirkondlik haigla ja teised TTO-d. Teiste TTO-de alla kuuluvad kõik teised tervisehoiuteenuseosutajad nagu perearstid, hambaarstid, eriarstid, kes pole haiglad. Tervishoiuteenuste korraldamise seadus (§ 22 lõike 4 alusel) määrab ära haigla liigi [26]:

- Piirkondlik haigla: tegemist on kolmanda taseme haiglaga, mis peab suutma tagada tervishoiuteenuseid väga suurele elanikkonnale ja tänase seisuga on neid Eestis kolm [27]

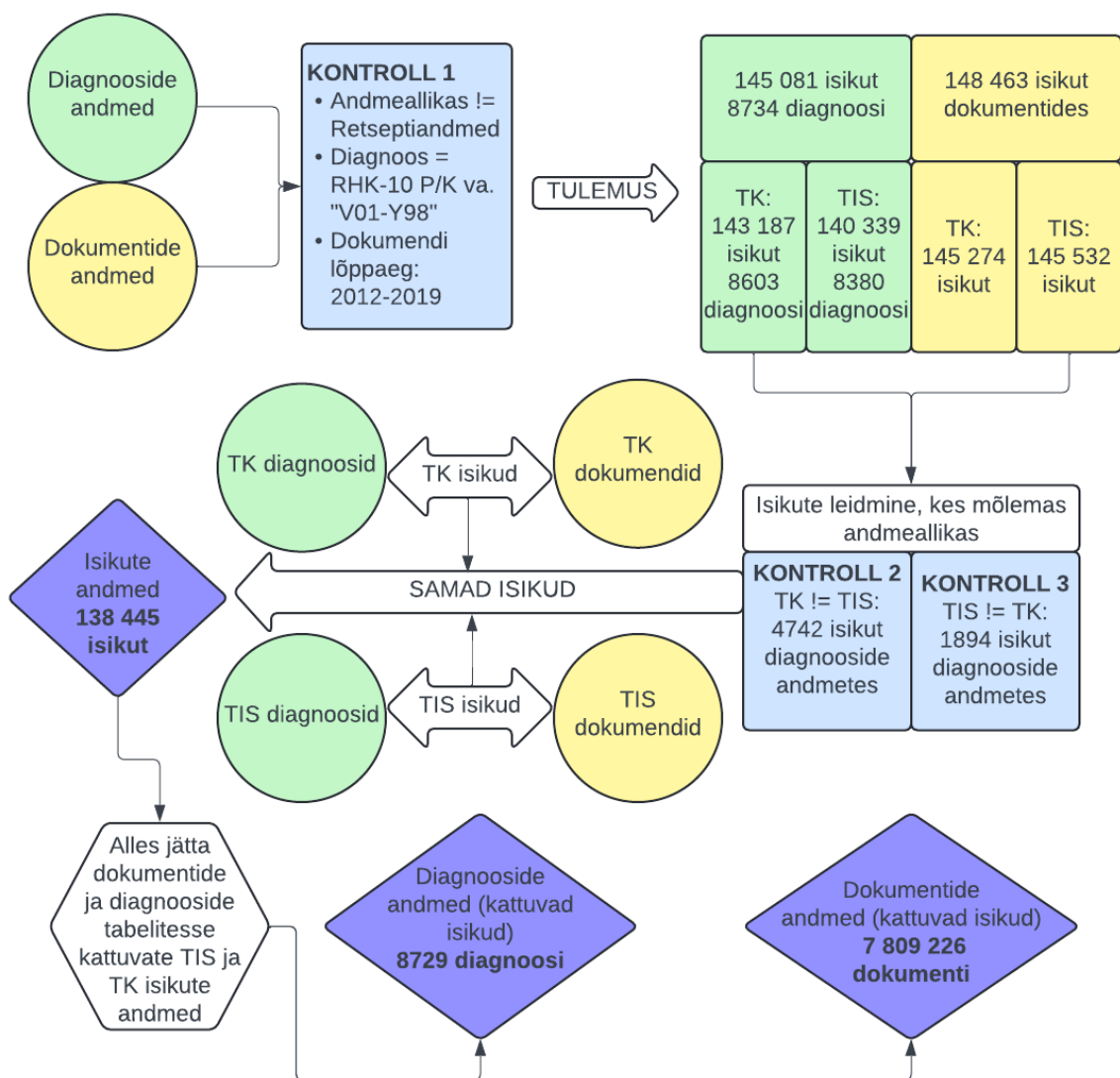
- Keskhaigla: üldjuhul peab keskhaigla suutma tagada tervishoiuteenuseid 100 000-200 000 inimesele. Keskhaiglas osutatakse üldhaigla tervishoiuteenuseid ja peale selle veel näiteks ortopeedia teenuseid. Eestis tegutseb neli keskhaiglat [27]
- Üldhaigla: üldjuhul peab üldhaigla suutma tagada tervishoiuteenuseid 50 000-100 000 inimesele. Üldhaigla pädevusse kuuluvad sisehaigused, üldkirurgia ja sünnitusabi. Eestis on selliseid haiglaid 12 [27]
- Kohalik haigla: kohalikuks haiglaks nimetatakse haiglaid, mis asuvad kõrgema etapi haiglatest kuni 70 km kaugusel ja kus kohalik haigla peab suutma tagada tervishoiuteenuseid kuni 40 000 inimesele. Sellistes haiglates üldjuhul ei ole olemas ööpäevaringset erakorralise kirurgia valveteenust. Selliseid on Eestis 1 [27]

Tervise Arengu Instituudi tervishoiuteenuse osutajate andmete (andmebaasis tabel TTO10) [28] alusel oli 2022. aastal Eestis kokku 1532 iseseisvat tervishoiuasutust. Antud töös on sees 1279 erinevat tervishoiuteenuse osutajat.

4.2. Andmete töötlus

Töö andmetabelitega toimus andmebaasirakenduses Datagrip (v.2023.2.3), kus oli võimalik teostada erinevaid SQL päringuid ning luua uusi töö jaoks vajaminevaid tabeleid. Andmete statistiliseks töötamiseks ja tulemuste visualiseerimiseks kasutati Jupyter Notebooki (v.6.14.12).

OMOP CMD kujul andmed sisaldasid andmeid, mida antud töös Tervisekassa raviarvete ja tervise infosüsteemi andmete võrdlemiseks vaja ei ole. OMOP-i andmemudelist olid kasutuses põhiliselt kaks tabelit “*visit_occurrence*” ja “*condition_occurrence*”. Esimene neist sisaldas andmeid haiguslugude dokumentide kohta ja teine andmeid haiguslugude diagnooside kohta – edaspidi kasutan esimese tabeli viitamiseks nimetust “dokumentide tabel” ja teise viitamiseks “diagnooside tabel”. Töös jäeti vaatluse alt välja teised potentsiaalselt RHK-10 diagnoose sisaldavad OMOP CDM andmemudeli tabelid *procedure_occurrence*, *observation*, *measurements* ning keskenduti *condition_occurrence* tabelis olevale infole.



Joonis 2. Tervisekassa raviarvete ja tervise infosüsteemi dokumentide andmete eeltöötlus protsess, kus lillad kujundid näitavad lõppanalüüsis olevate isikute, diagnooside ja dokumentide arvu.

Joonisel 2 on näidatud käesoleva töö andmete eeltöötlus protsessi. Nii dokumentide kui diagnooside tabelitest said välja filtreeritud retseptiandmed. Töös kasutati ainult RHK-10 diagnoosikoodidega andmeid, mille tüübiks oli märgitud kas põhidiagnoos või kaasuv diagnoos. Antud töös ei vaadatud välispõhjuseid (RHK-10 peatükk „V01-Y98“). Lisaks olid osade dokumentide kuupäevad varasemad kui 2012 või hilisemad kui 2019 ja seega sai lisatud veel üks kitsendus, kus dokumendi lõppaeg peab olema 2012 ja 2019 vahel. Peale filtreerimist jäi diagnoosi andmetesse 145 081 erinevat isikut, kellest 68 182 mehed ja 76 899 naised, ja 8734 erinevat RHK-10 diagnoosi. Samuti oli diagnooside ja dokumentides olevate isikute arv erinev, kus dokumentide andmetes oli 148 463 isikut, millest 145 274 olid Tervisekassa

raviarvete andmed ja 145 532 olid tervise infosüsteemi andmed. Andmestikku alles jäänud patsientide kohta oli kokku 7 834 342 erinevat dokumenti diagnooside kohta, millest 5 128 611 olid Tervisekassa dokumendid ja 2 705 731 olid TIS-i dokumendid. Tervisekassa raviarvete diagnooside andmetesse jäi 143 187 erinevat isikut ja tervise infosüsteemi andmetesse 140 339 erinevat isikut ning vastavalt 8603 erinevat diagnoosi ja 8380 erinevat diagnoosi. Isikute arvu erinevused diagnoosid ja dokumentide vahel võivad tulla sellest, et osadele dokumentidele oli märgitud diagnoos, mis OMOP andmemudelil läks tabelitesse *procedure_occurrence*, *observation* või *measurements* ning jäi vaatluse alt välja. Dokumentide andmetest on töös kasutusel TTO nimi, TTO tüüp, dokumendi algus- ja lõpp kuupäevad ja tervishoiuteenuse tüüp (ambulatoorne või statsionaarne). Kuna dokumentide ja diagnooside andmetes olevate isikute ja diagnooside arv ei ole sama Tervisekassa raviarvete ja tervise infosüsteemi andmetes, näitab see, et on isikuid ja diagnoose, mida on ühes andmeallikas ja teises pole. Antud töö kontekstis on oluline võrrelda Tervisekassa raviarvete ja tervise infosüsteemi andmeid, nii et mõlemas andmeallikas oleksid samad inimesed, kes esinevad mõlemas andmestikus. Kui võrrelda Tervisekassa raviarvete ja tervise infosüsteemide diagnooside tabelis olevaid isikuid, siis on 4742 isikut, kes olid Tervisekassa raviarvete diagnooside tabelis, aga polnud tervise infosüsteemi andmete diagnooside tabelis. Kui teha sama võrdlus dokumentide andmetega, siis oli 2373 isikut, kes olid Tervisekassa raviarvete dokumentide andmetes, aga mitte tervise infosüsteemi dokumentide andmetes. Kui teha võrdlus vastupidi, siis oli 1894 isikut diagnooside tabelis ja 3075 isikut dokumentide andmetes, kes olid tervise infosüsteemi andmetes, aga mitte Tervisekassa raviarvete andmetes.

Lisaks oli erinevus ka erinevate diagnooside arvus kahe andmeallika vahel. Kuna RHK-10 ei pruugi olla võimalikult täpse diagnoosiga diagnooside andmetes, kasutasin diagnooside arvude võrdlemiseks kolmandat ehk jaotiste taset (nt. A00). Edasises analüüsis vaatlen diagnoose RHK-10 kõige laiemal klassifikaatoriga ehk peatükkide lõikes (nt. A00-B99), kus on 22 erinevat peatükki, sest järgmine RHK-10 tase sisaldab 267 alampeatükki, mida on keeruline visuaalselt esitleda.

Antud töös on vaja võrrelda Tervisekassa raviarvete ja tervise infosüsteemi andmeid, nii et mõlemas andmestikus vaatleksime samu isikuid. Isikute tabelisse arvutasin juurde vanuserühma kasutades CASE lauset isikute liigitamiseks vanuserühmadesse nende sünniaasta alusel. See funktsioon arvutab iga inimese vanuse 2019. aasta 1. jaanuari seisuga ja määrab seejärel arvutatud vanuse isiku sünniaastast alusel vanuserühma. Vanuserühmad määrati 5-aastaste sammudega vahemikus 0-4 kuni 80-84, kusjuures kõik 85-aastased või vanemad

liigitati vanuserühma “85+”. Lisaks kasutasin kitsendust, et jätta alles ainult need isikud, kes vastavad kitsenduspäringutes konkreetsetele kriteeriumitele:

- Ühendasin Tervisekassa raviarvete diagnooside ja dokumentide tabelid, kus sisse jäid ainult RHK-10 põhidiagnoosid ja kaasuvad diagnoosid ja võtsin välja kõik erinevad isikud, kes esinesid nii diagnooside kui dokumentide tabelis.
- Ühendasin tervise infosüsteemi andmete diagnooside ja dokumentide tabelid, kus sisse jäid ainult RHK-10 põhidiagnoosid ja kaasuvad diagnoosid ja võtsin välja kõik erinevad isikud, kes esinesid nii diagnooside kui dokumentide tabelis.
- Kasutasin INTERSECT operaatorit, et kombineerida kahe erineva alampäringu tulemusi, nii et lõpptulemusse kaasatakse ainult mõlema alampäringu ühised isikud.

See jättis isikute tabelisse alles ainult inimesed, kes olid nii Tervisekassa raviarvete kui tervise infosüsteemi andmetes. Diagnooside ja dokumentide tabelisse jäid alles nende isikute andmed, kes olid uues isikute tabelis.

Uues isikute tabelis oli lõpuks järel 138 445 erinevat isikut, kellest 64 456 mehed ja 73 989 naised. Uues diagnooside tabelis oli alles 8729 erinevat diagnoosi, millest Tervisekassa raviarvetel oli 8600 erinevat diagnoosi ja tervise infosüsteemil 8376 erinevat diagnoosi. Uues dokumentide tabelis jäi alles 7 809 226 diagnoosiga seotud dokumenti, millest Tervisekassa raviarvete dokumente oli 5 107 311 ja TIS-i dokumente 2 701 915.

Lisaks eristati mõlema andmeallika puhul ravitüüpe: statsionaarne ravi ja ambulatoorne ravi. Statsionaarse ravi alla loeti kõik tervise infosüsteemi statsionaarsed epikriisid ja Tervisekassa raviarvete puhul statsionaarsed raviarved. Muud dokumenditüübid, mis esinesid tervise infosüsteemi andmetes, nagu ambulatoorsed epikriisid, saatekirjad ja saatekirjade vastused, loeti ambulatoorse ravi alla. Tervisekassa raviarvete puhul liigitati hambaravi, perearstiabi, päevaravi, ambulatoorne ja õendusabi ambulatoorse ravi alla.

Üheks töö eesmärgiks oli samuti vaadata, kui palju on täielikult kattuvaid andmeid, st. dubleerivaid diagnoosidega seotud andmeid Tervisekassa raviarvete ja Tervise Infosüsteemi andmete vahel ühiste andmetunnuste alusel. Täielikult kattuvad andmed pidid täitma kriteeriumeid:

- Erinev andmeallikas: Tervisekassa raviarve või tervise infosüsteemi andmed
- Sama isik

- Sama dokumendi algusaeg
- Sama dokumendi lõppaeg
- Sama põhidiagnoos
- Sama TTO

Täielikult kattuvatesse andmetesse jäi lõpuks 134 079 isikut, 7299 erinevat diagnoosi ja 3 161 908 dokumenti.

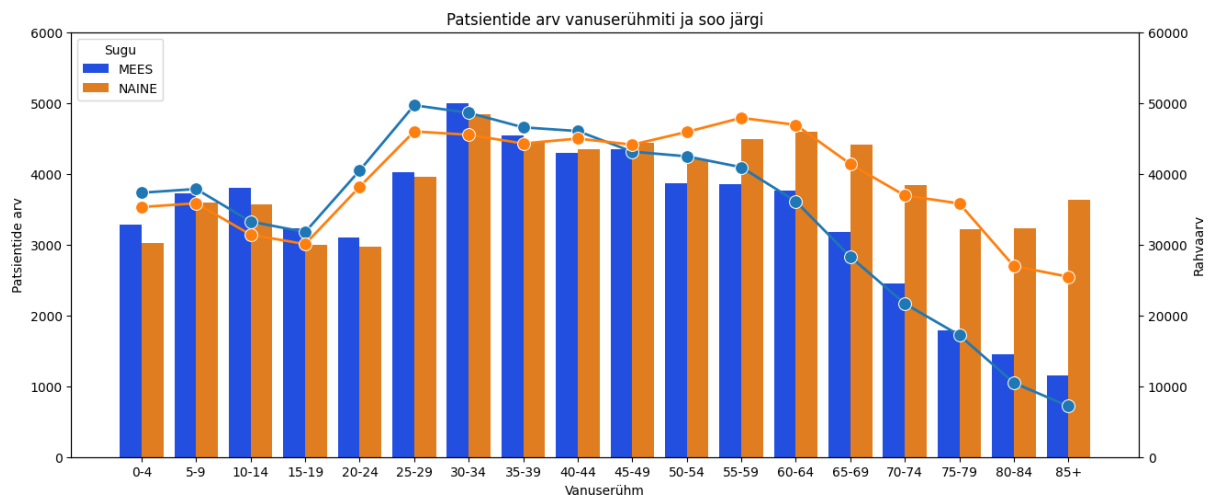
Kuna andmete eeltöötuse käigus näis, et raviarvetega seotud isikuid, diagnoose ja dokumente on rohkem kui TIS-i andmeid, siis protsentuaalsete osakaalude arvutamiseks on kasutatud töös valemit, mis näitab % kattuvust raviarvete osas TIS-i andmetest:

$$P(\%) = \frac{\textit{Raviarved} - \textit{tervise infosüsteemi andmed}}{\textit{Raviarved}} \times 100$$

5. Tulemused ja arutelu

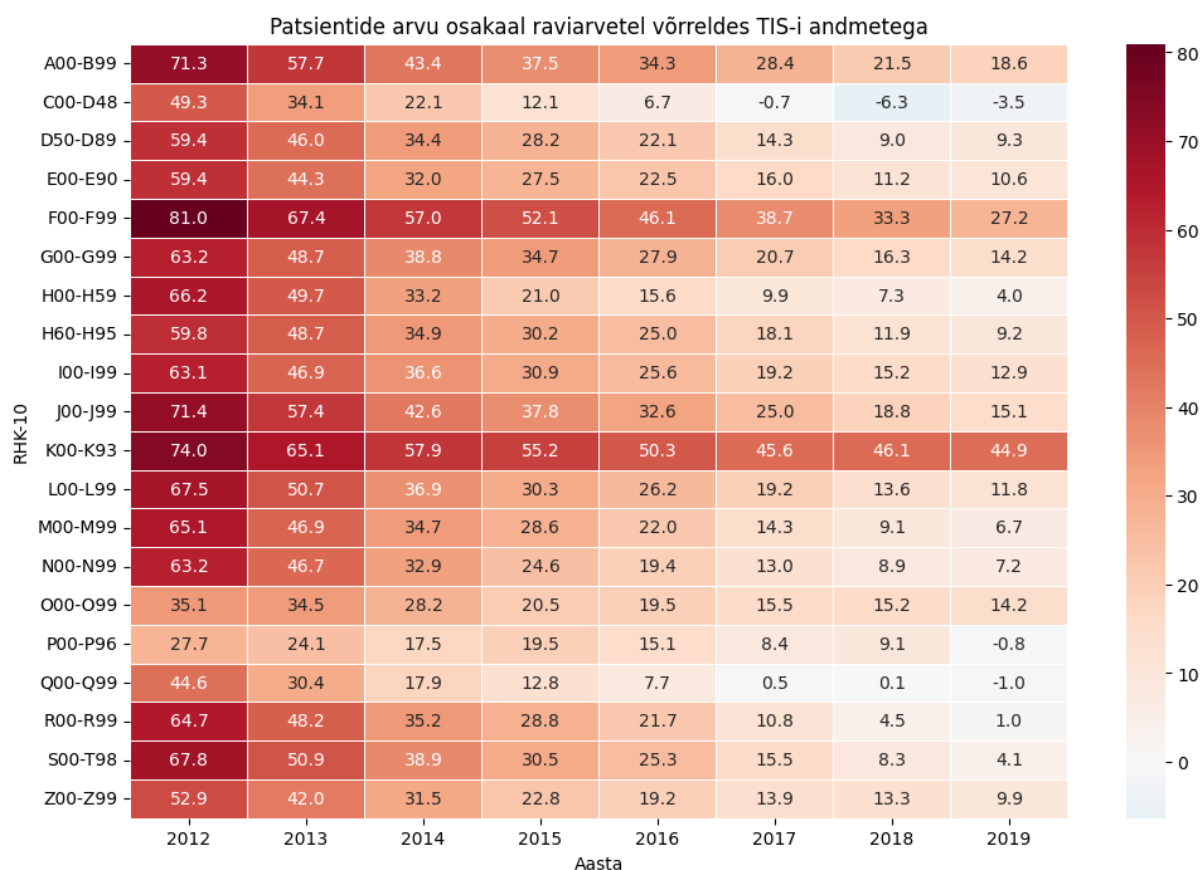
5.1. Patsientide võrdlus

Eelmises peatükis kirjeldati, millised andmed ja kuidas töödeldi, et oleks võimalik võrrelda Tervisekassa raviarveid ja Tervise Infosüsteemi andmeid. Peale andmete töötlust jäi andmestikku alles 138 445 patsienti, kellest 64 456 olid mehed ja 73 989 olid naised.



Joonis 3. Patsientide arv vanuserühmiti ja soo järgi näitab analüüsis sees olevaid patsientide arve soo ja vanuserühma järgi. Joonisele on lisatud ka populatsiooni kõver, mis on Statistikaameti „Rahvaarv ja rahvastiku koosseis“ RV021 [29] 2012-2019 keskmine inimeste arv vanuserühmiti. Joonisel on alles patsiendid, kes ei olnud surnud.

Joonisel 3 on näha, et üldjuhul populatsiooni kõver on samas suhtes patsientidega, kes käivad arsti juures. Lisaks näeme, et antud analüüsis on esindatud kõik vanuserühmas olevad inimesed. Näiteks näeme, et 0-39 vanuserühmades on meeste kohta rohkem terviseandmeid ja samas on meeste keskmine arv kõrgem 0-39 vanuserühmades. Alates vanuserühmast 45-49 on naise populatsioonis rohkem, mis põhjendab ka naiste suuremat osakaalu terviseandmetes, aga samas isikud vanuserühmades 20-29 käivad vähem arsti juures kui teised. Üheks põhjuseks miks vanemates vanuserühmades on naised rohkem, on see, et mehed elavad üldjuhul vähem kui naised. Seda kinnitab ka Statistikaameti oodatava eluea numbrid, mis näitab inimese sünnihetkel keskmiselt elada jäänud aastate arvu juhul kui suremus ei muutu. Statistikaamet on arvutanud 2019. aastal näiteks meeste oodatavaks elueaks 74,4 aastat ja naiste oodatavaks elueaks 82,7 aastat [30].



Joonis 4. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega näitab protsentuaalselt kui palju rohkem oli patsiente raviarvetes võrreldes elektrooniliste terviseandmetega RHK-10 peatükkide lõikes. Kui arv on negatiivne, siis oli TIS andmetes rohkem inimesi RHK-10 peatükkide lõikes kui Tervisekassa raviarvetel.

Vaadeldes Tervisekassa raviarvete ja Tervise Infosüsteemi andmeid huvitab meid, kas ja kui palju erinevad patsientide arvud RHK-10 peatükkide järgi. Joonisel 4 on näha, et üldjuhul elektrooniliste terviseandmete patsientide arv võrreldes raviarvetega tõuseb igal aastal, mis on tingitud sellest, et ajas on paranenud dokumentide edastamine TIS-i TTO-de poolt [31], aga patsientide erinevus jääb siiski suureks ka 2019. aastal järgmistes RHK-10 peatükkides:

- **A00-B99: Teatavad nakkus- ja parasiithaigused (18,6%)**
- **F00-F99: Psüühika- ja käitumishäired (27,2%)**
- **K00-K93: Seedeelundite haigused (44,9%)**

Nakkus- ja parasiithaiguste (A00-B99) diagnoosi alampeatükke vaadeldes on näha, et erinevused detailsemalt vaadates lisas 1 tulevad välja diagnoosides A00-A09 (soolenakkushaigused), A20-A28 (kindlad (teatavad) bakterzoonoosid), A65-A69 (muud spiroheetoosid e spiroheethaigused), A70-A74 (muud klamüdioosid e klamüüdiahaigused), A90-

A99 (lüljalgselevitatavad viiruspalavikud ja hemorraagilised viiruspalavikud), B00-B09 (naha- ja limaskestakahjustusega viirushaigused), B25-B34 (muud viirushaigused), B65-B83 (helmintiaasid e nugiussahaigused), mille puhul Tervisekassa raviarvetes on rohkem andmeid kui TIS-is. TTO-d on kohustatud esitama nakkushaigusi nii NAKIS-e infosüsteemi kui ka TIS-i [5: 22], mis tähendab TTO-de jaoks suuremat ajakulu, nii et võib juhtuda, et TTO saadab nakkushaiguste andmed ainult NAKIS-sse ja mitte TIS-i. Kusjuures on ka diagnoose, kus patsiente on rohkem TIS-i andmetes kui Tervisekassa andmetes 2016-2019 vahel: A15-A19 (tuberkuloos), A50-A64 (peamiselt suguliselt levivad nakkushaigused), A80-A89 (kesknärvisüsteemi viirushaigused), B15-B19 (viirushepatiit), B50-B64 (algloomahaigused), B85-B89 (pedikuloos, akariaas ja muud infestatsioonid e nugalisnakkused), B90-B94 (nakkus- ja parasiithaiguste jääknähud) ja B99-B99 (muud nakkushaigused). Teatavad nakkus- ja parasiithaiguste üldist erinevust võib põhjustada nakkus- ja parasiithaiguste kalli raviga [32: 34], mille puhul TTO-d on motiveeritud saatma rohkem raviarveid Tervisekassale kui TIS-i.

Psüühika ja käitumishäirete (F00-F99) alampeatükkide täpsema vaatluse tulemusena näib nagu lisas 2 näha, et enamus psüühika- ja käitumishäirete puhul on Tervisekassa andmetes patsiente rohkem kui TIS-i andmetes ja erinevused jäävad suureks 2019. aasta lõpuni. Samas jääb silma, et 2018. aastal on TIS-is rohkem patsientide andmeid kui Tervisekassa andmetes RHK-10 alampeatükis F99-F99 (täpsustamata psüühikahäire). Seda erinevust võib põhjendada sellega, et arst ei taha edastada patsiendi kohta käivaid tundlikke andmeid detailselt [5: 11] ja liigitab psüühikahäire täpsustamata koodi alla. Psüühika- ja käitumishäirete üldine statistika ei pruugi näidata tegelikku olukorda ühiskonnas. Üheks põhjuseks võib olla patsientide soov privaatsuseks ja nad ei ole tahtnud nõ. ametlikku diagnoosi saada. Samas võib see ka olla arstide praktika, kes ei soovi tundliku sisuga haiguslugusid TIS-i edastada [5: 11]. Kahe andmeallika andmete erinevuse vähenemisele võis kaasa aidata see, et 2011.a. sai valmis perearstidele suunatud depressiooni ravijuhis, st. patsiendid võisid saada abi juba perearstilt, mitte ei pidanud minema erialaspetsialisti juurde, kelle juurde võivad olla väga pikad järjekorrad. Lisaks jõustus 2016. aastal “E-konsultatsiooni saatekirja ja selle vastuse nõuded perearsti poolt patsiendi suunamisel psühhiaatri e-konsultatsiooni tervise infosüsteemi vahendusel” [33], mille puhul perearsti loodud saatekiri eriarstiabisse tehakse TIS-i vahendusel ja kõik selle saatekirjaga seotud toimingud nagu diagnoos, analüüsid, ravi jm on TIS-i andmetes olemas.

Seedeelundite haiguste (K00-K93) erinevusi tekitavad põhiliselt RHK-10 koodid K00-K14 (suuõõne-, süljenäärme- ja lõualuuhaigused) nagu lisas 3 näha, kusjuures 2018-2019 vahel on TIS-i edastatud kordades rohkem kui Tervisekassale andmeid RHK-10 koodides K35-K38

(ussripikuhaigused), K40-K46 (songad), K70-K77 (maksahaigused) ja K80-K87 (sapipõie-, sapiteede ja kõhunäärme-haigusseisundid). Lisaks on nähtav patsientide arvu erinevus ka K20-K31 (söögitoru-, mao- ja kaksteistsõrmikuhaigused) diagnoosides, kus 2019. aasta lõpuks on K20-K31 diagnoosiga patsiente 23,8% rohkem Tervisekassa andmetes kui TIS-i andmetes. Ülejäänud seedeelundite haigused on enam-vähem samad peale 2016. aastat kahe andmeallika alusel. RHK-10 koodide K00-K14 erinevus on tingitud sellest, et Tervisekassale edastatakse laste (alla 19 aasta) hambaravi arved, mis antud uuringus on olemas ja TIS-i hambaravikaart, mida antud uuringu andmete hulgas ei olnud. Mõningane kattuvus võib olla tingitud sellest, et mõned hambaarstid võisid olla koostanud ambulatoorse epikriisi, mis oleks antud juhul andmestikus olemas [33: 10]. Andmestikus ei ole täiskasvanu hambaraviga seotud arveid ja dokumente. Lisaks on Riigikontroll leidnud hambaravi pakkuvate TTO-dega suheldes, et tollane epikriisivorm hambaravi andmete saatmiseks ei sobinud TTO-dele ja oodati paremat andmevormi, millega hambaravi andmeid saata [5: 11]. Samas ei pruugi esimesed aastad peegeldada õiget erinevust Tervisekassa raviarvete ja TIS-i andmete vahel, sest on olnud juhtumeid, kus hambaravi teenust osutavad TTO-d petavad riigilt raha välja [34]. Aastate jooksul on parandatud kontrolle, et pettureid kätte saada.

Samas näeme ka joonisel 4, et on RHK-10 peatükke, kus raviarvetel on vähem patsiente kui tervise infosüsteemis, aga kuna need on haruldased ja üksikud juhtumid, siis käesolevas töös neid edasi ei analüüsita:

- C00-D48: Kasvajad (-3,5%)
- P00-P96: Perinataal- e sünniperioodis tekkivad teatavad seisundid (-0.8%)
- Q00-Q99: Kaasasündinud väärarendid, deformatsioonid ja kromosoomianomaaliad (-1%)

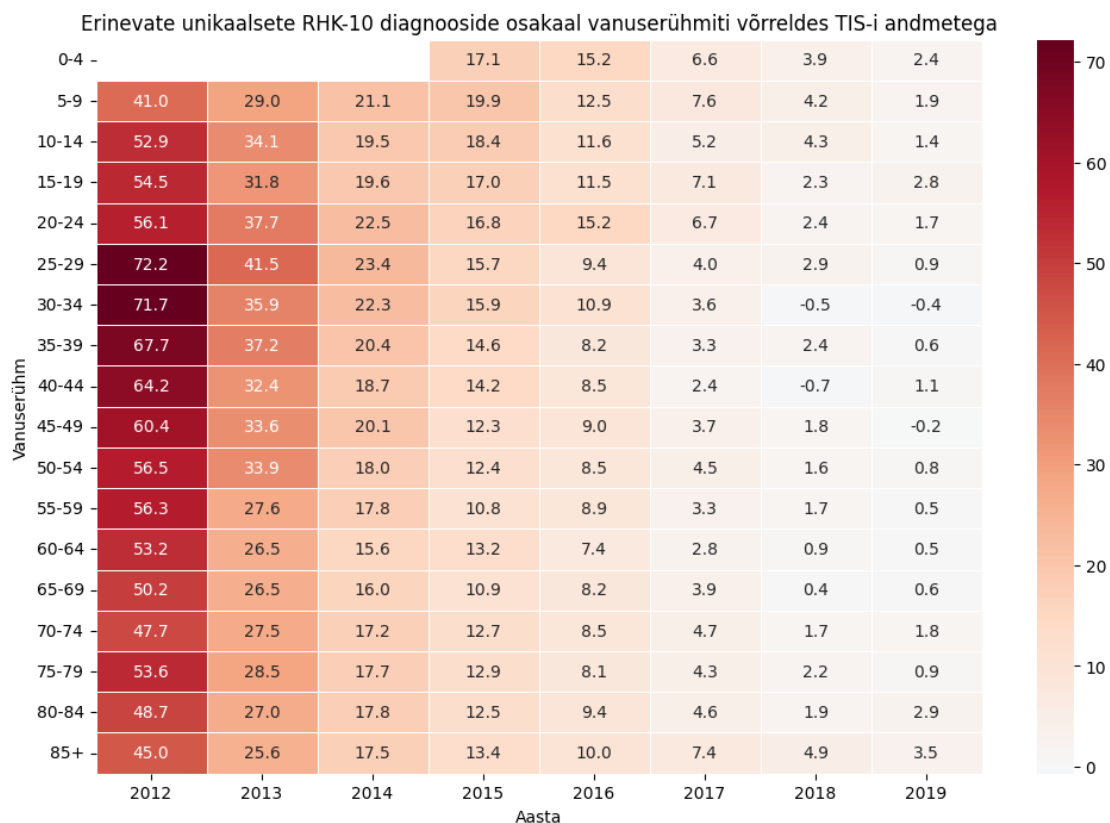
5.2. Diagnooside võrdlus

Peale andmete töötlust jäi alles andmetesse 8729 erinevat diagnoosi, millest Tervisekassa raviarvete alusel 8600 diagnoosi ja tervise infosüsteemi andmete alusel 8376 diagnoosi. Erinevate diagnooside arvu arvutamiseks võrreldi diagnoosikode nii nagu nad andmestikus

olid, st. ei muudetud diagnoosi koodide täpsust. Need numbrid juba näitavad, et on diagnooside arvude erinevus kahe andmeallika vahel. Kui vaadata täpsemalt põhi- ja kaasuvaid diagnoose, siis jagunevad need järgnevalt nagu tabelis 1 näha:

Tabel 1. Diagnoosi tüüpide jagunemine andmeallika alusel

| ANDMEALLIKAS | DIAGNOOSI TÜÜP | DIAGNOOSIDE ARV |
|-----------------------------|-----------------|-----------------|
| Tervisekassa raviarved | Põhidiagnoos | 8036 |
| Tervisekassa raviarved | Kaasuv diagnoos | 7921 |
| Tervise infosüsteemi andmed | Põhidiagnoos | 7727 |
| Tervise infosüsteemi andmed | Kaasuv diagnoos | 7674 |

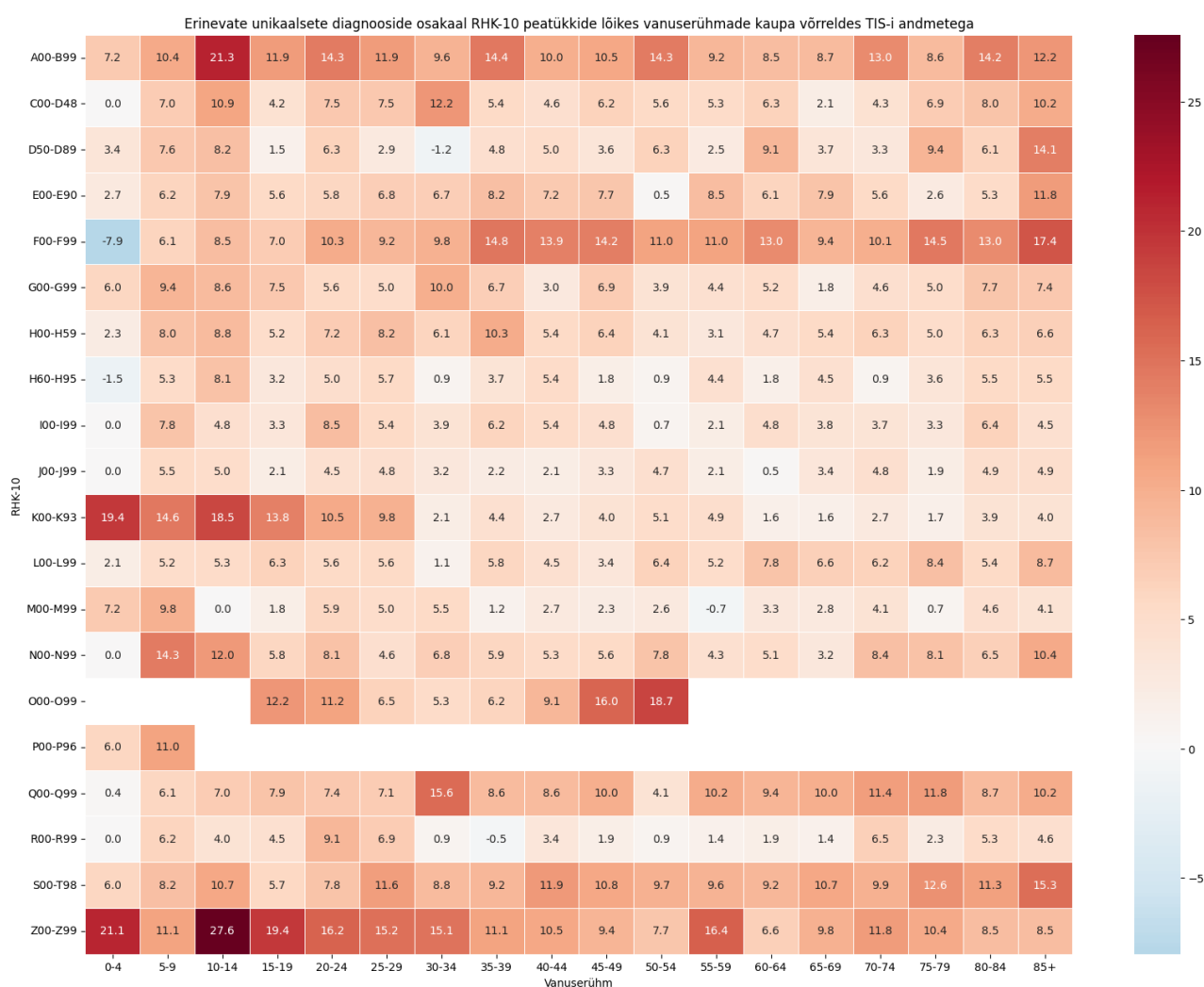


Joonis 5. Erinevate unikaalsete RHK-10 diagnooside osakaal vanuserühmiti võrreldes TIS-i andmetega näitab erinevate RHK-10 diagnooside osakaalu vanuserühmades ja aastate lõikes Tervisekassa raviarvete ja TIS-i andmete vahel. 2012-2014 aastatel on puudu 0-4 vanuserühm, sest vanus sai arvutatud 2019. aasta algusest.

Joonisel 5 on näha, kuidas 2012 ja 2013 võisid erinevate diagnooside arvud lausa 40-70% ulatuses erineda. Esimeste aastate erinevust võib põhjendada sellega, et 2012. aastal alles arendati TIS-i. Lisaks 2012. aasta lõpuks edastasid ainult 34% asutusi andmeid TIS-i, kuigi enamused TTO-d olid selleks ajaks juba ühinenud TIS-ga [5: 10]. Üldjuhul on aga erinevate

diagnooside arvud kahe andmeallika vahel 2019. a peaaegu ühtlustunud. Näeme ka, et kahe andmeallika erinevused vähenevad alates 2014. aastat, mida võis põhjustada seaduses määratletud kontroll Terviseameti poolt TIS-i andmete suhtes [35]. Alates 2014. aasta kontrollib Terviseamet TIS-i andmehõivet ja TTO-de andmete edastamist ning vajadusel tehakse ettekirjutusi või määratakse TTO-le rahatrahv nõuete mitte kinnipidamisest.

Tervisekassa raviarvete andmetes oli 1 diagnoos, mida ei olnud Tervise Infosüsteemi andmetes, ja vastupidi, Tervise Infosüsteemi andmetes oli 17 diagnoosi, mida ei olnud Tervisekassa raviarvete andmetes. Kuna tegemist on haruldaste ja harva esinevate haigustega, siis käesolevas töös neid eraldi ei vaadata.



Joonis 6. Erinevate unikaalsete diagnooside osakaal RHK-10 peatükkide lõikes vanuserühmade kaupa võrreldes TIS-i andmetega näitab vanuserühma ja RHK-10 peatükkide lõikes kui palju protsentuaalselt oli erinevaid diagnoose rohkem Tervisekassa raviarvetel kui TIS-i andmetes. Kui joonisel on tühi valge kast, siis järelikult oli tegemist haruldaste juhtumitega, mis antud jooniselt eemaldati. Haruldase juhtumina arvestati diagnoose, mille esinemine oli väiksem kui 15 vanuserühmas.

Joonis 6 toob välja erinevused Tervisekassa raviarvete ja tervise infosüsteemi andmete vahel vanuserühmiti. Diagnoosides O00-O99 (rasedus, sünnitus ja sünnitusjärgne periood) ja P00-P96 (perinataal- e sünniperioodis tekkivad teatavad seisundid) olid üksikud juhtumid kas väga noorte seas või vanemate inimeste seas, mis said antud jooniselt eemaldatud. O00-O99 diagnoosid on pandud ka patsientidele, kes on vanuserühmas 50-54, aga kuna vanus arvutati 2019. aastast, võisid patsiendid diagnoosi saamise ajal olla nooremad kui 50. Samas kui tegu on sünnitusega 50-54 vanuserühmas, võivad antud protsendid olla loogilised, sest vanemate naiste sünnitamine võib põhjustada nii emale kui lapsele edasisi komplikatsioone. Näeme, et ka Q00-Q99 (Kaasasündinud väärendid, deformatsioonid ja kromosoomianomaaliad) puhul on Tervisekassa raviarvetes rohkem andmeid kui TIS-i andmetes ja seda just vanemates vanuserühmades, mida võib põhjendada sellega, et lapse kohta käivad Q00-Q99 diagnoosid pandi vanemate andmete alla või tegemist on täiskasvanu eas leitud Q00-Q99 diagnoosiga. Põhjust, miks on O00-O99, P00-P96 ja Q00-Q99 diagnoose rohkem Tervisekassa andmetes kui TIS-i andmetes, antud töös ei leitud.

Samas näeme ikkagi üldiselt, et Tervisekassa raviarvetel on rohkem erinevaid diagnooside andmeid vanuserühmiti:

- **Z00-Z99: Tervise seisundit mõjustavad tegurid ja kontaktid terviseteenistusega** suuremad erinevused on vanuserühmades 0-4, 10-14, 15-19, 20-24, 25-29, 30-34 ja 55-59, mida põhjendab sagedasem külastus perearsti juurde tervise seisundi kontrolliks ja määratletud perearstiabi kvaliteedisüsteemi nõuete kinnipidamisega [36].
- **K00-K93: Seedeelundite haigused**, kus suurem erinevus on 0-29 aastaste seas, mis on tingitud sellest, et laste hambaravi arve on antud andmestikus olemas, aga TIS-i edastatud hambaravikaarte ei ole. Seda toetab tõsiasi, et K00-K93 erinevuse tekib peamiselt alapeatükis K00-K14 (suuõõne-, süljenäärme- ja lõualuuhaigused), kus on näiteks hambasööbija ja hammaste arengu häirete diagnoosid, mida tihti pannakse lastele.
- **F00-F99: Psüühika- ja käitumishäired**, mille puhul erinevus on pigem suurem keskealistel ja vanematel patsientidel, kes käivad rohkem arsti juures vaimsete häirete puhul, sest vanemad inimesed võivad abi saada eriarstiabist vaimsete häirete puhul, samas kui nooremate puhul tegeleb vaimsete häiretega perearst. Erinevus võib tulla patsientide soovist privaatsusele ja leida ravi väljaspool perearstiabi [37]. Juhul kui patsiendid lähevad eriarstiabisse abi otsima, ei pruugi nemad andmeid TIS-i edastada.

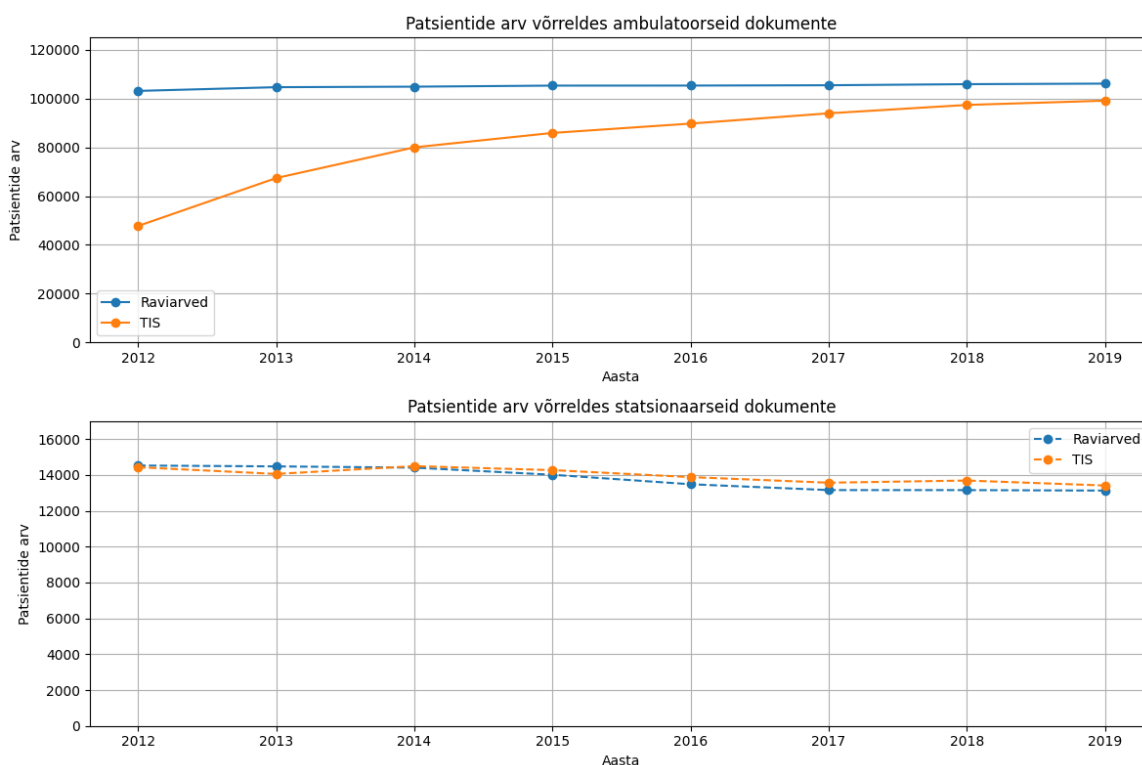
2012. aastal edastas TIS-i ainult 74% eriarstiabi teenust osutavatest haiglatest epikriise [5: 10]. Lisaks kõikidest TIS-ga liitunud eriarstiabi teenuse osutajatest oli ainult 8% asutustest mingi osa andmeid TIS-i saatnud [38: 11]. Joonisel 6 on samuti näha, et F00-F99 diagnoose on rohkem TIS-i andmetes 0-4 vanuserühmas. F00-F99 vahemikus jaotisi täpsemalt vaadates jäi silma, et TIS-i andmetes on erinevaid diagnoose 0-4 vanuserühmas rohkem kui Tervisekassa andmetes F07 (isiksus- ja käitumishäired ajuhaigusest, -kahjustusest või -düsfunksioonist), F41 (muud ärevushäired), F43 (rasked stressreaktsioonid ja kohanemishäired), F45 (somatoformsed häired), F50 (söömishäired), F94 (lapse või nooruki suhtlemishäired). Üldiselt oleneb psüühika ja käitumishäirete andmete TIS-i saatmine arstist endast, sest paljud ei taha põhimõtte pärast andmeid TIS-i saata [5: 11].

5.3. Statsionaarne vs. ambulatoorne võrdlus

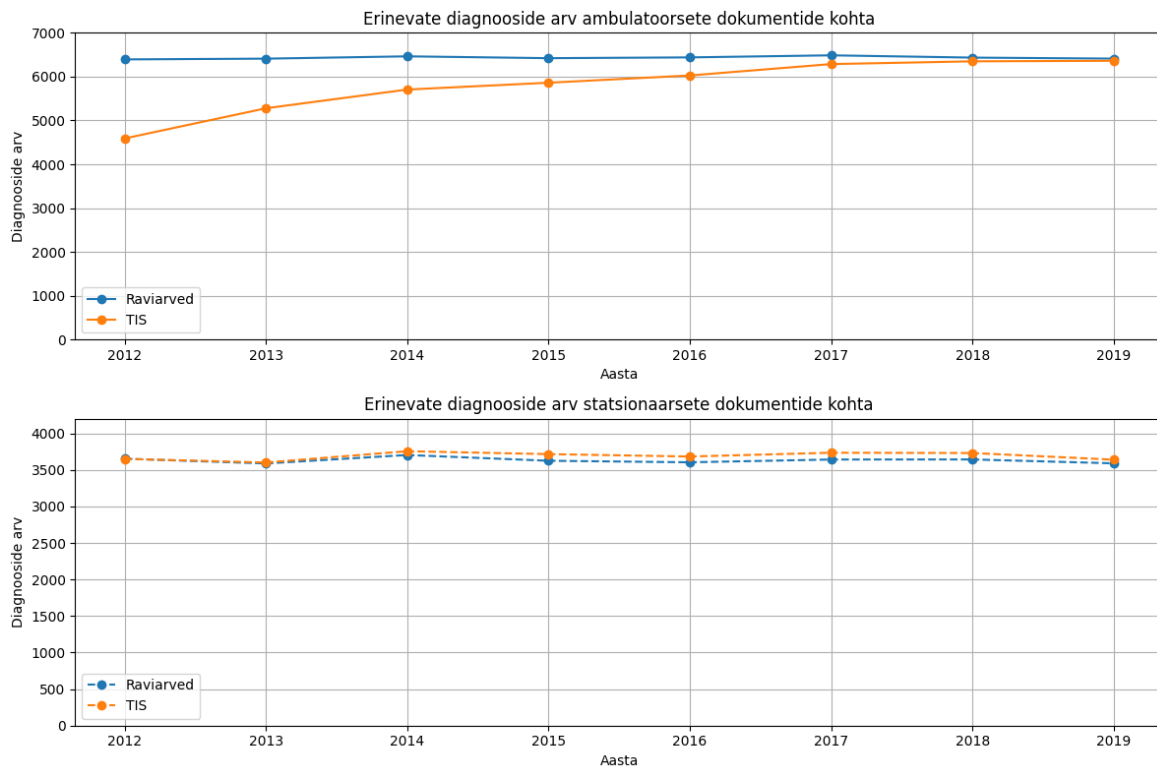
Statsionaarseid ja ambulatoorseid haigusjuhtumeid on võrreldud ennem Riigikontrolli poolt, kus võeti aluseks 2012. aasta Tervisekassa raviarved ja TIS-i andmed, mis on ka siinses töös sees osaliselt. Riigikontrolli terviseandmete aruandes Riigikogule 2014. aastal toodi välja, et statsionaarsete ja ambulatoorsete epikriiside vähesus TIS-is on TTO-de teadmatusest põhjustatud, sest TTO-d ei ole ühel meelel, millal peaks epikriisi koostama. Aruandes on seletatud, et eriti problemaatiline on ambulatoorsete epikriiside esitamine, sest epikriisi koostamise aeg sõltub sellest, millal TTO arvab, et peaks esitama. Nii võib juhtuda, et TIS-i jõuab mitme raviarve kohta üks TIS-i epikriis või jõuab TIS-i iga raviarve kohta epikriis. Analüüsi käigus leiti, et statsionaarseid epikriise võrreldes Tervisekassa raviarvetega oli 99% ulatuses olemas, samas ambulatoorseid epikriise oli ainult 48% ulatuses võrreldes Tervisekassa raviarvetega [5: 10]. Selle tulemusel esitatakse järjepidevalt andmeid nii elektroonilistes terviseandmetes kui raviarvetes, mis näitab nagu ka joonisel 7, 8, 9 näha, et statsionaarse ravi puhul on andmed ühtlustunud kahe andmeallika vahel. Arvestades, et statsionaarne ravi tähendab patsiendi jäämist haiglasse, pikemat patsiendi jälgimist ja erinevate protseduuride tegemist, siis statsionaarse ravi puhul on tavaliselt teenuste hüvitamismäärad kõrgemad, rangemad arveldusnõuded ja kliiniliselt omavad tähtsust statsionaarse ravi raviandmed rohkem väärtust kui ambulatoorse ravi raviandmed [35]. Selle tulemusel kontrollitakse ja järgitakse nii

TIS-i edastavate dokumentide kui raviarvete statsionaarse ravi andmete kodeerimisstandardeid ja dokumenteerimistavasid.

Ambulatoorne ravi on andmekoosseisu osas erinevam kui statsionaarne ravi, kus laias valikus tervishoiuteenuseid pakutakse erinevates kohtades nagu kliinikutes, arstide kabinettides ja haiglate ambulatoorsetes osakondades. See andmekoosseisu mitmekesisus võib põhjustada dokumenteerimistavade ja kodeerimistavade erinevust TTO-de vahel kuna on palju erinevaid detaile, mida täita iga erineva ambulatoorse ravi puhul. Osad TTO-d võivad olla harjunud esitama hoopis teisi andmeid kui teised TTO-d. Lisaks võib ambulatoorse ravi korral võrreldes statsionaarse raviga olla rohkem rõhku diagnostilisel kodeerimisel kui protseduurilisel kodeerimisel [39].

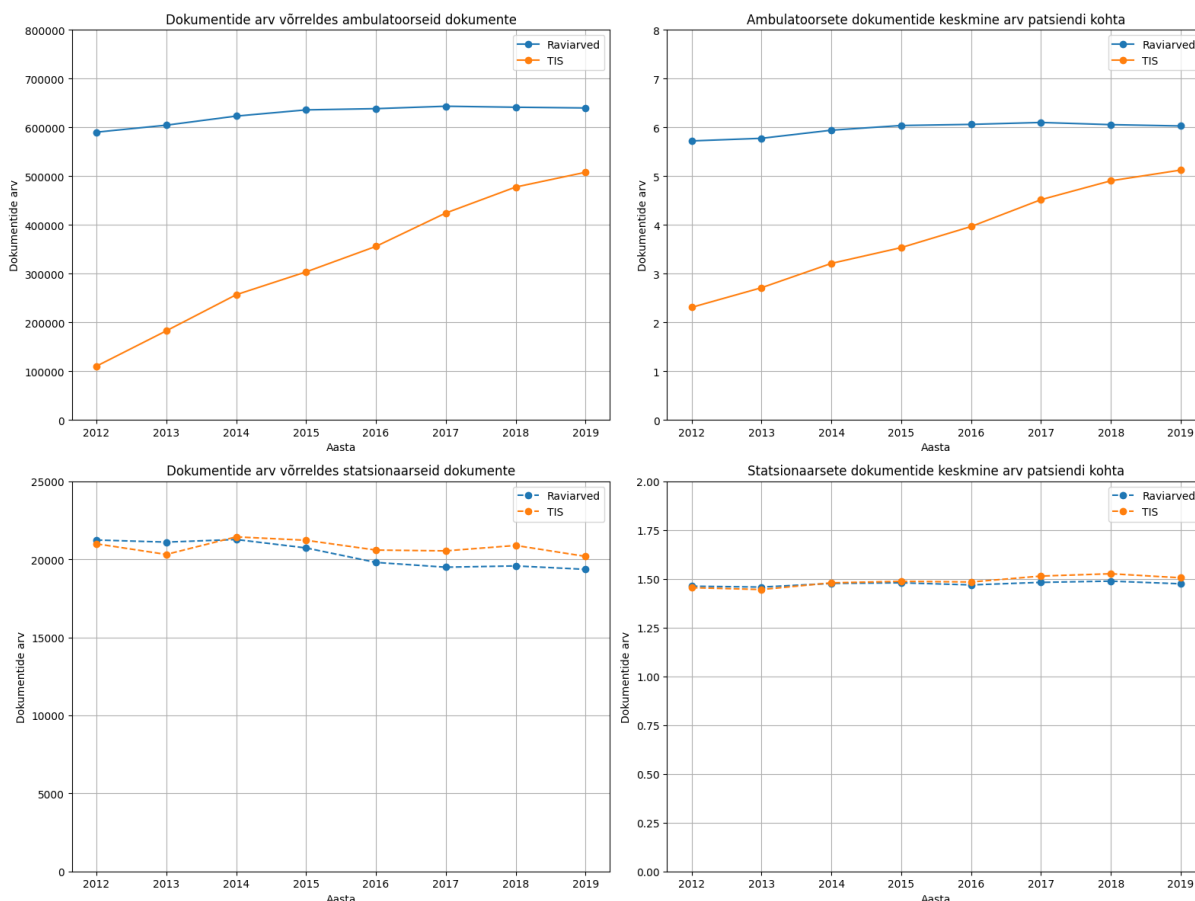


Joonis 7. Patsientide arv võrreldes ambulatoorseid ja statsionaarseid dokumente näitab patsientide arvu ambulatoorsete ja statsionaarsete dokumentide korral Tervisekassa raviarvete ja TIS-i andmete suhtes.



Joonis 8. Erinevate diagnooside arv ambulatoorsete ja statsionaarsete dokumentide kohta näitab erinevate diagnooside arvu ambulatoorsete ja statsionaarsete dokumentide korral Tervisekassa raviarvete ja TIS-i andmete puhul.

Joonisel 7 näeme, et statsionaarse ravi puhul on elektroonilised terviseandmed ja raviarved enam-vähem ühtlased, aga erinevus on ambulatoorse ravi puhul, kus aastate jooksul erinevus raviarvetega küll väheneb, aga vahe on ikkagi märgatav. 2012. a TIS-i ambulatoorseid andmeid on uuritud tervishoiuteenuse osutajate kohustuse võtmes, kus leiti, et 2012.a. I kvartalis olid esitanud ainult 31,6% tervishoiuteenuse asutajaid andmeid. Samas on aastate jooksul tervise infosüsteemiandmed paranenud tarkvara paranemise ja kasutamisele võtmisega, kvaliteedikontrollide lisamisega ja uute dokumentide tüüpide kasutusele võtmisega [40]. Lisaks leiti Rahvusarhiivi poolt tehtud terviseandmete arhiiviväärtuslikkuse hindamises, et peaaegu 100% statsionaarsetest haigusjuhtumitest edastatakse TIS-i, samas kui ambulatoorseid haigusjuhtumeid edastatakse umbes 50-60% [12: 9]. Joonisel 8 on näha, et isegi kui erinevate diagnooside arv oli aastate algul ambulatoorsetes juhtumites erinev ja Tervisekassa andmetes oli rohkem erinevaid diagnoose, siis 2019. aasta lõpuks on erinevate diagnooside andmed kahe andmeallika vahel ühtlustunud. Samas näeme, et statsionaarsete juhtumite osas on erinevate diagnooside arv peaaegu alati sama olnud.

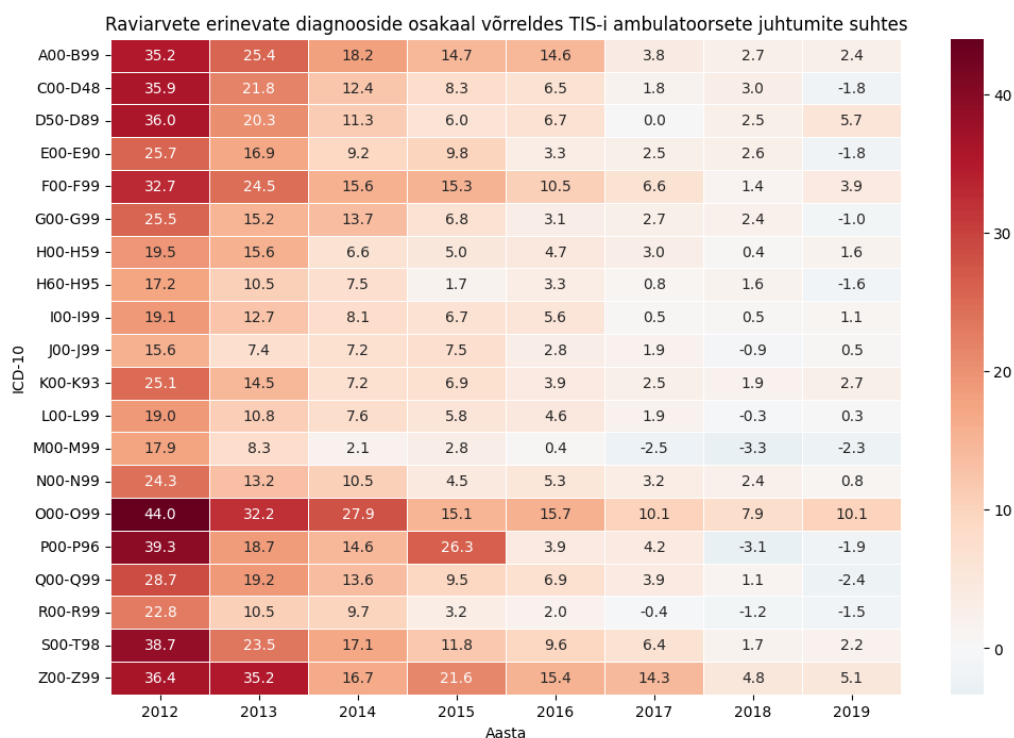


Joonis 9. Statsionaarsete ja ambulatoorsete dokumentide arv ning nende keskmine dokumentide arv patsiendi kohta näitab dokumentide arvu ja dokumentide keskmist arvu patsiendi kohta Tervisekassa ja TIS-i andmete vahel.

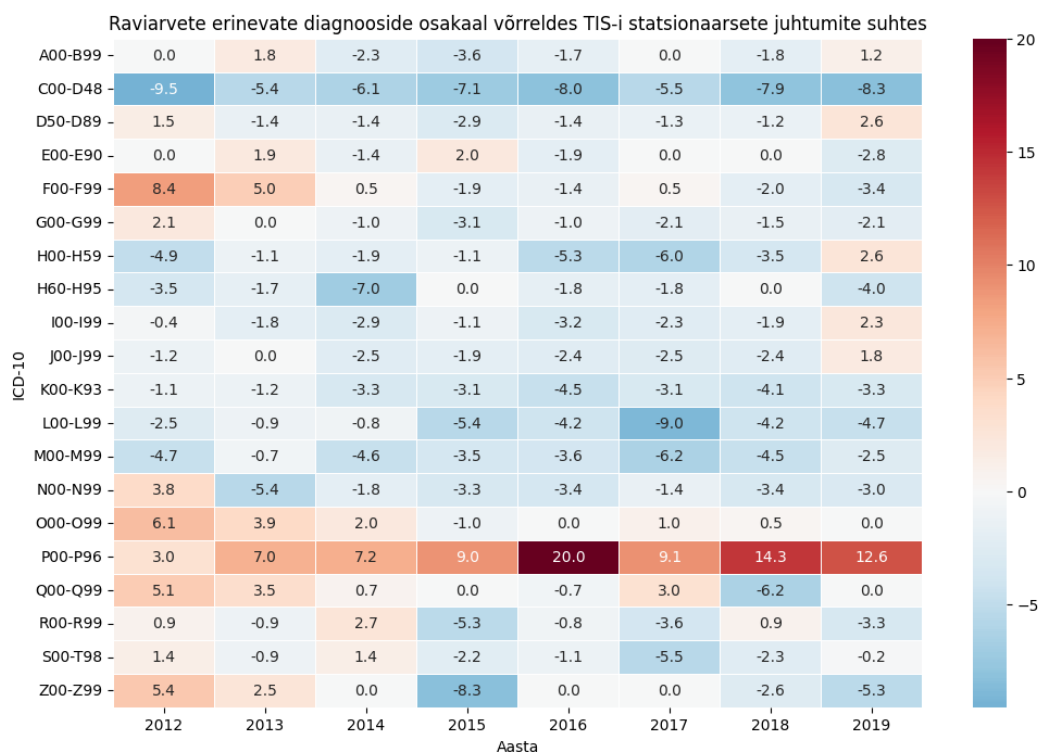
Elektroonilised terviseandmed ja raviarved täidavad erinevaid eesmärgi tervishoiusüsteemis, mis võib põhjustada dokumentide arvu erinevuse patsiendi kohta. Näeme joonisel 9, et ambulatoorse ravi dokumentide keskmine arv patsiendi kohta oli TIS-i andmetes 2012.aastal 2-3 korda madalam kui Tervisekassa andmed. Aastatega on vahe Tervisekassa ambulatoorse ravi dokumentide keskmine arv patsiendi kohta vähenenud. Põhjuseks võib olla asjaolu, et aastaid tagasi oli elektrooniliste terviseandmete saatmine kompaktse struktuuriga, kus saadeti võimalikult palju andmeid ühe korraga samal vormil. Tänapäevaks on tekkinud tervise infosüsteemi edastavate dokumentide andmekoosseisu kohta seaduses kohustus, kus on määratletud dokumentide saatmine ja nende andmekoosseis [9].

Nagu oleme juba mujal töös maininud, raviarvete andmed luuakse peamiselt tervishoiuteenuste arvelduse ja nende hüvitamise eesmärgil. Selle tulemusena sisaldavad need üksikasjalikult teavet patsiendile osutatud tervishoiuteenuste kohta nagu protseduurid ja muid tervishoiuteenuse osutamist tõendavaid dokumente, kus igaüks neist elementidest genereerib

eraldi dokumendi. Raviarved tuleb edastada riikliku ravikindlustuse asutusele iga kord kui osutatakse tervishoiuteenuseid ja esitatakse arveid. Selle tulemusena võib tekkida mitu dokumenti sama haigusjuhtumi kohta kui osutatakse mitut tervishoiuteenust või kui patsiendil on haigus, mis vajab pidevalt ravi jm tervishoiuteenuse osutaja teenuseid. Tervishoiuteenuse osutajad edastavad raviarved Tervisekassale kindlas vormis [41], kus on määratletud reeglid, mida TTO-d peavad täitma. Näiteks kui raviarve saata ja tegemist ei ole üldarstiabiga, siis arvel peab olema märgitud ainult üks lepingu eriala. Juhul kui patsient käib mitme erineva arsti juures, tähendaks see igakord uue raviarve saatmist.



Joonis 10. Raviarvete erinevate diagnooside osakaal võrreldes TIS-i ambulatoorsete juhtumite suhtes näitab erinevate RHK-10 diagnooside osakaalu ambulatoorsete juhtumite suhtes, kus Tervisekassa raviarvetega seotud diagnoose on protsentuaalselt rohkem kui TIS-i andmetega seotud diagnoose.



Joonis 11. Raviarvete erinevate diagnooside osakaal võrreldes TIS-i statsionaarsete juhtumite suhtes näitab erinevate RHK-10 diagnooside osakaalu statsionaarsete juhtumite suhtes, kus TIS-i andmetes on üldjuhul rohkem erinevaid diagnoose kui Tervisekassa raviarvetes.

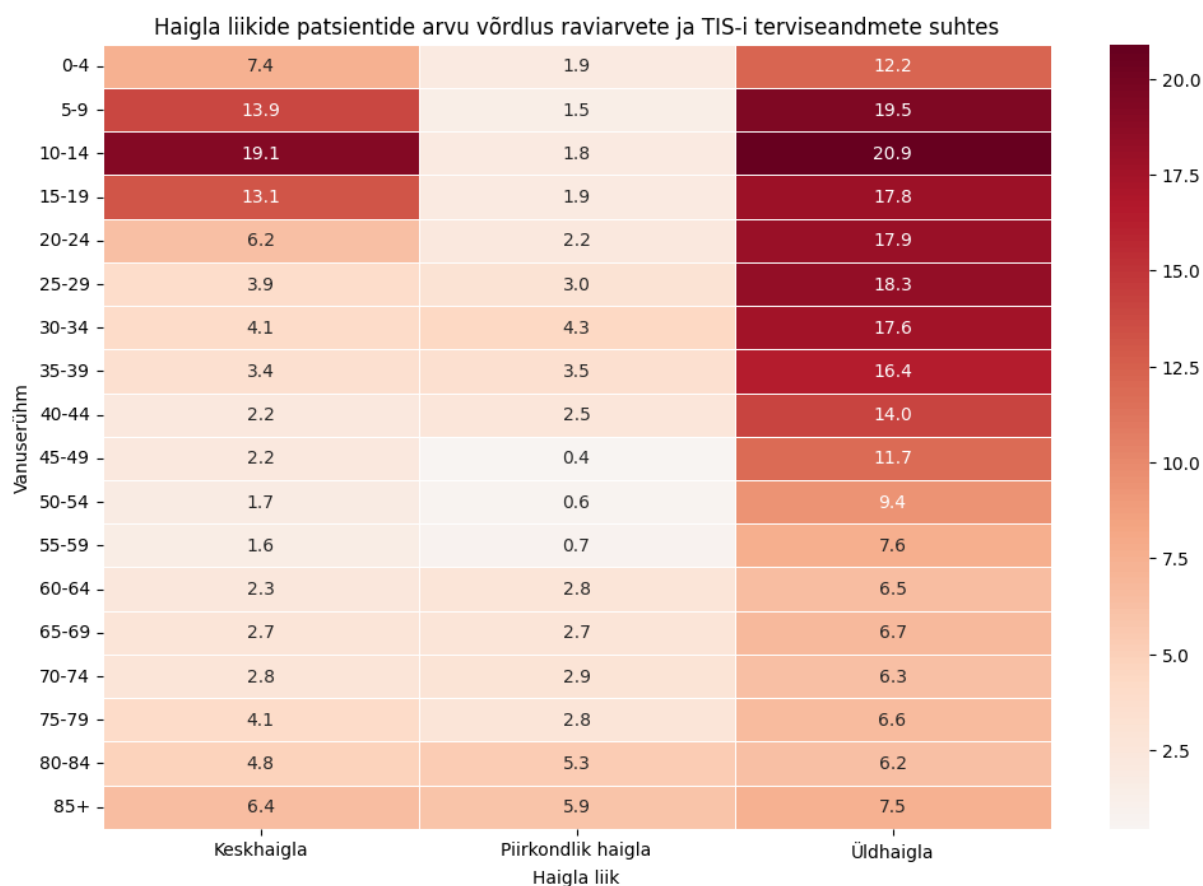
Joonistel 11 ja 12 näeme, et Tervisekassa raviarvete alusel olevate erinevate diagnooside arv on suurem ambulatoorsete juhtumite osas. Samas on statsionaarsete juhtumite osas TIS-is rohkem erinevaid diagnoose kui Tervisekassa raviarvete andmetes. Statsionaarsete juhtumite puhul on protsendilised erinevused väiksemad TIS-i ja Tervisekassa vahel. Näeme küll suuremaid protsendilisi erinevusi P00-P96 puhul, aga kuna tegemist on üksikjuhtumitega, siis näib protsendiline erinevus suur. Käesolevas töös ei leitud võimalikku põhjust, miks on Tervisekassa statsionaarsete juhtumite raviarvetes rohkem P00-P96 diagnoos kui TIS-i omades. Üldiselt näeme, et statsionaarsete juhtumite erinevus kahe andmeallika vahel on väiksem kui ambulatoorsete juhtumite puhul, mida leidis ka Riigikontroll 2012. aasta Tervisekassa ja TIS-i andmetes [5: 10].

5.4. TTO tüüpide ja haigla liigi võrdlus

Tervishoiuteenuse osutajad, kes on haiglad, peavad järgima tervishoiuteenuste korraldamise seadust, et pakkuda vastavaid tervishoiuteenuseid. Siinses analüüsis olid 1279 tervishoiuteenuse osutaja andmed, millest 1166 esinesid Tervisekassa raviarvete andmetes ja 921 esinesid Tervise Infosüsteemi andmetes. Võrreldes TTO-de tüüpide esinemist Tervisekassa raviarvete ja Tervise Infosüsteemi andmete vahel, näeme tabelis 2 et TTO-de arv on sama üldhaigla, keskhaigla, piirkondliku haigla suhtes. See näitab, et haiglad täidavad TIS-i andmete edastamise kohustust, aga erinevus on ülejäänud TTO-des, kes võivad olla küll liitunud TIS-ga, aga sinna andmeid ei saada.

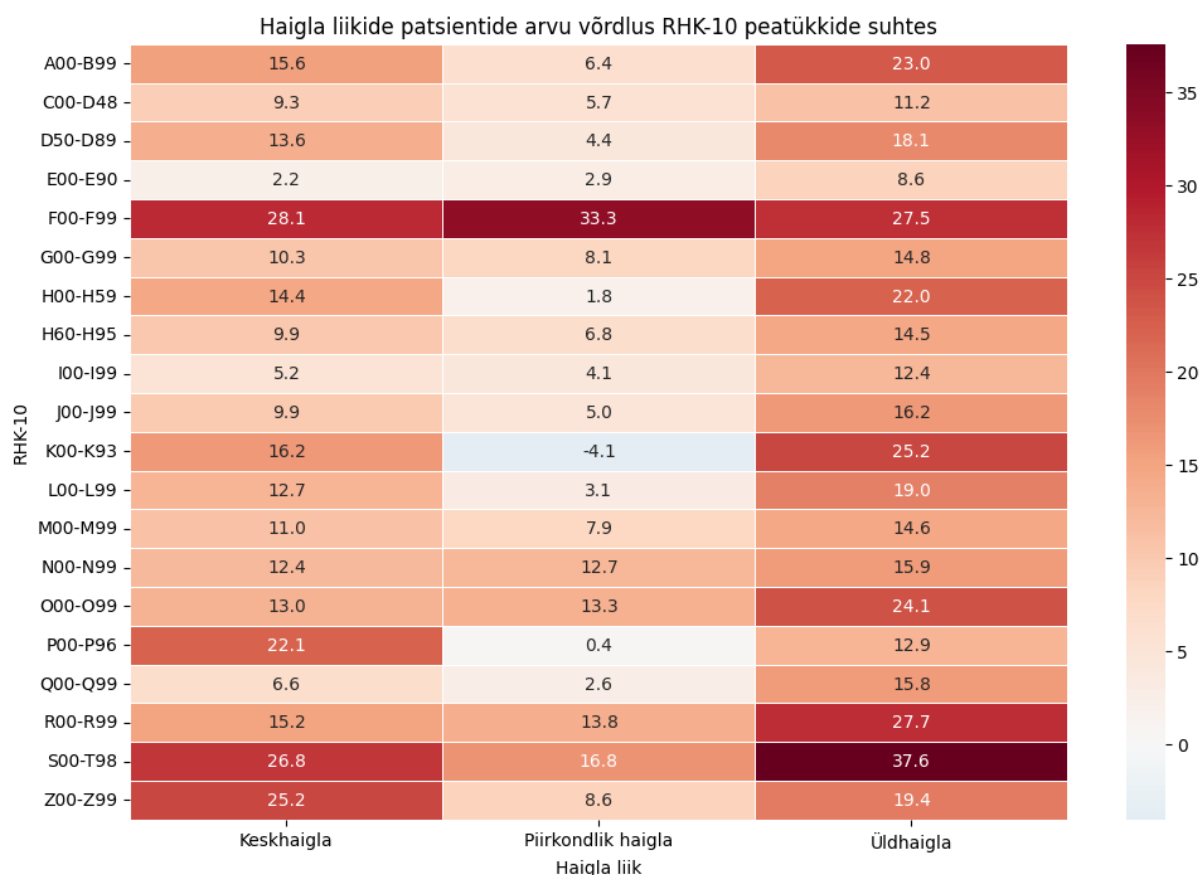
Tabel 2. Tervishoiuteenuse osutaja tüüpide TTO-de arv

| TERVISHOIUTEENUSE OSUTAJA TÜÜP | TTO-de ARV | TERVISEKASSA RAVIARVETEL | TERVISE INFOSÜSTEEMI ANDMETEL |
|---------------------------------------|-------------------|---------------------------------|--------------------------------------|
| Teised TTO-d | 1259 | 1146 | 901 |
| Üldhaigla + Kohalik haigla | 12 | 12 | 12 |
| Keskhaigla | 4 | 4 | 4 |
| Piirkondlik haigla | 3 | 3 | 3 |



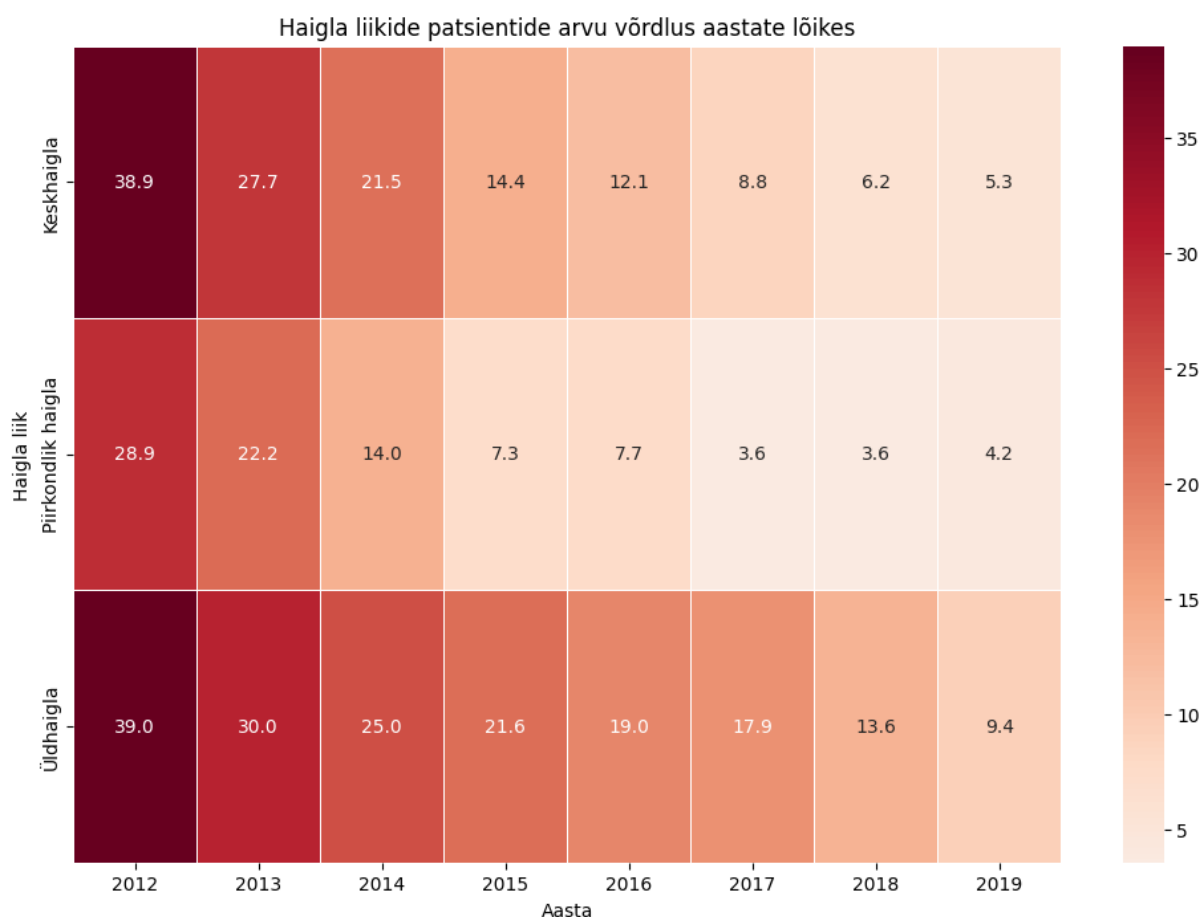
Joonis 12. Haigla liikide patsientide arvu võrdlus raviarvete ja TIS-i terviseandmete suhtes näitab Tervisekassa raviarvete ja TIS-i andmete patsientide arvu haigla liikide suhtes, mille puhul üldjuhul on Tervisekassa raviarvete patsientide arv suurem haigla liikide suhtes kui TIS-i andmetes.

Joonisel 12 on võrreldud patsientide arvu Tervisekassa raviarvete ja TIS-i andmetes haigla liigi suhtes (lisas 4 on leitavad haigla liikide alla kuuluvad haiglad). Kõige sarnasemad patsientide arvud on piirkondlikus haiglas ja kõige suuremad erinevused on üldhaigla andmetes. Keskhaigla puhul näeme, et suured erinevused on vanusevahemikus 0-19, mida võib põhjendada sellega, et paljud tervishoiuteenused on alaealistele ravikindlustusega kaetud. Üldhaigla andmete puhul on suuremad erinevused nooremate patsientide suhtes ja erinevus Tervisekassa raviarvete ja TIS-i andmete vahel väheneb patsientide vanuse suurenemisega. Üldhaigla andmete erisust Tervisekassa ja TIS-i vahel võib põhjendada see, et üldhaiglad on väiksemad haiglad kui näiteks keskhaiglad ja piirkondlikud haiglad, kellel võib olla vähem rahalisi vahendeid, et vajalikku tervise infosüsteemi üles ehitada. Antud töös ei saa seda edasi uurida, sest ei ole teada, millised olid haiglate IT-süsteemid ja mis ulatuses oldi liidestatud TIS-ga.



Joonis 13. Haigla liikide patsientide arvu võrdlus RHK-10 peatükkide suhtes näitab Tervisekassa raviarvete andmete osakaalu võrreldes Tervise Infosüsteemi andmetega patsientide arvu suhtes.

Kohaliku haigla patsientide andmed näitavad ka RHK-10 peatükkide suhtes kõige teineteisega lähedasemaid andmeid kahe andmeallika vahel. Siiski on suur ülekaal Tervisekassa raviarvete andmetel RHK-10 peatükis F00-F99, kus Tervisekassa raviarvete alusel on andmeid umbes 20% rohkem kui TIS-i andmeteid. Nagu oleme antud töös juba maininud, F00-F99 erinevust võib tekitada arstide otsus mitte saata selliseid andmeid TIS-i. Juhtumid, kus TIS-i andmete alusel on patsiente rohkem RHK-10 peatükis kui Tervisekassa raviarvetel, tulevad joonisel 13 välja. Näeme, et TIS-i andmete patsientide arv on suurem ligi 22% RHK-10 peatükis H00-H59: Silma- ja silmamanuste haigused. Nagu nägime eelmisel joonisel, näeme ka siin, et üldhaiglate ja keskhaiglate puhul on andmeid rohkem Tervisekassa raviarvetelt kui Tervise Infosüsteemi andmetelt.



Joonis 14. Haigla liikide patsientide arvu võrdlus aastate lõikes näitab Tervisekassa raviarvete patsientide osakaalu protsentuaalselt TIS-i andmete patsientide suhtes haigla liigi ja aastate lõikes.

Joonisel 14 on näha, et keskhaigla ja üldhaigla puhul oli Tervisekassa ja TIS-i patsientide arvu erinevus aastate alguses suur, aga aastate jooksul on vahe väiksemaks jäänud. Piirkondliku haigla puhul on erinevus kahe andmeallika vahel olnud väiksem ja 2019. aasta lõpuks on see vahe peaaegu olematu. Erinevus aastate jooksul on läinud paremaks, sest aastate jooksul on haiglate võrgustumine paremaks läinud. Lisaks oli paljudel haiglatel 2012. ja 2013. aastatel tehnoloogilised probleemid TIS-ga ühendamisel [5: 16-18]. Üheks Eesti e-tervise eesmärgiks [23: 15] on haiglate võrgustumine, mis tähendab ühiste infosüsteemide kasutusele võtmist haiglali.

5.5. Kattuvate patsientide ja diagnooside võrdlus

Eesti E-tervise arengukava näeb ette, et andmed oleksid ühekordsed ja taaskasutatavad [23]. Kattuvate andmete all on mõeldud andmeid, mille puhul Tervisekassa raviarved ja TIS-i andmed kuuluvad samale patsiendile samas kui RHK-10 põhidiagnoos, dokumendi kuupäevad ja TTO andmed on samad, mis mõlemas andmeallikas on olemas, st. dubleerivad diagnooside lõikes.

Kattuvatesse andmetesse jäi alles 134 079 patsienti (97% algandmetest), 3 161 809 dokumenti (41 % algandmetest), 7299 erinevat diagnoosi (84%). Algandmete all mõeldi andmeid, mis jäid alles kui jäeti alles nende isikute andmed, kes olid nii Tervisekassa raviarvete kui TIS-i andmetes. Alles jäänud dokumentide arv on küll ainult 41%, aga enamus nendest kuuluvad Tervisekassale, kes saadab patsiendi ja haigusjuhtumi kohta rohkem raviarveid kui TIS-i saadetakse epikriise. Tervishoiuteenuse lõikes jagunesid kattuvad andmed järgnevalt nagu tabelis 3 näha:

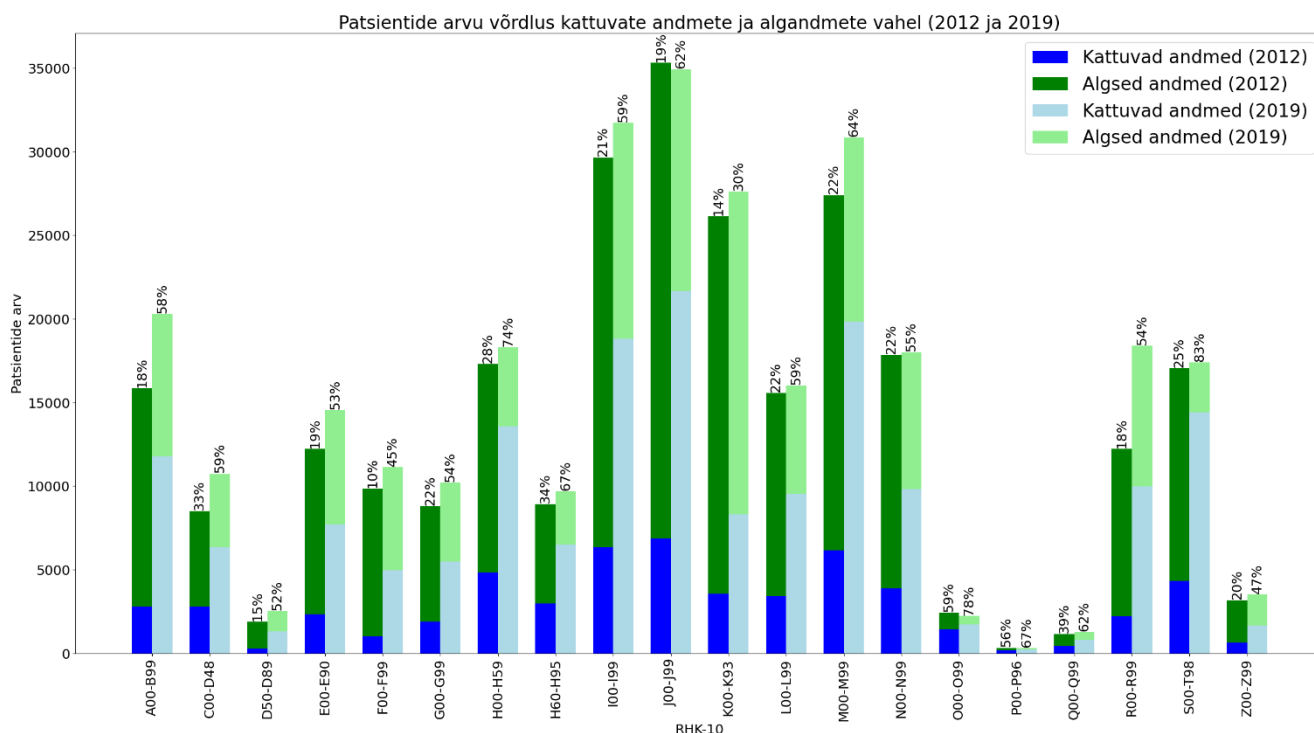
Tabel 3. Kattuvate andmete ravitüübi järgi dokumentide, diagnooside ja patsientide arv

| ANDMEALLIKAS | THT TÜÜP | DOKUMENTIDE ARV | DIAGNOOSIDE ARV | PATSIENTIDE ARV |
|--------------|---------------|-----------------|-----------------|-----------------|
| Raviarved | Statsionaarne | 132 947 | 4411 | 60 702 |
| Raviarved | Ambulatoorne | 1 427 728 | 6902 | 131 357 |
| TIS | Statsionaarne | 141 857 | 4531 | 62 757 |
| TIS | Ambulatoorne | 1 459 277 | 6904 | 130 893 |

Patsientide arv statsionaarse ja ambulatoorse ravi puhul peaks olema sama Tervisekassa ja TIS-i andmetes, aga ühel juhul on märgitud tervishoiuteenuse tüübiks statsionaarne ja teise puhul ambulatoorne, kusjuures nendel juhtudel on sama patsient, sama diagnoos, samad dokumendi kuupäevad ja sama TTO. Põhjust, miks sellistel juhtudel tervishoiuteenuse tüüp Tervisekassa ja TIS-i andmete vahel sama ei ole, antud töös ei leitud.

Kui võtta arvesse, et antud töös kattuvates andmetes on palju nõ. dubleerivaid andmeid diagnooside lõikes Tervisekassa ja TIS-i andmetes, siis 2019. aasta lõpuks ei olnud täidetud andmete ühekordne kogumine ja nende taaskasutamine. Kuna raviarvetel ja TIS-i andmetes on osa andmetunnuseid samad, siis andmete ühekordse kogumise, edastamise ja taaskasutamise põhimõttel oleks vaja läbi mõelda e-tervise infosüsteemid ja nende koostalitlusvõime. Kui

hakati planeerima 2004. aastal e-tervist, siis arutleti Tervisekassa raviarvete ja TIS-i edastavate andmete sidumist ja võrdlemist [5: 24], aga 2019. aasta lõpuks ei olnud see realiseerinud.



Joonis 15. Patsientide arvu võrdlus 2012. a ja 2019. a kattuvate andmete (sama patsiendi ID, RHK-10 põhidiagnoos, dokumendi algus- ja lõpuaeg, tervishoiuteenuse osutaja) ja kõigi isikute vahel, kellel esines vastaval aastal vastav RHK-10 diagnoos. Protsent näitab, kui suur osa isikutest kattus.

Kattuvate andmete puhul on oluline vaadata, kuidas on andmete kattuvus aastate jooksul muutunud. Joonisel 15 on näha alg- ja kattuvaid andmeid 2012 ja 2019 aasta andmete kohta. Näeme, et 2012. aastal oli vähem isikuid RHK-10 diagnooside lõikes, aga samas andmete kattuvus oli väiksem. 2019. aastal on andmeid rohkem kui 2012. aastal, aga samas on ka andmete kattuvus suurem. Joonisel 15 on näha, et 2012. aastal oli andmete kattuvus võrreldes algandmetega väike. Näiteks F00-F99 puhul oli andmete kattuvus ainult 10%, st. ainult 10% isikutest olid samade andmetunnustega (põhidiagnoos, kuupäevad, TTO) nii Tervisekassa kui TIS-i andmetes. F00-F99 puhul näeme, et 2019. aasta lõpuks on kattuvus ainult 45%. Samas on näha, et S00-T98 (vigastused, mürgistused ja teatavad muud välispõhjuste toime tagajärjed) peatüki puhul ei ole andmeid väga juurde tulnud, aga samas 2019. aasta lõpuks kattusid 83% isikute andmeid Tervisekassa ja TIS-i vahel. K00-K93 peatükk, mis on ka teistes analüüsi osades välja tulnud, joonisel 15 on kõige vähem kattuvam peatükk, sest 2019. aastal kattuvad ainult 30% andmeid Tervisekassa ja TIS-i vahel. Sellised suured erinevused näitavad, et K00-K93ja F00-F99 andmete edastamine TIS-i ei ole paranenud.

6. Järeldused

Käesoleva töö käigus tuli välja, et Tervisekassa raviarvetes on rohkem andmeid kui TIS-i andmetes. Aastate jooksul on andmete erinevused küll väiksemaks läinud, aga teatud diagnooside rühmades on erinevused siamaani suured. Isegi kui TTO-del on kohustus saada patsiendiga seotud terviseandmeid elektrooniliselt TIS-i, siis seda ei tehta alati. Tervisekassale saadetakse üldjuhul alati raviarveid, sest TTO-del on motivatsiooni andmeid saada, et saada tagasi ravikindlustuslepingu alusel patsiendile osutatud tervishoiuteenuste osutamise eest hüvitist. Samas kui TIS-i andmete edastamine on oluline kliinilisel eesmärgil.

2012 - 2019. aastate jooksul edastas 1166 TTO-d raviarveid Tervisekassale, kellest 78% edastas ka andmed TIS-i. 78% edastajate seas olid esindatud kõik haiglad ja erinevus tuli sisse teiste TTO-de liikides. Tervisekassa raviarvete andmetes oli 8600 erinevat diagnoosi ja nendest 97% olid ka TIS-i andmetes, mille erinevuse võisid põhjustada TTO-de poolt RHK-10 kooditäpsuse kasutamine. Kui üldiselt oli Tervisekassa raviarvetes rohkem patsiente RHK-10 diagnoosidega ja lisaks ka erinevaid RHK-10 diagnoose kui TIS-i andmetes, siis silmapaistvalt suured erinevused olid just A00-B99 (teatavad nakkus- ja parasiithaigused), F00-F99 (psüühika- ja käitumishäired) ja K00-K93 (seedeelundite haigused) diagnoosides, mille puhul Tervisekassa andmetes on rohkem andmeid kui TIS-i andmetes. A00-B99 erinevusi võivad põhjustada teiste meditsiiniregistrite olemasolu, kuhu TTO-d on kohustatud andmeid saatma ja võib juhtuda, et nad ei taha duplikaat andmeid TIS-i edastada. Lisaks võib nakkus- ja parasiithaiguste ravi olla kallis, mille puhul TTO-d soovivad teenuse hüvitamise eest raha saada. F00-F99 erinevusi võib põhjendada arstide praktikaga, kes ei soovi tundliku sisuga andmeid TIS-i edastada. Peale selle võib veel põhjuseks olla Riigikontrolli [5] poolt leitud fakt, et eriarstiabi TTO-d on küll TIS-ga liidestunud, aga TTO tüüpidest saadavad nemad kõige vähem andmeid. K00-K93 puhul oli andmeid rohkem nooremate inimeste seas, mida võib põhjendada sellega, et laste ravikindlustuse poolt hüvitatud hambaravi andmed edastatakse Tervisekassale. Riigikontroll [5] leidis hambaraviteenuse osutajatega suheldes, et hambaravi andmete TIS-i edastamise jaoks ei ole head epikriisi. 2015. aastal loodi hambaravikaart, mille andmeid peaksid TIS-i saatma TTO-d, aga mille andmeid antud töö analüüsis polnud.

Antud töös vaadeldi ka andmete erisust ambulatoorse ja statsionaarse ravi alusel. Tervisekassa raviarvetel oli üldjuhul rohkem andmeid kui TIS-i andmetes, aga statsionaarse ravi puhul oli andmeid TIS-is rohkem. Erisus ambulatoorse ravi puhul tuleb välja samuti A00-B99, F00-F99

ja K00-K93 diagnoosides, mille põhjuseid oleme antud töös juba vaadanud. Lisaks vaadeldi ka andmete erisust haigla liikide vahel, mille analüüsis tuli välja, et enam-vähem samad andmed Tervisekassa ja TIS-i vahel olid piirkondliku haigla puhul, aga kuna piirkondlikud haiglad on kõige suuremad haiglad, kellel on kasutusel oma infosüsteemid ja kellel on rohkem raha kui teistel haiglatel, siis nende puhul on loogiline, et Tervisekassa andmetes on rohkem andmeid.

Kuna Eesti riigis on andmete ühekordse sisestamise ja taaskasutamise põhimõtted enamuses valdkondades, siis sai uuritud ka nõ. kattuvate andmete probleemi Tervisekassa ja TIS-i andmetes. Kahe andmeallika vahel on ainult osa tunnuseid sama, nii et andmete kattuvust vaadeldi diagnooside lõikes, kus samal patsiendil pidi olema sama põhidiagnoos, sama dokumendi kuupäev, sama TTO ja sama andmete kombinatsioon pidi olema nii Tervisekassa kui TIS-i andmetes. Selle analüüsi käigus leiti, et 97% patsientidel on kattuvad andmed Tervisekassa ja TIS-i vahel, millest 84% diagnoose on sama TTO ja samade kuupäevadega nii Tervisekassa ja TIS-i andmetes. Kattuvaid andmeid Tervisekassa ja TIS-i diagnooside andmetes on üsna suur osa, mis näitab, et vähemalt osaliselt dubleeritakse terviseandmeid Eestis.

6.1 Töö tugevused, nõrkused, edasiarendusvõimalused

Siimaani on antud töö sarnastel teemadel uuritud Tervisekassa ja TIS-i andmeid agregeeritud andmete kujul, mis ei ole võimaldanud vaadata isikustatud ja kahe andmeallika vahel ühendatud andmeid. Näiteks siimaani ei ole vaadatud isikupõhiselt kui suures osas kattuvad Tervisekassa raviarvete ja TIS-i saadetud epikriiside andmed. Käesolevas töös on üritatud leida võimalikke põhjuseid, aga kuna on näha, et TIS-i edastatakse vähem andmeid kui Tervisekassale, siis puudu on TIS-i andmete edastajate poolset kommentaari ja põhjused. Analüüsi käigus tuli kokkuvõtvalt välja, et TIS-i andmete edastamise kohustust ei täideta 100% ulatuses ja probleemid on teatud diagnooside lõikes. Erinevused on aastatega küll vähenenud, aga oleks vaja ikkagi uurida, miks teatud TTO-d või teatud epikriise TIS-i ei saadeta.

Antud töös vaadeldi üldiselt andmete erinevusi diagnooside, patsientide, tervishoiuteenuse tüüpide, TTO liikide tunnuste lõikes, mis tõi välja, et oleks vaja edasi uurida andmete erisust A00-B99, F00-F99 ja K00-K93 RHK-10 peatükkides. Analüüsi käigus toodi välja, et 2019. aasta lõpuks oli Tervisekassa andmetes 45% rohkem K00-K93 peatükis patsiente, 28% F00-

F99 peatükis ja 19% A00-B99 peatükis võrreldes TIS-i andmetes olevate patsientide arvuga, mis on kordades suuremad % kui teised diagnooside peatükid. Lisaks oleks huvitav uurida, miks on statsionaarsete juhtumite puhul Tervisekassa ja TIS-i andmete erinevus väike ning miks on erinevus just ambulatoorsetes juhtumites.

Tulemuste viimases peatükis vaadeldi andmete kattuvuse probleemi, mille tulemusel oli näha, et aastate jooksul kattuvad aina rohkem diagnoosidega seotud andmed Tervisekassa ja TIS-i andmete vahel. Kuna Eesti riigil on eesmärk koguda andmeid ühekordselt ja neid taaskasutada, siis andmete kattuvus antud töös näitab, et seda probleemi peaks edasi uurima. Peaks leidma viisi, kuidas oleks võimalik koguda selliseid andmeid korra ja mis ajendaks ka TTO-sid andmeid TIS-i edastama. Näiteks oleks huvitav uurida, kuidas oleks võimalik korraldada seda, et TTO-d saadavad raviarved TIS-i koos muude epikriisi andmetega ja sealt vastavalt saab Tervisekassa kätte oma raviarvete andmed. Lisaks oleks huvitav uurida, miks 2019. aasta lõpuks andmete kattuvus siiski suur on. Kuna erinevus võib tulla sellest, et TTO-d saadavad sama haigusloo korras erineva detailsusega RHK-10 koodi Tervisekassa raviarvel ja TIS-i dokumendil, siis üheks uuritavaks probleemiks oleks vaadelda, kas ja mis TTO-d saadavad sama haigusloo korral erineva detailsusega RHK-10 koodi.

7. Kokkuvõte

Antud magistritöös võrreldi Tervisekassa raviarvete ja tervise infosüsteemi epikriiside andmeid patsientide, RHK-10 diagnooside, ravitüüpide ja tervishoiuteenuse osutajate lõikes. Töös kasutati 2012-2019. aasta andmeid, mille edasises analüüsis kasutati nende patsientide andmeid, kes olid nii Tervisekassa kui TIS-i diagnooside andmetes. Töö käigus selgus, et üldjuhul on Tervisekassa andmetes rohkem erinevaid patsiente, erinevaid RHK-10 diagnoose ja rohkem TTO-de andmeid. Kuna TTO-d saavad hüvitatud tervishoiuteenuste eest Tervisekassalt raha raviarve alusel, siis TTO-del on rahaline motivatsioon, et alati raviarveid Tervisekassale saata. Selline motivatsioon puudub TIS-i andmete saatmisega, sest TIS-i andmete saatmine on andmete edastamine patsiendi kohta patsiendile endale kui ka teistele arstidele.

Kõige suurem andmete erinevus oli 2012. aastal, mida võib põhjendada TIS-i jätkuva arendamisega sel aastal, osade dokumentide mitte-saatmise kohustusega ja Terviseameti kontrolli puudumisega TTO-de üle. Alates 2012. aastast vähenes andmete erinevus Tervisekassa ja TIS-i andmete vahel ning 2019. aasta lõpuks oli ka näha üksikuid juhtumeid, kus TIS-i andmeid oli rohkem kui Tervisekassa andmeid. Samas mõned silmapaistvad erinevused Tervisekassa ja TIS-i vahel olid alles 2019. aasta lõpuks RHK-10 koodides A00-B99 (teatavad nakkus- ja parasiithaigused), F00-F99 (Psüühika- ja käitumishäired) ja K00-K93 (seedeelundite haigused), mida võib põhjendada arstide praktikaga mitte saata tundliku sisuga andmeid TIS-i, Tervisekassale edastatud ravikindlustatute raviarvete osakaal, mitme meditsiiniregistri olemasolul ei saadeta samu andmeid TIS-i või hambaravi korral andmete edastamine erinevate dokumentide kaudu. Peale selle leiti, et statsionaarsete juhtumite puhul on Tervisekassa ja TIS-i andmete hulk peaaegu sama ja erinevus tuleb andmete vahel sisse just ambulatoorsete juhtumite puhul. TTO-de lõikes olid kõik Eesti haiglad esindatud nii Tervisekassa kui TIS-i andmetes ja TTO-de erinevus tuli sisse teiste TTO-de (perearstiabi, eriarstiabi jne) puhul, mille puhul tuli välja, et isegi kui TTO-d on ametlikult liitunud TIS-i süsteemidega, siis kõik TTO-d sinna andmeid ei saada.

Kuna Eesti riigil on andmete osas andmete ühekordse kogumise ja taaskasutamise eesmärk e-tervise strateegias, siis antud töös vaadati ka kui palju on nõ. kattuvaid andmeid kahe andmeallika vahel. Selleks loodi eraldi andmestik, kuhu jäeti sisse Tervisekassa ja TIS-i andmed, kus tegemist oli sama patsiendiga, sama diagnoosiga, samadel kuupäevadel ja sama

TTO-ga. Kattuvates andmetes oli sees 97% patsiente, 84% erinevaid diagnoose ja 41% dokumente võrreldes põhianalüüsis olevate andmetega. Need % näitavad, et suur osa andmeid kattuvad kahe andmeallika vahel ja siamaani ei ole täidetud andmete ühekordne edastamine ja taaskasutamine.

Viidatud kirjandus

- [1] Glynn R. J., Jin Y., Lin K. J., Mandl K. D., Murphy S. N., Rosenthal G. E., Schneeweiss S. External Validation of an Algorithm to Identify Patients with High Data-Completeness in Electronic Health Records for Comparative Effectiveness Research. *National Library of Medicine*, 2020, 2
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7007793/> (13.05.2024)
- [2] Balicer R. D., Einav L., Flaks-Manov N., Shadmi E., Shir T., Zeltzer D. Prediction Accuracy With Electronic Medical Records Versus Administrative Claims. *Medical Care Official Journal of the Medical Care Section, American Public Health Association*, 2019, 7
https://journals.lww.com/lww-medicalcare/abstract/2019/07000/prediction_accuracy_with_electronic_medical.10.aspx (13.05.2024)
- [3] Anderson E., Smirnov K., Sokurova D. Perearstiabi patsientide pöördumiskordade arvu ja haigusjuhu kestuse prognoosid Eesti Haigekassa ning tervise infosüsteemi andmete põhjal, Tervise Arengu Instituut, 2019
<https://www.tai.ee/et/valjaanded/perearstiabi-patsientide-poordumiskordade-arvu-ja-haigusjuhu-kestuse-prognoosid-eesti> (13.05.2024)
- [4] RAHVASTIKU TERVISE ARENGUKAVA 2020–2030, 2019
[https://www.emcdda.europa.eu/system/files/attachments/13553/Ministry%20of%20Health%2C%20Estonia%20\(2020\)%20Estonia's%20National%20Health%20Plan%202020-30.pdf](https://www.emcdda.europa.eu/system/files/attachments/13553/Ministry%20of%20Health%2C%20Estonia%20(2020)%20Estonia's%20National%20Health%20Plan%202020-30.pdf) (13.05.2024)
- [5] Riigikontrolli aruanne Riigikogule, Tallinn, 17. jaanuar 2014. Riigi tegevus e-tervise rakendamisel
https://egov.nik.gov.pl/EE/2014/eHealth/E-health_est.pdf (13.05.2024)
- [6] Tervishoiusüsteemi korraldus Eestis (24.04.2024), Tervisekassa
<https://www.tervisekassa.ee/kontaktpunkt/arstiabi-euroopa-liidu-kindlustatule-eestis/tervishoiususteemi-korraldus-eestis> (13.05.2024)
- [7] SNOMED CT Managed Service - Eesti laiendi versioonimärkmed - May 2022
<https://confluence.ihtsdotools.org/pages/viewpage.action?pageId=146872656> (13.05.2024)
- [8] TEHIK Teabekeskus
<https://teabekeskus.tehik.ee/et/standardid> (13.05.2024)
- [9] Tervise infosüsteemi edastatavate dokumentide andmekoosseisud ning nende säilitamise tingimused ja kord (26.05.2023). *Riigi Teataja*
<https://www.riigiteataja.ee/akt/128062018010> (13.05.2024)
- [10] Tervisekassa nõudis mullu raviasutustelt tagasi ligi 700 000 eurot, *ERR* (25.04.2024)
<https://www.err.ee/1609323531/tervisekassa-noudis-mullu-raviasutustelt-tagasi-ligi-700-000-eurot> (13.05.2024)

- [11] Tervisekassa andmekogu pidamise põhimäärus (22.03.2024), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/119032024008> (13.05.2024)
- [12] HINDAMISOTSUS nr 97, Tervishoiuteenuse osutamist tõendavate isikupõhiste meditsiinidokumentide arhiiviväärtuse väljaselgitamine (16.12.2015), *Rahvusarhiiv*
https://www.ra.ee/wp-content/uploads/2016/11/HO_97_tervishoiudokumendid_16.12.2015.pdf (13.05.2024)
- [13] McCallie Jr. D., Overhage J. M. Physician Time Spent Using the Electronic Health Record During Outpatient Encounters: A Descriptive Study. *Annals of Internal Medicine*, 2020, 1
<https://www.acpjournals.org/doi/10.7326/m18-3684> (13.05.2024)
- [14] Raviarvete saatmine (29.04.2024), Tervisekassa
<https://www.tervisekassa.ee/raviarvete-saatmine> (13.05.2025)
- [15] TEHIK Standardite ja klassifikaatorite publitseerimiskeskus
<https://pub.e-tervis.ee/> (13.05.2024)
- [16] TEHIK Tarkvara meditsiinasutustele
<https://www.tehik.ee/tarkvara-meditiinasutustele> (13.05.2024)
- [17] Strateegilise arenguplaani töötas välja Riigikantselei juures juulist 2014 novembrini 2015 tegutsenud rakkerühm. E-tervise visioon 2025 E-tervise strateegiline arenguplaan 2020
https://www.sm.ee/sites/default/files/content-editors/eesmargid_ja_tegevused/Eesti_e_tervise_strateegia/e-tervise_strateegia_2020.pdf (13.05.2024)
- [18] Võlaõigusseadus (01.01.2024), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/108072011021> (13.05.2024)
- [19] Tervishoiuteenuste korraldamise seadus (21.03.2024), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/TTKS> (13.05.2024)
- [20] Isikuandmete kaitse seadus (01.11.2023), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/104012019011> (13.05.2024)
- [21] Tervise infosüsteemi põhimäärus (10.09.2023), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/106122016011> (13.05.2024)
- [22] Ravikindlustuse seadus (01.04.2023), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/111032023081#para30lg1> (13.05.2024)
- [23] Eesti e-tervise valitsemise raamistik, Hetkeolukorra analüüs (20.05.2021)
https://www.sm.ee/sites/default/files/content-editors/Tervishoid/e-tervis/eesti_e-tervise_valitsemisraamistiku_i_vaheraport_-_hetkeolukorra_analuus.pdf (13.05.2024)
- [24] Oja M, Tamm S, Mooses K, Pajusalu M, Talvik H-A, Ott A, Laht M, Malk M, Lõo M, Holm J, Haug M, Šuvalov H, Särg D, Vilo J, Laur S, Kolde R, Reisberg S. Transforming Estonian health data to the Observational Medical Outcomes Partnership (OMOP) Common Data Model: lessons learned. *JAMIA Open*, Volume 6, Issue 4, 2023,

<https://academic.oup.com/jamiaopen/article/6/4/ooad100/7459333> (13.05.2024)

[25] Eesti Sotsiaalministeerium. Rahvusvaheline haiguste ja nendega seotud terviseprobleemide statistiline klassifikatsioon, Kümnes väljaanne, 2. köide, Instruktsioonide käsiraamat, 1996
https://www.sm.ee/sites/default/files/content-editors/eesmargid_ja_tegevused/Tervis/E-tervis_ja_e-tervisetoend/kaesiraamat.pdf (13.05.2024)

[26] Haigla liikide nõuded (26.05.2023), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/101062016007> (13.05.2024)

[27] Haiglavõrgu arengukava (08.04.2023), *Riigi Teataja*
<https://www.riigiteataja.ee/akt/830528> (13.05.2024)

[28] Tervise Arengu Instituut, Tervisestatistika ja terviseuuringute andmebaas. Tervishoiuteenuse osutajad
https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_04THressursid_01TTosutajad/?tablelist=true (13.05.2024)

[29] Eesti Statistikaamet, Statistika andmebaas, RV021: Rahvastik soo ja vanuserühma järgi, 1. jaanuar
https://andmed.stat.ee/et/stat/rahvastik_rahvastikunaitajad-ja-koosseis_rahvaarv-ja-rahvastiku-koosseis/RV021 (13.05.2024)

[30] Eesti Statistikaamet, Statistika andmebaas, RV045: Oodatav eluiga sünnimomendil ja elada jäänud aastad soo ja vanuse järgi
https://andmed.stat.ee/et/stat/rahvastik_rahvastikunaitajad-ja-koosseis_demograafilised-pehinaitajad (13.05.2024)

[31]] Tervise Arengu Instituut, Tervisestatistika ja terviseuuringute andmebaas. TTO60: E-tervisega liidestunud ja andmeid esitanud iseseisvad tervishoiuasutused õigusliku vormi, teenuseosutaja ja omaniku liigi järgi
https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_04THressursid_01TTosutajad/TTO60.px/ (13.05.2024)

[32] Eesti Haigekassa 2019. aasta majandusaasta aruanne
https://www.tervisekassa.ee/sites/default/files/tagasiside_aruanded/HK-aastaruanne-2019-final-online.pdf (13.05.2024)

[33] Liivlaid H. Hambaravi andmed tervise infosüsteemis, Tervise Arengu Instituut, 2018
https://www.tai.ee/sites/default/files/2021-03/154416834850_Hambaravi_andmed_tervise_infosusteemis_.pdf (13.05.2024)

[34] Hussar K. Hambaarstide liit soovib saada õigusi petturitest hambaarstide tegevuse tõkestamiseks, ERR (27.08.2018)
<https://www.err.ee/856548/hambaarstide-liit-soovib-saada-oigusi-petturitest-hambaarstide-tegevuse-tokestamiseks> (13.05.2024)

[35] Zabel J. Key Differences between Inpatient and Outpatient Medical Billing (31.07.2023)

<https://medicalbillingcompany.medium.com/key-differences-between-inpatient-and-outpatient-medical-billing-a02c2fe14060> (13.05.2024)

[36] Perearsti kvaliteedisüsteem (13.03.2024)

<https://www.tervisekassa.ee/partnerile/raviasutusele/perearstile/perearsti-kvaliteedisusteem> (13.05.2024)

[37] Lakoma M. D., Lu C. Y., Madden J. M., Rusinak D., Soumerai S. B. Missing clinical and behavioral health data in a large electronic health record (EHR) system. *Journal of the American Medical Informatics Association*, Volume 23, Issue 6, November 2016, Pages 1143-1149

<https://academic.oup.com/jamia/article/23/6/1143/2399287> (13.05.2024)

[38] E-konsultatsiooni saatekirja ja selle vastuse nõuded perearsti poolt patsiendi suunamisel psühhiaatri e-konsultatsiooni tervise infosüsteemi vahendusel, *Riigi Teataja*
https://www.riigiteataja.ee/akti/1291/2201/7054/SOM_m67_lisa20.pdf (13.05.2024)

[39] Ravijuhtude kodeerimise hindamine Haigekassa andmebaasi ning ravidokumentide põhjal. Aruanne koostööpartneritele, Eesti Haigekassa, Tervishoiuökonomika osakond, 2007

https://www.tervisekassa.ee/files/est_raviasutusele_drg/Ravijuhtude_kodeerimise_hindamise_raport1.pdf (13.05.2024)

[40] Eigo N., Liivlaid H., Poolakese A., Rätsept M., Tilk R. Tervise infosüsteemi ja regulaarse tervisestatistika andmete võrdlus III analüüs, Tervise Arengu Instituut, 2013

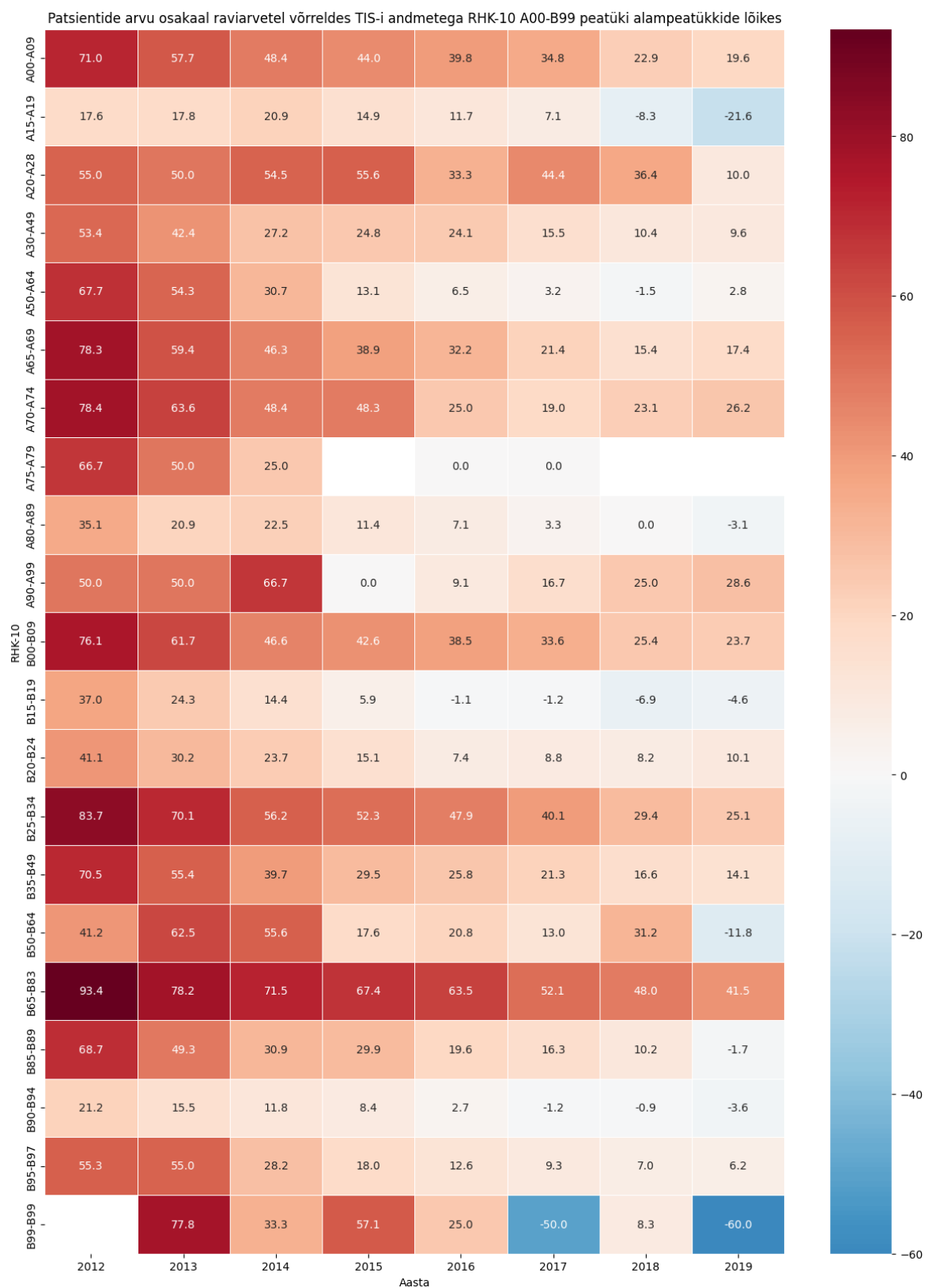
<https://dspace.ut.ee/server/api/core/bitstreams/c804626a-0247-40e2-9188-afbd9ce26ef1/content> (13.05.2024)

[41] Raviarvete ja lepingute andmevahetusteenused, Eesti Haigekassa, Kindlustatud Isikute Register, 2023

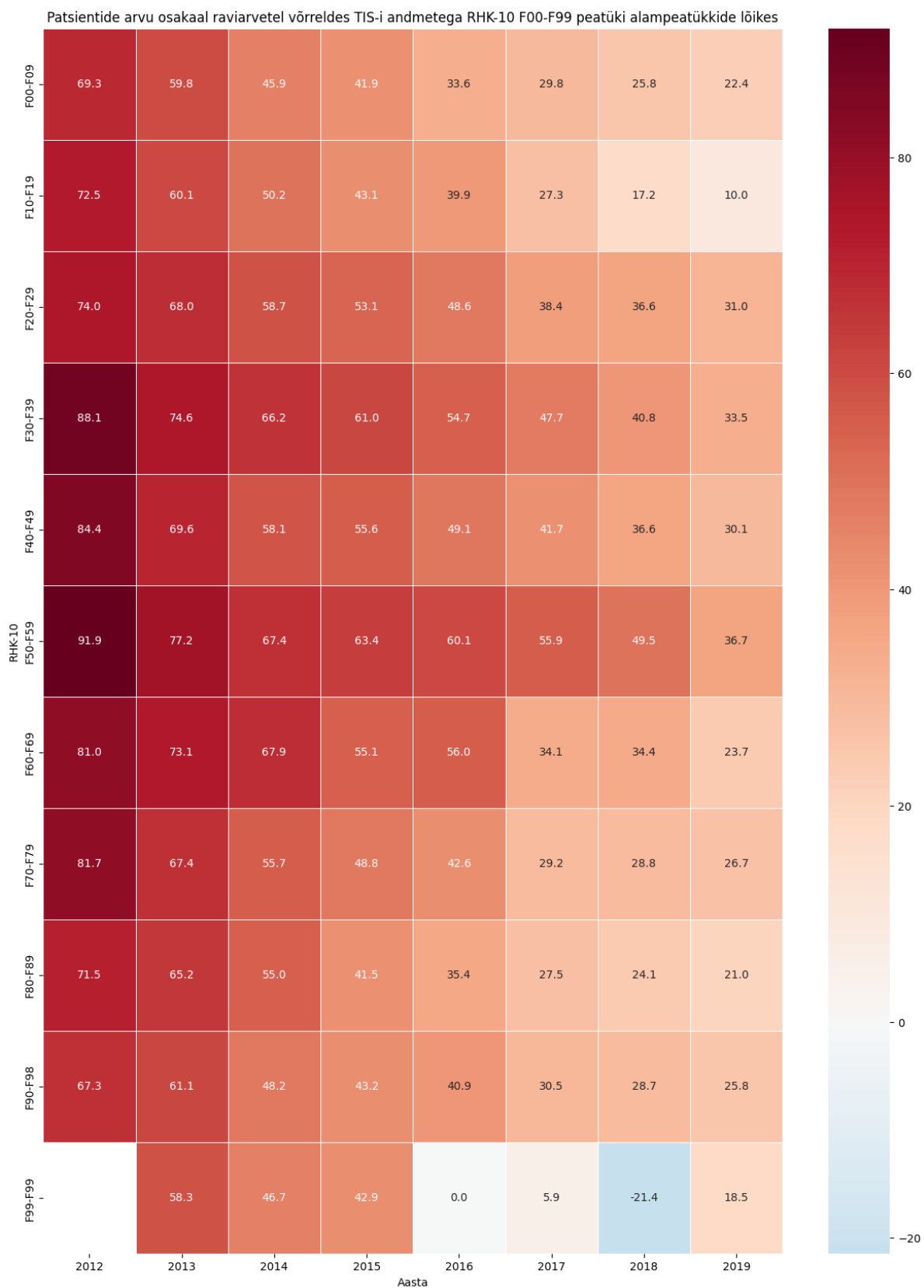
[https://tervisekassa.ee/sites/default/files/EHK_RTA_teenused_20232501%20\(1\).pdf](https://tervisekassa.ee/sites/default/files/EHK_RTA_teenused_20232501%20(1).pdf) (13.05.2024)

Lisad

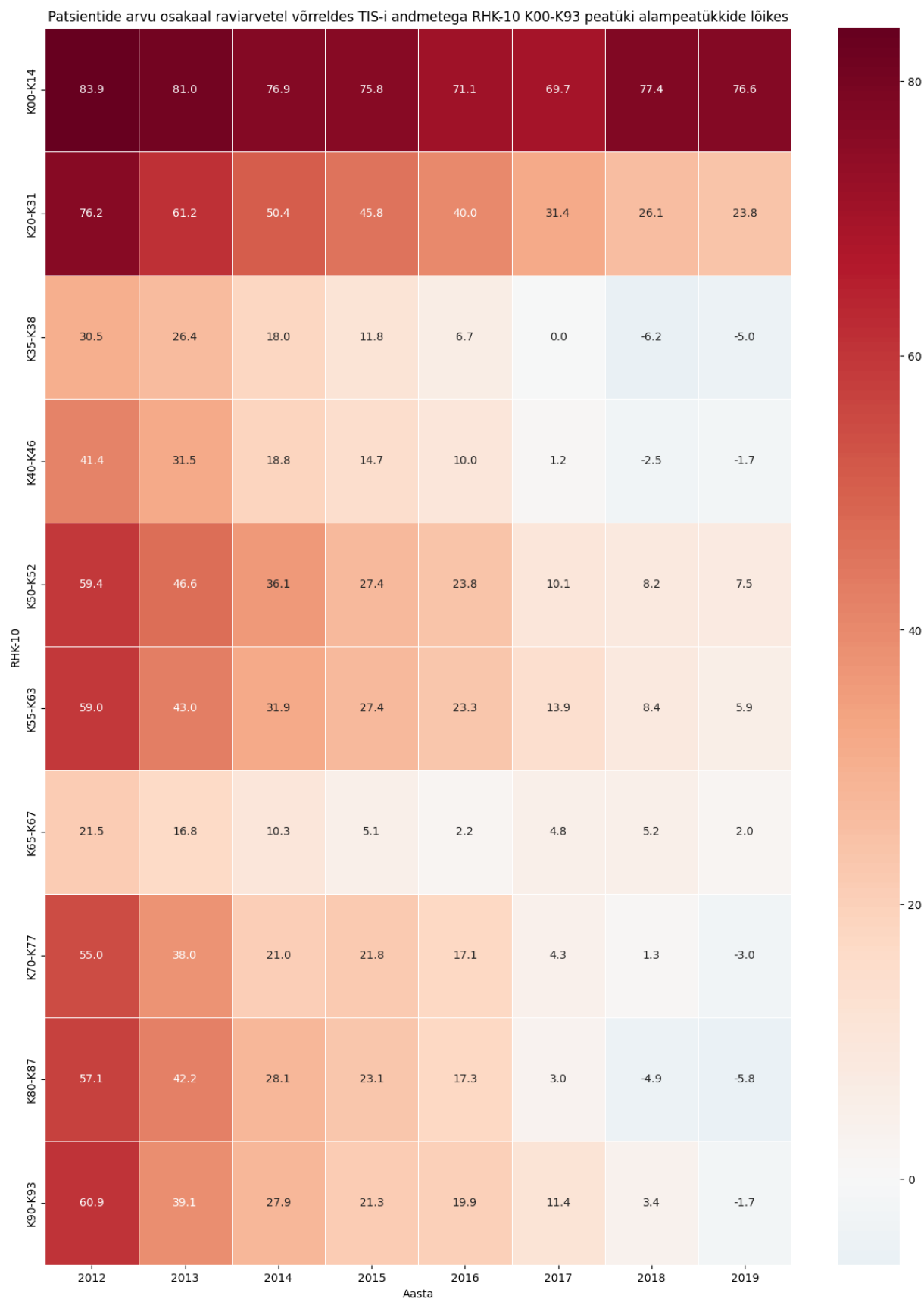
Lisa 1. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 A00-B99 peatüki alampeatükkide lõikes



Lisa 2. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 F00-F99 peatüki alampeatükkide lõikes



Lisa 3. Patsientide arvu osakaal raviarvetel võrreldes TIS-i andmetega RHK-10 K00-K93 peatüki alampeatükkide lõikes



Lisa 4. Haigla liikidesse kuuluvad haiglad

Haiglavõrgu arengukava kohaselt kuuluvad haigla liikidesse [27]:

Piirkondlikud haiglad:

- 1) sihtasutus (edaspidi SA) Põhja-Eesti Regionaalhaigla;
- 2) SA Tartu Ülikooli Kliinikum;
- 3) SA Tallinna Lastehaigla.

Keskhaiglad:

- 1) aktsiaselts (edaspidi AS) Ida-Tallinna Keskhaigla;
- 2) AS Lääne-Tallinna Keskhaigla;
- 3) SA Ida-Viru Keskhaigla;
- 4) SA Pärnu Haigla.

Üldhaiglad:

- 1) AS Järvamaa Haigla;
- 2) Kuressaare Haigla SA;
- 3) SA Läänemaa Haigla;
- 4) AS Rakvere Haigla;
- 5) AS Lõuna-Eesti Haigla;
- 6) SA Narva Haigla;
- 7) SA Viljandi Haigla;
- 8) AS Valga Haigla;
- 9) SA Hiiumaa Haigla;
- 10) AS Põlva Haigla;
- 11) SA Rapla Maakonnahaigla.

Kohalikud haiglad:

- 1) SA Jõgeva Haigla;
- 2) AS Põlva Haigla;
- 3) SA Rapla Maakonnahaigla.

Lisa 5. *Litsents*

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Kristiina Miller,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose *Tervise infosüsteemi ja Tervisekassa andmete võrdlus perioodil 2012–2019 patsientide, diagnooside, ravitüübi ja teenuseosutajate lõikes*,

mille juhendajad on Marek Oja ja Sirli Tamm,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kristiina Miller

15.05.2024