



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

**PSYKOSOSIAALISTEN
RISKITEKIJÖIDEN
KIRJAAMINEN
TYÖTERVEYSHUOLLON
TERVEYSTARKASTUKSISSA**

Koneoppimisen mahdollisuuksia
automaattiseen tekstintunnistukseen

Leena Uronen



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

PSYKOSOSIAALISTEN RISKITEKIJÖIDEN KIRJAAMINEN TYÖTERVEYSHUOLLON TERVEYSTARKASTUKSISSA

Koneoppimisen mahdollisuuksia automaattiseen
tekstintunnistukseen

Leena Uronen

Turun yliopisto

Lääketieteellinen tiedekunta
Hoitotieteen laitos
Hoitotiede
Hoitotieteen tohtoriohjelma

Työn ohjaajat

Professori Sanna Salanterä
Hoitotieteen laitos
Turun yliopisto

Professori Jaakko Hartiala
Kliininen fysiologia ja isotooppilääke-
tiede
Turun yliopisto

Esitarkastajat

Dosentti Outi Kanste
Hoitotieteen ja terveyshallintotieteen tut-
kimusyksikkö
Oulun yliopisto

Dosentti Päivi Leikkola
Yhteiskuntatieteiden tiedekunta, ter-
veystieteiden yksikkö
Tampereen yliopisto

Vastaväittäjä

Professori Tarja Kvist
Hoitotieteen laitos
Itä-Suomen yliopisto

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä.

ISBN 978-951-29-8935-5 (PRINT)

ISBN 978-951-29-8936-2 (PDF)

ISSN 0355-9483 (Print)

ISSN 2343-3213 (Online)

Painosalama, Turku 2022

“If you can dream it, you can do it”.
- *Walt Disney*

TURUN YLIOPISTO

Lääketieteellinen tiedekunta

Hoitotieteen laitos

Hoitotiede

LEENA URONEN: Psykososiaalisten riskitekijöiden kirjaaminen työterveys-
huollon terveystarkastuksissa

Koneoppimisen mahdollisuuksia automaattiseen tekstintunnistukseen

Väitöskirja, 109 s.

Hoitotieteen tohtorihjelma

Toukokuu 2022

TIIVISTELMÄ

Tämän tutkimuksen tavoitteena oli kehittää ja arvioida työterveyshoitajien terveystarkastusten kirjaamista työn psykososiaalisten riskitekijöiden osalta. Tavoitteena oli kartoittaa koneoppimisen mahdollisuuksia tunnistaa psykososiaalisia riskitekijöitä työterveyshuollon terveystarkastuksien kirjauksissa. Tutkimus toteutettiin kaksivaiheisena. Ensimmäisessä vaiheessa tarkasteltiin työterveyshoitajien näkemyksiä terveystarkastusten sisällöstä ja kirjaamisesta (n=17, Delphi-paneeli kaksi kierrosta, laadullinen sisällön analyysi, tilastollinen analyysi). Toisessa vaiheessa koostettiin sanoja ja sanapareja terveystarkastusten kirjauksien työn psykososiaalisista tekijöistä PrimaEF standardin, PAS1010, pohjalta. Aineistona oli 15 443 terveystarkastusasiakirjaa yhdeltä kansalliselta työterveyspalvelujen tuottajalta. Tutkimuksessa verrattiin annotoituja (n=196 henkilön asiakirjat) sanoja ja sanapareja Työkykyindeksiin ja B-lausuntoihin liittyen eläkkeisiin, sairauspoissaoloihin ja kuntoutukseen (n=7078 asiakirjaa, tekstinlouhinnan puoliohjattu malli). Lopuksi arvioitiin koneoppimisen soveltuvuutta työn psykososiaalisten riskitekijöiden kirjausten automaattiseen tunnistamiseen (n=100, koneoppiminen ohjaamaton ja ohjattu NER). Tulokset toisesta vaiheesta analysoitiin tilastollisin menetelmin.

Tulosten perusteella työterveyshoitajilla ei ollut yhteistä näkemystä terveystarkastusten sisällöstä, myös mielipiteet terveystarkastusten kirjaamisesta vaihtelivat huomattavasti. Psykososiaalisten riskitekijöiden interventiona (95 %) työterveyshoitajat konsultoivat lääkäriä. Työterveyshoitajien terveystarkastuksissa kirjaamat psykososiaaliset riskitekijät korreloivat merkittävästi Työkykyindeksiin sekä eläkkeeseen, pitkittyneeseen poissaoloon tai kuntoutukseen tehtyyn B-lausuntoon. Koneoppimisella oli mahdollista löytää työn psykososiaalisia riskitekijöitä terveystarkastusten vapaan tekstin kirjauksissa.

Tutkimuksen perustella tarvitaan terveystarkastuksiin psykososiaalisten riskitekijöiden kirjaamiseen yhtenäinen tapa. Automaattisesti tunnistava työkalu tukisi työterveyshoitajan ja moniammatillisen tiimin päätöksentekoa ja työkyvyn lukutaitoa.

AVAINSANAT: terveystarkastus, psykososiaaliset riskitekijät, sähköinen kirjaaminen, koneoppiminen, työterveyshuolto

UNIVERSITY OF TURKU

Faculty of Medicine

Department of Nursing Science

Nursing Science

LEENA URONEN: Recording of Psychosocial Risk Factors in Health Checks of Occupational Health Care

Machine learning capabilities for automatic text recognition

Doctoral Dissertation, 109 pp.

Doctoral Programme in Nursing Science

May 2022

ABSTRACT

The aim of this study was to develop and evaluate the recording of occupational health examinations by occupational health nurses in relation to psychosocial risk factors at work. The aim was to explore the potential of machine learning to identify psychosocial risk factors in the recording of occupational health examinations by occupational health nurses. The study was conducted in two phases. In the first phase, the views of occupational health nurses on the content and recording of health checks were examined (n=17, Delphi panel two rounds, qualitative content analysis, statistical analysis). In the second phase, words and word pairs were compiled on the psychosocial aspects of the work of health check records based on the PrimaEF standard, PAS1010. The data set consisted of 15 443 health examination records from one national occupational health service provider. We compared annotated (n=196 individual documents) words and word pairs from the Work Ability Index and B statements related to pensions, sickness absence and rehabilitation (n=7078 documents, semi-guided model of text mining). Finally, the applicability of machine learning for the automatic identification of records of psychosocial risk factors at work was evaluated (n=100, machine learning unsupervised and supervised NER). The results of the second phase were analysed using statistical methods.

The results showed that occupational health nurses did not have a common view on the content of health assessments, and their opinions on the recording of health assessments also varied considerably. For psychosocial risk factor interventions (95%), occupational nurses consulted a doctor. The psychosocial risk factors recorded by occupational nurses during health examinations were significantly correlated with the Work Capacity Index and the B statement for pension, extended absence or rehabilitation. Machine learning was able to identify psychosocial risk factors at work in the free-text recordings of health examinations.

The study suggests that a uniform way of recording psychosocial risk factors is needed. An automatic identification tool would support the decision making and work literacy of the occupational health nurse and the multidisciplinary team.

KEYWORDS: health check, psychosocial risk factors, electronic documentation, machine learning, occupational health care

Sisällysluettelo

Lyhenteet	8
Osajulkaisuluettelo.....	9
1 Johdanto.....	10
2 Kirjallisuuskatsaus	11
2.1 Kirjallisuuskatsauksen toteutus	11
2.2 Työterveyshuollon terveystarkastuksen sisältöalueet	13
2.3 Työn psykososiaaliset riskitekijät.....	17
2.4 Terveystarkastusten sähköinen kirjaaminen.....	19
2.5 Terveystarkastusten kirjaamisen jatkohyödyntäminen koneoppimisen avulla.....	22
2.6 Yhteenveto kirjallisuuskatsauksesta	24
3 Tavoitteet	25
4 Tutkimusasetelma, aineiston keruu ja analysointi.....	27
4.1 Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältöalueista ja kirjauksesta.....	28
4.2 Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla	30
4.3 Työn psykososiaalisten riskitekijöiden mallinnus ja tunnistus tekstin louhinnalla terveystarkastusten kirjauksista.....	33
4.4 Työn psykososiaalisten riskitekijöiden tunnistaminen vapaan tekstin kirjauksesta koneoppimisen avulla	35
4.5 Tutkimuksen eettinen pohdinta.....	38
5 Tulokset	39
5.1 Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältöalueista ja kirjauksesta.....	39
5.2 Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla	45
5.3 Työn psykososiaalisten riskitekijöiden mallinnus ja tunnistus tekstin louhinnalla terveystarkastusten kirjauksista.....	50
5.4 Työn psykososiaalisten riskitekijöiden tunnistaminen koneoppimisen avulla vapaan tekstin kirjauksista.....	52
5.5 Yhteenveto tutkimustuloksista	55

6	Pohdinta.....	56
6.1	Tulosten pohdinta.....	56
6.2	Tutkimuksen pätevyys ja luotettavuus	60
6.2.1	Tiedon keruun pätevyys ja luotettavuus.....	60
6.2.2	Tutkimusmenetelmien pätevyys ja luotettavuus.....	61
6.2.3	Tulosten pätevyys ja luotettavuus.....	62
6.3	Ehdotuksia tulevia tutkimuksia varten.....	62
6.4	Tutkimuksen tulosten hyödyntäminen.....	63
7	Johtopäätökset.....	65
	Kiitokset.....	67
	Lähdeluettelo	69
	Osajulkaisut	75

Lyhenteet

AUC	Area Under the Curve
BERT	Bidirectional Encoder Representations from Transformers
CINAHL	Cumulative Index to Nursing and Allied Health Literature
ERSV	Erityinen sairastumisen vaara
EU	Euroopan Unioni
HTTK	Hyvä Työterveyshuoltokäytäntö
IE	Information extraction, tietojen poiminta
ILO	International Labour Organization, Kansainvälinen työjärjestö
Korpus	Kokoelma, joukko tekstejä tai dokumentteja
NER	Named entity recognition, nimettyjen kokonaisuuksien automaattinen tunnistus
n-grammi	n-merkin pituinen peräkkäinen jatkuva jakso tekstissä
NLP	Natural Language Processing, luonnollinen kielenkäsittely
NLTK	Natural Language Toolkit
PrimaEF	Psychosocial Risk Management Excellence Framework
PubMed	biolääketieteellinen hakukone ja viitetietokanta
ROC	Receiver Operating Characteristics, ROC-käyrä
SAS	Statistical Analysis System
SD	Standard Deviation
SPSS	Statistical Package for Social Sciences
THL	Terveysten ja hyvinvoinnin laitos
TKI	Työkykyindeksi
VNa	Valtioneuvoston asetus

Osajulkaisuluettelo

Väitöskirjan yhteenveto-osa perustuu seuraaviin alkuperäisjulkaisuihin, joihin viitataan tekstissä roomalaisilla numeroilla I–IV.

- I Uronen L, Hartiala J, Salanterä S. The content and documentation of health check-ups in occupational health care. *International Journal of Occupational Health and Public Health Nursing* 2014, 1(1) 87–97.
- II Uronen L, Heimonen J, Puukka P, Martimo K-P, Hartiala J, Salanterä S. Documentation of psychosocial factors with Work Ability Index in health check-ups. *Occupational Medicine* 2017, 67(2) 151–154.
- III Uronen L, Moen H, Teperi S, Martimo K-P, Hartiala J, Salanterä S. Towards automated detection of psychosocial risk factors with text mining. *Occupational Medicine* 2020, 70(3), 203–206.
- IV Uronen L, Salanterä S, Hakala K, Hartiala J, Moen H. Combining supervised and unsupervised named entity recognition to detect psychosocial risk factors in occupational health checks. *International Journal of Medical Informatics*, 2022,160, 104695.

Artikkelien käyttöön väitöskirjan osajulkaisuna on saatu kustantajien lupa.

1 Johdanto

Työterveyshuollossa terveystarkastuksia tehtäessä on kirjausten perusteella vaikeaa saada selkeää kuvaa työntekijän työkyvystä jatkuvan potilaskertomuksen perustella. Tieto terveystarkastuksissa tehdyistä toimenpiteistä on sirpaleista. Tämän tutkimuksen aikana työkyvyn käsite laajeni. Työkyky alettiin nähdä fyysisenä, psyykkisenä, sosiaalisena toimintakykynä ja myös osaamisena, arvoina, motivaationa ja asenteina (Ilmarinen ym. 2006, Pehkonen 2018). Työterveyshuollon terveystarkastukset vaikuttivat olevan irrallinen prosessi työkyvystä. Myös terveystarkastusten vaikuttavuutta alettiin kyseenalaistamaan työpaikoilla tämän tutkimuksen alkuvaiheessa.

Terveystarkastuksen vaikuttavuutta on haastavaa arvioida tieteellisesti. Tutkimusten mukaan niiden avulla pystytään seulomaan työkykyä ja tukemaan työhön paluuta pitkältä sairauspoissaololta. (Leino ym. 2014, Sauni & Leino 2007.) Terveystarkastuksia on työterveydessä tehty koko sen historian ajan altisteperusteisesti (melu, pölyt, hitsausuurut yms.). Altisteiden määrä työpaikoilla vähenee hyvän työturvallisuuden ja teknologian avulla. Työelämässä on tunnistettu psykososiaalisia riskitekijöitä, jotka kuormittavat kaikkia aloja. (Liira ym., 2014, Vartia, 2012.) Euroopassa 25 % työntekijöistä kokee työhön liittyvää stressiä suurimman osan työajastaan. Yli 50 % poissaoloista on yhteydessä psykososiaalisiin riskitekijöihin kuten työhön liittyvään stressiin, kiusaamiseen ja työväkivaltaan (Bernal ym. 2015, Eurofund 2014, Slany ym 2014). Näiden kuormitustekijöiden yhtenäiseen kirjaamiseen terveystarkastuksissa ei kirjallisuuden mukaan ole kansallisesti sovittua käytäntöä.

Suomessa työterveyshuollon terveystarkastuksia ja niiden kirjaamista on tutkittu vähän. Louhi-hanke, myöhemmin Ikitik-konsortio, johon myös tämä tutkimus kuuluu, pyrki koneoppimisen ja tekoälyn avulla lounimaan tietoa potilaskertomusten vapaasta tekstistä (Suominen ym. 2007). Tällä tutkimuksella haluttiin selvittää, voitaisiinko psykososiaalisten riskitekijöiden kirjaamista kehittää ja voidaanko niitä kirjauksista automaattisesti tunnistaa. Tämä tutkimus toteutettiin neljänä osatutkimuksena vuosina 2005–2022. Aineistot kerättiin yksityisten työterveyshuollon palveluntuottajan tiedostoista. Aineistot koostuivat sähköisiin potilastietojärjestelmiin kirjaamista työterveyshuollon terveystarkastuksista.

2 Kirjallisuuskatsaus

2.1 Kirjallisuuskatsauksen toteutus

Kirjallisuuskatsaus koostuu neljästä osa-alueesta. Ensimmäiseksi tarkastellaan työterveyshuollon terveystarkastuksiin kohdistuvaa tutkimusta. Toisena osa-alueena tarkastellaan työn psykososiaalisia tekijöitä ja etenkin riskitekijöitä käsittelevää tutkimusta. Kolmannessa osa-alueessa paneudutaan hoitotyön, etenkin työterveyshoitajien, sähköisen kirjaamisen tutkimuksiin. Viimeisenä alueena tarkastellaan vapaan tekstin kirjausten luonnollisen kielen käsittelyä ja koneoppimista. (Kuvio 1.)



Kuvio 1. Kirjallisuuskatsauksen osa-alueet.

Kirjallisuuskatsauksen tavoitteena on kuvata työterveyshuollon terveystarkastuksia, psykososiaalisia riskitekijöitä, niiden kirjaamista sekä jatkohyödyntämistä. Tutkimusraportteja haettiin useassa vaiheessa useilla eri hakulauseke- ja hakusanayhdistelmillä. Työterveyshuollon terveystarkastuksia käsittelevä kansainvälinen tieteellinen kirjallisuus ja tutkimukset ovat vähäisiä johtuen ainakin osittain työterveyshuollon toteuttamisessa olevista maakohtaisista eroista. Kirjallisuushakuja tehtiin käytämällä PubMed (Medline), CINAHL, Cochrane Library ja Web of Science- tietokantoja. Myös kirjaston asiantuntijoita käytettiin apuna, koska hakutuloksia oli niukasti. Lisäksi julkaisuista etsittiin pääkäsitteisiin liittyvää kirjallisuutta PubMedin avulla sisältäen (Medline) ja CINAHL-tietokannat. Hakua täydennettiin myös asiankuuluvien organisaatioiden verkkosivuilta (Työterveyslaitos, Finlex, Kela) ohjeista, kansallisesta lainsäädännöstä ja Internetistä. Manuaalisia hakuja tehtiin myös

viiteluetteloista mukana olleista julkaisuista. Haku rajattiin vuosiin 1990–2019 tutkimusvaiheittain ja mukaan otettiin vain ne vertaisarvioidut tieteelliset alkuperäisjulkaisut, jotka olivat kirjoitettu englanniksi tai suomeksi. Käsitteet ja hakutermit ovat esitetty tarkemmin taulukossa 1 ja kirjallisuuskatsaukseen sisällytetty artikkelit taulukossa 2.

Taulukko 1. Tieteellisessä tietokannassa käytetyt hakutermit.

Käsitteet	Hakutermit
terveystarkastus	health check, health check-up, health promotion AND (((occupatio*)) OR (occupational *)) OR (occupatio*[MeSH Terms]) OR (occupational *[MeSH Terms])
sähköinen kirjaaminen	electronic patient record system, electronic healthcare records, nursing documentation AND (((((psychosocial)) AND (risk factors)) AND (documentation))) AND (((((occupatio*)) OR (occupational *)) OR (occupatio*[MeSH Terms]) OR (occupational *[MeSH Terms]))
työn psykososiaaliset tekijät	(((psychosocial)) AND (risk factors)) AND (((((occupatio*)) OR (occupational *)) OR (occupatio*[MeSH Terms]) OR (occupational *[MeSH Terms]) OR work)
tekstin louhinta	text mining, data mining, natural language processing AND (((((occupatio*)) OR (occupational *)) OR (occupatio*[MeSH Terms]) OR (occupational *[MeSH Terms]) OR (nursing),

Taulukko 2. Kirjallisuuskatsaukseen sisällytettyjen artikkelien haun tulokset.

Käsitteet	PubMed/ Medline Kaikki/ artikkelit	Cinahl Kaikki/ artikkelit	Cochrane Library Kaikki/ artikkelit	Web of Science Kaikki/ artikkelit	Mukaan otettu artikkelit
terveystarkastus	7/2	4/1	2/0	5/1	4
sähköinen kirjaaminen	19/5	22/6	10/4	8/5	19
työn psyko-sosiaaliset tekijät	13/8	17/6	8/4	11/5	23
tekstin louhinta	42/7	10/3	6/3	36/11	24

2.2 Työterveyshuollon terveystarkastuksen sisältöalueet



Työterveyshuollon terveystarkastuksen sisältöalueet

Työterveyshuolto edistää yhdessä työpaikan ja työntekijän kanssa terveyttä, työkykyä ja turvallisuutta työssä. Suomessa ennaltaehkäisevä työterveyshuolto on lain mukaan pakollinen yrityksessä, jossa on yksikin työntekijä (Työterveyshuoltolaki 1383/2001). Työterveyshuoltoa ohjaa lain lisäksi myös ”Hyvä Työterveyshuoltokäytäntö” eli HTTK (Uitti, 2014). Terveystarkastukset ovat merkittävää työterveyshuollon toimintaa. Suomen työterveyshuolloissa tehdään yli miljoona (1,4 miljoona vuonna 2019) terveystarkastusta vuodessa. Tämä luku on pysynyt suhteellisen samana jo usean vuoden ajan. Työterveyshuollon piirissä on yli 1,9 miljoonaa suomalaista työkäistä, joita työterveyshuollon terveystarkastukset koskevat (Kelan Työterveyshuoltotilasto 2019.) Näiden terveystarkastusten tavoitteena on tukea työntekijän työkykyä ja terveyttä. Noin 30 % terveystarkastuksista on lakisääteisiä. Tällöin työssä on erityinen sairastumisen vaara. Erityinen sairastumisen vaara (ERSV) määritellään Valtioneuvoston asetuksessa (VNa 1485/2001) seuraavasti: ”ERSV aiheutuu sellaisista työoloista, joissa fyysisen, kemiallisen tai biologisen tekijän aiheuttamana voi todennäköisesti seurata sairaus, liiallinen altistuminen tai vaara lisääntymisterveydelle. Myös yötyö ja erityinen väkivallan uhka työssä voivat aiheuttaa erityistä sairastumisen vaaraa.” Muut terveystarkastukset ovat yleensä työn aloittamiseen, ikään, työvuosiin tai työkykyyn suunnattuja.

Terveystarkastus tehdään (Työterveyshuoltolaki 1383/2001, 7 §):

- ” työntekijän iän, sukupuolen, fysiologisen tilan sekä työstä johtuvien terveyden vaarojen ja haittojen sitä edellyttäessä
- tarvittaessa työaikajärjestelyjen perusteella

- uusien käyttöön otettavien teknologioiden, menetelmien ja aineiden terveysvaikutusten sitä edellyttäessä
- tarvittaessa työntekijän terveydentilan toteamiseksi ja tarvittavan jatko-seurannan suunnittelemiseksi altistumisen loputtua
- työn terveydellisten erityisvaatimusten perusteella
- tarpeen vaatiessa työtehtävien olennaisesti muuttuessa tai työuran eri vaiheissa
- työntekijän terveydentilan, työkyvyn tai työhön liittyvien sairauksien ja oireiden sitä edellyttäessä
- työssä selviytymisen mahdollisuuksien arvioimiseksi ja tukemiseksi sekä tarvittaessa työn sopeuttamiseksi työntekijän voimavaroihin, terveydellisiin edellytyksiin ja työ- ja toimintakykyyn
- tarvittaessa ennen palvelussuhteen päättymistä terveyden ja työkyvyn arvioimiseksi ja työllistymisen terveydellisten edellytysten selvittämiseksi sekä suunnitelman laatimiseksi terveyden ja työnkyvyn ylläpitämisestä ja jatkotoimiin ohjaamisesta.”

Työterveysshuollossa terveystarkastukset ovat tavoitteellisia. Terveystarkastuksen tavoitteet ja sisältö suunnitellaan yhteistyössä yrityksen kanssa, työterveysshuollon toimiessa suunnittelussa sisällön asiantuntijana huomioiden altisteet, työn riskit ja työkyvyn tuen tarpeet. Terveystarkastuksen tekee yleensä työterveysshuoltoon päte- vöitynyt työterveysshoitaja konsultoiden moniammatillista tiimiä (Vna 708/2013).

Asetuksessa (708/2013) terveystarkastuksen merkitystä korostetaan työntekijän voimavaroja lisäävänä, jatkuvana ja suunnitelmallisena työkyvyn seurannan väli- neenä. Lisäksi asetuksessa painotetaan työssä selviytymisen mahdollisuuksien arvi- oimista ja tukemista sekä työn sopeuttamista työntekijän voimavaroihin, terveyteen ja työ- ja toimintakykyyn. Terveystarkastusten tavoitteena on näin mahdollisimman varhainen tarpeellisen hoidon ja kuntoutustarpeen selvittäminen, hoitoon ja kuntou- tukseen ohjaaminen sekä hoidon ja kuntoutuksen seurannan suunnittelu, työkyvyt- tömyyden ehkäisy ja työhön paluun tuki (Leino ym. 2014). Tavoitteet on esitelty taulukossa 3. Terveystarkastuksen sisällöistä on ERSV-tarkastuksiin annettu suosi- tuksia (Karvala ym. 2019). Muissa nk. vapaaehtoisissa terveystarkastuksissa sisältö ja tavoitteet sovitaan toimintasuunnitelmassa yhdessä yrityksen kanssa. Terveystar- kastuksiin ei ole kansallisesti määritelty järjestystä tai rakennetta kirjallisuuden pe- rusteella.

Työterveysshuollon terveystarkastukset koetaan pääsääntöisesti hyödyllisinä. Tutkimuksessa työntekijät arvostivat terveystarkastuksen yksilöllistä toteutustapaa. Tutkituista jopa 75 %:n mielestä terveystarkastuksissa käsiteltiin hänelle itselleen

tärkeitä asioita. Tutkittavat kokivat, että suurin vaikutus terveystarkastuksilla oli terveyteen ja elintapoihin sekä jonkin verran työ- ja toimintakykyyn. (Leino ym. 2014.) Raution ym. (2000) tutkimuksessa todettiin, että työterveyshuollon terveystarkastuksissa painopistealueina olivat fyysinen kunto ja yleinen terveydentila. Terveystarkastuksessa laaditaan terveystarkastus suunnitelma yhdessä työntekijän kanssa tulevista työkykyä edistävästä jatkotoimista. Henkilökohtaisen terveystarkastus suunnitelman avulla voidaan arvioida terveystarkastuksen vaikutusta työkyvyn tukemiseen. Terveystarkastus suunnitelman avulla voidaan seurata sovittujen jatkotoimenpiteiden toteutumista sovittuun aikaväliin. (Liira ym. 2014, Manninen ym. 2007.)

Työterveyshuollon terveystarkastuksen tavoitteena on varhaisessa vaiheessa havaita terveysongelmat sekä työn ja työkykyisyyden riskitekijät. Terveystarkastuksessa työntekijä on aktiivinen toimija. (Manninen ym. 2007.) Työntekijä tuo vuorovaikutustilanteeseen oman työnsä, terveytensä ja elämäntilanteensa ja työterveyshuolto työn ja terveyden asiantuntemuksensa. Ohjauksen ja neuvonnan tavoitteena on lisätä sekä yksittäisen työntekijän että koko työpaikan työterveysosaamista. Palmgrenin (2008) tutkimusten perusteella terveystarkastuksen kirjauksissa on harvoin mainittu ohjausta ja neuvontaa. Työterveyshuollon terveystarkastuksien vaikuttavuudesta on vain vähän tutkimusta (Leino ym. 2014, Sauni ja Leino 2007). Tieteellistä näyttöä on saatu elämäntapariskien vähentämisessä (Martimo & Antti-Poika 2000, Uitti ym. 2007, Viitasalo ym. 2011). Yleisistä terveystarkastuksista (ei työterveyshuollossa) tehdyssä tutkimuksessa todettiin, etteivät ne vähentäneet sairastumattomuutta tai kuolleisuutta (Krogsbøll ym. 2012).

Taulukko 3. Terveystarkastusten pää- ja osatavoitteet (mukailtu Leino ja Rautio 2014).

Työhön liittyvien sairauksien ja oireiden ehkäisy – terveellisen ja turvallisen työympäristön sekä toimivan työyhteisön kehittäminen	
Työpaikkaan liittyvät:	<ul style="list-style-type: none"> - Tiedon kerääminen henkilöstön terveydestä ja työkyvystä sekä työssä esiintyvistä terveysongelmista - Palautteen antaminen terveyttä edistävien toimenpiteiden suunnittelemiseksi työyhteisössä - Terveyttä uhkaavien kuormitustekijöiden vähentäminen
Työntekijään liittyvät:	<ul style="list-style-type: none"> - Työntekijän sairastumisen ehkäisy työssä olevien altisteiden ja kuormitustekijöiden vaikutuksesta - Sairauksien oireiden lisääntymisen ja sairauksien pahenemisen estäminen - Työntekijöiden auttaminen siinä, että he voivat itse edistää omalla toiminnallaan työpaikan terveellisyyttä, turvallisuutta ja työyhteisön toimivuutta
Työntekijän terveyden ja työ- ja toimintakyvyn tukeminen	
<p>Mahdollisimman varhainen toteaminen:</p> <ul style="list-style-type: none"> - terveydentilan ja työkyvyn muutokset - alkavat sairaudet - merkit jaksamisongelmista ja uhkaavasta työkyvyn laskusta - tarvittaessa arvio työkyvystä, työrajoitteista tai työkuormituksesta 	
<p>Kartoittaminen:</p> <ul style="list-style-type: none"> - hoidon tarve - hoitoon tai kuntoutukseen ohjaus - työntekijän työhön paluun ja työssä jatkamisen tukeminen mahdollisista sairauksista huolimatta 	
<p>Terveyttä ja työkykyä edistävän terveystarkastuksen laatiminen yhdessä työntekijän kanssa:</p> <ul style="list-style-type: none"> - terveyttä edistäviin valintoihin rohkaiseminen - työntekijän aktivoiminen oman terveytensä sekä työ- ja toimintakyvynsä ylläpitämiseen ja - työntekijän aktivoiminen työ- ja toimintakyvyn liittyvien tekijöiden tunnistamiseen 	
<p>Seuranta ja arviointi:</p> <ul style="list-style-type: none"> - työntekijän fyysinen ja psyykinen terveys ja toimintakyky - työntekijän elintapavalinnat - työntekijän terveyttä tukevat voimavarat 	

2.3 Työn psykososiaaliset riskitekijät



Kansainvälinen työjärjestö (ILO) määrittelee työn psykososiaaliset tekijät vuorovaikutukseksi seuraavien tekijöiden välillä: työympäristö, työn sisältö, organisaation tila, työntekijöiden pätevyys ja tarpeet, kulttuuri, tekijät voivat havaintojen ja kokemuksen kautta vaikuttaa terveyteen, työn suorituskykyyn ja työtyytyväisyyteen. Negatiivinen vuorovaikutus ammatillisiin olosuhteisiin ja inhimilliset tekijät voivat johdattaa emotionaalisiin häiriöihin, käyttäytymisongelmiin sekä biokemiallisiin ja neurohormonaalisiin muutoksiin, mikä lisää henkisten ja fyysisten sairauksien riskejä. Haitallisia vaikutuksia voi olla odotettavissa myös työtyytyväisyyteen ja työn suorituskykyyn. (International Labour Office 2016.)

Työn psykososiaalisilla tekijöillä tarkoitetaan niitä työhön, työympäristöön ja työn organisointiin liittyviä tekijöitä, jotka vaikuttavat työntekijän terveyteen, työkykyyn ja työhyvinvointiin (Leka ym. 2003). Tunnistettavia psykososiaalisia tekijöitä ovat yrityksen ilmapiiri, johtaminen sekä sen oikeudenmukaisuus, työn vaikutusmahdollisuudet, ihmissuhteet ja roolit organisaatiossa (Leka ja Kortum 2008). Psykososiaalisia riskitekijöitä on kaikissa ammateissa ja niiden vaikutus työntekijöiden terveyteen, turvallisuuteen ja hyvinvointiin on dokumentoitu tieteellisesti viime vuosikymmeninä (Leka ym. 2015, Bliese ym. 2017). Tietyissä ammateissa ja ammattiryhmissä psykososiaalisia riskitekijöitä on enemmän. Työtä saatetaan tehdä epäsuotuisissa olosuhteissa, josta aiheutuu negatiivisia seurauksia, kuten ammattitaudit, heikko suorituskyky ja työtapaturmat. (Cendales ym. 2016.)

Psykososiaaliset riskit ovat niitä työn suunnittelun, hallinnan, sosiaalisen tai organisatorisen ympäristön näkökohtia, jotka voivat aiheuttaa psykologisia tai fyysisiä

haittoja (Eurofund 2015). Työn psykososiaalisia riskitekijöitä ovat esimerkiksi kohtuuton työmäärä ja jatkuva kiire, liian vaikeat työtehtävät ja suuri vastuu sekä samanaikaisuuden paineet työssä. Toisaalta työ voi olla myös yksitoikkoista ja tapahutumaköyhää tai liian helppoa. Tutkimusten mukaan työn psykososiaaliset kuormitustekijät myös liittyvät vuorovaikutustilanteiden häiriöihin, puutteisiin tai niiden heikkoon laatuun. (Agency for Safety ja Health at Work 2007, Leka ym. 2015, Vahtera ym. 2000, Vartia 2012.)

Työn psykososiaalisten riskien hallinnalla on myös liiketaloudellista merkitystä. Riskien hallinnalla tuetaan työntekijöiden hyvää terveyttä, työkykyä sekä hyvinvointia ja samalla parannetaan organisaation tuottavuutta. (Cox ym. 2000, Frick ja Zwetsloot 2007.) Työhön liittyvät psykososiaaliset riskit voivat aiheuttaa stressiä, sairauspoissaoloja ja varhaiseläkkeelle siirtymistä (European Agency for Safety ja Health at Work 2008, Leka ym. 2003, Slany ym. 2013). Tutkimuksissa työn psykososiaalisista riskitekijöistä on todettu vahvaa näyttöä siitä, että korkeat työn vaatimukset, heikko työn johtaminen, heikko työoverituki ja esimiehen tuki sekä epäoikeudenmukaisuus ennustivat stressin esiintyvyyttä (Niuwenhuijsen ym. 2010). Rasakas fyysinen työkuormitus, vaarallinen altistuminen, korkeat psyykkiset vaatimukset, syrjintä, kiusaaminen, työn matala vaatimustaso sekä työ- ja yksityiselämän epätasapaino olivat voimakkaimmin yhteydessä sairauspoissaoloihin (Laaksonen ym. 2010, Niedhammer ym. 2012). Suomalaisessa tutkimuksessa suurimmat sairauspoissaolojen riskit liittyivät työn hallintaan, työn vaatimuksiin, työn negatiivisiin muutoksiin sekä huonoon sosiaaliseen tukeen (Vahtera ym. 2000). Työkyvyttömyyseläkkeiden on tutkimuksissa todettu olevan yhteydessä fyysiseen työmäärään, tietokone-taitoihin ja heikkoon työn hallintaan (Christensen ym. 2008, Lahelma ym. 2012).

Psykososiaaliset riskitekijät vaikuttavat merkittävästi työterveyteen ja -turvallisuuteen Euroopassa. Työperäinen stressi on sairausloman yleisin syy. Se vaikuttaa yli 40 miljoonaan työntekijään Euroopan Unionin (EU) alueella. Tutkimukset ovat osoittaneet työperäisellä stressillä olevan yhteyttä 50–60 prosentissa kaikista menetyksistä työpäivistä. (Cox ja Griffiths 2010.) Kaikilla työn psykososiaalisilla riskitekijöillä, erityisesti vaarallisilla altisteilla, työn vaatimuksilla, sosiaalisilla suhteilla, työpaikalla esiintyvällä väkivallalla ja vuorotyöllä, on vahva yhteys pitkään sairauspoissaoloon (Laaksonen ym. 2010, Slany ym. 2014). Työn psykososiaaliset riskitekijät vaikuttavat myös varhaiseläkkeelle siirtymiseen ja poissaoloihin. Heikko työnhallinta ja organisaation epäoikeudenmukaisuus voivat lisätä varhaiseläkkeelle siirtymistä (Heponiemi ym. 2008, Suodican ym. 2013). Noin 10–15 % työkyvyttömyyseläkkeistä johtuu psykososiaalisista työympäristöön liittyvistä tekijöistä (Christensen ym. 2008). Fyysinen työmäärä, näyttöpäätetyö ja heikko työnhallinta ovat riskejä työkyvyttömyyseläkkeelle (Christensen 2008, Lahelma ym. 2012).

Prima EF konsortio on luonut kattavaa ehkäisy politiikkaa psykososiaalisten riskitekijöiden tunnistamiseen ja ennaltaehkäisyyn EU:ssa sekä estämään pitkäaikaisia

sairauspoissaoloja. Prima EF tarjoaa työkaluja ja toimintatapoja psykososiaalisten riskien ehkäisy politiikkaan sekä jäsenmaiden käytäntöjen edistämiseksi kansallisella ja yritystasolla. EU:n jäsenvaltioiden kaikilla työntekijöillä pitäisi periaatteessa olla pääsy työterveyshuoltoon. (Leka ym. 2015, Leka ja Cox, 2011, Leka ja Kortum 2008.) Prima EF konsortio on kehittänyt työn psykososiaalisten riskien hallintaa koskevan eurooppalaisen viitekehysten sekä mallin, PAS 1010. Riskien hallinnasta on tehty myös opas työpaikkojen käyttöön, joka tiivistää psykososiaalisten riskien hallinnan eurooppalaisen kehyksen (PRIMA-EF) keskeiset osa-alueet. Opas antaa apuvälineitä sekä arviointimenetelmiä psykososiaalisten riskien hallintaan työpaikoille ja työnantajan edustajille. PAS 1010- mallissa keskitytään erityisesti työperäisen stressin ja työpaikalla tapahtuvan väkivallan ehkäisyyn. Psykososiaaliset riskitekijät on luokiteltu 13 alaluokkaan (työn hallinta, työympäristö ja välineet, organisatiokulttuuri ja toiminta, sosiaaliset suhteet, rooli organisaatiossa, työuran kehitys, työn ja muun elämän, työpaikkaväkivalta, häirintä, työpaikkakiusaaminen) (Leka ja Cox 2011). Taulukko ja tarkempi erittely tässä tutkimuksessa käytetyistä edellä mainituista 13 psykososiaalisten riskitekijöiden alaluokista on esitelty osajulkaisussa II.

2.4 Terveystarkastusten sähköinen kirjaaminen



Terveystarkastusten sähköinen kirjaaminen

Työterveyshuollon potilasasiakirjojen laadinnassa noudatetaan samoja määräyksiä kuin laissa potilaan asemasta ja oikeuksista. Se sisältää ohjeet potilasasiakirjojen laatimisesta ja säilyttämisestä, tietojen salassapidosta, potilasmuistutuksesta, kantelusta ja valituksesta. (L785/92.) Laki yksityisyyden suojasta työelämässä (L759/04) täydentää edellä mainittua lakia. Lain mukaan terveydentilaa koskevia tietoja saavat

käsitellä vain ne henkilöt, jotka tekevät työsuhdetta koskevat päätökset näiden tietojen perusteella. Hyvän Työterveyshuoltokäytännön mukaisesti työterveyshuollossa voidaan kerätä vain niitä tietoja, joilla on merkitystä työntekijän terveyden tai työolojen terveellisyyden tai työterveyshuollon toteuttamisen kannalta. Tällöin terveystarkastuksissa kerätään vain sellaista tietoa, joka on tarpeellista terveystarkastuksen ennalta asetettujen tavoitteiden saavuttamiseksi. (Uitti 2014.)

Suomessa tuli voimaan vuonna 2007 laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä. Tämän perusteella rakentui myös sosiaali- ja terveydenhuollon valtakunnalliset tietojärjestelmäpalvelut, Kanta-palvelut. Kanta-palvelut mahdollistavat paremman päätöksenteon sekä sujuvamat hoitoketjut, koska potilastiedot ovat laajasti ammattilaisten käytettävissä organisaatiosta tai lääketieteen alasta riippumatta. (L159/2007.) Kanta-palveluiden eli potilastiedon arkiston toimivuuden edellytyksenä on, että terveydenhuollon ammattilaiset käyttävät potilastietojen kirjaamiseen yhteisesti sovittuja käsitteitä ja luokituksia. Potilastiedon arkiston avulla varmistetaan parhaiten kansallisesti yhtenäinen kirjaamistapa. (Häyriinen ym. 2011.) Potilastiedon arkistossa käytettäviä tietorakenteita määrittellen Terveyden ja hyvinvoinnin laitoksen (THL) johdolla. Työterveyshuollossa kaikki ammattiryhmät kirjaavat käynnit, myös sairaanhoitokäynnit TYÖ-näkymälle. Työterveyshuollon ydintietojen määrittely on potilastiedon arkistossa vielä osittain kesken. (THL 2018b.) Potilastiedon arkistoa kehitetään tulevaisuudessa tukemaan myös työterveyshuollon kirjaamistarpeita (Nissinen ym. 2018).

Sähköinen kirjaamisjärjestelmä tarjoaa mahdollisuuden tukea kliinistä päätöksentekoa, kehittää ja analysoida tietokantaa sekä parantaa hoitoa ja kommunikointia (de Groot 2019, Mahler 2016, Rosenbloom ym. 2011). Sähköinen kirjaamisjärjestelmä varmistaa paremmin tietosuojan toteutumisen, sillä tietojen käytöstä jää aina käyttäjän lokimerkintä tietojärjestelmään (Entzeridou ym. 2018). Sähköisen potilaskertomuksen käyttö edellyttää, että keskeiset tiedot tallennetaan potilasasiakirjoihin yhteisesti sovitussa, rakenteellisessa muodossa. Rakenteellisella kirjaamisella tarkoitetaan etukäteen sovittujen kirjattavien tietojen määrittelyä, ainakin otsaketasolla kirjaaminen on määritelty. Näiden otsakkeiden alle voidaan kirjata joko vapaamuotoisesti tai olemassa olevia nimikkeitä ja luokituksia hyödyntäen. (Ensio ja Saranto 2004, Mäkelä-Bengs ja Vuokko 2014.) Näitä yhdenmukaisia rakenteellisia tietoja sanotaan ydintiedoiksi. Nämä ydintiedot on määritetty myös Kanta-palvelussa. Potilasasiakirjan sähköisessä kirjaamisessa on huomioitava tiedon sisäisen siirron ja tiedon luovutustarpeiden sekä yksityisyyden suojan vaatimukset. Potilasasiakirjaan jää sähköisessä kirjaamisjärjestelmässä aina myös vapaan tekstin kirjaamisen tila. (Sosiaali- ja terveysministeriö 2004a.)

Hoidon kirjaaminen potilaskertomukseen ja erilaisiin hoitolomakkeisiin on keskeinen osa päätöksentekoa hoitotyössä. Kirjaamisen tavoitteena on hoidon joustava eteneminen, hoitoa koskevan tiedon saannin ja hoidon jatkuvuuden turvaaminen,

sekä hoidon toteutuksen ja sen seurannan varmistaminen. (Lauri ja Salanterä 2002.) Tutkimuksissa on todettu, että hoitajan on tärkeää ymmärtää päätöksentekoprosessi hoitotyössä, jotta hän tunnistaa onnistuneet päätökset ja voi näin kehittää omaa toimintaansa. Tämän on taas osoitettu parantavan potilaiden hoitoa ja hoitajien kykyä toimia potilaiden kanssa yhteistyössä. (Buckingham 2000, Ensio 2008, Lauri ja Salanterä 2002, O'Neill ym. 2005.) Hoitotyön kirjaamisen tulisi todentaa myös järjellinen ja kriittinen ajattelu, josta on seurannut päätöksenteko ja toimenpiteet (Blair ja Smith 2012). Tämän lisäksi potilaskertomukseen kirjattujen tietojen tulee säilyä lukukelpoisina ja muuttumattomina niille laissa määritetyn säilytysajan. Tämä varmistaa oikeusturvan potilaalle ja ammattilaiselle. (Häyrynen ym. 2008, Wilbanks ja Moss 2018.)

Rakenteelliseen elektroniseen tietokantaan yhdenmukaisesti kirjaaminen parantaa kirjaamisen laatua. Asiakirjojen yhdenmukaistaminen hoitoprosessin kanssa, vaikiintuneen käsitteistön käyttö sekä käyttäjäystävällisten muotojen ja järjestelmien käyttäminen ovat edellytykset korkealaatuiselle hoitotyön kirjaamiselle. Hyvin suunniteltu sähköisen kirjaamisen sisällön rakenne tuottaa järjestelmällisempää, yhtenäisempää ja tarkempaa kirjaamista. Rakenteelliset valmiit hoitopolut sähköisessä kirjaamisessa vähentävät virheitä ja parantavat hoitotyön laatua, koska esimerkiksi lääkitys kirjautuu lähes automaattisesti eikä asiakirjoja ole kadoksissa. (de Groot 2019, Kelley ym. 2011, Mahler ym. 2007, Mykkänen 2016, Roosenblom ym. 2011, Saranto ym. 2014, Stengel ym. 2004.) Suomessa, työterveyshuollossa terveystarkastukset kirjataan moniammatillisesti sähköiseen potilastietojärjestelmään. Osa työterveystarkastusten kirjaamisesta on rakenteisessa muodossa, mutta suurin osa työterveyshoitajan kirjauksesta on vapaata, kertovaa tekstiä. (Leino ym. 2014.)

Työterveyshuollossa on todettu puutteita työhön ja työolosuhteisiin liittyvien tietojen kirjaamisessa. Osa-alueet, jotka löytyivät terveystarkastuksista kirjattuina, olivat terveydentila, terveys ja turvallisuus, fyysinen tarkastus ja hoitotyön arviointi. Työhistorian osa-alueita ei kirjattu 25 %:ssa tiedostoihin eikä myöskään tietoa nykyisestä työtilanteesta. Kirjaukset olivat puutteellisia myös ammatista, altistumisesta, työprosessista, työympäristöstä tai työajoista. (Silveira ja Marin 2005.) Terveystarkastuksen sisällön ja kirjauksen tulisi olla täsmällistä ja jäsenneltyä tukeakseen terveydenhuollon ammattilaisia päätöksenteossa sekä jatkotoimenpiteissä työntekijän ja työyhteisön työkyvyn sekä työhyvinvoinnin parantamiseksi. Terveystarkastusten kirjaamista pitäisi pystyä hyödyntämään jatkokäynnillä sekä jatkohoittoon. Leino ym. (2014) totesivat tutkimuksessaan terveystarkastuksien kirjaamisen olevan puutteellista estäen vaikuttavuuden arvioinnin. Tämä osin siksi, ettei terveys-suunnitelmia kirjata, joten seuranta ja arviointi on haasteellista. Suosituksena Leino ym. (2014) toteavat: ”Työterveyshuollon terveystarkastusten kirjaamiskäytäntöjä tulee kehittää siten, että ne palvelevat kokonaisvaltaista lähestymistä ja työntekijöiden tarpeista lähtevää ja voimavaroja tukevaa toimintaa”.

2.5 Terveystarkastusten kirjaamisen jatkohyödyntäminen koneoppimisen avulla



Terveystarkastusten kirjaamisen jatkohyödyntäminen koneoppimisen avulla

Vapaan tekstin kirjaaminen on pääasiallinen terveystarkastusten kirjaamismuoto sähköisissä työterveyshuollon potilastietojärjestelmissä. Vapaan tekstin kirjaaminen on kuvaileva ja kertova tapa ilmaista käsitteitä ja tapahtumia, mutta asettaa haasteita hakutoiminnoissa, yhteenvedoissa, päätöksenteon tukemisessa tai tilastollisessa analyysissä. Terveystarkastusten tietojärjestelmiin tallennettua suurta määrää vapaan kirjaamisen tekstitietoja voidaan hyödyntää luonnollisella kielenkäsittelyllä (Natural Language Processing, NLP). NLP on muotoilua ja tutkimusta laskennallisesti tehokkaista mekanismeista luonnollisen kielen kautta tapahtuvaan viestintään (Clark ym. 2013).

NLP-teknologiat tarjoavat ratkaisun vapaamuotoisen tekstitiedon muuntamiseen rakenteelliseen esitysmuotoon. NLP on monitieteinen laskennallinen kielitiede, joka keskittyy luonnollisen kielen teknisiin näkökohtiin. NLP sisältää kirjoitetun ja puhutun kielen sekä luonnollisen kielen laskennallista mallintamista, analysointia ja luomista. Se antaa käsityksen ihmiskielen ja tietotekniikan luonteesta käsittäen useita tilastollisia lähestymistapoja, minkä vuoksi myös tilastot ja koneoppiminen ovat välttämättömiä NLP:ssä. (Clark ym. 2013.)

NLP tutkimus keskittyy rakentamaan laskennallisia malleja luonnollisen kielen ymmärtämiseksi (Carbonell ja Hayes 1992). Viimeisten vuosikymmenien aikana on tutkittu ja testattu erilaisia NLP-tekniikoita, jotka sopisivat potilaskertomuksiin (Linguistic String Project, MedLEE, Mplus). Tutkimuksia NLP:stä ja koneoppimisesta löytyy enemmän lääketieteestä mm. patologiasta ja radiologiasta (Haug 1997, Hrip-

sak 2002, Sager 1987) sekä jonkin verran hoitotieteestä esim. tehohoidosta ja koti-
hoidosta (Moen ym. 2019, Suominen ym. 2007, Topaz, 2019). Lupaavaa tutkimusta
on esim. Nimlemner- menetelmästä, jossa louhinta voidaan toteuttaa ilman nimettyä
tietojoukkoa. Tällainen menetelmä sopii hyvin aloille, joissa NLP tekniikan kehitys
on vähäisempää kuten esim. hoitotiede. (Topaz. 2019.)

Tietojen poiminta (Information extraction, IE) on NLP:n alakäsite. IE vaatii tyy-
pillisesti jonkin verran "esikäsitteilyä", kuten esim. oikeinkirjoituksen tarkistuksen,
asiakirjan rakenteen analyysin, lauseen jakamisen, arvon määrittämiseen tai sanatun-
nisteiden erottamisen. IE sisältää siis tiedon ennalta määriteltyjen tietotyyppien purka-
miseen tekstistä. Asiayhteyteen liittyvät piirteet, kuten kielteisyys, ajallisuus ja ta-
pahtuman kohteen tunnistaminen ovat ratkaisevan tärkeitä poimittujen tietojen tark-
kaan tulkintaan. (DeJong 1982, Spyns 1996.)

Nimettyjen kokonaisuuksien automaattinen tunnistus (Named entity recognition,
NER) on alataso edellä mainittuun tietojen poimintaan, IE. NER tunnistaa nimettyjä
kokonaisuuksia, entiteettejä, ilmaisevat lausekkeet kuten sairaudet, lääkkeet tai po-
tilaiden nimet vapaan tekstin kirjauksissa. Kliinisesti merkittävien entiteettien tun-
nistaminen tekstistä on tärkeä osa-alue kaikille NLP tekniikoille. Jotkut entiteetit
pystytään tunnistamaan yksistään rakenteisista kuvioista, esim. henkilötunnus, mutta
suurin osa vaatii sääntöjä. Säännöt auttavat tunnistamaan potilastiedoista uutta tietoa
ja tiedon automaattinen poiminta mahdollistaa tutkimusten mukaan potilaiden hoi-
tokirjauksien parantamisen. Eniten tutkimuksia kliinisten entiteettien tunnistami-
sesta on englannin kielellä ja vain vähän muilla kielillä esim. ruotsi ja suomi. Sään-
töpohjaiset NER-järjestelmät voivat olla erittäin tehokkaita, mutta vaativat toimiak-
seen käsityötä. Koneoppimisen lähestymistavat voivat onnistuneesti poimia entiteet-
tejä, mutta vaativat toimiakseen suuria nimikoituja kokonaisuuksia. Koneoppimisen
edut ovat kuitenkin siinä, että ne eivät vaadi ihmisen intuitiota ja ne voidaan koulut-
taa uudelleen. (Meystre ym 2008, Skeppstedt ym. 2014.)

Annotointi eli nimikointi on puhe- tai muun aineiston kuvaamista, luokittelua ja
jäsentelyä systemaattisella tavalla. Lennes ja Ahjoniemi (2005) ovat määritelleet an-
notaation seuraavasti: ”Nimikoinnilla (engl. labeling) eli annotoinnilla (annotation >
to annotate) tarkoitetaan sitä, että digitaalisesti tallennettuun puhenäytteeseen liite-
tään systemaattisesti nimikkeitä (labels) eli annotaatioita (annotations) eli symbolisia
kuvauksia näytteen sisällöstä.” Annotointi voidaan tehdä sana-, sanapari- sekä lau-
setasolla ja se on osa koneoppimista vapaan tekstin kirjauksista. Koneoppiminen
(machine learning) nimensä mukaisesti opettaa tietokoneille löytämään tiedoista
sille annettuja tehtäviä. Tässä tietokonetekniikan osa-alueessa käytetään tilastotie-
teen menetelmiä ja sillä luokitellaan tai ennustetaan ilmiöitä. Koneen opetuksessa
käytetään sille kulloinkin määriteltyä laajaa opetusaineistoa. (Ailisto ym. 2018.)

Työn psykososiaalisia riskitekijöitä havaitsevasta työkalusta, olisi hyötyä ter-
veystarkastuksissa, jotta työkyvyn alenemisen uhka huomattaisiin ajoissa. Tekstin

louhintatekniikat ja NLP- toimintamallit voivat tulevaisuudessa mahdollistaa tällaisten työkykyyn vaikuttavien riskitekijöiden automatisoidun varhaisen tunnistamisen. Semanttista analyysiä tukevien NLP-avaintehtävien eteneminen on lisääntynyt. Monimutkaisillakin semanttisilla tietomalleilla varustettujen työkalujen on ollut mahdollista luoda tarkoin määrätty ja koottu joukko kirjoitetun kielen tekstejä (korpus), mikä on tukenut suuresti uusien työkalujen ja lähestymistapojen kehittämistä. Muiden kuin englanninkielisten tutkimusten määrä myös kasvaa jatkuvasti. NLP-menetelmiä on käytetty menestyksekkäästi todellisissa kliinisissä tehtävissä, kuitenkin tutkimuksen ja käytännön työkalujen kehittämisen välillä on vielä aukko. Automatisaatiota kyllä tarvitaan, sillä käsin annotointi on kallista ja aikaa vievää. (Meystre ym. 2008, Skeppstedt ym. 2014.)

2.6 Yhteenveto kirjallisuuskatsauksesta

Kirjallisuuskatsauksen yhteenveto on esitetty kuviossa 2.

	<p>Työterveyshuollon terveystarkastuksen sisältöalueet</p> <p>Työterveyshuollon terveystarkastusten hoitotieteellinen tutkimus on vähäistä. Myös tieteellisesti arvioitujen terveystarkastusten ja terveys suunnitelmien jatkotoimenpiteitä sekä vaikuttavuutta koskevien tutkimusten määrä on edelleen pieni.</p>
	<p>Työn psykososiaaliset riskitekijät</p> <p>Työntekijän kokemien psykososiaalisten haittojen kartoittamiseen ja tukemiseen olevien tutkimusten määrä on vähäinen.</p>
	<p>Terveystarkastusten sähköinen kirjaaminen</p> <p>Terveystarkastusten kirjaamisesta tehty tutkimus on vähäistä sekä kansallisesti että kansainvälisesti. Lisää tutkimusta tarvitaan terveystarkastusten vaikuttavuudesta työkyvyn tukemisessa.</p>
	<p>Terveystarkastusten kirjaamisen jatkoohjodyntäminen koneoppimisen avulla</p> <p>Suomenkielistä tutkimusta vapaan tekstin koneoppimisesta työterveyden terveystarkastusten kirjauksissa ei ole tehty.</p>

Kuvio 2. Kirjallisuuskatsauksen yhteenveto.

3 Tavoitteet

Tämän tutkimuksen tavoitteena oli kehittää ja arvioida työterveyshoitajien työterveyshuollon terveystarkastusten kirjaamista, erityisesti työn psykososiaalisten riskitekijöiden osalta. Tutkimuksen tavoitteena oli myös kartoittaa koneoppimisen mahdollisuuksia tunnistaa psykososiaalisia riskitekijöitä työterveyshuollon terveystarkastuksien kirjauksissa. Tutkimus toteutettiin kaksivaiheisena.

Ensimmäisessä vaiheessa tavoitteena oli vahvistaa teoreettista tietoa työterveyshuollon terveystarkastusten sisällöstä ja kirjauksesta (osatutkimus 1). Toisessa vaiheessa tavoitteena oli kartoittaa työn psykososiaalisten riskitekijöiden kirjaamista työterveyshuollon terveystarkastuksissa (osatutkimus 2–4). Tutkimuksen yhteenvedo on esitetty kuviossa 3.

Ensimmäisen osatutkimuksen tutkimuskysymykset olivat:

- 1) Mitkä ovat työterveyshuollon terveystarkastuksen sisältöalueet?
- 2) Miten terveystarkastuksen sisältöalueet tulisi työterveyshoitajien mielestä kirjata?

Toisen osatutkimuksen tutkimuskysymykset olivat:

- 1) Miten työterveyshoitajat kirjaavat työn psykososiaalisia tekijöitä työterveyshuollon terveystarkastuksissa?
- 2) Miten työterveyshoitajat käyttävät terveystarkastuksissa Työkykyindeksiä (TKI), joka mittaa työntekijän kykyä hallita työn vaatimuksia?

Kolmannen osatutkimuksen tutkimuskysymykset olivat:





- 1) Voidaanko tekstinlouhinnan avulla automaattisesti havaita työn psykososiaaliset riskitekijät terveystarkastuksen vapaassa kirjauksessa, joista on tehty lääkärintuomio varhaiseläkettä, pitkää sairauslomaa tai kuntoutusta varten?
- 2) Ennustavatko kirjatut työn psykososiaaliset riskitekijät varhaiseläkettä tai työkyvyn alenemista?

Neljännän osatutkimuksen tutkimuskysymys oli:

- 1) vaikuttaako BERT-NER tai Qexp, jompikumpi tai molemmat lähestymistavat yhdessä lupaavilta käytettäväksi järjestelmässä, joka tunnistaa automaattisesti psykososiaalisia riskitekijöitä työterveystarkastusten vapaasta kirjauksesta?

4 Tutkimusasetelma, aineiston keruu ja analysointi

Osatutkimukset esitetään kootusti kuviossa 3.

	<p>Osatutkimus 1.</p> <p>Tavoite: Kuvata työterveyshuollon terveystarkastusten sisältöalueita ja rakennetta sekä kirjausta työterveyshoitajien näkökulmasta</p> <p>Aineisto: Työterveyshoitajat (17)</p> <p>Menetelmä: Delphi-paneeli (kaksi kierrosta; haastattelu ja kyselylomake).</p> <p>Analyysi: Laadullinen sisällön analyysi, tilastollinen analyysi</p>
	<p>Osatutkimus 2.</p> <p>Tavoite: Analysoida, miten terveystarkastuksia suorittavat työterveyshoitajat kirjaavat työn psykososiaalisia tekijöitä ja käyttävät työkykyindeksiä (TKI)</p> <p>Aineisto: 196 henkilön terveystarkastusasiakirjat</p> <p>Menetelmä: 13 psykososiaalisen tekijän (PAS1010) annotointi (kaksi kierrosta), kahden peräkkäisen terveystarkastusasiakirjan ja TKI-pisteiden analyysi</p> <p>Analyysi: Tilastollinen analyysi</p>
	<p>Osatutkimus 3.</p> <p>Tavoite: Arvioida tekstinlouhinnan mahdollisuuksia psykososiaalisten riskien varhaiseen havaitsemiseen ja tekijöitä terveystarkastusten kirjauksista, joka voi viitata varhaiseläkkeelle, pitkittyneeseen sairauslomaan tai kuntoutukseen suositteleviin lääketieteellisiin lausuntoihin.</p> <p>Menetelmä: 7078 henkilön terveystarkastusasiakirjan tekstin louhinta puoliohjatulla mallilla</p> <p>Analyysi: Tilastollinen analyysi</p>
	<p>Osatutkimus 4.</p> <p>Tavoite: Arvioida voidaanko psykososiaalisten riskitekijöiden tunnistamisen automatisoida työterveyshuollon kirjauksissa luonnollisen kielen käsittelyllä (NLP).</p> <p>Menetelmä: 7078 henkilön terveystarkastusasiakirjaa verrattiin ohjaamaton ja ohjattu nimetyn koneoppimisen entiteettitunnistuksen (NER) havaitsemia psykososiaalisia riskitekijöitä, 100 asiakirjan tarkistus</p> <p>Analyysi: Tilastollinen analyysi</p>

Kuvio 3. Osatutkimusten kuvaus.

4.1 Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältöalueista ja kirjauksesta

Ensimmäisen osatutkimuksen tavoitteena oli kartoittaa työterveyshuollon terveystarkastusten sisältöalueet ja kirjausten rakennetta. Tämä kuvaileva laadullinen tutkimus pyrki selvittämään työterveyshoitajien mielipiteitä aiemmin vähemmän tutkitulta alueelta. Tutkimusmenetelmänä käytettiin Delphi-menetelmää, joka soveltuu tutkimuksiin, joiden tehtävänä on mitata asiantuntijaryhmän arvioita tehtävistä päätöksistä, arvioida tärkeysjärjestystä, tehdä ennustuksia tai ratkaista ongelmia. Menetelmän tarkoituksena on saavuttaa luotettava konsensus, yhtenäinen ymmärrys, asiantuntijaryhmän mielipiteestä. Delphi-menetelmän tunnuspiirteitä ovat anonymiteetti, toistaminen kunnes konsensus on saavutettu ja palautteen anto haastateltaville, sekä yhteisen ymmärryksen saavuttaminen. Tutkijan on Delphi-metodissa tärkeää tietää, milloin tutkimus tulee lopettaa. (Burns ja Grove 2020, Hasson ym. 2000, Hyttinen ja Aavarinne 2000, Keeney ym. 2004, Kennedy 2004, Polit ja Beck 2017.) Tässä tutkimuksessa toteutettiin kaksi kierrosta, ensimmäinen toteutettiin haastattelulla ja toinen tehtiin kyselylomakkeella. Delphi-menetelmässä paneeliin suositellaan tyypillisesti kahdelle tai kolmelle kierrokselle 15–40 asiantuntijaa. (Hyttinen ja Aavarinne 2000, Kennedy 2004.)

Tähän tutkimukseen kutsuttiin kaksikymmentä työterveyshoitajaa kahdelta eri yksityiseltä suomalaiselta työterveyshuollon palveluntuottajalta, heistä kolme kieltäytyi. Asiantuntijapaneeli koostui tämän jälkeen 17 työterveyshoitajasta, joilla oli työkokemusta 1–35 vuotta, keskimäärin 15,4 vuotta. Työkokemuksen minimivaatimuksena tutkimukseen osallistumisessa oli vähintään vuoden työterveyshuollon työkokemus. Kaikki osallistuvat työterveyshoitajat olivat tehneet työssään terveystarkastuksia.

Ensimmäisellä kierroksella haastattelussa asiantuntijapaneelin osallistujia pyydettiin tunnistamaan ja kommentoimaan työterveyshuollon terveystarkastuksen sisältöä ja tarkastuksessa tehtäviä toimenpiteitä. Työterveyshoitajia pyydettiin haastattelussa myös ottamaan kantaa terveystarkastuksen yleisimpiin sisältöalueisiin sekä siihen, miten ne tulisi kirjata. Kysymyslomake koostui seitsemästä kysymyskonnaisuudesta, joissa oli useita alakohtia. Kaksi kysymyksistä oli taulukkomuotoon valmiiksi asetettuja valintatehtäviä, jotka ensin haastateltiin ja tämän jälkeen haastattelun osallistujat vielä rastittivat mielipiteensä suoraan haastattelurunkoon. Aiheet perustuvat hyvän työterveyshuollon periaatteisiin ja terveystarkastusten kirjaamisen sisältöalueisiin näiden periaatteiden pohjalta, kirjaamisvaihtoehtoihin ja minimitiedoston sisältöön. Terveystarkastuksen sisältöalueet olivat mukailtu Hyvän Työterveyshuoltokäytännön sekä Työterveyslaitoksen suositusten perusteella (Karhula 2006, Manninen ym. 2007). Haastattelut kestivät keskimäärin 60 minuuttia ja ne tehtiin jokaisen omassa työhuoneessa ja nauhoitettiin.

Ensimmäisen kierroksen puolistrukturoidun haastattelun jälkeen tutkija kuunteli nauhoitukset ja kirjoitti haastattelut sanatarkasti. Aineistoa kertyi kaikkiaan 154 sivua kirjoitettuna tekstinä. Analysointi aloitettiin lukemalla aukikirjoitettu aineisto useampaan kertaan, jotta kokonaiskuva aineistosta hahmottuisi tutkijalle. Tämän jälkeen termit ja sanayhdistelmät ryhmiteltiin seitsemän kysymyskokonaisuuden ja tutkimusongelmien mukaan eri värejä käyttäen. Värikoodatut termit ja sanayhdistelmät luokiteltiin ja niiden frekvenssi aineistossa laskettiin. Seuraavassa vaiheessa analyysissä yhdistettiin vastauksia, joiden sisältöalueet liittyivät toisiinsa. Erittelyn aikana tutkija luki alkuperäisaineiston useaan kertaan, jotta muodostetut luokat kuvaisivat mahdollisimman tarkasti ja kattavasti alkuperäisilmauksia. Tästä analyysistä muodostettiin kyselylomake, joka perustui työterveyshoitajien ensimmäisen kierroksen vastauksiin.

Toisella Delphi-kierroksella käytettiin kyselylomaketta, joka oli muodostettu edellisen kierroksen haastattelun perusteella. Työterveyshoitajien mielipiteet pyrittiin pitämään mahdollisimman muuttumattomina ja kaikki ensimmäisen kyselykierroksen kysymykset olivat mukana toisella kierroksella. Neuvonnasta ja ohjauksesta tehtiin uusi kysymys toiseen kierrokseen, vaikka työterveyshoitajat eivät olleet maininneet sitä ensimmäisellä kierroksella. Neuvonta ja ohjaus on oleellinen osa terveystarkastusta (Karhula 2006, Manninen ym. 2007). Kyselylomake lähetettiin kaikille ensimmäisen kierroksen asiantuntijapaneeliin osallistuneille työterveyshoitajille ($n = 17$). Heille lähetetyssä saatteessa kerrottiin, että heillä oli myös mahdollisuus vaihtaa mielipiteensä ensimmäiseltä kierrokselta. Kyselylomake oli rakenteinen ja sisälsi Osgoodin asteikon, jossa 0 tarkoitti täysin eriävää mielipidettä ja 10 tarkoitti täysin samaa mieltä olemista (Burns ja Grove 2020). Kysymyksiä oli yhteensä 103 ja vastaamiseen meni aikaa keskimäärin 20 minuuttia. Kyselylomakkeita palautettiin 14 kappaletta ($n = 14/17$). Vastausprosentti kummallakin Delphi-kierroksella oli 85 %.

Toisen kierroksen aineisto analysoitiin SPSS 12.0 ohjelmalla ja avoimet kysymykset laadullisella sisällön analyysillä. Avoimet kysymykset luokiteltiin ja värikoodattiin. Värikoodatut termit ja sanayhdistelmät pyrittiin luokittelemaan ylempiin luokkiin ja samalla laskettiin niiden frekvenssi aineistossa. Seuraavassa vaiheessa analyysissä yhdistettiin vastauksia, joiden sisältöalueet liittyivät toisiinsa. Tuloksista laskettiin myös normaalijakauma, keskihajonta (standard deviation, SD) ja nominaaliasteikkojen vastausten (nominal scaled questions) osalta prosentit frekvensseistä. Nominaaliasteikolla tulokset pystyttiin luokittelemaan ja laskemaan niiden esiintyvyyttä. Keskiarvo kuvaa ryhmän mielipidettä ja normaalijakauma sekä keskihajonta (standard deviation, SD) kuvaavat ryhmän yksimielisyyttä. Tässä tutkimuksessa yksimielisyyden (konsensuksen) raja-arvoksi oli määritelty 70 % (Hasson ym. 2000). Yksimielisyys arvioitiin laskemalla keskiarvo Osgood-asteikon vastauksista.

4.2 Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla

Toinen osatutkimus oli taannehtiva (retrospektiivinen) rekisteritutkimus ja se toteutettiin yhden suomalaisen valtakunnallisen työterveyshuollon palveluntuottajan rekisteristä. Osatutkimuksen tiedosto koostui 7956 yksilön terveystarkastuksien kirjauksien tietokokonaisuudesta vuosina 2002–2007. Jokaisella tutkimukseen valitulla yksilöllä tuli olla ainakin kaksi terveystarkastusta kirjattuna terveystietoihin.

Suomessa on kehitetty Työkyvyindeksi (TKI), jonka avulla arvioidaan ihmisen työkykyä hänen oman näkemyksensä mukaan. Vastaaaja antaa vastauksia työkykyä mittaaviin kysymyksiin, joille on annettu numeerinen arvo. Nämä arvot luokitellaan neljään alaluokkaan (huono, kohtalainen, hyvä ja erinomainen, Taulukko 4). (Tuomi ym. 1997.) TKI- kyselyjen vastauksia oli kirjattuna koko aineistossa yhteensä 6890 kappaletta. Nämä jakaantuivat arvoiltaan seuraavasti: huono 322 kpl, kohtalainen 1477 kpl, hyvä 3150 kpl ja erinomainen 1941 kpl. Aineistossa oli kuitenkin vain 2214 kpl TKI-kyselyjen vastauksia kirjattuna kaksi kertaa samalle terveystarkastetavalle. Lähtökohtana otokselle oli, että jokaiseen luokkaan (huono, kohtalainen, hyvä ja erinomainen) tulisi yhtä monta asiakirjaa. Lisäksi samalla henkilöllä tuli olla kaksi TKI:n arviointia kirjattuna, jotta voitiin katsoa muutosta yksittäisen henkilön TKI:ssä. Nämä samalle henkilölle kaksi TKI-kyselyä jakaantuivat arvoiltaan seuraavasti: huono 72 kpl, kohtalainen 423 kpl, hyviä 1060 kpl ja erinomainen 659 kpl. Tämän seurauksena valittiin terveystarkastuksien kirjauksista tähän tutkimukseen satunnaisesti 200 henkilöä, eli jokaiseen TKI-kyselyn neljään luokkaan 50 henkilön terveystarkastuskirjaukset. Mukaan otettavien terveystarkastuskirjausten kriteereinä oli lisäksi, että kirjauksen kohteena olevat henkilöt olivat vähintään 18-vuotiaita ja heidän terveystarkastuksen tietoihinsa oli kirjattuna TKI-pisteet sekä ainakin kaksi terveystarkastusta. Tässä tutkimuksessa käytimme kolmea luokkaa: huono, kohtalainen ja hyvä, johon yhdistimme erinomaiset.

Taulukko 4. Työkyvyindeksi (TKI) arvojen luokitus (Tuomi ym. 1997)

Arvo	Luokka	Selitys
7–27	Huono	työkyvyn palauttamisen tarve
28–36	Kohtalainen	työkyvyn edistämisen tarve
37–43	Hyvä	työkyvyn vahvistuksen tarve
44–49	Erinomainen	työkyvyn vahvistuksen tarve

Annotoitavat työn psykososiaaliset tekijät määriteltiin Euroopan Prima EF konsortion luoman luokituksen perusteella (PAS 1010-standardi) (Työterveyslaitos 2008), joka on yleisesti saatavilla oleva standardi psykososiaalisen riskien hallintaan ja

määrittelee 13 psykososiaalisista tekijää, joita olisi työpaikoilla seurattava säännöllisesti (Taulukko 1, artikkeli II). Käsitönä tapahtuneessa annotoinnissa jokaiseen psykososiaaliseen tekijään liittyvään sanaan, lauseeseen tai tekstikappaleeseen tehtiin merkintä. Jokainen annotoitu merkintä kirjattiin myös erikseen joko riskinä tai ei-riskinä mainittuun ryhmään.

Työterveyshoitajien kirjaamista 200 henkilön terveystarkastuksista annotoitiin käsityönä työn psykososiaalisia tekijöitä. Annotointia suoritti kolme työterveyshoitajaa siten, että yksi työterveyshoitajista annotoi kaikki 200 (annotaattori 1) ja kaksi työterveyshoitajaa (annotaattorit 2 ja 3) annotoivat hänen kanssaan ristiin 100 terveystarkastuksen kirjausta. Työterveyshoitajat olivat tehneet kirjallisen konsensuksen niistä tekijöistä, jotka annotoidaan sekä miten nämä tekijät luokitellaan riskitekijöiden tai yhdistettyyn neutraaleiden ja positiivisten tekijöiden luokkaan. Annotaattoriparin ja -ryhmän yksimielisyyttä arvioitiin Cohenin kappa-arvolla (Cohen 1960). Tämä laskettiin ottamalla huomioon tekstimerkit (eli sanat) kohteina, joista kukin kuului täsmälleen yhteen luokkaan. Ensimmäisellä annotointikierroksella annotaattorien 1 ja 2 yksimielisyys kappa-arvo oli 0,41 ja annotaattorien 1 ja 3 yksimielisyysarvo oli 0,49, nämä arvot ovat vain kohtalaisia.

Annotointiryhmän jäsenet keskustelivat virheistä ja erimielisyyksistä sekä tuottivat yhdessä uuden päivitetyn annotointiohjeen. Virheiksi ja poistettaviksi sovittiin esimerkiksi:

- Työuran pituus, työvuodet yrityksessä ja työelämässä sekä työhistoria, sillä nämä ovat historiaa sekä taustatietoa, eivät psykososiaalisia riskitekijöitä.
- Ammattikuvaukset ja työtehtävät nimikkeinä (esimerkiksi hitsaaja, kassa, johdon assistentti) sekä niiden kuvaukset poistettiin. Nämä ovat taustatietoa, eivätkä työtehtävien sisällöt ole yhteneväiset.
- Työssä käytetyt koneet ja koneiden nimet luetteluita otettiin annotoinnista pois. Koneet sinällään eivät ole psykososiaalinen riskitekijä, vaan siihen vaikuttavat suojaimet, käyttö, ergonomia ja työolosuhteet.
- Työajan jaksotukset ja päivät otettiin pois, samoin työvuorojen jaksotukset ja selitykset esim. ”7 yötä peräkkäin”, ”kaksi iltaa ja sitten viisi vapaata”, tai ”j.t. vkonloppu vapaa”. Psykososiaalisiksi riskitekijöiksi luokiteltiin ainoastaan vuorotyö (2-vuoro tai 3-vuoro), yötyö, ylityö.
- Erilliset lääkärin toteamat oireet ja sairaudet eivät ole työn psykososiaalisia riskitekijöitä.

Uuden ohjeen pohjalta toteutettiin toinen annotointikierros, jonka toteuttivat jo aiemmin mukana ollut annotaattori 1 sekä uusi, ei aiemmin mukana ollut, annotaattori 4. Toisen kierroksen jälkeen yksimielisyyden Cohenin kappa-arvo oli 0,72, jota voidaan pitää riittävänä (Cohen 1960).

Lopullisessa aineiston tarkastelussa todettiin neljällä henkilöllä olevan terveystarkastuksissa liian pieni arviointiväli (6 kk tai vähemmän) ja heidän terveystarkastustensa kirjaukset jätettiin tämän osatutkimuksen ulkopuolelle. Osatutkimuksen 2 otoksen lopullinen määrä oli siten 196 henkilön kirjaukset; näistä 136 oli miehiä ja 60 naisia. Keski-ikä ensimmäisen terveystarkastuksen aikana oli 49,7 vuotta ja toisen terveystarkastuksen aikana 52,4 vuotta (vaihteluväli 21–63 vuotta). Kahden terveystarkastuksen aikaväli oli keskimäärin 2,68 vuotta (1–10 vuotta).

Tutkimuksessa kartoitettiin, olivatko työterveyshoitajien kirjaamat psykososiaaliset tekijät yhteydessä Työkykyindeksiin (TKI). TKI:ä on käytetty Suomessa lähes 30 vuoden ajan osana työterveystarkastuksia. TKI-lomake voidaan määritellä työkaluna, jolla mitataan työntekijän kykyä hallita työterveyttä ja työn henkisiä vaatimuksia (Ilmarinen 2001). Yleensä työntekijä täyttää TKI-lomakkeen ennen terveystarkastusta ja arvioi omat psykososiaaliset riskitekijänsä sekä omaa selviytymistään työssä. Työkalussa tekijät, jotka liittyvät heikkoon työkykyyn, ovat vapaa-ajan puute, heikko fyysinen aktiivisuus, tuki- ja liikuntaelimestön ongelmat, korkea ikä, lihavuus, korkeat henkisen työn vaatimukset, autonomian puute, huono fyysinen työympäristö ja suuri fyysinen kuormitus (Tuomi ym. 1997, taulukko 3). Tutkimukset ovat osoittaneet, että matala TKI- arvo ennakoii työkyvyn menetystä ja sairauspoissaoloa. TKI antaa tietoa yksittäisen työntekijöiden psykososiaalisista riskitekijöistä. (van den Berg ym. 2009.) Tässä tutkimuksessa TKI luokat hyvä ja erinomainen yhdistettiin, ja tutkimuksessa käytettiin kolmea TKI- luokkaa.

Jokaisella psykososiaalisen tekijän 13 osa-alueella muodostettiin työterveyshoitajien kirjaamisen mukaan kyllä/ei- muuttuja. Tämän jälkeen kullekin työntekijälle tehtiin yhteenveto riski ja ei- riski- merkintöjen muuttujista. Tulokset esitettiin keskiarvoina ja keskihajontoina (SD) sekä frekvensseinä että prosentteina. Miehiä ja naisia verrattiin Studentin t-testillä. Ensimmäisessä ja toisessa terveystarkastuksessa luokitellun TKI:n yhteys psykososiaalisiin tekijöihin arvioitiin Kappa-kertoimella ja muutoksen symmetria Bowkerin symmetriatestillä. TKI:n muutos ja merkintöjen määrä ensimmäisestä terveystarkastuksesta toiseen tarkistettiin parittaisella t-testillä. TKI: n ja riskitekijöiden sekä ei-riskijöiden välinen yhteys analysoitiin kaksisuuntaisella Studentin t-testillä. Spearmanin luokitusjärjestystä käytettiin kuvaamaan TKI:n ja psykososiaalisten riski- sekä ei-riskitekijöiden määrien välistä yhteyttä. Analyysissä käytettiin SAS for Windows 9.3 -ohjelmistoa. Tilastollisen merkitsevyyden rajana käytettiin alle 0,05 p-arvoa.

4.3 Työn psykososiaalisten riskitekijöiden mallinnus ja tunnistus tekstin louhinnalla terveystarkastusten kirjauksista

Kolmas taannehtiva (retrospektiivinen) osatutkimus perustui samaan potilasasiakirjarekisteriin kuin osatutkimus 2. Tutkimusaineistoon ei otettu mukaan potilasasiakirjoja, joiden terveystarkastuskirjauksissa ei ollut lainkaan työterveyshoitajan tekstiä. Aineiston ulkopuolelle jätettiin myös alle 18-vuotiaista asiakkaista tehdyt kirjaukset. Tutkimukseen otettiin mukaan kirjaukset, joissa samalle henkilölle oli kirjattu vähintään kaksi terveystarkastuskirjausta (pois n=878). Tämän jälkeen aineisto koostui 7078 työntekijän terveystarkastusasiakirjoista, joita oli yhteensä 15 443. Tutkimus toteutettiin valtakunnallisen työterveyshuollon palvelutuottajan asiakirjoista Suomessa vuosina 2002–2007.

Aineistossa 46 % terveystarkastuksista oli lakisääteisiä määräaikaistarkastuksia. Vapaaehtoisia, ikään perustuvia ikäryhmätarkastuksia, oli 38 % terveystarkastuksista. Terveystarkastuksista oli ”muu terveystarkastus” nimikkeellä 11 %, ”kuntoutustarvearvio” nimikkeellä oli 2 % ja 1,6 % ”vajaakuntoisen terveystarkastuksen” nimikkeellä. Terveystarkastuksen syytä ei ollut mainittu 0,03 % terveystarkastuksissa. Terveystarkastuksien kirjauksissa naisia oli 2409 henkilöä ja miehiä 4679 henkilöä.

Tutkittavat olivat syntyneet vuosien 1938–1987 välillä. Tutkimuksen ikäjakama aineistossa on esitetty taulukossa 5.

Taulukko 5. Kolmannen osatutkimuksen terveystarkastettavien syntymävuodet.

Terveystarkastettujen syntymävuosi	Terveystarkastettujen henkilöiden lukumäärä
1930-luvulla	2
1940-luvulla	1480
1950-luvulla	2882
1960-luvulla	1839
1970-luvulla	713
1980-luvulla	170

Sanat ja lauseet annotoitiin käsityönä listaksi (osatutkimus 1) kullekin psykososiaaliselle riskitekijälle Prima EF standardin mukaan (Työterveyslaitos 2008). Sen jälkeen jäljelle jääneistä kirjauksista etsittiin lisää riskitekijöitä puoliohjatulla (semi-supervised) mallilla. Malli on yhdistelmä ohjaamatonta koneoppimista, tiedoston hakukonetta sekä manuaalista suodatusta. Tässä puoliohjatussa mallissa käytettiin hakukoneena Apache SOLR-hakualusta (lucene.apache.org/solr) koneoppimisen asiantuntijan johdolla. Tutkimuksessa käytettiin word2vec työkalua kehittämään se-

manttisen sana-avaruuden mallin terveystarkastusten kirjauksista. Word2vec edustaa kutakin erillistä sanaa tietyllä numeroluettelolla, a-vektorilla (Mikolov ym. 2013). Tässä tutkimuksessa malli koulutettiin käyttämällä Word2vec-työkalupakettia koko tietojoukossa käyttäen skip-gram -mallia, hierarkkista softmaxia sekä 300-ulotteisuutta. Kutakin alkuperäisten riskitekijöiden luettelon sanaa varten mallia käytettiin tunnistamaan ja ehdottamaan muita sanoja, joilla on samanlainen merkitys. Tämä on verrattavissa aiemmin käytettyyn interaktiiviseen nopeaan sanavaratutkimukseen (Topaz ym, 2019). Monisanaisten lauseiden kohdalla havaitsimme, että useimmissa tapauksissa sanat olivat negatiiveja tai prepositioita (esim. "Ei tietoja"). Tutkimuksessa käytettiin yksinkertaista heuristiikkaa, joka olettaa, että samanlaiset merkitykset voidaan tunnistaa korvaamalla yksinkertaisesti viimeinen sana samanarvoisilla sanoilla. Jokainen eroteltu yksittäisen sanan maininta riskitekijästä käytettiin yksittäisinä kyselyinä syöttämään semanttiselle mallille semanttisesti samankaltaisia sanoja (käytimme kosinin samankaltaisuuden kynnystä 0,6). Apache SOLR-hakualustaa käytettiin tarkistamaan, että poimittuja ehdokkaita on tietojoukossa. Lauseita, joita ei näin löydetty, hylättiin.

Viimeiseksi arvioitiin käsityönä luodut sanojen ja lauseiden synonyymit ja suodatettiin pois ne, jotka eivät selvästi vastanneet PAS 1010 liittyvää psykososiaalista kategoriaa riskitekijöistä. Tällä tavoin oli mahdollista tunnistaa uusia sanoja ja lauseita kuvailemaan eri riskitekijöitä. Koko terveystarkastustiedostosta tunnistettua laajennettua riskitekijäluetteloa käytettiin luomaan riskitekijöistä taulukko, joka näyttää montako kertaa kukin riskitekijä on mainittu terveystarkastuskirjauksessa.

Tähän tutkimukseen otettiin mukaan yksilöity tieto kaikista terveystarkastuksista ennen aineistossa olevaa ensimmäistä B-lausuntoa. Kaikki terveystarkastukset B-lausunnon tekemisen jälkeen jätettiin huomioimatta. Aineisto sisälsi täten havaintoja työntekijöistä, joille oli kirjoitettu B-lausunto eläköitymisestä, sairauspoissaolosta tai kuntoutuksesta. Aineistossa huomioitiin myös ikä, sukupuoli ja terveystarkastusten määrä sekä psykososiaaliset riskitekijät ja näiden riippuvuutta tutkittiin khiin neliotestillä. Tilastolliset analyysit toteutettiin SAS-ohjelman versiolla 9.4 Windowsille (SAS Institute Inc., Cary, NC, USA). Riippumattomien ja riippuvien muuttujien yhteyksien merkitys määritettiin logistisella regressiolla. Muuttujat ilmoitettiin mediaaneina ja lukumäärinä. Tilastollinen merkitsevyystaso asetettiin $p = 0,05$. Kaikki testit olivat kaksisuuntaisia.

4.4 Työn psykososiaalisten riskitekijöiden tunnistaminen vapaan tekstin kirjauksesta koneoppimisen avulla

Neljännessä osatutkimuksessa käytettiin samaa aineistoa kuin osatutkimuksissa 2 ja 3. Tiedot koostuvat siis 15 443 terveystarkastusasiakirjasta 7078 työntekijältä. Terveystarkastusasiakirjat on laadittu yhdessä kansallisessa, yksityisessä työterveyshuollon palveluntuottajalla vuosina 2002–2007. Tietojoukko suodatettiin valitsemalla vähintään 18-vuotta täyttäneiden terveystarkastuskirjaukset.

Osana aikaisempaa tutkimusta (osatutkimus 2) kerättiin satunnaisotoksella valituista 196 terveystarkastuskirjauksista luetteloa kaikista yksilöllisistä sanoista ja lausekkeista, jotka liittyvät PrimaEF:n PAS1010 luokituksen mukaisiin psykososiaalisiin riskitekijäluokkiin. PAS1010 koostui seuraavista 13 luokista: *työn sisältö, työmäärä ja työmäärä, työaika-työmuoto, työn hallinta, työympäristö ja välineet, organisaatiokulttuuri ja toiminta, sosiaaliset suhteet, rooli organisaatiossa, työuran kehitys, työn ja muun elämän yhteensovittaminen, työpaikka väkivalta, häirintä ja kiusaaminen* (Leka, Cox 2008). Tässä tutkimuksessa otimme huomioon vain negatiiviset maininnat, joihin viitataan psykososiaalisina riskitekijöinä. Aineisto esikäsiteltiin suomen kielen NLTK-sanakirjaimella (Bird ym. 2009).

Tutkimuksessa käytimme kahta koneoppimisen Nimettyjen entiteettien tunnistus eli NER-lähestymistapaa, ohjattua ja ohjaamatonta, psykososiaalisten riskitekijöiden havaitsemiseksi työterveyshuollossa kirjattujen terveystarkastusten asiakirjoissa. Ensimmäinen lähestymistapa perustui perinteisempään ohjattuun NER-lähestymistapaan, joka on koulutettu jäljittelemään käsityönä annotoituja tietoja. Kutsumme sitä BERT-NER lähestymistavaksi. Nimettyjen entiteettien tunnistus (NER) on perustehtävä tiedon poiminnassa, ja sillä on kyky havaita maininnat alueeseen liittyvistä kokonaisuuksista (Hakala ja Pyysalo 2019). BERT on lyhenne sanoista Bidirectional Encoder Representations from Transformers, joka on alun perin Googlen kehittämä luonnollisen kielen käsittelyn menetelmä. Se on suunniteltu esiopettamaan syvät kaksisuuntaiset maininnat nimeämättömästä tekstistä. Tämä tarkoittaa, että se tarkastelee kaikilla käsittelytasoiltaan mainintaa käsiteltävän tekstikohdan molemmilta puolilta ja tuottaa näin parempia ennakoitintuloksia. Mallin kaksisuuntaisuus perustuu kykyyn tunnistaa sanojen asiayhteyttä koskevia suhteita. BERT koulutetaan kätkemällä lauseesta sanoja ja antamalla sen arvata puuttuva osa (masked language modeling) (Devlin ym. 2019).

Toinen ohjaamaton lähestymistapamme käytti valvottua NER-menetelmää, jota on täydennetty valvomattomalla parafrasointimenetelmällä. Tässä tutkimuksessa suomensimme sen muotoilumenetelmäksi. Sana parafrasi on peräisin latinalaisesta termistä parafrasis, joka on vapaasti käännettynä "lisäilmaisu, muunnelma". Muotoilumenetelmää käytetään etsimään lyhemmin sanoja ja lauseita, jotka valvottu NER-menetelmä on mahdollisesti ohittanut. Nämä sanat ja lauseet poikkeavat kielellisesti

vain vähän alkuperäisistä, käsityönä tehdyistä annotoinneista, sanoilla ja lauseilla tulee kuitenkin olla samanlainen merkitys. Kutsumme sitä tässä tutkimuksessa NER-Qexp- lähestymistavaksi. Tämän tutkimuksen tavoitteena oli selvittää, kumpi lähestymistavoista tai mahdollisesti molemmat osoittautuisivat toimivaksi järjestelmässä, joka tunnistaisi automaattisesti psykososiaaliset riskitekijät työterveystarkastusten vapaan tekstin kirjauksista.

Ohjattu NER – BERT-NER

Ensimmäinen tutkimuksessa käytetty lähestymistapa, NER-menetelmä, perustuu ennalta koulutettuun BERT-malliin sekä ehdollisiin satunnaiskentän löydöksiin (Hakala ja Pyysalo, 2019, Devlin ym., 2019, Lafferty, 2001). Tätä menetelmää kutsutaan nimellä BERT-NER. Menetelmä ”esikoulutettiin” suomenkielisessä korpuksessa (Virtanen, 2019) ja hienosäädettiin koneoppimisen asiantuntijan toimesta sen jälkeen aiemmin kuvatuille - 196 terveystarkastusasiakirjalle (artikkeli II).

Menetelmä merkitsi automaattisesti aiemmin annotoimattomien tekstitietojen jokaisen sanan tai lauseen O-tarralla (ei riskitekijällä) tai yhdellä 13 riskitekijäryhmästä. Esimerkki löytyy artikkelista IV kuvasta 1, jossa Bert-NER-löydökset ovat merkitty punaisella. Tutkimuksessa menetelmää mitattiin F1-pistemäärän avulla. Kyseinen F1-mittari valittiin, sillä se huomioi tarkkuuden (precision) ja herkkyuden (recall). Ohjattu NER-malli saavutti 76 % F1-pistemäärän (tarkkuus = 75,73 %, herkkyys = 76,56 %).

Ohjaamaton NER – Qexp

Ohjaamaton NER perustuu julkaisussa Moen ym. 2019 esitettyyn menetelmään. Kutsumme tätä menetelmää nimellä QExp. Sen tavoitteena on lyhyemmin etsiä asiankuuluvia psykososiaalisia riskitekijöitä, jotka BERT-NER-menetelmä on mahdollisesti ohittanut. Syötteenä on kysely, joka tässä tutkimuksessa oli yksittäinen riskitekijäesimerkki (sana tai lause). Tämän jälkeen menetelmä suorittaa kyselyn uudelleenkirjoituksen (parafrasoin) löytääkseen lisää sanoja ja lauseita kohdennetuista tekstitiedoista, jolla on samankaltaisia merkityksiä kuin kyselyssä.

Tutkimuksessa käytetty Qexp-lähestymistapa perustuu kolmen komponentin yhdistelmään: 1) semanttinen sanan n-grammin vektorimalli, joka on koulutettu word2vec-työkalulla (Mikolov ym, 2013), 2) tilastollinen kielimalli, joka on koulutettu tietojoukon KenLM -työkalulla (Heafield, 2011) ja 3) asiakirjojen hakukone Apache SOLR muodossa (Apache SOLR, 2004). Menetelmä luo ensin joukon uskottavia lauseita eli ehdokkaita kyselylle. Tämä tehdään kirjoittamalla kyselyn vektorisuus (t) ja etsimällä sitten sanan n-grammi (n-merkin pituinen peräkkäinen jatkuva jakso tekstissä), jotka ovat lähellä semanttisessa vektoriavaruudessa. Nämä n-

grammit yhdistetään muodostamaan lause-ehdokkaita. Tässä menetelmässä tilastollinen kielimalli auttaa hylkäämään sanoja ja lauseita (eli sanasarjoja), jotka ovat todennäköisesti järjettömiä. Seuraavaksi lauseet luokitellaan niiden samankaltaisuuden mukaan kyselyyn (jälleen käyttämällä semanttista mallia), ja lopuksi hakukone tarkistaa, mitkä sana- ja lause -ehdokkaat todellisuudessa ovat kohdekorpuksessa. Vastaavat lauseet on merkitty ja niille on annettu sama psykososiaalisen riskitekijän luokkatunniste kuin kyselylle. Koska tämä menetelmä toimii ilman valvontaa, toisin kuin BERT-NER, sille voidaan yksinkertaisesti antaa luetteloita esimerkkipraaseista hakulausekkeina. Tässä tutkimuksessa ohjaamaton NER toteutettiin koneoppimisen asiantuntijan avulla. Esimerkki löytyy artikkelista IV, kuvassa 1, jossa QExp -menetelmällä löydetty psykososiaaliset riskitekijät on merkitty keltaisella.

Tiesimme aineistomme sisältävän joitain ilmeisiä virheitä (vääriä positiivisia tuloksia). Tarkastimme aineiston aihealueen asiantuntijan käsityönä tekemällä arvioinnilla. Näitä virheellisiä merkintöjä, eli löydettyjä virheellisiä termejä, käytettiin virhelistan muodostamiseen niiden suodattamiseksi tutkimuksessamme.

Lopullinen arviointi

Ohjatun ja ohjaamattoman NER-menetelmän vertailuun poimimme satunnaisesti tietokoneella 100 terveystarkastusasiakirjaa. Nämä eivät olleet samoja asiakirjoja, jotka sisältyivät alussa annotoituihin 196 terveystarkastusasiakirjan tietojoukkoon. Koska 196 asiakirjaa oli annotoitu aiemmin ja annotoinnin perusteella tiedettiin, että 100 terveystarkastusasiakirjan analysointi oli riittävä tuomaan esiin psykososiaalisten riskitekijöiden eri luokat ja niissä esiintyvät variaatiot.

Tässä tutkimuksessa testattiin ja arvioitiin kaksi lähestymistapaa:

Lähestymistapa 1: käytimme ohjattua BERT-NER-mallia

Lähestymistapa 2: käytimme ohjaamatonta NER-mallia, Qexp

Molemmissa lähestymistavoissa arvioitiin:

Oikea riskitekijä tunnistettu ja luokiteltu oikein (todelliset positiiviset).

Riskitekijä tunnistettu, mutta, sillä on väärä riskitekijäluokka.

Ei-riskitekijät (väärät positiiviset).

Puuttuvat riskitekijät (väärät negatiiviset).

Tutkimustulosten arvioinnissa käytettiin ”confusion matrix” taulukkoa, jolle ei löydy virallista suomennosta. ”Confusion Matrix ” on suorituskyvyn mittausta koneoppimisen luokitteluongelmalle, jossa tulos voi olla kaksi tai useampaa luokkaa. Tämä on kuvattu tarkemmin artikkelissa IV, taulukoissa 2 ja 3. ”Confusion matrix” on erittäin hyödyllinen mittaamaan herkkyyttä (Recall), tarkkuutta (Precision), väitteen paikkaansa pitävyyttä (Accuracy). Olisimme halunneet käyttää tulosten arvioinnissa ROC-käyrää.

ROC-käyrä (receiver operating characteristic) on luokitusongelmien suorituskyvyn mittaaminen. ROC on todennäköisyyskäyrä ja monesti siihen liitettävä AUC (Area Under The Curve) kuvaa erotettavuuden astetta tai mittaa. Käyrä kertoo, kuinka paljon malli pystyy erottamaan luokkia. Roc-käyrän käyttö ei tässä tutkimuksessa kuitenkaan ollut mahdollista, koska vastaukset olivat kaksijakoisia (dikotomisias).

4.5 Tutkimuksen eettinen pohdinta

Tutkimusprosessi pohjautui hyvään tieteelliseen käytäntöön noudattaen tarkkuutta ja huolellisuutta tutkimuksessa ja tulosten käytössä ja tallentamisessa. Tutkimukselle ja aineistolle on haettu ja saatu lupa Terveys- ja sosiaaliministeriöstä (STM/4347/2009) sekä Varsinais-Suomen sairaanhoitopiirin eettisen toimikunnan puoltava lausunto tutkimuksen eettisyydestä (77/2009). Lisäksi tutkimukselle on haettu ja saatu tutkimuslupa yksityisiltä työterveyden palveluntuottajilta. Sosiaali- ja terveysministeriöön on kuvattuna tämän tutkimuksen tietosuojaja- ja tietoturvallisuustiedot.

Ensimmäisessä osatutkimuksessa noudatettiin hyvää tieteellistä käytäntöä huomioiden ihmiseen kohdistuvan tutkimuksen eettiset periaatteet (Kohonen ym., 2019). Delphi-menetelmän eettisyydestä huolehdittaessa tulee ottaa huomioon nimettömyys, luottamuksellisuus ja osanottajan oikeudet (Hasson ym., 2000). Haastattelututkimuksessa oli mahdollista saavuttaa vain osittainen nimettömyys (quasi-anonyymiteetti). Osallistujat tiesivät muiden työterveyshoitajien osallistumisesta, mutta eivät heidän nimiään. Osallistujat eivät myöskään tieneet toistensa vastauksia. Ennen haastattelua asiantuntijapaneelin jäsenet saivat sekä suullista että kirjallista tietoa tutkimuksen kulusta, luottamuksellisuudesta, tutkimuksen tarkoituksesta ja haastateltavien oikeuksista keskeyttää tutkimus missä vaiheessa tahansa. Osallistujat allekirjoittivat kirjallisen suostumuksen. Toisen kierroksen kysymyslomake antoi mahdollisuuden miettiä kysymyksiä ilman haastattelijaa.

Toisessa, kolmannessa ja neljännessä osatutkimuksessa käytettiin terveydenhuollon rekisteriä, joka oli deidentifioitu eli henkilö ja yksilöintitiedot oli poistettu ennen sen luovuttamista tutkijoille. Terveystarkastettavien henkilöiden sukupuoli ja syntymävuosi olivat tutkijoiden tiedossa. Tutkimus oli osa suurempaa Louhi/Ikitik-hanketta. Kaikki tutkijat allekirjoittivat salassapitosopimuksen ja saivat ohjeistuksen aineiston tietoturvallisen käyttöön. Deidentifioitu tutkimusaineisto tallennettiin Turun yliopiston palvelimelle, joka toimii tässä tutkimuksessa rekisterinpitäjänä. Palvelimen käyttäjät ovat allekirjoittaneet salassapitosopimuksen. Louhi/Ikitik-hankkeen tietojen käsittelyn rekisteriselostetta noudatettiin myös tässä tutkimuksessa. Rekisteriselosteeseen kuvattiin, miten aineistoa käsitellään, miten se on tietosuojattu sekä kenellä on lupa käyttää aineistoa. Tutkimuksen päätyttyä aineisto tuhotaan tietoturvasuhteita noudattaen.

5 Tulokset

5.1 Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältöalueista ja kirjauksesta

Terveystarkastuksen sisältö

Ensimmäisessä osatutkimuksen haastattelussa ilmeni, että työterveyshoitajat käsittivät terveystarkastuksen sisällön vaihtelevasti. He kertoivat muokkaavansa terveystarkastuksen sisällön yksilöllisesti, ottaen huomioon työntekijän henkilökohtaisen tilanteen ja ongelmat. Kaikkien työterveyshoitajien tavoitteena oli laajasti tukea työntekijän terveyttä ja työkykyä yhdessä työntekijän kanssa. Työterveyshoitajien ensimmäisellä kierroksella haastattelussa mainitsevat terveystarkastuksen sisältöalueet on lueteltu sekä konsensuksen saaneet listattu tärkeysjärjestyksessä (Taulukko 6). Molempien kierrosten sisällöt on kuvattu tarkemmin artikkelissa I, taulukko 2.

Taulukko 6. Työterveyshoitajien mielestä työterveyshuollon terveystarkastuksiin kuuluvat sisältöalueet ensimmäisellä Delphi- kierroksella.

Työterveyshoitajien mainitsema terveys-tarkastuksen osa-alue (n = 17)	Kuuluu sisältöalueeksi (%) Kierros 1
Altisteet	100
Verenpaine BMI	82
Työolosuhteet	76
Rokotukset	76
Fyysisten työolojen selvitys	71
Liikunta	65
Tupakointi	59
Näkö	59
Kuulo	59
Päihteet	59
Ravinto	59
Henkinen työn kuormitus	53
Spirometria	53
Uni/nukkuminen	53
Työkyky	53
Voimavarat työssä tai muut mittari	47
Diabetesriskityyppi	41
Mielialan selvittäminen	41
Vyötärönympäryys	29

Vastaajat korostivat terveystarkastuksen ensisijaisesti kohdistuvan työperäisten sairauksien sekä mahdollisten altisteiden aiheuttamien oireiden tunnistamiseen. Tämän takia työterveyshoitajat kokivat, että terveystarkastusta tehtäessä on tärkeää tuntea työpaikka sekä tietää työstä, työtehtävästä ja työympäristöstä. Työterveyshoitajat kertoivat, että heillä on oltava tietoa työpaikan vaaroista, työssä esiintyvistä altisteista sekä työn turvallisuusasioista. Kaikki työterveyshoitajat (n = 17) arvioivat tärkeäksi työn fyysisten olosuhteiden sekä altisteiden (melu, pöly, kemikaalit, ergonomia) tunnistamisen.

Kaikki työterveyshoitajat (n = 17) kertoivat terveystarkastuksen tärkeimmäksi tavoitteeksi yksilön työkyvyn ylläpitämisen ja edistämisen. Työterveyshoitajat kokivat tärkeäksi selvittää työntekijän oman arvion ja kokemuksen sen hetkisestä terveydentilasta ja hyvinvoinnista. Kaikki työterveyshoitajat määrittivät terveystottumukset (liikunta, ravitsemus, nukkuminen, tupakointi, päihteet) sekä kertoivat mittaavansa painon, vyötärön ympäryksen sekä verenpaineen. Työterveyshoitajat ker-

toivat kartoittavansa rokotustarpeen. Jotkut hoitajat (n = 7) mainitsivat myös sosiaalisen elämän, elämäntilanteen ja harrastukset tärkeiksi osa-alueiksi, ja kiinnittivät huomiota parisuhteeseen ja perhe-elämään.

Työterveyshoitajan kertoman mukaan (n = 13) nykyinen työelämä vaatii yhä enemmän henkisiä voimavaroja ja siksi työn psykososiaalisten tekijöiden tarkastelu terveystarkastuksissa on tärkeämpää. Työterveyshoitajat kertoivat myös yrittävänsä selvittää työyhteisön toimivuutta koskevia asioita. Tällaisissa tapauksissa he arvioivat työyhteisön toimintaa, esimiehen tukea ja johtamiskulttuuria sekä työntekijän työtyytyväisyyttä. Lähes kaikki työterveyshoitajat (n = 15) arvioivat henkistä työkuormaa kyselylomakkeen avulla. Henkisen hyvinvoinnin ja työstressin arviointiin he käyttivät Työkykyindeksiä (TKI). Muutama työterveyshoitaja mainitsi kartoittavansa työntekijän työn tekemisen osaamistasoa sekä työhön saatua perehdytystä.

Ensimmäisellä Delphi-kierroksella kukaan työterveyshoitajista ei maininnut terveystarkastuksen kirjattavana sisältöalueena tietojen antoa, neuvontaa tai ohjausta. Tutkija lisäsi neuvonnan toisen kierroksen kyselylomakkeeseen. Tämän jälkeen lähes kaikki työterveyshoitajat ilmoittivat, että he pitävät neuvontaa ja ohjausta tärkeänä terveystarkastuksen sisältöalueena. Tämä neuvonta koski lähinnä työterveyttä ja -turvallisuutta, mutta työterveyshoitajat antoivat myös tietoa ergonomiasta ja työhön liittyvistä altisteista sekä terveystottumuksista. Työterveyshoitajat pitivät myös erittäin tärkeänä, että he tekevät terveystarkastuksen asiakkaan kanssa. He mainitsivat, että allergiat, perussairaudet ja nykyiset lääkkitykset haastatellaan terveystarkastuksen yhteydessä. Muutama työterveyshoitaja mainitsi arvioivansa osteoporoosin ja diabeteksen riskiä. Työterveyshoitajat kokivat työkykyrajoitteisen työntekijän seurannan ja tukemisen vaikeana, eikä sitä mainittu sisältöalueena.

Toisen Delphi-kierroksen kyselylomakkeen jälkeen työterveyshoitajien yhteisymmärrykseen (70%) perustuen työterveyden terveystarkastus sisältää seuraavat osa-alueet: 1) altistuminen työhön liittyville vaaroille, (2) BMI-mittaus, (3) fyysisten työolojen kartoitus, (4) työhyvinvoinnin kartoitus, 5) terveystottumusten kartoitus, 6) henkisen kuormituksen kartoitus (7) neuvonta, (8) verenpaineen mittaaminen, (9) kuulon tutkimus, (10) tupakointia koskeva neuvonta, (11) altisteita koskeva neuvonta, 12) ravitsemusta koskeva neuvonta, 13) rokotusten tarkistus ja täydentäminen, (14) unta koskeva neuvonta, (15) työkyvyn kartoitus, (16) kroonisten, aiempien sairauksien kartoitus, (17) lääkityksen kartoitus, (18) suojaimien kartoitus ja neuvonta suojaimien käytön tarpeellisuudesta, (19) allergioiden kartoitus, (20) terveystarkastuksen laatiminen, (21) liikuntaa koskeva neuvonta, (22) diabetesriskin kartoitus ja (23) henkisen hyvinvoinnin kartoitus. Kaikki työterveyshoitajien mainitsemat sekä yksimielisyyden saaneet terveystarkastuksen sisältöalueet on lueteltu tärkeysjärjestyksessä artikkelissa I (Artikkeli I, Taulukko 2).

Terveystarkastuksen kirjaaminen

Haastattelujen ensimmäisellä kierroksella työterveyshoitajia pyydettiin kertomaan, mitä ja miten he kirjaavat terveystarkastuksista potilastietojärjestelmään. Nämä haastattelujen tulokset kirjattiin toisen kierroksen kyselylomakkeeseen ja toisen Delphi- kierroksen kyselylomakkeessa työterveyshoitajat ottivat kantaa väittämiin kirjauksesta 10-kohtaisella Osgood-asteikolla, jossa 0 tarkoitti työterveyshoitajan olevan täysin eri mieltä ja 10 täysin samaa mieltä. Molempien kierrosten yhteenveto on esitetty Taulukossa 6.

Työterveyshoitajien mielestä terveystarkastuksissa kerättyjä ja kirjattuja tietoja ei ollut helppo käyttää kiireisen työpäivän aikana. Lähes kaikki työterveyshoitajat kertoivat, että nykyinen kirjaamisjärjestelmä ei ole yhteneväinen. Työterveyshoitajien kirjoittama vapaa, kertova teksti on yksilöllinen, kirjoittajalähtöinen eikä sillä ole yhteistä muotoa tai järjestystä. Jokaisella työterveyshoitajalla on kirjauksessa oma tyyliinsä, taitonsa ja tottumuksensa. Lisäksi kirjauksessa käytetty kieli sisältää lyhenteitä. Puolet työterveyshoitajista pitivät kirjauksia liian vähäisinä, kun taas puolet työterveyshoitajista katsoi, että terveystarkastuksissa oli kirjattu liikaa tietoa.

Osa työterveyshoitajista ilmoitti, että he kirjaavat yhteenveton terveystarkastuksesta. Lähes kaikki työterveyshoitajat kertoivat tekevänsä terveystarkastuksen asiakkaan kanssa, mutta vain osa heistä kirjasi sen potilasasiakirjoihin. Työterveyshoitajien mielestä tärkeintä on parantaa asiakirjojen yhteenvetoa ja terveystarkastusta. Hoidon jatkuvuus on vaikeaa, jos yhteenvetoa ja terveystarkastusta ei selkeästi kirjata. Työterveyshoitajilta kysyttiin molemmilla kierroksilla, mitkä osa-alueet tulisi aina kirjata terveystarkastuksista. Suurimman yksimielisyyden kirjauksista saivat: työhön liittyvät altisteet, terveysongelmat työssä, terveystarkastus, tarkka työn kuvaus ja terveystottumukset lyhyesti. Tarkemmat tiedot on esitetty taulukossa 7.

Taulukko 7. Terveystarkastusten sisältöalueet, jotka työterveyshoitajien mielestä tulee aina kirjata.

Terveystarkastusten sisältöalueet	0	1	2	3	4	5	6	7	8	9	10	n=14	keski-arvo	sd
Työhön liittyvät altisteet										1	13	14	9,93	0,267
Terveysongelmat työssä						1			1	3	9	14	9,29	1,383
Terveysuunnitelma									2	6	6	14	9,29	0,726
Tarkka työn kuvaus							1		1	4	7	13	9,23	1,166
Terveystottumukset lyhyesti								1	2	4	7	14	9,21	0,975
Annettu ohjaus ja neuvonta							1		4	5	4	14	8,79	1,122
Yhteenveto terveystarkastuksesta					1	1		1	1	5	5	14	8,50	1,912
Työyhteisön toimivuus						1		3	3	3	4	14	8,36	1,499
Johtaminen			1			2	1	3	6	1		14	6,86	1,834

Työterveyshoitajat olivat tyytymättömiä terveystarkastuksen kirjauksissa siihen, miten työtä, työhistoriaa ja työtehtävää kirjattiin eli tarkkaa työn kuvausta koskeviin kirjauksiin. He arvioivat myös, että työkykyyn liittyvistä terveysongelmista kirjataan huonosti. Terveystottumuksiin liittyviä terveysongelmia oli helpompi löytää kirjauksista. Työterveyshoitajien mukaan työyhteisön toiminnasta ja johtamisesta ei ole kirjattu tarpeeksi. Työterveyshoitajat kertoivat, että he antavat neuvontaa koko terveystarkastuksen aikana, mutta kirjaavat sitä harvoin. Yksityiskohtaisemmat tulokset on lueteltu taulukossa 3, artikkelissa I.

Työterveyshoitajilta kysyttiin myös, mitkä terveystarkastuksen osa-alueet tule heidän mielestään kirjata vapaana tekstinä ja mitkä rakenteisena tietona. Ensimmäisen kierroksen vastauksissa oli paljon hajontaa, eikä haastattelussa osattu ottaa kantaa niin selkeästi. Toinen kirjallinen kysymyksen asetteleu toi tiettyihin sisältöalueiden kirjauksiin konsensuksen (>70 %). Tarkempi erittely on esitetty taulukossa 8.

Taulukko 8. Terveystarkastusten sisältöalueet ja niiden kirjaaminen vapaana tekstinä tai rakenteisena työterveyshoitajien näkemys, yhteenveto Delphi-kierroksista.

Terveystarkastuksen sisältöalueet	n1/ n2	Kierros 1	Kierros 2	Kierros 1	Kierros 2
		Vapaa teksti %	Vapaa teksti %	Rakenteinen %	Rakenteinen %
Elämän tilanne	17/14	18	100		
Parisuhde, perhe	17/14	29	100		
Vapaa-aika	17/14	18	100		
Sosiaalinen elämä	17/14	18	100		
Johtaminen	17/14	12	100		
Suvun sairaudet	17/14	30	100		
Gynekologikäynnit	17/13	6	93		
Allergiat	17/14	24	93		8
Työtyytyväisyys	17/13	6	93		
Ammatillinen osaaminen	17/13	6	86		8
Krooniset sairaudet	17/14	29	86		15
Työolosuhteet	17/14	65	79	24	14
Osteoporoosiriski	17/13	6	79		14
Suojaimet	17/13	6	79		7
Fyysiset työolosuhteet	17/14	41	72	35	14
Lääkitys	17/13	18	71		23
Koettu työkyky	17/13	6	71		14
Neuvonta ja ohjaus	17/11	0	71		7
BMI	17/14			53	93
Työhyvinvointikyselyt	17/14	6	7	53	86
Rokotukset	17/14	6		53	86
Työkyky	17/14	18		47	86
Näkö	17/14			53	79
Spirometria	17/14			53	79
Uni	17/14	35	21	53	71
Liikunta	17/14	18	14	53	71
Verenpaine	17/14			24	7
Työn altisteet	17/14			12	21
Kuulo	17/14			47	29
Diabetesriski	17/14	6		18	29
Terveysuunnitelma	17/11	6	64		7
Henkinen työkuorma	17/14	29	29	53	64
Päihteet	17/14	18	21	47	64
Vyötärön ympärys	17/14		7	41	57
Mieliala	17/13	41	36	41	50
Tupakointi	17/14	12	7	35	29
Ravitsemus	17/13	12	14	41	43

5.2 Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla

Työn psykososiaalisten tekijöiden kirjaaminen

Toisessa osatutkimuksessa 196 tutkittavan terveystarkastusten kirjauksissa tunnistettiin työterveyshoitajien kirjaamina kaikkiaan 3681 Prima EF:n, PAS1010 mukaisesti määriteltä työpsykososiaalista tekijää. Sekä työn psykososiaalisia riskitekijöitä että ei-riskitekijöitä havaittiin terveystarkastusten kirjauksissa. Näistä lähes puolet (1734, 47 %) oli psykososiaalisia riskitekijöitä, loput olivat positiivisia tai neutraaleja psykososiaalisia tekijöitä. Kirjattujen psykososiaalisten tekijöiden profiilit on esitelty tarkemmin kuviossa 1, artikkeli II. Ensimmäisen ja toisen terveystarkastuksen kirjausten analyysin tulokset olivat samankaltaisia. Kirjaukset erosivat enemmän muiden kuin riskitekijöiden osalta ensimmäisessä ja toisessa terveystarkastuksessa. Työpaikkaväkivallasta oli kirjattuna vain yhdessä terveystarkastuksessa ja työpaikkakiusaamisesta oli kirjattu kolmessa terveystarkastuksessa. Häirintää ei mainittu terveystarkastuksien kirjauksissa lainkaan.

Kirjausten perusteella työn riskitekijöiden seuranta oli terveystarkastuksissa heikkoa. Vain kymmenessä prosentissa kirjauksista työterveyshoitajat oli kirjanneet toisessa terveystarkastuksessa samalle henkilölle saman työn psykososiaalisen tekijän, vaikka se oli ollut työn psykososiaalinen riskitekijä aikaisemmassa terveystarkastuksessa tai kirjallisesti ottanut kantaa edellisessä kirjauksessa esitettyyn työn psykososiaaliseen riskitekijään. Näiden kirjauksien yleisin psykososiaalinen riskitekijä liittyi työympäristöön & välineisiin. Jatko-toimenpiteenä lähes kaikissa (98 %) terveystarkastuksissa todetuille psykososiaalisille riskitekijöille työterveyshoitajat ohjasivat työntekijän työterveyslääkärin vastaanotolle.

Työn ja muun elämän yhteensovittaminen oli selvästi yleisimmin kirjattu psykososiaalinen tekijä (Kuvio 1, artikkeli II) Työn ja muun elämän yhteensovittamisesta työterveyshoitajat kirjasiivat mm seuraavia riskitekijöitä: "heikkous työpäivän jälkeen", "ei resursseja mihinkään fyysiseen harrastukseen työpäivän jälkeen", "yksinäisyys", "avioero" tai "ongelmia perhe-elämässä. Yleisimmin kirjattu työn psykososiaalinen riskitekijä oli työympäristö & välineet. Nämä kirjaukset sisälsivät riskitekijöitä kuten melu, huono valaistus ja huono ergonomia.

Taulukossa 3 on verrattu sukupuolten välisiä eroja psykososiaalisten tekijöiden esiintyvyydessä terveystarkastusten kirjauksissa. Miehillä eniten kirjattuja psykososiaalisia riskitekijöitä terveystarkastuksissa olivat työaikamuoto sekä työympäristö & välineet. Naisilla oli työn psykososiaalisina riskitekijöinä enemmän mainintoja sosiaalisista suhteista, urakehityksestä ja organisaatiokulttuurista ja toiminnasta. Myös ei-riskitekijöissä oli joitakin tilastollisesti merkitseviä eroja sukupuolten välillä. Sosiaaliset suhteet, organisaatiokulttuuri ja toiminta, työn hallinta ja työympäristö & välineet mainittiin useammin naisille terveystarkastusten kirjauksissa. (Taulukko 9).

Taulukko 9. Työterveyshuollon terveystarkastusten kirjauksista annotoinnilla havaitut psykososiaaliset tekijät (riski- ja ei-riskitekijät) sukupuolen mukaan. Keltaisella on merkitty tilastollisesti merkitsevät erot.

Psykososiaaliset tekijät sukupuolen mukaan	Kaikki n=196	Miehet n=136	Naiset n=60	
Riski/ei-riskitekijä	%	%	%	p
Riski_ organisaatiokulttuuri & toiminta 1	12	10	18	0.08
Riski_ Rooli organisaatiossa_1	6	5	8	0.52
Riski_ Sosiaaliset suhteet_1	11	6	23	0.0004
Riski_ Työaikamuoto 1	29	38	8	<0.0001
Riski_ Työn hallinta_1	8	7	10	0.4
Riski_ Työn ja muun elämän yhteensovittaminen_1	41	38	48	0,19
Riski_ Työmäärä ja työtahti_1	36	32	43	0,14
Riski_ Työn sisältö_1	39	35	48	0.08
Riski_ Työpaikkakiusaaminen_1	1	0	2	0.31
Riski_ Työpaikkaväkivalta_1	1	0	2	0,31
Riski_ Työuran kehitys_1	10	7	17	0.047
Riski_ Työympäristö & välineet_1	42	46	32	0.06
Ei-riski_ Organisaatiokulttuuri & toiminta_1	9	4	18	0.001
Ei-riski_ Rooli organisaatiossa_1	11	12	8	0.47
Ei-riski_ Sosiaaliset suhteet_1	20	16	28	0.049
Ei-riski_ Työaikamuoto 1	22	23	22	0.86
Ei-riski_ Työn hallinta_1	27	28	25	0.67
Ei-riski_ Työn ja muun elämän yhteensovittaminen_1	87	85	92	0.18
Ei-riski_ Työmäärä ja työtahti_1	22	21	23	0.75
Ei-riski_ Työn sisältö_1	12	12	13	0.76
Ei-riski_ Työpaikkakiusaaminen_1	0	0	0	-
Ei-riski_ Työpaikkaväkivalta_1	1	1	0	1
Ei-riski_ Työuran kehitys_1	10	10	10	0.95
Ei-riski_ Työympäristö & välineet_1	38	35	45	0.16
Riski_ organisaatiokulttuuri & toiminta 2	11	7	18	0.02
Riski_ Rooli organisaatiossa_2	5	4	8	0.18
Riski_ Sosiaaliset suhteet_2	11	7	20	0.01
Riski_ Työaikamuoto 2	30	39	10	<0.0001
Riski_ Työn hallinta_2	39	35	47	0.12
Riski_ Työn ja muun elämän yhteensovittaminen_2	39	35	47	0.13
Riski_ Työmäärä ja työtahti_2	35	32	43	0.11
Riski_ Työn sisältö_2	37	34	43	0.2
Riski_ Työpaikkakiusaaminen_2	1	0	2	0.31
Riski_ Työpaikkaväkivalta_2	1	1	2	0.52
Riski_ Työuran kehitys_2	11	10	13	0.43
Riski_ Työympäristö & välineet_2	44	49	35	0.08
Ei-riski_ Organisaatiokulttuuri & toiminta_2	12	10	17	0.15
Ei-riski_ Rooli organisaatiossa_2	9	1	5	0.22
Ei-riski_ Sosiaaliset suhteet_2	31	21	52	<0.0001
Ei-riski_ Työaikamuoto 2	37	37	37	0.99
Ei-riski_ Työn hallinta_2	36	31	48	0.02
Ei-riski_ Työn ja muun elämän yhteensovittaminen_2	87	87	87	0.99
Ei-riski_ Työmäärä ja työtahti_2	27	29	23	0.44
Ei-riski_ Työn sisältö_2	22	21	23	0,75
Ei-riski_ Työpaikkakiusaaminen_2	1	0	2	0,31
Ei-riski_ Työpaikkaväkivalta_2	3	3	3	1
Ei-riski_ Työuran kehitys_2	22	22	23	0,84
Ei-riski_ Työympäristö & välineet_2	42	38	53	0,04

Työkyvyn ennustaminen psykososiaalisten tekijöiden kirjaamisen perusteella

Kaikkiin (N = 196) tähän osatutkimukseen mukaan otettuihin terveystarkastuksien kirjauksiin sisältyi Työkykyindeksi eli TKI:n arvo. TKI:n mittauksien keskiarvo ensimmäisessä terveystarkastuksessa oli miehillä 36,5 pistettä ja naisilla 35,2 pistettä (28 – 36 = kohtalainen työkyky). Toisessa terveystarkastuksessa TKI:n keskiarvo oli miehillä 36,2 pistettä ja naisilla 37,3 pistettä (37 – 49 = hyvä työkyky), muutos TKI:ssä oli – 0,3 miehille (ei tilastollisesti merkitsevä) ja + 2,1 naisille (tilastollisesti merkitsevä, $p < 0,01$). Tarkempi erittely on esitetty taulukossa 10.

Taulukko 10. Henkilöiden lukumäärä TKI-luokissa ensimmäisessä ja toisessa terveystarkastuskirjauksessa

		TKI terveystarkastus 2			
		Huono	Kohtalainen	Hyvä/ erinomainen	Yhteensä, n (%)
TKI terveystarkastus 1	Huono (7–27)	21	20	6	47 (24,0)
	Kohtalainen (28–36)	4	23	24	51 (26,0)
	Hyvä/Erinomainen (37–49)	2	13	83	98 (50,0)
	Yhteensä, n (%)	27 (13,8)	56 (28,6)	113 (57,6)	196

Symmetriatesti, $p = 0.0012$, Kappa 0,42

Kirjattuja psykososiaalisia riskitekijöitä oli ensimmäisessä terveystarkastuksessa 2,35 tekijää ja toisessa terveystarkastuksessa 2,37 tekijää. Ensimmäisen ja toisen terveystarkastuksen tai miesten ja naisten välillä ei ollut merkitseviä eroja. Ei-riskitekijöitä oli kirjattu huomattavasti enemmän toisessa terveystarkastuksessa (3,28) kuin ensimmäisessä terveystarkastuksessa (2,58) (Taulukko 10).

Korrelaatiot kirjattujen TKI:n sekä riski- ja ei-riskitekijöiden lukumäärien välillä olivat -0,24 ($p = 0.0007$) ja 0,20 ($p = 0,005$) ensimmäisissä terveystarkastuksissa. Korrelaatio riskitekijöiden lukumäärän ja TKI:n välillä oli tilastollisesti merkitsevä ($r_s = -0,16$, $p < 0,05$) toisessa terveystarkastuksessa. Miesten ja naisten välillä ei ollut huomattavia eroja. Tilastollisesti merkitseviä yhteyksiä havaittiin TKI:n ja tiettyjen psykososiaalisten riskitekijöiden välillä, kuten työn hallinta, työympäristö & välineet, työuran kehitys, rooli organisaatiossa ja työn sisältö.

TKI:n alemmassa luokassa työterveyshoitajat kirjasivat enemmän työn psykososiaalisia riskitekijöitä. Sama päti myös toisin päin, eli jos TKI luokka oli erinomainen/hyvä, työterveyshoitajat kirjasivat enemmän ei-riskitekijöitä. Kirjattujen psykososiaalisten riski- ja ei-riskitekijöiden lukumäärien välillä oli myös tilastollisesti merkitseviä korrelaatioita ensimmäisen terveystarkastuksen kirjauksessa ja TKI-

mittauksissa sekä TKI-mittauksissa toisen terveystarkastuksen kirjauksen osalta ($r_s = -0,19$, $p = 0,008$ ja $r_s = 0,18$, $p = 0,012$, Taulukko 11).

Tässä osatutkimuksessa tarkasteltiin myös kaikkia PAS 1010-standardin työn psykososiaalisia osa-alueita (13 kategoriaa) ja TKI-luokan välisiä yhteyksiä. Jos psykososiaalisia riskitekijöitä esiintyi, TKI oli matala, jos taas riskitekijöitä ei havaittu, oli TKI korkea (Artikkeli II). Tilastollisesti merkitseviä yhteyksiä havaittiin seuraavissa riskitekijöissä: työn hallinta, työympäristö & välineet, työuran kehitys, rooli organisaatiossa ja työn sisältö. Seuraavilla ei-riskitekijöillä oli selvästi yhteys TKI-luokkaan: työn hallinta, työaikamuoto, työmäärä ja työtahti, työn ja muun elämän yhteensovittaminen sekä rooli organisaatiossa. Työn hallinta oli tilastollisesti merkittävä sekä riskitekijöissä että ei-riskitekijöissä molemmissa terveystarkastuksissa. Työympäristö & välineet liittyvät riskitekijöiden kirjaukset olivat tilastollisesti merkittävästi yhteydessä TKI-mittauksiin miesten molemmissa terveystarkastuksissa. Muut yksityiskohdat ja erot miesten ja naisten välillä on eriteltyinä taulukossa 9.

Taulukko 11. Työkykyindeksin keskiarvot sekä annotoitujen psykososiaalisten tekijöiden lukumäärät sukupuolittain sekä niiden kirjaaminen ensimmäisessä ja toisessa terveystarkastuksessa suhteessa Työkykyindeksiin

	Kaikki (N=196)	Miehet (N=136)	Naiset (N=60)	Miehet vs. naiset
	Keskiarvo (SD)	Keskiarvo (SD)	Keskiarvo (SD)	P
TKI (Työkykyindeksi)				
1. terveystarkastus	36,1 (8,2)	36,5 (8,2)	35,2 (8,0)	0,30
2. terveystarkastus	36,5 (7,7)	36,2 (7,9)	37,3 (7,1)	0,35
p, muutos	0,32	0,49	0,003	
Annotoitujen riskitekijöiden lukumäärä				
1. terveystarkastus	2,35 (1,44)	2,24 (1,32)	2,60 (1,66)	0,14
2. terveystarkastus	2,37 (1,60)	2,27 (1,51)	2,60 (1,78)	0,19
p, muutos	0,87	0,84	1,0	
Annotoitujen ei-riskitekijöiden lukumäärä				
1. terveystarkastus	2,58 (1,46)	2,46 (1,48)	2,85 (1,39)	0,09
2. terveystarkastus	3,28 (1,59)	3,08 (1,55)	3,73 (1,60)	0,008
p, muutos	<. 0001	<. 0001	0,0002	

Psykososiaalinen tekijä (riski/ei-riski)	Annotaatioita kirjauksissa		Ei merkintöjä		Annotaatiot vs.ei merkintöjä
	TKI		TKI		
ENSIMMÄINEN TERVEYSTARKASTUS	N	Keskiarvo (SD)	N	Keskiarvo (SD)	P
Kaikki (n = 196)					
Riski/Työn hallinta	15	29,8 (6,1)	181	36,6 (8,1)	0,002
Ei-riski/Työaikamuoto	44	33,9 (8,5)	152	36,7 (8,0)	0,04
Ei-riski/Työn hallinta	53	38,4 (7,5)	143	35,2 (8,2)	0,02
Ei-riski/Työmäärä ja työtahti	43	40,5 (6,7)	153	34,9 (8,1)	<. 0001
Miehet (n = 136)					
Riski/Työn hallinta	9	29,6 (6,9)	127	37,0 (8,1)	0,008
Riski/ Työympäristö & välineet	63	34,4 (8,1)	73	38,3 (8,0)	0,006
Ei-riski/Työn hallinta	38	38,9 (7,4)	98	35,6 (8,4)	0,03
Ei-riski/Työn ja muun elämän yhteensovittaminen	115	37,1 (8,1)	21	33,0 (8,4)	0,03
Ei-riski/Työmäärä ja työtahti	29	40,7 (6,9)	107	35,4 (8,2)	0,002
Naiset (n = 60)					
Riski/Rooli organisaatiossa	5	26,4 (6,6)	55	36,0 (7,7)	0,009
Riski/Työn sisältö	29	33,0 (7,3)	31	37,2 (8,2)	0,04
Ei-riski/Rooli organisaatiossa	5	27,2 (6,1)	55	35,9 (7,8)	0,02
Ei-riski/Työaikamuoto	13	31,2 (7,8)	47	36,3 (7,8)	0,04
Ei-riski/Työmäärä ja työtahti	14	40,1 (6,6)	46	33,7 (7,8)	0,008
TOINEN TERVEYSTARKASTUS					
Kaikki (n = 196)					
Riski/Työuran kehitys	21	30,5 (9,3)	175	37,2 (7,1)	0,0001
Riski/ Työympäristö & välineet	87	34,9 (7,9)	109	37,8 (7,2)	0,01
Ei-riski/Työn hallinta	71	38,4 (7,2)	125	35,4 (7,8)	0,009
Miehet (n = 136)					
Riski/Työuran kehitys	13	26,0 (7,2)	123	37,2 (7,2)	<. 0001
Riski/ Työympäristö & välineet	66	34,7 (8,2)	70	37,5 (7,4)	0,04
Naiset (n = 60)					
Ei-riski/Organisaatiokulttuuri ja toiminta	10	32,5 (5,3)	50	38,2 (7,1)	0,02
Ei-riskit/Työn hallinta	29	39,4 (6,4)	31	35,3 (7,3)	0,02

5.3 Työn psykososiaalisten riskitekijöiden mallinnus ja tunnistus tekstin louhinnalla terveystarkastusten kirjauksista

Kolmanteen osatutkimukseen sisältyi yhteensä 7078 henkilön terveystarkastuskirjauksia, joita oli kaikkiaan 15 443 kirjauskertaa. Terveystarkastusasiakirjojen analyysissä havaittiin yhteensä 26 309 PrimaEF:n PAS 1010 mukaista psykososiaalista riskitekijää. Psykososiaalisista riskitekijöistä ei ollut lainkaan mainintaa 1171 henkilön (16.7 %) terveystarkastusasiakirjoissa. Tarkempi erittely riskitekijöistä (PAS 1010 -kategorioittain) on esitetty taulukossa 1, artikkelissa III. Tarkempi erittely annotointien määrästä on esitetty taulukossa 12. Häirintä-luokkaan ei terveystarkastuksissa ole kirjattu yhtään mainintaa. Useimmin kirjatut psykososiaaliset riskitekijät olivat työmäärä ja työtahti -kategoriassa (10 663 mainintaa), työuran kehitys (4659) ja organisaatiokulttuuri ja toiminta (3019). Työmäärä ja työtahti -kategoriassa yleisimmät riskitekijät olivat työmäärä, kuormitus ja uupumus. Urakehityksessä oli mainintoja YT-neuvotteluista, päätöksistä ja eläkkeelle siirtymisestä. Organisaatiokulttuurissa ja toiminnassa terveystarkastuksiin oli kirjattu ongelmia työyhteisössä, epävarmuutta ja tuen puute esimieheltä.

Taulukko 12. Psykososiaalisten riskitekijöiden annotointien määrä työterveyshuollon terveystarkastuksissa kategorioittain.

Psykososiaaliset riskitekijät-kategoria	Riskitekijä mainittu (%)	Vaihtelu väli	Kaikkiaan (N=15 667)	Keskiarvo	Mediaani	Keskiahajonta
Työympäristö & välineet	91.70	1–45	14366	8.02	7.00	5.50
Työaikamuoto	87.82	1–10	13759	2.15	2.00	1.53
Työn ja muun elämän yhteensovittaminen	67.03	1–12	10502	1.66	1.00	1.73
Työmäärä ja työtahti	64.86	1–12	10162	1.54	1.00	1.64
Työn sisältö	64.19	1–8	10057	1.31	1.00	1.38
Työuran kehitys	62.79	1–13	9838	1.30	1.00	1.42
Organisaatiokulttuuri ja toiminta	60.24	1–15	9438	1.24	1.00	1.45
Työn hallinta	58.95	1–10	9236	1.17	1.00	1.33
Sosiaaliset suhteet	52.66	1–9	8251	0.93	1.00	1.15
Rooli organisaatiossa	41.52	1–6	6505	0.59	0.00	0.84
Kiusaaminen	0.74	1–2	586	0.04	0.00	0.20
Väkivalta työssä	0.70	1–3	110	0.01	0.00	0.09
Häirintä	0	0	0	0	0	0

Lääkärin kirjaamia B-lausuntoja aineistossa havaittiin ja arvioitiin kaikkiaan 1820 kappaletta, 883 oli miehille ja 937 naisille. Taulukossa 10 on tarkempi erittely lukumääristä. Ensimmäisen B-lausunnon saaneiden keski-ikä aineistoissa oli 42,5 vuotta. Ensimmäisen terveystarkastuksen jälkeen tehty keskimääräinen seuranta-aika aineiston perusteella oli 2,3 vuotta. Eläkettä varten B-lausunnon saaneiden keski-ikä oli 56 vuotta (n= 102), sairauspoissaoloja varten B-lausunnon saaneiden keski-ikä oli 54 vuotta (n = 657) ja kuntoutuksia varten B-lausunnon saaneiden keski-ikä oli 52 vuotta (n = 1061) (taulukko 13).

Taulukko 13. Lääkärin kirjaamien B-lausuntojen määrä työterveyshuollon terveystarkastuksissa kategorioittain, muuttujina lausunnon saajan ikä ja sukupuoli.

Muuttuja		Eläke	Kuntoutus	Sairauspoissaolo
N		102	1061	657
Ikä (mediaani)		56.5	52	54
Sukupuoli	Miehet (%)	47 (46)	459 (43)	377 (57)
	Naiset (%)	55 (54)	602 (57)	280 (43)

Lääkäri oli kirjannut yhteensä 68 henkilölle B- lausunnon eläkettä varten. Yleisimmin terveystarkastuksiin kirjattuja psykososiaalisia riskitekijöitä eläkettä varten B-lausunnon saaneiden ryhmässä olivat työmäärä ja työtahti (keskiarvo: 1 001), työuran kehitys (keskiarvo: 0,895), organisaatiokulttuuri ja toiminta (keskiarvo: 0,770) ja työn hallinta (keskiarvo: 0,456). Näitä oli kirjattu useammin henkilöille, joilla on B- lausunto eläkettä varten kuin niille, joilla ei ollut B-lausuntoa ($p \leq 0,0001$). Erot olivat tilastollisesti merkitseviä ($P \leq 0,0001$). Sosiaaliset suhteet ($p \leq 0,0001$), työaikamuoto ($P \leq 0,0001$) ja työn ja muun elämän yhteensovittaminen ($P = 0,0027$) mainittiin harvoin, mutta ero työntekijöihin ilman kirjattua B- lausuntoa eläkettä varten olivat tilastollisesti merkitseviä. Kun kaikki psykososiaaliset riskitekijät otettiin huomioon, työmäärä ja työtahti sekä työuran kehitys ja työn hallinta olivat selkeitä riskitekijöitä B-lausunnon kirjaamiseen eläkettä varten ($P \leq 0,0001$). Työmäärä ja työtahti todettiin riskitekijäksi jopa 17,4 kertaa useammin terveystarkastuskirjauksissa henkilöille, jotka olivat saaneet B-lausunnot eläkettä varten verrattuna niihin, jotka eivät olleet saaneet B-lausuntoa (Artikkeli III, Taulukko 2).

Kaikkiaan 882 henkilöllä oli kirjattuna B-lausunto lääkinnällistä tai ammatillista kuntoutusta varten. Yleisimmin kirjattuja psykososiaalisia riskitekijöitä kuntoutusta varten B-lausunnon saaneilla oli työmäärä ja työtahti (keskiarvo: 1,068). Muita psykososiaalisia riskitekijöitä olivat organisaatiokulttuuri ja toiminta (keskiarvo: 0,604) ja työuran kehitys (keskiarvo: 0,343). Näitä työn psykososiaalisia riskitekijöitä oli kirjattu enemmän niissä terveystarkastuksissa, joissa oli kirjattuna B-lausunto kuntoutusta varten kuin ei B-lausuntoja sisältävissä kirjauksissa. Harvemmin mainittuja

psykososiaalisia riskitekijöitä olivat työn sisältö (keskiarvo: 0,009), työympäristö & välineet (keskiarvo: 0,108), sosiaaliset suhteet (keskiarvo: 0,084), työaikamuoto (keskiarvo: 0,162), työn hallinta (keskiarvo: 0,104) sekä työn ja muun elämän yhteensovittaminen (keskiarvo: 0,153). Kaikissa edellä mainituissa luokissa ero B-lausunnon kuntoutusta varten saaneiden ja ei B-lausuntoa saaneiden välillä oli tilastollisesti merkitsevä. Samoin tilastollisesti merkitseviä vaikkakin harvemmin terveystarkastuksiin kirjattuja olivat: organisaatiokulttuuri ja toiminta ($P \leq 0.0001$) sekä työ määrä ja työtahti ($P \leq 0.0001$) (Artikkeli III, Taulukko 2).

Aineistossa oli 328 henkilöä saanut pitkää sairauspoissaoloa koskevia B-lausuntoja. Työn psykososiaaliset riskitekijät, joita oli kirjattu eniten B-lausuntoihin pitkissä sairauspoissaoloissa, olivat työ määrä ja työtahti (keskiarvo: 0,784). Tämän lisäksi B-lausunnon sairauspoissaoloa varten saaneilla oli psykososiaalisina riskitekijöinä kirjattu organisaatiokulttuuri ja toiminta (keskiarvo: 0,439), työuran kehitys (keskiarvo: 0,282), työn hallinta (keskiarvo: 0,230) ja työn sekä muun elämän yhteensovittaminen (keskiarvo: 0,221). Työpaikkaväkivalta ja -kiusaaminen oli kirjattu vain yhdelle henkilölle, joten niitä ei voitu analysoida. Myös sosiaaliset suhteet oli harvoin kirjattu riskitekijä, mutta silti tilastollisesti merkitsevä. Työaikamuoto $p=0,5614$ sekä työympäristö & välineet eivät olleet tilastollisesti merkitseviä henkilöillä, joilla oli B-lausunto pitkissä sairauspoissaoloissa tai B-lausuntoa ei ollut. Kaikissa muissa luokissa psykososiaalisia riskitekijöitä esiintyi tilastollisesti merkitsevästi enemmän henkilöillä, joille oli kirjoitettu B-lausunto pitkistä sairauspoissaoloista kuin niillä, joilla ei B-lausuntoa ollut kirjoitettu ($P \leq 0.0001$) (Artikkeli III, taulukko 2).

5.4 Työn psykososiaalisten riskitekijöiden tunnistaminen koneoppimisen avulla vapaan tekstin kirjauksista

Tässä osatutkimuksessa käytettiin NER- menetelmää automatisoimaan psykososiaalisten riskitekijöiden havaitsemisen työterveystarkastusten asiakirjojen vapaasti kirjatusta tekstistä. Yhdistimme ohjatun NER-menetelmän (BERT-NER), joka oli koulutettu käsityönä annotoidun tietojoukon avulla, ja uuden ohjaamattoman muotoilumenetelmän, eli parafrasimenetelmän (NER-QExp). Tulokset on esitetty artikkelissa IV, taulukossa 1.

Lähestymistavassa 1 arvioitiin ainoastaan ohjatun BERT-NER-mallin, joka oli koulutettu annotoidulle aineistolle. Lähestymistapa 1 avulla löydettiin 296 oikeaa annotointia psykososiaalista riskitekijöistä. Annotoinneista 23 oli väärässä PAS1010-luokassa. Esimerkiksi lause "ei kestä tätä työtä" oli merkitty urakehitysluokkaan, vaikka sen olisi pitänyt kuulua työn sisältöön. Tällä menetelmällä löytyi

yhteensä 39 väärää annotointia. Tämä lähestymistapa ei tunnistanut puuttuvia annotointeja, joita oli 179 riskitekijää. Tulosten erittely on esitetty taulukossa 14.

Lähestymistavassa 2 yhdistimme BERT-NER- mallin (joka on koulutettu annotoidulle tietojoukolle) QExp-menetelmään. Tutkimuksessa annoimme sille alkupe- räiset merkinnät eli kaikki ne kyselyjen tekstialueet, jotka annotointimenetelmä mää- ritti riskitekijöiksi. Lähestymistapa 2:n avulla löydettiin yhteensä 351 oikein annotoi- tua riskitekijää, lisäksi 32 oli oikein annotoitu, mutta väärään PAS1010-luokkaan ja 146 oli annotoitu väärin. Menetelmä löysi 55 uutta riskitekijää oikeissa luokissa (yh- teensä 351) ja kahdeksan oikeaa riskitekijää, jotka olivat väärissä luokissa (yhteensä 31). Järjestelmä merkitsi lisäksi virheellisesti 107 psykososiaalista riskitekijämainin- ta. Tämä menetelmä ei tunnistanut 116 riskitekijää, ns. puuttuvia annotointeja. Useimmiten riskitekijät, joita ei löytynyt, olivat työn ja muun elämän yhteensovitta- misen -luokassa. Tyypillinen esimerkki väärästä annotoinnista oli "matkatyö: ei", jossa matkatyö on oikein eli riskitekijä, mutta ei tässä tapauksessa, kun sanaan on liitoksissa negaatio eli ei-sana. Nämä tulokset on kuvattu taulukossa 14.





Taulukko 14. Ohjatun BERT-NER- ja QExp-mallin avulla löydettyjen psykososiaalisten riskitekijöi- den määrä PAS1010-luokittain

Riskitekijät BERT-NER	Oikea	Väärä	Ei riskitekijä	Puuttuva	Riskitekijät QExp	Oikea	Väärä	Ei riskitekijä	Puuttuva
Työn sisältö	40	3	8	16	Työn sisältö	45	5	10	9
Väkivalta työssä	0	0	0	3	Väkivalta työssä	0	0	0	3
Kiusaaminen työssä	0	0	0	0	Kiusaaminen työssä	0	0	1	0
Työuran kehitys	6	3	5	6	Työuran kehitys	8	3	5	4
Työympäristö & välineet	106	2	5	61	Työympäristö & välineet	129	4	37	36
Häirintä	-	-	-	-	Häirintä	-	-	-	-
Organisaatiokulttuuri ja toiminta	10	1	0	4	Organisaatiokulttuuri ja toiminta	10	1	1	4
Rooli organisaatiossa	0	1	1	23	Rooli organisaatiossa	6	5	10	13
Sosiaaliset suhteet	6	0	0	4	Sosiaaliset suhteet	7	0	13	3
Työaikamuoto	14	5	6	5	Työaikamuoto	15	5	17	4
Työn hallinta	11	1	1	13	Työn hallinta	10	1	4	14
Työn ja muun elämän yhteensovittaminen	38	3	7	21	Työn ja muun elämän yhteensovittaminen	50	3	35	9
Työmäärä ja työtahti	65	4	6	23	Työmäärä ja työtahti	71	5	13	16
yhteensä	296	23	39	179	yhteensä	351	32	146	129

Molemmissa lähestymistavoissa löytyi selvästi eniten oikeita riskitekijöitä PAS1010-luokan ”työympäristöstä & -välineistä”. Tähän luokkaan kuuluvat muun muassa laitteiden riittämätön saatavuus, soveltuvuus tai huolto; huonot ympäristöolot, kuten tilan puute, huono valaistus, liiallinen melu. Toiseksi eniten riskitekijöitä löytyi luokasta ”työmäärä ja työtahti”. Tässä luokassa on työn ylikuormitusta tai alikuormitusta, työn korkeita aikapaineita ja jatkuvia työn määräaikoja. Kolmanneksi eniten riskitekijöitä tunnistettiin ”työn sisältö” -luokasta. Riskitekijöitä ovat esimerkiksi työn vaihtelevuuden puute tai lyhyet työkierrot, pirstoutunut tai merkityksetön työ, taitojen vajaakäyttö, suuri epävarmuus, jatkuva asiakastyö.

Tarkistuksena yhden asiantuntijan tekemään arviointiin, teimme rinnakkaisarvioinnin, jossa toinen tutkija luki satunnaisotannalla otetut 10 asiakirjaa. Näissä asiakirjoissa oli 165 merkintää. Yksi virhe löydettiin molemmista asiakirjoista. Virheprosentti oli 2,4 % kaikista 165 merkinnästä.

5.5 Yhteenveto tutkimustuloksista

	<p>Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältöalueilta ja kirjauksesta Työterveyshoitajien Delphi-paneeli.</p>
<p>Työterveyshoitajilla ei ollut yhtenäistä käsitystä terveystarkastuksen sisällöstä. Kaikkien työterveyshoitajien tavoitteena oli yhteistyössä työntekijän kanssa tukea työkykyä ja terveyttä.</p> <p>Terveystarkastuksen kirjaaminen sai tässä tutkimuksessa ristiriitaisia tuloksia</p> <p>Terveystarkastuksessa annettua neuvontaa, terveystarkastusta tai jatkosuunnitelmaa eivät kaikki työterveyshoitajat kirjanneet, kuitenkin niiden kirjaamista pidettiin tärkeänä.</p>	
	<p>Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla.</p>
<p>Työterveyshoitajien terveystarkastuksissa kirjaamat psykososiaaliset riskitekijät korreloivat tässä tutkimuksessa merkittävästi Työkykyindeksiin.</p> <p>Yleisimmin kirjatut psykososiaaliset riskitekijät liittyivät työn ja muun elämän yhteensovittamiseen, työympäristö & välineisiin, työmäärään ja työtahtiin sekä työaikamuotoon.</p> <p>Työterveyshoitajien yleisin jatkotoimenpide havaittuihin psykososiaalisiin riskeihin oli työterveyslääkärin konsultaatio.</p> <p>Tutkitut terveystarkastuksen asiakirjat eivät tue hoidon jatkuvuutta.</p>	
	<p>Työn psykososiaalisten riskitekijöiden mallinnus ja tunnistus tekstin louhinnalla terveystarkastusten kirjauksista.</p>
<p>Tekstin louhinta voi tunnistaa psykososiaaliset riskitekijät, jotka kuvataan työterveyshuollon terveystarkastuskirjauksien vapaassa tekstissä.</p> <p>Työterveystarkastuksissa kirjatut psykososiaaliset riskitekijät osoittivat merkittävää korrelaatiota eläkkeitä, pitkittynyttä sairauslomaa tai kuntoutusta varten kirjoitettuihin B-lausuntoihin.</p> <p>Työmäärä ja työtahti sekä työuran kehitys olivat yleisimmin kirjattuja psykososiaalisia riskitekijöitä.</p>	
	<p>Työn psykososiaalisten riskitekijöiden tunnistaminen vapaan tekstin kirjauksesta koneoppimisen avulla.</p>
<p>Koneoppimisella on mahdollista löytää työn psykososiaalisia riskitekijöitä terveystarkastusten vapaan tekstin kirjauksissa.</p> <p>BERT-NER: n (valvottu NER) ja QExp: n (kyselyn laajennus / muotoilu) yhdistelmä näyttää muodostavan sopivimman järjestelmän psykososiaalisten riskitekijöiden tunnistamiseksi.</p> <p>Menetelmä tunnistasi parhaiten työympäristö & välineet-luokan riskitekijöitä.</p>	

Kuvio 4. Yhteenveto tutkimustuloksista tutkimuskysymyksittäin.

6 Pohdinta

6.1 Tulosten pohdinta

Työterveyshoitajien kuvaus työterveyshuollon terveystarkastuksen sisältö-alueista ja kirjauksesta

Ensimmäisessä osatutkimuksessa osoitettiin, että työterveyshoitajien mielipiteet terveystarkastuksen sisältöalueista erosivat yllättävän paljon. Työterveyshoitajat kertoivat pyrkivänsä terveystarkastuksessa kartoittamaan työolosuhteet ja työtyytyväisyyden. Yhteisymmärrys kuitenkin saavutettiin siitä, että työntekijän terveyden ja työkyvyn tukeminen on keskeinen sisältöalue ja tavoite terveystarkastuksessa. Terveystarkastuksen tulisi olla työkykyä tukeva, voimaannuttava ja dynaaminen. Sen tulisi olla vuoropuhelua työntekijän ja työterveyshoitajan välillä työn turvallisuudesta, työkyvystä ja terveydestä (Leino ym. 2014). Jokainen työterveyshoitaja antoi selkeän kuvan siitä, miten he itse toteuttavat terveystarkastuksen sisällön. Terveystarkastuksen sisältöä ei tutkimuksessa saatu yleisesti määriteltä, mutta jokainen työterveyshoitaja määritteli sen oman työkokemuksensa ja ammattitaitonsa perusteella. Terveystarkastuksen sisällön yleinen yhteinen määrittely huomioiden työpaikkakohtaiset erovaisuudet sekä tämän pohjalta yhteneväinen kirjaus tukisi työterveyshoitajan työtä työkyvyn ja työtyytyväisyyden kartoituksessa ja arvioinnissa.

Ensimmäisen Delphi-kierroksen jälkeen työterveyshoitajat eivät maininneet neuvontaa ja ohjausta terveystarkastuksen sisältöalueeksi. Tämä oli yllättävää, koska neuvonta ja ohjaus nähdään yhtenä tärkeimmistä sisältöalueista ennaltaehkäisevässä työterveyden hoitotyössä. Neuvonnan lisäys toisen kierroksen kyselylomakkeeseen sisältöalueeksi osoitti, että sitä pidettiin tärkeänä. Aiempien tutkimusten mukaan neuvonta ja ohjaus on kirjattu huonosti hoitotyössä (Allan, ja Englebright, 2000, Meadows, 2002), vaikka se on tärkeä osa työterveyshoitajan työtä. Oletettavasti se on harvoin kirjattu, koska se on niin automaattinen ja rutiininomainen osa työterveyshoitajan työtä. Neuvonnan ja ohjauksen ja sitä koskevien jatkotoimenpiteiden kirjaaminen auttaisivat turvaamaan terveystarkastettavan työkyvyn tukemista ja työ- ja toimintakyvyn ylläpidon jatkuvuutta. Kirjaamattomuus myös estää tunnistamasta osaa työterveyshoitajien tekemästä hoitotyöstä.

Tutkimuksessa löydettiin ristiriitainen tulos niistä sisältöalueista, jotka työterveyshoitajien mukaan tulisi terveystarkastuksista kirjata. Haastattelussa työterveyshoitajat pitivät terveystarkastuksen yhteenvedoa sisältöalueena, joka olisi kirjattava. Tästä huolimatta terveystarkastuksen yhteenvedon sisällöstä ei Delphi-metodilla päästy riittävään yhteisymmärrykseen. Työterveyshoitajat eivät myöskään esittäneet sitä terveystarkastuksen sisältöalueeksi. Työterveyshoitajien mukaan yhteenvedo ei ollut riittävä eikä selkeästi kirjattu. Hyvin kirjattu terveystarkastuksen yhteenvedo eli tarkastettavalle kirjallisesti annettu terveystarkastus suunnitelma varmistaisi toimenpiteiden jatkuvuutta ja laatua.

Tämän tutkimuksen tulokset ovat hyödyllisiä työterveyshoitajien sähköisen kirjaamisalustan kehittämisessä. Tutkimuksemme on yksi ensimmäisistä, joka on tunnistanut terveystarkastuksen kirjaamisen sisältöalueita. Sähköisen kirjaamisen avulla on mahdollista saada rakenteista ja jäsenneltyä tietoa kliinisen työn sekä johtamisen päätöksentekoa varten. Tällä tavalla voimme tulevaisuudessa mahdollisesti arvioida terveystarkastusten vaikuttavuutta sekä saada terveystarkastuksissa tehtyä työtä näkyvämmäksi. Terveystarkastukset ovat inhimillinen ja kustannustehokas tapa tukea työikäisen työkykyä ja terveyttä. Työterveyshuollon terveystarkastuksia ja interventioita sekä kirjauksia tulee kehittää vastaamaan nyky-yhteiskunnan vaatimuksia.

Työkyvyn ennustaminen työn psykososiaalisten riskitekijöiden avulla

Toisessa osatutkimuksessa keskeinen havainto oli, ettei psykososiaalisten tekijöiden osalta kirjaamisessa tunnistettu työkyvyn tukemisen jatkuvuutta. Kirjauksista ei löytynyt kannanottoa edellisen terveystarkastuksen löydöksiin eikä havaintoja edellisissä terveystarkastuksissa todettuihin psykososiaalisiin riskitekijöihin. Psykososiaalisiin riskeihin yleisimmin kirjatut tekijät olivat samat riskitekijät, jotka on todettu johtavan varhaiseläkkeelle aiemmissa tutkimuksissa (Ilmarinen 2001, Leka ym. 2003, Slany ym. 2013). Alhaisen työkykyindeksin (TKI) taustalla olevia syitä oli dokumentoitu todella vähän. Samoin suunnitelmia, jatkotoimenpiteitä ja seurantaan tunnistettiin vain muutamasta kirjauksesta liittyen matalaan TKI:iin tai edellisiin psykososiaalisiin riskitekijöihin. Useimmissa kirjauksissa, joissa edellinen TKI oli ollut matala, ei seuraavan terveystarkastuksen kirjauksessa ollut mainintaa aikaisemmasta matalasta TKI:stä tai siihen vaikuttaneista tekijöistä tai psykososiaalista riskitekijöistä. Terveystarkastuksissa mitattu heikko TKI löytyi kirjattuna vain neljänneksessä kaikista asiakirjoista. Kuitenkin työterveyshoitajat olivat kirjanneet psykososiaalisia riskitekijöitä eniten alimmissa TKI-luokissa. Tämä voi viitata siihen, että kun riskitekijät ovat ilmeisiä, ne myös kirjataan. Useimmin kirjatut riskitekijät olivat työn hallinta, työympäristö & välineet, urakehitys, rooli organisaatiossa sekä työn sisältö.

Tulokset osoittivat, että työterveyshoitajien kirjaamien psykososiaalisten tekijöiden määrä korreloi TKI:n pistemääriin. Työterveyshoitajat kirjasivat eniten työympäristöön & välineisiin sekä työn ja muun elämän yhteensovittamiseen liittyviä tekijöitä. Tuloksemme ovat aiempien tutkimusten mukaisia, jotka ovat osoittaneet, että työn hallinta, työn vaatimukset ja sosiaalinen tuki liittyvät sairauspoissaoloihin ja varhaiseläkkeeseen (Ilmarinen 2001, Leka ym. 2003, Slanyt ym. 2013). TKI oli parempi toisessa terveystarkastuksessa, mutta lisää tutkimusta tarvitaan sen selvittämiseksi, ovatko tehdyt terveystarkastukset ja sen jälkeiset mahdolliset tukitoimet nostaneet TKI:ä.

Tutkimuksen tulokset osoittivat myös, että työterveyshoitajilla ei ollut selkeää rakennetta psykososiaalisten tekijöiden kirjaamiseksi. Psykososiaalisten tekijöiden kirjaaminen terveystarkastuksissa oli epäjohdonmukaista, niukkaa ja sattumanvaraista. Tämä vaikeuttaa kirjauksien käyttöä jatkohoitoa tai toimenpiteitä koskevassa päätöksenteossa. Työterveyshoitajan kirjauksia oli eniten työympäristöä & välineitä sekä työn ja vapaa-ajan yhteensovittamista kuvaavissa riskitekijöissä. Työterveyshoitajat kirjasivat positiivisia tai neutraaleja psykososiaalisia tekijöitä enemmän kuin riskitekijöitä. Kymmenessä prosentissa terveystarkastuksista ei ollut mainittu lainkaan sanoja työ tai työkyky. Työterveyshoitajat kirjasivat todella vähän myös työpaikkaväkivallasta tai -kiusaamisesta, eivätkä kirjanneet lainkaan häirinnästä. Työterveyshoitajien kirjaamista ja työkyvyn jatkuvuuden seurantaa parantaisi, jos työterveyshoitajilla olisi kansallisesti sovittu, jäsenelty, rakenteinen kirjaamisalusta psykososiaalisille riskitekijöille.

Tässä tutkimuksessa 95 % kaikista psykososiaalisten riskitekijöiden ongelmilanteista ohjattiin jatkotoimenpiteenä työterveyslääkärille. Vain kahdessa kirjauksessa työterveyshoitaja ohjasi työntekijän yrityksen henkilöstöosaston tai esimiehen kanssa keskustelemaan. Näin pienen osan ohjaaminen työpaikalle ongelmien ratkaisemiseen voi johtua puutteellisesta yritysyhteistyöstä, työntekijän kielloista, yhteisten keinojen puutteesta ongelmien ratkaisemiseksi tai huonon yhteyden ja luottamuksen puutteesta johdon kanssa tai yrityksestä suojella työntekijää enemmiltä ongelmilta työpaikalla. Työterveyslääkärin vastaanottokäynnin lisäksi tarvitaan muita toimia työn psykososiaalisiin riskitekijöihin. Osa ongelmista on ratkaistavissa työntekijän omalla aktiivisuudella sekä työpaikan toimilla esim. työolojen tai työympäristön kehittämällä. Työympäristö & työvälineet, työn ja muuan elämän yhteensovittaminen, työn sisältö, työmäärä ja työtahdin riskitekijät olivat kuitenkin eniten kirjattuja psykososiaalisia riskitekijöitä ja ne olisivat ratkaistavissa organisaation tasolla. PAS 1010- standardia tulisi käyttää ohjeena yhdenmukaistamaan psykososiaalisten tekijöiden riskinhallintaa (Leka ym. 2015). Psykososiaalisia riskitekijöitä voidaan harvoin vähentää lääketieteellisin keinoin. Lisätutkimusta tarvitaan työterveyshuollon terveystarkastusten kirjaamisesta.

Työn psykososiaalisten riskitekijöiden mallintaminen ja tunnistaminen tekstin louhinnalla terveystarkastusten kirjauksista

Kolmannessa osatutkimuksessa arvioitiin tekstin louhinnan mahdollisuutta työn psykososiaalisten riskitekijöiden tunnistamiseksi työterveyshuollon terveystarkastuksien vapaiden tekstien kirjauksista, joista voitaisiin ennustaa työkyvyn menettämisen mahdollisuus lääkärin kirjaamien B-lausuntojen perusteella. Nämä B-lausunnot olivat kirjoitettuja eläkkeelle siirtymiseen, kuntoutukseen tai pitkäiseen sairauspoissaoloon. Psykososiaaliset riskitekijät kirjattiin useammin niissä terveystarkastuksissa, jotka johtivat lääketieteellisiin lausuntoihin, kuin niissä, joista ei B-lausuntoa kirjattu. Työmäärä ja työtahti, urakehitys sekä organisaatiokulttuuri ja toiminta mainittiin kaikkien lääketieteellisten lausuntojen luokkien yhteydessä, ja ne kaikki ennustivat lääketieteellisiä lausuntoja varhaiseläkkeelle, kuntoutukseen ja sairauslomaan.

Tutkimuksen mukaan psykososiaalisten riskitekijöiden kirjaaminen työterveyshuollon terveystarkastuksissa voisi ennustaa työkyvyn menetyksiä. Työn hallinta ja urakehitys dokumentoitiin huomattavasti useammin työntekijöillä, joilla oli lääkärin B-lausunto kirjattuna. Työn hallinta on tunnettu varhaiseläkkeelle siirtymisen riskitekijä (Christensen ym., 2008, Heponiemi ym., 2008, Laaksonen ym. 2010, Lahelma ym. 2012), mutta urakehitystä on tutkittu vähemmän psykososiaalisena riskitekijänä. Lausekkeita, jotka liittyivät urakehitykseen psykososiaalisena riskitekijänä, olivat ”ei voi edetä urallaan”, ”tekee hankalaa työtä”, ”osallistui YT- neuvotteluihin” tai ”raskas työ”. Näiden teemojen kansallisesti yhtenäinen rakenteellinen ja säännöllinen tarkastelu, joko terveystarkastuksen esikyselyssä tai itse terveystarkastuksessa, edistäisi varhaista tukea työkykyyn ja tukisi työterveysyhteistyössä jatkotoimenpiteitä.

Tulokset tukevat psykososiaalisten riskitekijöiden rakenteisen sähköisen kirjaamisen kehittämistä terveystarkastuksissa. PAS 1010 työkirja tarjoaa rakenteet, joka yhdenmukaistaa psykososiaalisten riskitekijöiden hallintaa ja johtamista. Se tarjoaa lainsäädäntöön perustuvat puitteet ja selvittää terminologiaa, jota voidaan käyttää perustana työn psykososiaalisten riskitekijöiden havaitsemiseksi terveystarkastuksissa. Se voisi myös selvittää näiden riskitekijöiden kirjaamista. Tässä tutkimuksessa psykososiaalisia riskitekijöitä ei mainittu 1171 työntekijän terveystarkastuksissa. Näiden mainintojen puuttuminen saattoi johtua myös siitä, ettei psykososiaalisten riskitekijöiden kirjaamiseen ole selkeää rakennetta, mutta silti niistä saatetaan keskustella tarkastuksen aikana. Tällä hetkellä psykososiaalisten riskitekijöiden tunnistaminen ja käyttö työkyvyn tukemiseksi riippuu työterveyshuollon ammattilaisten kyvystä kirjata ja koota tietoa eri kirjauksista.

Työn psykososiaalisten riskitekijöiden tunnistaminen koneoppimisen avulla vapaan tekstin kirjauksista

Neljännessä osatutkimuksessa testattiin ohjattua (BERT-NER) ja ohjaamatonta (NER QExp) lähestymistapaa. Näistä kahdesta menetelmästä osoittautui paremmaksi psykososiaalisten riskitekijöiden tunnistamiseksi vapaiden tekstien terveystarkastuksista ohjaamaton QExp lähestymistapa. Lähestymistapa tunnisti oikein 70 % psykososiaalisista riskitekijöistä terveystarkastuksista. Jos mukaan lasketaan vielä oikein tunnistetut, mutta vääriin luokkiin luokitellut riskitekijät, on tulos 76 %. Nämä tulokset ovat lupaavia ja ovat linjassa aiempien tutkimusten kanssa tunnistaa automaattisesti merkitsemättömät sanat ja ilmaukset kliinisessä kirjauksessa (Skeppstedt et al 2014, Topaz et al 2019).

Molemmat lähestymistavat havaitsivat yli 60 % työn psykososiaalisista riskitekijöistä. Tämä osoittaa, että koneoppimisen avulla on mahdollista tunnistaa psykososiaaliset riskitekijät. Lisätutkimuksilla ja lisätyöllä on mahdollista rakentaa järjestelmä tai työkalu, joka helpottaisi työterveyshoitajan sekä koko moniammatillisen tiimin jatkotoimenpiteitä ja päätöksentekoa. Työkalu antaisi työterveyshoitajalle tietoa riskitekijöiden määrästä, kun ne kirjoitetaan terveystarkastuksissa eri lauseisiin ja kohtiin sekä mahdollisesti herätteitä, jos riskitekijöitä olisi havaittavissa. Jatkokehityksen myötä työkalu voisi löytää tietoa kaikista aikaisemmista tiedoista ja siten auttaa ylläpitämään työkykyä ja edistämään työterveysalan ammattilaisten työkyvyn lukutaitoa, kun tieto ei ole hajallaan. Ihanteellisimmillaan tällainen työkalu korostaisi mahdollisista riskitekijöistä terveystarkastusten kirjaamisen aikana tai takautuvasti.

6.2 Tutkimuksen pätevyys ja luotettavuus

Seuraavissa kappaleessa pohditaan tämän tutkimuksen pätevyyttä eli validiutta sekä luotettavuutta eli reliabiliteettia.

6.2.1 Tiedon keruun pätevyys ja luotettavuus

Tässä tutkimuksessa on pätevyyden ja luotettavuuden suhteen joitain rajoituksia Ensimmäisessä osatutkimuksessa asiantuntijapanelistien 20 tutkittavan ryhmä pieneni 14 työterveyshoitajaan toisella kierroksella. Tämä heikentää osittain tutkimuksen luotettavuutta. Kuitenkin tämän tutkimuksen paneelin jäsenet valittiin huolella ja he edustivat työterveyshoitajia monipuolisesti.

Tutkimusaineiston terveystarkastusten kirjaamiset edustivat sen hetken todellisuutta, tutkimusaineisto on kerätty vuonna 2007. Aineisto on jonkin verran vanha. Tämän kaltaisen aineiston hankinta tutkimukseen on kallista ja tässä tutkimushankkeessa ei ollut mahdollista hankkia uutta aineistoa. Lisäksi haluttiin, että kaikissa

tutkimuksen vaiheissa käytettiin samaa aineistoa, jotta voitiin luotettavasti vertailla eri tuloksia. Tekstin louhintatutkimus on edelleen harvinaista hoitotieteen alalla ja yhden aineiston analysointikuntoon saattaminen on tutkimuksen alkuaikoina ollut erittäin aikaa vievä prosessi. Työterveyshuollon terveystarkastusten kirjaukset tehdään edelleen valtaosin vapaana tekstinä, vain pieni osa terveystarkastuksen löydöksistä ja kirjauksista on rakenteisessa muodossa.

Kirjaukset ovat yhden organisaation työterveyshoitajien tuottamia, ja niissä voi näkyä tämän yrityksen toimintatavat ja organisaatiokulttuuri. Tutkimusaineiston työterveyshoitajien kirjaukset edustavat laajaa maantieteellistä aluetta, kokemusta ja ikäjakaumaa kansallisesti, vaikka aineisto koostui yhden valtakunnallisen työterveyspalvelujen tuottajan tiedostoista Suomessa.

Tämän tutkimuksen aineisto oli riittävän suuri käsin tehtyyn annotointiin. Koneoppimisessa aineiston suuruudesta on hyötyä, jotta koneoppiminen tapahtuu luotettavammin. Vielä suurempi aineisto olisi lisännyt tutkimuksen luotettavuutta. Nyt kaikkia psykososiaalisia riskitekijöitä ei välttämättä esiintynyt tässä aineistossa, ja on mahdollista, että suuremmalla aineistolla myös harvemmin kirjatut psykososiaaliset riskitekijät, kuten häirintä, olisi tullut esiin.

6.2.2 Tutkimusmenetelmien pätevyys ja luotettavuus

Ensimmäisessä osatutkimuksessa käytetyssä Delphi-menetelmässä on tiettyjä heikkouksia, jotka voivat vaikuttaa sen tieteelliseen pätevyYTEEN. Tässä tutkimuksessa otoksen valintaa ja nimettömyyttä voidaan pitää hyväksyttävänä. Tutkimuksessa suoritettiin vain kaksi tiedonkeruukierrosta. Delphi-menetelmässä on suositeltavaa, että tutkimuksessa suoritetaan kolme tai jopa neljä kierrosta. Tämä tutkimus olisi hyötynyt yhdestä lisäkierroksesta, jotta työterveyshoitajat olisivat voineet tarkentaa mielipiteitään, toki osallistujat vähenivät joka kierroksella. Menetelmän käyttö oli hyödyllistä, koska se antoi tutkimukseen mahdollisuuden kerätä reaali maailmasta erilaisia asiantuntijalausuntoja ja -mielipiteitä. Delphi-menetelmän tulokset edustavat asiantuntijoiden mielipiteitä enemmän kuin kiistattomia tosiseikkoja (Burns ja Grove 2020, Keeney ym. 2004).

Yksi tutkimuksen rajoituksista on myös se, ettei psykososiaalisten riskitekijöiden kirjauksessa ole kansallisesti yhtenäistä käytäntöä. Käytimme kuitenkin annotointiin merkintäjärjestelmänä PAS 1010:tä, joka perustuu kattavaan psykososiaalisten riskien hallintamallien katsaukseen kaikkialla Euroopassa (Leka ym. 2008, Leka ym. 2011).

Manuaalisessa annotaatiossa heikkoutena voivat olla inhimilliset virheet ja annotoijan huomiokyky. Manuaalinen annotointi on hidasta ja monivaiheista (Meystre ym. 2008, Skeppstedt ym. 2014.). Tekstin louhinnalle tämä oli kuitenkin ainoa tapa

opettaa konetta tunnistamaan psykososiaaliset riskitekijät. Tätä virhemahdollisuutta vähennettiin useammalla annotoijalla, tarjoilla ohjeilla ja varmistuskierroksella.

Koneoppimisessa luotettavuutta pyrittiin lisäämään käyttämällä useita eri lähestymistapoja sekä näiden vertailua. Vain yksi asiantuntija tarkisti kaikki annotointilistat, ja toinen asiantuntija tarkisti vain 10 % tuloksista. Virheiden määrä näin tarkastettuna oli kuitenkin vähäinen (2.6 %).

6.2.3 Tulosten pätevyys ja luotettavuus

Delphi-menetelmää käytettäessä on huomioitava, että konsensuksen käsitys on epäselvä ja tutkija määrittelee sen tutkimuksessaan. Vaihtoehtoisesti on tutkimuksissa myös ehdotettu, että vastausten pysyvyyttä (stabiiliutta) kysymyskierroksilla voitaisiin pitää yhtä luotettavana konsensuksen indikaattorina kuin prosenttimittausta (Hasson ym. 2000, Hyttinen ja Aavarinne 2000, Kennedy 2004). Ensimmäisessä osatutkimuksessa vastaukset esitettiin prosentteina. Prosenttiosuuksina esittäminen vääristää helposti tuloksia näin pienen asiantuntijapanelin näkemyksissä (Hasson ym. 2000). Sisällön johdonmukaisuutta kierrosten välillä ei voitu laskea tilastollisesti pienen ryhmäkoon takia.

Toisessa osatutkimuksessamme heikkoutena oli, että vain harvoilla alkuperäisen tietojoukon terveystarkastettavilla oli TKI kirjattuna kahdessa peräkkäisessä terveystarkastuksessa. Tämä osoittaa, että TKI:ä ei terveystarkastuksessa käytetä järjestelmällisesti. Ehkä tämä oli myös syy siihen, miksi näiden kahden mittauksen välillä ei tunnistettu selkeitä muutoksia.

Terveystarkastusten kirjausten laatu, lyhenteet sekä sekavuus estivät tekemästä kehittyneempiä analyysejä. PAS 1010 -pohjaisten merkintöjen (annotaatioiden) luominen oli aluksi haastavaa kirjausten monisanisuuden takia. Toisella annotoijien koontumisella ja kirjausohjeiden päivityksellä sekä kehittämällä onnistuimme kuitenkin nostamaan merkintöjen välisen yksimielisyysarvon (Cohenin Kappa-arvo) 0,49:stä 0,72:een, jota voidaan pitää merkittävänä parannuksena (Cohen 1960).

6.3 Ehdotuksia tulevia tutkimuksia varten

Tämä tutkimus vahvisti hoitotyön kirjaamisen ja työn psykososiaalisten riskitekijöiden tutkimusta Suomessa. Silti on tarvetta tuleville lisätutkimuksille näistä aiheista:

- Nyt tehtyjä sanojen annotointia ja luokituksia tulee laajentaa jatkotutkimuksia varten. Tämä tarkoittaa uuden annotointitutkimuksen tekemistä. Osa psykososiaalisten riskitekijöiden sanoista ja sanapareista tai lauseista ei ollut mukana tässä aineistoissa. Psykososiaalisia riskitekijöitä on hyvä tutkia saman PAS1010-luokituksen mukaan.

- Terveystarkastusten vapaiden kirjausten luokittelua tulee jatkaa. Haastattelimme suomen kielen kielitieteilijää, jotta löytäisimme parhaiten koneoppimiseen soveltuvat lyhennykset ja kirjaustavat psykososiaalisiin riskitekijöihin. Jatkotutkimuksissa tulisi tehdä yhteistyötä kielitieteilijöiden kanssa.
- Tutkimusta tulisi jatkaa laajemmalla ja kattavammalla aineistolla. Isompi aineisto mahdollistaisi kaikkien työn psykososiaalisten kirjauksien muotojen löytymisen ja auttaisi työkalun luotettavuuden parantamista arviointia.

6.4 Tutkimuksen tulosten hyödyntäminen

Seuraavat vaikutukset käytännön hoitotyöhön ja päätöksentekoon voidaan esittää tämän tutkimuksen tulosten perusteella:

Hyödyntäminen hoitotyöhön

- Työn psykososiaalisten riskitekijöiden kirjaamista tulee yhtenäistää terveystarkastuksessa yhtenäisen kirjaamisrakenteen avulla.
- Työkyvyn lukutaidon mahdollisuutta tulee vahvistaa nykyisissä tietojärjestelmissä lisäämällä rakenteisuutta.
- Työn psykososiaalisiin kuormitustekijöihin ja muihin työkykyä alentaviin riskitekijöihin tulisi kehittää työkaluja hälyttämään työterveyshoitajalle riskitekijöiden esiintyessä.

Hyödyntäminen käytännössä

- Kansallisten rakenteisten kirjaamisalustojen kehittämistyötä tulee tukea ja jatkaa.
- Potilastietojen järjestelmätoimittajien tulisi tehdä yhteistyötä kansallisesti.
- Psykososiaalisten riskitekijöiden opastus ja neuvonta työpaikoilla ja niiden ennaltaehkäisyn jatkaminen.

Hyödyntäminen koulutuksessa

- Terveystarkastuksen ja terveystarkastuksen kirjaamiseen tarvitaan ohjausta ja opetusta peruskoulutuksesta alkaen ja sitä tulee jatkaa koko työuran ajan.
- Kirjauskoulutusta tulisi lisätä terveydenhoitajakoulutukseen.
- Kansallisten yhteisten kirjauskäytäntöjen koulutus työterveyshuoltoon tulisi suunnitella.

Hyödyntäminen johtamisessa

- Esihenkilön tulisi kartoittaa, miten työn psykososiaaliset riskitekijät työterveyshuollossa tunnistetaan ja kirjataan sekä mitä jatkotoimenpiteitä riskitekijöille toteutetaan.

- Kansallisesti tulisi käyttää yhtenäisiä työkaluja psykososiaalisten riskitekijöiden johtamisessa ja työterveysyhteistyössä, esim. PAS1010 oppaita, lomakkeita ja työkirjoja.

7 Johtopäätökset

Psykososiaaliset riskitekijät ovat olennainen osa työterveyshuollon terveystarkastusta työntekijän työkyvyn selvityksessä. Psykososiaaliset riskitekijät tulisi kansallisesti yhtenäisin tavoin kartoittaa ja kirjata. Työterveyshuollolla yksinään ilman yhteistyötä työntekijän tai työnantajan kanssa ei ole välineitä tai mahdollisuuksia vaikuttaa tai tehdä toimenpiteitä työn psykososiaalisiin riskitekijöihin kuten esimerkiksi työmäärän ja työtahtiin, urakehitykseen tai organisaatiokulttuuriin ja toimintaan liittyviin riskeihin. Siksi työyhteisön, esihenkilön ja työntekijän tiivis yhteistyö on välttämätöntä. Työterveyshuollon tulisi myös tarjota terveystarkastuksissa saatua arvokasta tietoa työnantajille ja työyhteisölle psykososiaalisista riskitekijöistä noudattaen samalla henkilökohtaista tietosuojaa ja varmistamalla työntekijän oikeuden yksityisyyden suojaan.

Terveystarkastuksissa ja muissa työterveyshuollon käyntitapahtumissa tulee työterveyshuollon ammattihenkilöiden rohkaista työntekijöitä omaan työhön liittyvien kysymysten, pulmien hallintaan ja ratkaisemiseen. Terveystarkastuksessa keskustelujen tulisi motivoida työntekijää aktiivisuuteen ja itseohjautuvuuteen työhön ja työkykyyn liittyvissä asioissa. Myös yritysten tulisi olla tietoisia siitä, kuinka psykososiaalista hyvinvointia tuetaan työpaikalla ja nähdä se investointina yrityksen tuottavuuteen.

Tutkimuksen tulokset osoittivat, että tekstin louhinnalla on mahdollista tunnistaa vapaan tekstin kirjauksista psykososiaalisia riskitekijöitä. Tällainen automaattinen tunnistus parantaisi potilastietojärjestelmän kirjauksien tulkintaa. Määriteltyjen sanojen ja lauseiden automaattinen poiminta ja hälyttäminen auttaisivat työterveyshuollon ammattilaisia päätöksenteossa ja jatkotoimenpiteissä. Psykososiaalisten riskitekijöiden tunnistaminen kaikista vapaan tekstin kirjauksista kaikilla työterveyshuollon käynneillä kaikilla ammattiryhmillä edistäisi jatkuvuutta työkyvyn tukemisessa. Näin saataisiin paremmin deidentifioitua tietoa työpaikoille terveystarkastusten ja muiden työterveyshuollon käyntien löydöksistä työyhteisön ja työympäristön tasolla yritysten toimien tueksi.

Yhteenvetona voidaan todeta, että psykososiaalisten riskitekijöiden kartoittaminen työterveyshuollon terveystarkastuksissa on tärkeää työkyvyn edistämiseksi.

Työkyvyn kehittymistä työuran aikana ei ole helppo havaita terveystarkastusten kirjauksissa, koska terveystiedot ovat hajanaiset. On selkeä tarve kehittää yhteinen kansallinen kirjaamistapa psykososiaalisiin riskitekijöihin sekä kehittää tekstinkäsittelytyökalu työkykyä uhkaavien työn psykososiaalisten ja muiden riskien havaitsemiseksi varhaisessa vaiheessa. Tekstin louhinta, digitalisointi ja tekoäly voivat auttaa tämän tavoitteen saavuttamisessa.

Kiitokset

Tämän väitöskirjan tekeminen on kestänyt aikamoisen ajanjakson. Monet ihmiset ovat tukeneet ja rohkaisseet minua tutkijaksi kasvamisen polulla. Haluan ilmaista suuret kiitokseni teille kaikille, vaikka en voi mainita teitä kaikkia erikseen.

Tämä tutkimus tehtiin Turun yliopiston hoitotieteen laitoksella. Kiitän lämpimästi laitoksen johtajaa, professori Helena Leino-Kilpeä mahdollisuudesta sekä kärsivällisestä ja pitkämielisestä tuesta tämän tutkimuksen aikana.

Olen sanomattoman suuret kiitokseni velkaa väitöskirjani ohjaajille professori Sanna Salanterälle ja professori Jaakko Hartialalle. Kiitos, kun uskoitte minuun tutkijana. Kiitos, kun uskoitte tähän tutkimukseen silloinkin, kun muut eivät siihen uskoneet. Sanna, arvostan suunnattomasti tekemääsi työtä hoitotieteen, digitaalisten ratkaisujen ja monitieteisen tutkimuksen hyväksi. Sinun luovuutesi ja viisautesi löysivät ratkaisuja, kun tutkimukseni joutui umpikujaan. Osaat lempeän lujasti kannustaa tekemään laadukasta tutkimusta. Tulen kaipaamaan yhteisiä keskustelujamme, jotka ohjasivat kasvuani vahvasta käytännön ajattelusta alkeelliseen tieteelliseen ajatteluun. Ilman sinun valoisaa uskoasi tähän tutkimukseen, se olisi jäänyt kesken. Jaakko, kiitos kuuluu yhtä lailla myös sinulle. Lähdit tutkimuksen alussa rakentamaan kanssani työterveydestä aivan uudenlaista näkökulmaa, psykososiaalisia tekijöitä, jotka siihen aikaan olivat määrittelemättä, mutta ovat tänä päivänä itsestään selvyys. Kiitän kannustavista kommentteistasi ja asiantuntemuksestasi työterveyshuollosta sekä tutkimuksesta. Näkökulmasi ja rauhallinen ohjauksesi auttoivat erityisesti artikkelien kirjoituksessa.

Kiitokset myös esitarkastajilleni dosentti Outi Kanstalle sekä dosentti Päivi Leikkolalle. Annoitte arvokasta palautetta sekä kannustavia kommentteja, kun jo itse olin sokeutunut tutkimukselleni ja tekstilleni.

Kiitos myös seurantaryhmän ainoalle jäsenelle, jota ei ole vielä kiitetty, Maija Huplille. Kommenttisi ja kannustuksesi tulivat aina tarpeeseen. Iso kiitos myös Kari-Pekka Martimolle työterveyden asiantuntijuudesta sekä silmiä avaavien kysymysten kysymisestä kirjoittaessamme artikkeleita. Samoin kiitän kanssatutkijoitani Hans Moenia, Kai Hakalaa ja Juho Heimosta. Kiitos opastuksesta tekstin käsittelyn ja koneoppimisen saloihin. Tehtävänne minun opettamisessani ei ollut helppo, mutta kärsivällisyytenne selittää ja kysyä oli aivan uskomaton. Kiitos myös tilastotieteen

osaajille Pauli Puukka, Elisa Löyttyniemi ja Simo Teperi, tulososio rakentui teidän avullanne. Kiitos myös annotointia kanssani tehneille Anne Mohnille, Minna Lai-
neenojalle ja Heidi Parisodille. Kiitos kuuluu myös kaikille entisen Louhi-hankkeen
nykyisen Ikitik-hankkeen tutkijoille, jotka olette olleet myötävaikuttamassa tähän
tutkimukseen.

Kiitos koko hoitotieteen laitoksen henkilökunnalle ja jatko-opiskelijoille. Heljä
Lundgren Laine, Elina Kontio, Johanna Olli, Erika Siirala, Laura-Maria Peltonen ja
Anna Axelin kiitos avustanne ja rohkaisustanne. Kiitos myös Mehiläiselle kärsiväl-
lisestä ja joustavasta suhtautumisesta opiskeluihini ja mahdollisuudesta käyttää sii-
hen aikaa. Erityiskiitos Kari Antilalle ja Ville Tuomolalle, teidän avullanne tämä
tutkimus oli mahdollinen. Kiitos ihan jokaiselle työkaverilleni lämpimistä sanois-
tanne ja kannustuksista. Erityiskiitos Jaana Malmikareelle työnohjauksestasi kahvi-
kupposten äärellä. Suurkiitos myös esihenkilöilleni Tapio Virralle ja Antti Miettiselle,
kun olette mahdollistaneet työn ja opiskelun yhteensovittamisen. Onhan teitä
pöriäisiä paljon, paljon muitakin, jotka jaksoitte minuun uskoa. Kiitos mm. Liisa,
Irmeli, Heli, Tiina, Leena, Tuuli, Sanna, Kati, Annika, Pauliina, Pirkko, Pirjo, Päivi.

En olisi tätä ikinä pystynyt tekemään ilman läheisten tukea. Olen niin etuoikeu-
tettu, että olen välittävien läheisten ja ystävien ympäröimä. Kiitos sisarukset, appi-
vanhemmat, sukulaiset ja kaikki ystävät. Kiitos tätin tyttö Jonsku, et antanut koskaan
nenän viistää maata. Erityisesti kiitos ystävälleni Marjolle lempeistä halauksista ja
nauruista, kun tiede ei sujunut. Kiitos myös vävylleni Joonakselle ja miniälleni Min-
tulle, jaksoitte aina kannustaa anoppianne. En löydä sanoja Pyry, Kamu ja Vauva,
miten kiitollinen olen, että olette elämässäni. Olette opettaneet minulle, että elämässä
tärkeintä on hetkessä eläminen, leikki ja isot halaukset.

Jäljellä on enää elämäni kolme tärkeintä ihmistä, joille kuuluu suunnattoman
suuri kiitos. Tämä hetki ei olisi koskaan koittanut ilman teitä. Kiitos Uronen ikuisesta
uskosta kykyihini, sadoista keitetyistä kahvikupillisista ja asioiden oikeaan mittasuhtei-
siin laittamisesta. Vielä on kiittämättä elämäni kahta suurinta ”saavutustani” eli
te, Mimmu ja Jörö. Teidän olemassaolonne antaa minulle supervoimia. Teiltä olen
oppinut, että muna on aina viisaampi kuin kana. Erityiskiitos Mimmulle koko kirjan
lukemisesta ja tarkastamisesta. Rakastan teitä kuuhun ja takaisin.

Turussa 15.6.2022

Leena Uronen

Lähdeluettelo

- Agency for Safety ja Health at Work. Experts forecast on emerging psychosocial risks related to occupational safety and health. European risk observatory report 2007. Luettavissa: <https://osha.europa.eu/en/tools-and-publications/publications/reports/7807118>.
- Ailisto, H. (toim.), Heikkilä, E., Helaakoski, H., Neuvonen, A & Seppälä T. Tekoälyn kokonaiskuva ja osaamiskartoitus. 2018. Luettavissa: <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160925/46-2018-Tekoalyn%20kokonaiskuva.pdf>
- Bernal D, Campos-Serna, J., Tobias A., Vargas-Prada, S., Benavides F. & Serra, C. 2015. Work-related psychosocial risk factors and musculoskeletal disorders in hospital nurses and nursing aides: a systematic review and meta-analysis. *Int J Nurs Stud Feb*, 52(2), s. 635-48.
- Bird, S., Klein, E & Loper, E. 2009. Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc, California.
- Blair W, Smith B. 2012. Nursing documentation: frameworks and barriers. *Contemporary Nurse*, 41(2), s.160-168.
- Bliese, P. D., Edwards, J. R., & Sonnentag, S. 2017. Stress and well-being at work: a century of empirical trends reflecting theoretical and societal influences. *J. Appl. Psychol*, 102, s. 389–402.
- Buckingham C. 2000. Classifying clinical decision making: a unifying approach. *Journal of Advanced Nursing*, 32(4), s. 981-989.
- Burns N, Grove S. 2020. The practice of nursing research. Conduct, critique ja utilisation. W.B. Saunders Company, Philadelphia.
- Carbonell JG, Hayes PJ. 1992. Natural Language Understanding. In: Shapiro SC, editor. Encyclopedia of Artificial Intelligence. Wiley, New Jersey, s. 660-77.
- Cendales, B., Useche, S. A., Gómez, V & Bocarejo, J. P. 2016. Bus operators' responses to job strain: an experimental test of the job demand-control model. *J. Occup. Health Psychol*, 22, s. 518–527.
- Christensen K, Fevile H, Labriola M, & Lund T. 2008. The impact of psychosocial work environment factors on the risk of disability pension in Denmark. *Eur J Public Health*, 18(3), s. 235-237.
- Clark A., Fox C. & Lappin S. 2013. The handbook of computational linguistics and natural language processing. John Wiley & Sons. New Jersey.
- Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), s.37-46.
- Cox T, Griffiths A & Rial E. 2000. Research on work-related stress. Luxembourg: Publications Office of the European Union. Luettavissa: <http://osha.europa.eu/en/publications/reports/203>.
- De Groot, K., Triemstra, M., Paans, W. & Francke, A. 2019. Quality criteria, instruments, and requirements for nursing documentation: A systematic review of systematic reviews. *Journal of Advanced Nursing*, 75(7), s. 1379-1393
- DeJong G.F. 1982. An Overview of the FRUMP System. In: Ringle WGLaMH, editor. Strategies for Natural Language Processing. Hillsdale, NJ: Lawrence Erlbaum, s. 149-76.
- Devlin, J, & al. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. Luettavissa: arXiv preprint arXiv:1810.04805.

- Ensio, A. 2008. Päätöksenteon prosessimalli hoitotyössä. Teoksessa Saranto, K., Ensio, A., Tanttu, K., Sonninen, A-L. (toim.). *Hoitotietojen systemaattinen kirjaaminen*. 2. uudistettu painos. Helsinki: WSOY Oppimateriaalit Oy.
- Ensio, A. ja Saranto, K. 2004. *Hoitotyön elektroninen kirjaaminen*. Sipoo: Suomen sairaanhoitajaliitto ry.
- Entzeridou, E., Markopoulou, E. & Mollaki, V. 2018. Public and physician's expectations and ethical concerns about electronic health record: Benefits outweigh risks except for information security. *Int J Med Inform*, Feb (110), 98-107
- Eurofund, 2015. Psychosocial work environment: Health and well-being at work - Q2 2014-Q1 2015. (EurWORK topical update). Luettavissa: <https://www.eurofound.europa.eu/publications/report/2015/eu-member-states/psychosocial-work-environment-health-and-well-being-at-work-q2-2014-q1-2015-eurwork-topical-update>.
- Eurofund, 2014. Psychosocial risks in Europe Prevalence and strategies for prevention.2014. Luettavissa <https://www.eurofound.europa.eu/publications/report/2014/eu-member-states/working-conditions/psychosocial-risks-in-europe-prevalence-and-strategies-for-prevention>
- European Agency for Safety and Health at Work. 2008. EU-OSHA Annual report 2007: bringing safety and health closer to European workers. Luettavissa: <https://osha.europa.eu/en/publications/eu-osha-annual-report-2007-bringing-safety-and-health-closer-european-workers>.
- Frick K, Zwetsloot G. 2007. From safety management to corporate citizenship - An overview of approaches to health management. *Work Health and Management Control*, 2007,99-134.
- Hakala, K., Pyysalo, S. 2019. Biomedical named entity recognition with multilingual BERT. Proceedings of the 5th Workshop on BioNLP Open Shared Tasks. Luettavissa: <https://aclanthology.org/D19-5709>.
- Hasson, F., Keeney S. & McKenna, H. 2000. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), s. 1008-1015.
- Haug, P., Christensen, L., Gundersen, M., Clemons, P., Koehler, S. & Bauer, K. 1997. A natural language parsing system for encoding admitting diagnoses. Proceedings: a conference of the American Medical Informatics Association. Proc AMIA Annu Fall Symp.,1997, 814-8.
- Heafield, K. 2011. KenLM: Faster and smaller language model queries. In Proceedings of the Sixth Workshop on Statistical Machine Translation, WMT '11, Stroudsburg, PA, USA. Association for Computational Linguistics, s. 187-197.
- Heponiemi, T., Kouvonen A., Vänskä J., Hallila H., Sinervo T. & Kivimäki M. 2008. Health, psychosocial factors, and retirement intentions among Finnish physicians. *Occupational Medicine*, 58, s. 406-412.
- Hripesak G, Austin J. & Alderson, P.O. 2002. Use of natural language processing to translate clinical information from a database of 889,921 chest radiographic reports. *Radiology*, 224, s.157-63.
- Hyttinen, H., Aavarinne, H. 2000. Delfi -menetelmä hoitotieteellisessä tutkimuksessa – esimerkkinä gerontologinen hoitotyö. *Hoitotiede*, 12(3), s. 121-128.
- Häyrinen, K., Harno, K. & Nykänen, P. 2011. Use of headings and classifications by physicians in medical narratives of EHRs – an evaluation study in a Finnish hospital. *Applied Clinical Informatics*, 2(2), s. 143-157.
- Häyrinen, K., Saranto, K. & Nykänen, P. 2008. Definition, Structure, Content, Use and Impacts of Electronic Health Records: A Review of the Research Literature. *International Journal of Medical Informatics*, 77(5), s. 291-304.
- Ilmarinen J., Gould, R., Järviskoski A. & Järvisalo, J. 2006. Työkyvyn moninaisuus. Teoksessa Raija Gould, Juhani Ilmarinen, Jorma Järvisalo ja Seppo Koskinen (toim.) Työkyvyn ulottuvuudet. Terveys 2000 -tutkimuksen tuloksia, Hakapaino Oy, Helsinki 2006, s. 17-34.
- International Labour Office. 2016. Psychosocial risks, stress and violence in the world of work. *International Journal of Labour Research*, Vol. 8, Issue 1-2.
- Karhula A-L. (Toim.) 2006. Terveystarkastukset työterveyshuollossa. Työterveyslaitos. Vammalan kirjapaino, Vammala.

- Karvala, K., Leino, T., Oksa, P., Santonen, T., Sainio, M., Latvala J. & Uitti, J. (Toim.) 2019. Altistelähtöinen työterveysseuranta. Työterveyslaitos. Helsinki: Duodecim.
- Keeney, S., Hasson, F. & McKenna, H. 2001. A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies*, 38, s. 195–200.
- Kelan työterveyshuoltotilasto 2019. Luettavissa: https://helda.helsinki.fi/bitstream/handle/10138/331665/Kelan_tyoterveyshuoltotilasto_2019.pdf?sequence=4&isAllowed=y.
- Kelley, T., Brandon, D. & Docherty, S. 2011. Electronic nursing Documentation as a Strategy to Improve Quality of Patient care. *Journal of Nursing Scholarship*, 43(2), s. 154-162.
- Kennedy, H. 2004. Enhancing Delphi research: methods and results. *Journal of Advanced Nursing*, 45 (5), s. 504-511.
- Kohonen, I., Kuula-Luumi A. & Spoof, S-K. (toimikunta), Tutkimuseettisen neuvottelukunnan julkaisu 3. 2019. Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakoarviointi Suomessa. Tutkimuseettinen neuvottelukunta. Helsinki. Luettavissa: https://www.tenk.fi/sites/tenk.fi/files/Ihmistieteiden_eettisen_ennakoarvioinnin_ohje_2019.pdf?_ga=2.56559851.1780503720.1640693274-1625180354.1632043877
- Krogsbøll, L. Jørgensen, K. & Larsen, C. 2012. General health checks in adults for reducing morbidity and mortality from disease. *Cochrane Database Syst Rev*; 2012:10. Luettavissa: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009009.pub2/full>.
- L 785/1992. Laki potilaan asemasta ja oikeuksista. Luettavissa: <http://www.finlex.fi>.
- L 759/2004. Laki yksityisyyden suojasta työelämässä. Luettavissa: <http://www.finlex.fi>.
- L159/2007. Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä. Luettavissa: <https://www.finlex.fi/fi>.
- Laaksonen, M., Pitkämäki, J., Rahkonen, O. & Lahelma, E. 2010. Work arrangements, physical working conditions, and psychosocial working conditions as risk factors for sickness absence: Bayesian analysis of prospective data. *Ann Epidemiol.*, 20, s. 332-338.
- Lafferty, J., McCallum, A. & Pereira, F. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. Luettavissa: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1162&context=cis_papers.
- Lahelma, E, Laaksonen, M., Lallukka, T., Martikainen, P., Pietiläinen, O. & Saastamoinen, P. 2012. Working conditions as risk factors for disability retirement: a longitudinal register linkage study. *BMC Public Health*, 12, s. 309-319.
- Lauri, S., Salanterä, S. 2002. Developing an instrument to measure and describe clinical decision making in different nursing fields. *J Prof Nurs*, 18(2), s. 93–100.
- Leino, T., Rautio, M., Kanervisto, M., Tilli, J. & Kaleva, S. 2014. Terveystarkastuskäytännöt työterveyshuollossa. Turkimusraportti. Työterveyslaitos 2014. Luettavissa: https://www.julkari.fi/bitstream/handle/10024/120383/Terveystarkastuskaytannot_suomalaisessa_tyoterveyshuollossa.pdf?sequence=1&isAllowed=y.
- Leka, S., Cox, T. 2008. PRIMA-EF Guidance on the European Framework for Psychosocial Risk Management. A Resource for Employers and Workers representatives. Luettavissa: http://www.primaef.org/uploads/1/1/0/2/11022736/prima-ef_brochure_english.pdf.
- Leka, S., Griffiths, A. & Cox, T. 2003. Work organization and stress; systematic problem approaches for employers, managers, and trade union representatives. Geneva: WHO. Luettavissa: <https://apps.who.int/iris/bitstream/handle/10665/42625/9241590475.pdf?sequence=1&isAllowed=y>.
- Leka, S., Jain, A., Iavicoli, S. & Di Tecco, C. 2015. An evaluation of the policy context on psychosocial risks and mental health in the workplace in the European Union: achievements, challenges, and the future. *BioMed Res. Int.*, 2015:213089.
- Leka S., Kortum E. 2008. A European framework to address psychosocial hazards. *J Occup Health*, 50(3), s. 294–6.
- Lennes, M., Ahjoniemi, S. 2005. Puheaineiston annotaatio eli nimikointi. Luettavissa: <https://lennes.github.io/puheen-annotaatio/>

- Liira, J., Rautio M. & Leino T. 2014. Terveystarkastuksen tekeminen. Teoksessa Jukka Uitti (toim.) Hyvä työterveyshuoltokäytäntö. Saarijärven Offset, Saarijärvi, s. 189–210.
- Mahler, C., Ammenwerth, E., Wagner, A., Tautz, A., Happek, T., Hoppe, B. & Eichstädter, R. 2007. Effects of a Computer-based Nursing Documentation System on the Quality of Nursing Documentation. *Journal of Medical Systems*, 31(4), s. 274–282.
- Manninen, P., Laine, V., Leino, T., Mukala, K. & Husman, K. (Toim.) 2007. Hyvä työterveyshuoltokäytäntö. Työterveyslaitos. Vammala: Vammalan kirjapaino.
- Martimo K-P, Antti-Poika M. 2000. Työterveyshuollon terveystarkastukset ja niiden tuloksellisuus. *Suomen lääkärilehti*, 55, s. 4631–4635.
- Meystre, S., Savova, G., Kipper-Schuler, K. & Hurdle, J. 2008. Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. In Geissbuhler A, Kulikowski C, editors. IMIA Yearbook of Medical Informatics 2008. *Methods Inf Med*, 47 suppl 1, s. 128-44.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. & Dean, J. 2013. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, s. 3111-3119.
- Moen, H., Peltonen, L-M., Suhonen, H., Matinelli, H-M., Mieronkoski, R., Telen, K. & Salanterä, S. 2019. An Unsupervised Query Rewriting Approach Using N-gram Co-occurrence Statistics to Find Similar Phrases in Large Text Corpora. In Proceedings of the 22nd Nordic Conference on Computational Linguistics, s. 131-139.
- Mykkänen, M., Miettinen, M. & Saranto, K. 2016. Standardized Nursing Documentation Supports Evidence-Based Nursing Management. *Studies in Health Technology and Informatics*, 225, s. 466-470.
- Mäkelä-Bengs, P., Vuokko, R. 2014. Potilaskertomuksen rakenteistamisen menetelmät Suomessa. Teoksessa: Hyppönen, H., Vuokko, R., Doupi, P. ja Mäkelä-Bengs, P. (toim.) Sähköisen potilaskertomuksen rakenteistaminen. Menetelmät, arviointikäytännöt ja vaikutukset. Raportti 31/2014. Helsinki: Terveiden ja hyvinvoinnin laitos.
- Niedhammer, I., Chastang, J. F., Sultan-Taïeb, H., Vermeylen, G., & Parent-Thirion, A. 2013. Psychosocial work factors and sickness absence in 31 countries in Europe. *The European Journal of Public Health*, 23(4), s. 622–629.
- Nissinen, S., Soini, S., Leino, T., Hakulinen, H. & Saranto K. 2018. Kanta-arkiston käyttökokemuksia työterveyshuollossa. *FinJeHeW*, 18,10(1).
- Niuwenhuijsen, N., Bruinvelde, D. & Frings-Dresen M. 2010. Psychosocial work environment and stress-related disorders, a systematic review. *Occup Med*, 60, s.277–86.
- O'Neill, E., Dluhy, N. & Chin, E. 2005. Modelling novice clinical reasoning for a computerized decision support system. *Journal of Advanced Nursing*, 49(1), s. 68–77.
- Palmgren, H., Jalonen, P., Jurvansuu, H., Kaleva, S. & Tuomi, K. 2008. Tietojen antaminen, neuvonta ja ohjaus (TANO) työterveyshuolloissa. Työterveyslaitos. Työ ja ihminen. Tutkimusraportti 35. Tampereen yliopistopaino Oy. Juvenes Print.
- Pehkonen, I. 2018. Työkykytalo. Työterveyslaitos. Luettavissa: <https://www.ttl.fi/tyoyhteiso/tyokykytalo/>
- Polit, D.F. & Beck, C. 2017. Nursing Research. Generating and Assessing Evidence for Nursing Practice. Wolters Kluwer Health, Philadelphia.
- Rautio, M., Piikivi, L. & Nikkonen M. 2000. Työterveyshuollon työkykyä ylläpitävän toiminnan painopistealueet ja vaikuttamismahdollisuudet. *Työ ja ihminen, työympäristötutkimuksen aikakauskirja*, 14, s. 485–495.
- Rosenbloom, S., Denny, J., Xu, H., Lorenzi, N., Stead, W. & Johnson, K. 2011. Data from clinical notes: a perspective on the tension between structure and flexible documentation. *J Am Med Inform Assoc*, 18(2), s.181-6.

- Rosenbloom, S., Stead, W., Denny, J., Giuse, D., Lorenzi, N., Brown, S. & Johnson, K. 2010. Generating clinical notes for electronic health record systems. *Applied Clinical Informatics*, 1, s. 232–243.
- Sager, N., Friedman, C. & Lyman, M. 1987. Medical Language Processing: Computer Management of Narrative Data. Computer Science. Addison Wesley, Reading, Massachusetts.
- Saranto, K., Kinnunen, U-M., Kivekäs, E., Lappalainen, A-M., Liljamo, P., Rajalahti, E. & Hyppönen, H. 2014. Impacts of structuring nursing records: a systematic review. *Scandinavian Journal of Caring Sciences*, 28(4), s. 629-647.
- Sauni, R., Leino, T. 2007. Työterveyshuollon terveystarkastusten vaikuttavuus. *Lääketieteellinen Aikakauskirja Duodecim*, 123(6), s.723–30.
- Silveira, D., Marin, H. 2005. Nursing documentation in occupational health. *Stud Health Technol Inform*, 116, s.143-8.
- Skeppstedt, M., Kvist, M., Nilsson, G.& Dalianis, H. 2014. Automatic recognition of disorders, findings, pharmaceuticals, and body structures from clinical text: An annotation and machine learning study. *Journal of Biomedical Informatics*, 49, s.148–158.
- Slany, C., Schütte, S., Chastang, J-F., Parent-Thirion, A., Vermeulen, G. & Niedhammer, I. Psychosocial work factor and long sickness absence in Europe. *Int J Occ. and Env. Health*, 20(1), s. 16–25.
- Sosiaali- ja terveystieteiden tutkimuskeskus. 2004. Sähköisten potilasasiakirjajärjestelmien toteuttamista ohjaavan työryhmän loppuraportti. Työryhmämuistioita 2994:18. Helsinki 2004 a. <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/71030/TRM200418.pdf?sequence=1>.
- Spyns, P. 1996. Natural language processing in medicine: an overview. *Methods Inf Med.*, 35 (4–5), s. 285–301.
- Stengel, D., Bauwens, K., Walter, M., Köpfer, T & Ekkernkamp, A. 2004. Comparaison of Handheld Computer-Assisted and Conventional Paper Chart Documentation of Medical Records. A Randomized Controlled Trial. *The Journal of Bone and Joint Surgery*, 86 (3), s. 553-560.
- Suodicani, P., Bonde, J-P., Olsen, K.& Gyntelberg, F. 2013. Job satisfaction and intention to quit the job. *Occupational Medicine*, 63, s. 96-102.
- Suominen ym. 2006. Theoretical Consideration of Ethics in Text Mining of Nursing. *Stud Health Technol Inform*, 22, s. 359-64.
- Suominen, H., Lehtikunnas, T., Back, B., Karten, H., Salakoski T.& Salanterä, S. 2007. Applying language technology to nursing documents: Pros and cons with a focus on ethics. *International Journal of Medical Informatics*, 76(2), s. 293–301.
- THL. 2018b. Kanta-sanasto. Kanta-palveluihin liittyviä keskeisiä käsitteitä termeineen, määritelmineen ja kielenkäyttöohjeineen, versio 1.0. Luettavissa: https://thl.fi/documents/920442/2920708/Kanta_sanasto_29062018.pdf/7859b9f8-7e67-4b96-b98b-78b81d10193e.
- Topaz, M., Murga, L., Gaddis, K., McDonald, M, Bar-Bachar, O., Goldberg, Y.& Bowles, K. 2019. Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approaches. *Journal of Biomedical Informatics*, 2019 Feb;90:103103. doi: 10.1016/j.jbi.2019.103103.
- Tuomi, K., Ilmarinen, J., Jahkola, M., Katajarinne, L. & Tulkki, A. 1997a. Työkykyindeksi, Työterveyshuolto 19, 2. korjattu painos, Työterveyslaitos, Helsinki.
- Työterveyshuoltolaki 1383/2001. Luettavissa: <https://www.finlex.fi/fi/laki/ajantasa/2001/20011383>.
- Työterveyslaitos. 2008. PRIMA-EF eurooppalaisia linjauksia psykososiaalisten riskien hallintaan. Opaskirja työnantajille ja työntekijöiden edustajille. Luettavissa: http://www.prima-ef.org/uploads/1/1/0/2/11022736/who_booklet_fin.pdf.
- Uitti, J. (toim.) 2014. Hyvä työterveyshuoltokäytäntö. Sosiaali- ja terveystieteiden tutkimuskeskus, Työterveyslaitos. Helsinki.
- Uitti, J., Sauni, R.& Leino, T. 2007. Työterveyshuollon vaikuttavuus asiakkaiden näkökulmasta. *Duodecim*, 123, s.723–30.

- Vahtera, J., Kivimäki, M., & Pentti, J. 2000. Effect of change in the psychosocial work environment on sickness absence: a seven year follow up of initially healthy employees. *J Epidemiol Community Health*, 54, s.484–493.
- Valtioneuvoston asetus hyvän työterveyshuoltokäytännön periaatteista, työterveyshuollon sisällöstä sekä ammattihenkilöiden ja asiantuntijoiden koulutuksesta. 2013. Luettavissa: <https://finlex.fi/fi/laki/alkup/2013/20130708>.
- Valtioneuvoston asetus terveystarkastuksista erityistä sairastumisen vaaraa aiheuttavissa töissä. 2001. Luettavissa: <https://www.finlex.fi/fi/laki/alkup/2001/20011485>.
- van der Berg, T., Elders, L., de Zwart, B. & Burdorf, A. 2009. The effects of work-related and individual factors on the Work Ability Index: a systematic review. *Occup Environ Med*, 66(4), s. 211–20.
- Vartia, M. ym. 2012. Psykososiaaliset tekijät suomalaisessa yhteiskunnassa. Sosiaali- ja terveystieteiden raportteja ja muistioita 2012:14. Luettavissa <http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74112/URN:NBN:fi-fe201504224815.pdf>
- Viitasalo, K., Helmiö, K., Härmä, M., Lindström, J., Peltonen, M., Puttonen, S. & Koho, A. 2011. Työterveyshuolto ehkäisee vuorotyön ja elintapojen terveysriskejä. Tyypin 2 diabeteksen seulonta ja ehkäisy ilmailualan työterveyshuollossa. Terveyden ja hyvinvoinnin laitos (THL). Raportti 66/2011. Helsinki.
- Wilbanks, B., Moss, J. 2018. Evidence-Based Guidelines for Interface Design for Data Entry in Electronic Health Records. *Computers, Informatics Nursing*, 36(1), s. 35–44.



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

ISBN 978-951-29-8935-5 (PRINT)
ISBN 978-951-29-8936-2 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)