

Eeva Lahti

**HENKILÖHISTORIAN TIETOKANNAT
JA ESIMERKKI NIIDEN KÄYTÖSTÄ
HISTORIAN TUTKIMUKSESSA**

- I HENKILÖHISTORIAN TIETOKANNAT**
- II PIENTEN LASTEN KUOLLEISUUDESTA
LAVANSAARELLA VUOSINA 1751-1899**

Historian laitos
Pro gradu-
tutkielma
25.1.2001

Jyväskylän yliopisto
Suomen historia
Jyväskylä

I HENKILÖHISTORIAN TIETOKANNAT

TIIVISTELMÄ

Lahti, Eeva Marjatta

Henkilöhistorian tietokannat/ Eeva Lahti

Jyväskylä: Jyväskylän yliopisto, 2001

104 s.

Tutkielma

Tässä tutkielmassa tarkastellaan historiallisten väestöluetteloiden, lähinnä kirkonkirjojen tietosisällön pohjalta rakennettuja tietokantoja. Tutkielman päätavoitteena on esittää malli, jota voidaan soveltaa henkilöhistorian tietokantojen arvioinnissa. Koska henkilöhistoriallisia tietokantoja rakennetaan palvelemaan tutkimusta, malli auttaisi tutkijoita ratkaisemaan, ovatko tietokannan tietosisältö ja hakujärjestelmät sellaisia, että niistä on hyötyä hänen tutkimukselleen.

Henkilöhistorian tietokantojen suunnitteluun ja rakentamiseen on perehdytty kirjallisuuden pohjalta. Yksityiskohtaisesti on tarkasteltu kolmea relaatiotietokantana toteutettua järjestelmää, jotka ovat ruotsalainen Demografiska databasen, suomalainen Karjala-tietokanta sekä amerikkalainen Gillespie-järjestelmä. Edellä mainittuja tietokantoja on tutkimuksessa kuvattu ja verrattu yhdenmukaisesti. Tietokantoja on arvioitu erityisesti tietosisällön sekä hakujen perustana toimivien rakenteiden ja hakujärjestelmien osalta. Tutkimuksen keskeisin tulos on, että tietokantojen tietosisältöä voidaan yhdenmukaisesti verrata. Tietokantojen toisistaan poikkeavat tavoitteet sekä erilaiset hakujen perustana toimivat rakenteet ja toteutukseen käytetyt ohjelmistot vaikeuttavat hakujärjestelmien yhdenmukaista vertaamista. Henkilöhistorian tietokantojen rakentaminen edellyttää niiden suunnittelijoilta huolellista perehtymistä historiallisen lähdeaineiston erityispiirteisiin. Hakujärjestelmiä suunniteltaessa ei ole mahdollista tietää eri tieteen aloja edustavien käyttäjien kaikkia tarpeita. Käyttäjille olisikin tietokannan käytöstä saatujen kokemusten perusteella varattava tilaisuus hakujärjestelmien täydentämiseen.

AVAINSANAT: historiallinen väestöluettelo, kirkonkirja, linkitys, nimien normalisointi, henkilöhistorian tietokanta, tietosisältö, hakujärjestelmä

Sisällys

1. Johdanto	1
2. Henkilöhistorian lähteet	5
2.1 Kirkonkirjat	6
2.1.1 Historiakirjat	8
2.1.2 Rippikirjat	9
2.2 Kirkonkirjojen käyttömenetelmiä	13
2.3. Kirkonkirjat tieteellisessä tutkimuksessa	16
3. Kirkonkirja-aineiston tallentamiseen liittyviä erityispiirteitä	21
3.1 Nimien normalisointi	23
3.1.1 Kielen fonetiikkaan perustuvat menetelmät	24
3.1.2 Kirjainten aseman tutkimiseen perustuvat menetelmät	30
3.2 Linkitys	33
3.2.1 Linkitysmuuttujat	35
3.2.2 Linkityksen lähestymistavat	37
3.2.3 Bouchardin linkitysmenetelmä	40
3.2.4 Demografiska Databasenin linkitysmenetelmä	43
4. Historiallisiin väestöluetteluihin perustuvia tietokantoja	47
5. Kirkonkirjoihin perustuvia relaatiotietokantoja	53
5.1 Relatiotietokantamäärittelyt	54
5.2 Tietokantojen tavoitteet	56
5.3 Järjestelmän kehitystyön vaiheet ja lähdeaineiston tietojen tallennus	58
5.4 Lähdeaineiston tietosisältö	63
5.5 Tietokantojen käytön tehostaminen	68
5.5.1 Demografiska Databasenin väestötietokannan taulukot	69
5.5.2 Karjala-tietokannan käyttäjäjärjestelmän taulukot	78
5.5.3 Gillespie-järjestelmän käyttäjille kehitetyt taulukot	80
5.6 Hakujärjestelmät	83
5.6.1 Demografiska Databasenin lähdetiedostojen hakujärjestelmä	83
5.6.2 Demografiska Databasenin väestötietokannan hakujärjestelmä	85
5.6.3 Karjala-tietokannan hakujärjestelmä	87
5.6.4 Gillespie-järjestelmän hakujärjestelmä	90
6. Yhteenveto	94
Lähteet	100
Liitteet	104

1 JOHDANTO

Historian tutkimuksen lähdeaineistoon kuuluu väestöluetteloita, joihin on kirjattu laajoihin väestönosiin kohdistettuja hallinnollisia toimenpiteitä ja jotka sen tähden sisältävät runsaasti merkintöjä. Väestöluetteloiden arvo historiallisina lähteinä on merkittävä, sillä ne antavat tietoa tavallisesta kansasta, josta muuten on säilynyt niukasti lähdeaineistoa.¹ Luetteloiden hyödyntäminen tutkimuksessa on työlästä ja aikaa vievää. Tietotekniikan kehittymisen mukana syntyivät yritysten liiketaloudelliset tietokannat. Ne herättivät kiinnostusta eri tieteen alojen tutkijoiden keskuudessa. Myös historiantutkijat alkoivat suunnitella väestöluetteloiden merkintöjen siirtämistä koneella luettavaan muotoon, tietokannoiksi, joiden tietoja useat tutkijat voisivat hyödyntää omissa tutkimuksissaan.² Ensimmäiset tuloksiin johtaneet projektit henkilö-historiallisten tietokantojen rakentamiseksi alkoivat 1970-luvulla.

Henkilöhistorian tietokantoja käsittelevä kirjallisuus koostuu pääasiassa tietokantojen rakentamista varten muodostettujen projektien kuvauksista. Tällaisia projekteja tarkastellaan esimerkiksi seuraavissa tämän tutkimuksen lähteissä Attack, Bateman, Gregson 1992, Bouchard 1986, Bouchard 1992, Danell 1985, Fogelvik, Sperlings 1982, Gutmann 1976, Gutmann, Fliess, Holmes, Fairchild, Teas 1989, Karjalan väestötieteellisen tietokannan loppuraportti 1989, Leboutte, Alter, Gutmann 1987, Lemming 1983, Schofield 1992, Smets 1987, Stenflo, Sundin 1987 ja Winchester 1992. Keskeisimpiä tarkastelun kohteita näissä kuvauksissa ovat samaan menneisyyden henkilöön liittyvien tietojen yhdistäminen eli linkitys ja nimien oikeinkirjoitusasun yhtenäistäminen eli nimien normalisointi. Linkitystä tarkastellaan lähteissä Bouchard 1986, Bouchard 1992, De Brou, Olsen 1986,

¹ Thaller 1991

² Winchester 1992

Danell 1985, Guth 1976, Schofield 1992 ja Winchester 1992. Näissä lähteissä esitellään erilaisia menetelmiä, joilla pyritään varmistamaan, että historiallisista lähteistä löydetty tiedot liittyvät samaan menneisyyden henkilöön. Lisäksi niissä on esitetty käsityksiä siitä, mitkä alkuperäislähteiden tiedot parhaiten soveltuvat henkilön tietojen yhdistämiseen. Kirjallisuudessa on myös tarkasteltu kahta toisistaan poikkeavaa linkityksen lähestymistapaa: onko lähdeaineistosta muodostettava mahdollisimman suuri määrä linkkejä eli samaan henkilöön liittyvien tietojen yhdistelmiä, jolloin väärin linkkien mahdollisuus lisääntyy vai onko pyrittävä mahdollisimman suureen varmojen linkkien määrään, jolloin taas osa lähdeaineiston tiedoista jää käyttämättä.

Nimien normalisointiin on kehitetty sekä kielen fonetiikkaan että nimien kirjainten asemaa tarkastelevia menetelmiä, joita käsitellään seuraavissa tämän tutkimuksen lähteissä Atask, Bateman, Gregson 1992, Bouchard 1986, Bouchard 1992, Guth 1976, De Brou, Olsen 1992 ja Sippu 1985. Siirrettäessä historiallista tietoa tietokoneella luettavaan muotoon kohdataan ongelmia, joita yleisellä tasolla käsitellään lähteissä Thaller 1987 ja Thaller 1991. Erityisesti kirkonkirjojen osalta näitä ongelmia tarkastellaan lähteissä Danell 1985 ja Karjalan väestötieteellisen tietokannan loppuraportti 1989.

Henkilöhistorian tietokantoja rakentavien projektien kuvauksissa on esitelty erilaisia tallennuksessa käytettyjä ohjelmistoja; niiden etuja ja haittoja. Projektit on toteutettu etupäässä relaatiotietokannan hallintajärjestelmiä käyttäen kuten käy ilmi seuraavista lähteistä Atask, Bateman, Gregson 1992, Bouchard 1986, Bouchard 1992, Danell 1985, Fogelvik, Sperlings 1982, Guth 1976, Gutmann, Fliess, Holmes, Fairchild, Teas 1989, Karjalan väestötieteellisen tietokannan loppuraportti 1989, Lemming 1983, Schofield 1992 ja Winchester 1992. Henkilöhistorian tietokantojen rakentamisessa on kuitenkin alettu kiinnostua myös muista tietokannan hallintajärjestelmistä. Lähteissä Thaller 1987 ja Thaller 1991 on esitelty semanttiseen verkkoon ja lähteessä Price, Gray 1994 oliosuuntautuneeseen tietokannan hallintajärjestelmään perustuva tietokanta. Tässä tutkimuksessa

ei ole yksityiskohtaisesti tarkasteltu muita kuin relaatiotietokannan hallintajärjestelmiin perustuvia henkilöhistorian tietokantoja.

Tämän tutkimuksen tavoitteena on antaa yleiskuva henkilöhistorian tietokantoihin tallennetuista tiedoista. Siksi tutkimuksessa esitellään erilaiset historialliset väestöluettelot ja niiden tietosisältöä. Yksityiskohtaisesti perehdytään yhden historiallisen väestöluettelolajin nimittäin kirkonkirjojen rakenteeseen ja tietosisältöön. Tähän valintaan on päädytty, koska eri väestöluettelot rakenteeltaan muistuttavat toisiaan; ne ovat luetteloita, joihin henkilönimiä käyttäen on kirjattu erilaisia viranomaisten toimenpiteitä. Tietosisällöltään ne eroavat alkuperäisen tehtävänsä mukaisesti. Yhtenä syynä on myös se, että kirkonkirjat useissa henkilöhistoriallisissa tietokannoissa muodostavat rakentamisen lähtökohdan. Lyhyesti esitellään kirkonkirjojen tietoja käyttävää tutkimusta ja kirkonkirjojen käytön menetelmiä. Historiantutkijoille on tärkeää, että eri väestöluetteloihin kirjatut, samaa henkilöä koskevat tiedot voidaan löytää ja rakentaa henkilöstä yhtenäinen elämätkuva. Tutkimuksessa kuvataan ja arvioidaan menetelmiä, joilla eri lähteiden samaan henkilöön liittyvät tiedot yhdistetään ja myös menetelmiä, joilla nimien kirjoitusasuja yhtenäistetään.

Tutkimuksessa annetaan lyhyt selonteko eräistä henkilöhistorian tietokannoista. Tietokantoihin on tallennettu erilaisten historiallisten väestöluetteloiden merkintöjä, eikä niitä ole kaikkia ole toteutettu relaatiotietokantoina. Selonteon tarkoituksena on tuoda esiin kiinnostus, jota tutkijoiden keskuudessa tunnetaan automaattista tietojenkäsittelyä kohtaan tutkimuksen apuvälineenä. Lisäksi halutaan antaa yksityiskohtaisempi kuva henkilöhistoriallisten tietokantojen päämääristä, tietosisällöstä sekä tiedon haku- ja käsittelymahdollisuuksista. Siksi tarkastellaan ja verrataan keskenään kolmea relaatiotietokantaa (Demografiska Databasen, Karjala-tietokanta ja Gillespie-järjestelmä). Näihin tietokantoihin on tallennettu pääasiassa kirkonkirjojen merkintöjä. Kirkonkirjoja lähteinään käyttäville tutkijoille on tärkeää, että tietomalli pystyy välittämään lähteiden tiedot oikeina ja täydellisinä sekä säilyttämään mahdollisimman paljon niiden

alkuperäisestä todistusvoimasta. Esimerkkietokantojen valinta johtuu osittain yksityiskohtia selvittävän kirjallisuuden puuttumisesta, mutta toisaalta tietokannat ovat myös edustavia. Ruotsalaisen Demografiska Databasenin rakentaminen aloitettiin 1970-luvun alkupuolella ja se on pitkäikäisin näistä järjestelmistä.³ Muut kaksi suomalaista Karjala-tietokanta⁴ ja amerikkalainen Gillespie-järjestelmä⁵ on rakennettu 1980-luvun lopulla. Näiden esimerkkien avulla selvitetään henkilöhistoriallisten tietokantojen tavoitteita sekä suunnittelu- ja rakentamistapojen tehtäviä. Lisäksi tarkastellaan, missä muodossa lähteiden tieto siirretään tietokoneella luettavaan muotoon ja minkälaisia rakenteita tallennuksessa on käytetty. Tallennukseen käytetyt tietueet ovat laajoja ja ne sisältävät lähteiden merkinnät paljolti samassa muodossa kuin alkuperäisissä asiakirjoissa. Tehokkaan ja joustavan käytön takaamiseksi on jokaisessa esimerkkietokannassa jatkettu kehitystyötä muokkaamalla lähdetietueista käytön kannalta edullisempia rakenteita ja hakuja tukevat hakujärjestelmät.

Luku 2 esittelee kirkonkirjojen syntyhistoriaa ja tietosisältöä. Lisäksi siinä selvitetään kirkonkirjojen käyttömenetelmiä ja tieteenaloja, jotka hyödyntävät kirkonkirjojen tietoja tutkimuksessaan. Luku 3 käsittelee kirkonkirjojen tallentamiseen liittyviä erityispiirteitä. Tarkemmin siinä esitellään ja arvioidaan nimien kirjoitusasun yhtenäistämiseen sekä samaan henkilöön liittyvien tietojen yhdistämiseen kehitettyjä menetelmiä. Luvussa 4 esitellään yleisesti eräitä väestöluetteloihin perustuvia tietokantoja. Luvussa 5 keskitytään yhdenmukaisesti kuvaamaan ja vertaamaan kolmea esimerkkietokantaa (Demografiska Databasen, Karjala-tietokanta ja Gillespie-järjestelmä). Siinä tarkastellaan näiden tietokantojen tavoitteita, tietosisältöä, lähdemerkintöjen siirtoa tietokoneelle sekä hakujen tehostamiseksi muokattuja relaattorakenteita ja hakujärjestelmiä.

³ Danell 1985

⁴ Karjalan väestötieteellisen tietokannan loppuraportti 1989

⁵ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

2 HENKILÖHISTORIAN LÄHTEET

Henkilöhistoriallinen tutkimus määrittellään historiatieteessä yksityisen henkilön tai henkilöryhmien elämän yleispiirteiseksi tai yksityiskohtaiseksi selvitykseksi.⁶ Tutkimuksen päälähteitä ovat historialliset väestöluettelot, jotka syntyivät toteuttamaan oman aikakautensa hallinnollisia tavoitteita.⁷ Väestöluettelot sisältävät merkintöjä henkilöistä, joihin valtiovallan tai kirkon edustajat virkansa puolesta kohdistivat erilaisia toimenpiteitä. Maallista hallintoa palvelivat esimerkiksi verotusta, väestölaskentaa, vaaleja, asepalvelua ja poliisitointa koskevat henkilöluettelot. Ruotsi-Suomessa saivat 1500-luvulla valtionhallinnon tehostamisen seurauksena alkunsa veronkannon asiakirjasarjat: voudintilit ja läänintilit.⁸ Vouti oli maaherran alainen virkamies, jonka tehtävänä oli verojen kokoaminen alueeltaan. Tilikirjoissa oli merkintä siitä, miten paljon veroja kunkin talouden tai henkilön tuli maksaa ja myös kerätyt määrät. Uskonnolliset yhdyskunnat alkoivat varhain luetteloida seurakunnan jäsenille suoritettuja kirkollisia toimituksia kuten kasteita, vihkimisiä ja hautaamisia.⁹ Vanhimmat säilyneet kasteluettelot ovat Italiasta 1300-luvulta. Myöhemmin niitä tavataan eri puolilta Eurooppaa. Katolinen kirkko velvoitettiin pitämään luetteloita kasteista ja vihityistä 1500-luvun lopulla. Englannissa ovat luettelot muutamien katkoksin säilyneet 1610-luvulta pääosassa seurakuntia ja osassa jopa 1560-luvulta.¹⁰ Ruotsi-Suomessa kirkolliset luettelot eli kirkonkirjat yleistyivät 1600-luvulla.¹¹ Historiallisten väestöluetteloiden avulla on mahdollista seurata yksittä-

⁶ Hyppönen, Luttinen, 1988, s. 18

⁷ Karskela 1987, s. 26

⁸ Hyppönen, Luttinen 1988, s. 132-136

⁹ Karskela 1987, s. 73

¹⁰ Schürer 1987

¹¹ Karskela 1987, s. 28-29, Hyppönen, Luttinen 1988, s. 18-21

sen ihmisen elämän kulkua syntymästä kuolemaan, mutta niiden henkilötiedot rajoittuvat parhaimmillaankin vain suppeisiin tilastollisiin taustatietoihin. Siksi henkilöhistoriallinen tutkimus käyttää apuna myös muuta elämäkerrallista asiakirja-aineistoa: tuomio-, lainhuudatus- ja perukirjoja sekä esimerkiksi muistelmia, päiväkirjoja ja kirjeitä.

Tietotekniikan kehitys on lisännyt historiallisten väestöluetteloiden hyödyntämistä tutkimuksessa.¹² Aikaisemmin laiminlyötyjen, runsaasti lähdemerkintöjä sisältävien, historiallisten väestöluetteloiden tietoja voidaan tietotekniikan avulla yhdistää laajoiksi, runsaasti tapauksia sisältäviksi henkilöhistoriallisiksi tietokannoiksi. Tässä työssä tarkastellaan yksityiskohtaisesti kirkonkirjojen merkinnöistä rakennettuja tietokantoja, koska usein laajojen, historialliseen lähdeaineistoon perustuvien tietokantojen rakentaminen aloitetaan juuri näistä luetteloista. Aluksi esitellään kirkonkirjojen tietosisältöä, jotta olisi mahdollista ymmärtää esiteltävien henkilöhistoriallisten tietokantojen rakenteellisia ratkaisuja.

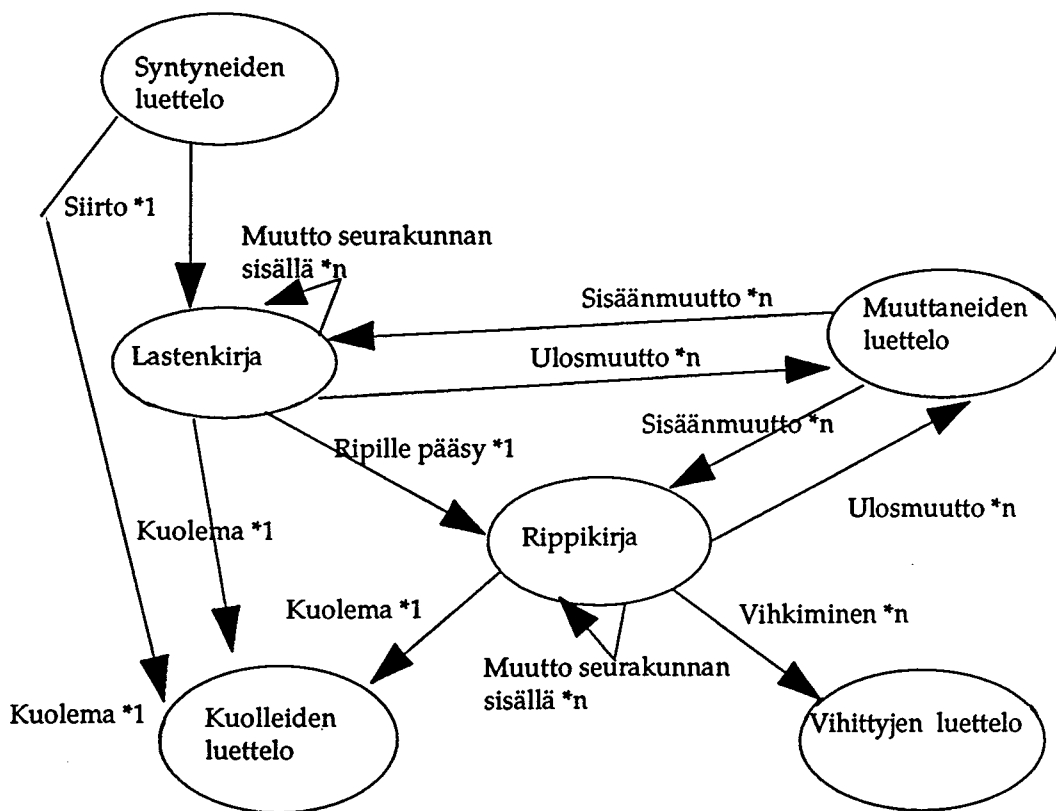
2.1 Kirkonkirjat

Kirkonkirjat koostuvat historiakirjoista ja rippikirjoista. Historiakirjoja ovat syntyneiden ja kastettujen, kuulutettujen ja vihittyjen sekä kuolleiden ja haudattujen luettelot. Myöhemmin kehittyivät vielä muuttaneiden luettelot. Syntyneiden ja kastettujen, vihittyjen sekä kuolleiden ja haudattujen luettelojen tiedot eivät paljon eroa toisistaan eri maissa. Suomalaisia ja ruotsalaisia rippikirjoja vastaavia asiakirjoja ei löydy muualta.¹³ Historian tutkimuksen kannalta on tärkeää, että seurakuntien papisto on merkinnyt henkilöt perhekunnittain ja siten antaa valmiina nämä tiedot, jotka muiden maiden väestöluetteloista on etsittävä

¹² Best 1991, Genet, Zampolli 1992

¹³ Kälvemarm 1977, Winchester 1992

erilaisia menetelmiä käyttäen. Seurakuntien luettelot sidottiin arkistointia varten kirjojen muotoon. Kuva 1 havainnollistaa henkilöiden merkitsemistä eri kirkonkirjoihin, merkintöjen syytä ja maksimilukumäärää.¹⁴ Ellipsit kuvaavat eri kirjoja, nuolet niiden välisiä yhteyksiä ja merkinnän aiheuttavaa tapahtumaa. Merkintä on tehty sekä nuolen alku- että loppupäässä olevaan kirkonkirjaan. Nuolen vieressä oleva numero ilmoittaa, kuinka monta kertaa kyseinen tapahtuma voi henkilölle sattua ja aiheuttaa merkinnän kirkonkirjaan.



Kuva 1 Kirkonkirjat, niiden väliset yhteydet ja suhteet

¹⁴ Rouhiainen 1983

2.1.1 Historiakirjat

Historiakirjat eli kronologiset luettelot seurakunnassa syntyneistä ja kastetuista, kuulutetuista ja vihityistä, kuolleista ja haudatuista olivat vanhimpia kirkollisia luetteloita. Myöhemmin alettiin kirjata myös seurakuntaan ja sieltä pois muuttaneita. Historiakirjat yleistyivät Suomessa 1600-luvun puolivälistä lähtien, sillä vuoden 1686 kirkkolaki määräsi seurakuntien papit pitämään näitä luetteloita ja mitä niihin tuli merkitä. Alunperin historiakirjat oli tarkoitettu vain seurakunnan sisäiseen käyttöön, mutta 1700-luvun puolivälistä lähtien alettiin Ruotsi-Suomessa järjestelmällisesti kerätä tilastotietoja maan väestöstä ja silloin seurakuntien papisto sai tehtäväkseen laatia historiakirjojen pohjalta tilastoja syntyneistä, vihityistä ja kuolleista sekä myöhemmin myös muuttaneista ja rokotetuista. Tämä tehtävä johti kirkollisen väestökirjanpidon täsmentymiseen.¹⁵

Historiakirjat sisälsivät tapahtuman vuosittaisen järjestysnumeron sekä välttämättömät nimi-, päivämäärä- ja asuinpaikkatiedot. Muuttaneiden luetteloissa ilmoitettiin lähtöpaikkakunnan ja tulopaikan lisäksi joskus vain perheen päämies, toisinaan lueteltiin koko talous palkollisia myöten. Syntyneiden ja kastettujen luetteloihin merkittiin ainakin myöhemmin vanhempien ikä ja kummit sekä tieto siitä, oliko kysymyksessä avioliitossa tai avioliiton ulkopuolella syntynyt lapsi.¹⁶ Liite 1 esittää Korpilahden seurakunnan syntyneiden ja kastettujen luettelon sivua vuodelta 1791. Luettelon pääotsakkeena on latinankielinen sana *Nati* eli syntyneet ja vuosiluku, jota kirjaukset koskevat. Syntyneet lapset on merkitty kuukausittain. Kuukauden latinankielinen nimi on merkitty sarakkeiden yläpuolelle otsikoksi. Sarakkeen 1 juokseva numero ilmoittaa seurakunnassa tuona vuonna syntyneiden järjestysnumeron. Sarakkeeseen 2 on merkitty syntymän ja sarakkeeseen 3 kasteen päivämäärä. Sarakkeessa 4 on lapselle

¹⁵ Karskela 1987, s. 29-30, 73, Kälvemark 1977

¹⁶ Karskela 1987, s. 73-80, Hyppönen, Luttinen 1988, 84-85

annetut nimet ja sarakkeissa 5 ja 6 isän ja äidin nimet sekä yleensä vain isän ammatti. Isän nimen alapuolella ovat miespuolisten ja äidin nimen alla naispuolisten kummien nimet. Sarake 7 ilmoittaa vanhempien kotipaikan. Esimerkiksi tapahtumanumero 63 kertoo, että syyskuun 3. päivänä on syntynyt itsellinen Anders Pählinpojalle ja Hedwig Andersintyttärelle poika, joka on kastettu syyskuun 4. päivänä ja saanut nimekseen Maths. Kummeja oli neljä, kaksi miestä ja kaksi naista. Vanhempien kotipaikka oli Painaja.

Kuolleiden luetteloihin merkittiin kuolleen henkilön tai hänen vanhempiansa sääty tai ammatti sekä kuolinsyy. Kuolinsyyntä määritteli tavallisesti pappi kuolleen omaisten antamien tietojen perusteella. Liitteessä 2 on Korpilahden kuolleiden luettelon sivu vuodelta 1791. Pääotsikkona on latinankielinen sana *Denati* eli kuolleet. Kuten syntyneiden luettelossakin kuukausien latinankieliset nimet on merkitty otsikoiksi. Ensimmäisessä sarakkeessa on vuoden aikana kuolleiden seurakuntalaisten järjestysnumero, seuraavissa sarakkeissa kuolleen henkilön kotipaikka, ammatti, siviilisääty ja nimi. Jos kysymyksessä on lapsi, ilmoitetaan myös isän nimi ja ammatti. Seuraavissa sarakkeissa ovat kuolintapahtuman päivämäärä, hautauspäivä, kuolleen henkilön ikä sekä kuolinsyy. Tapahtumanumerossa 62 on tieto, että Painajassa on itsellinen Anders Pählinpojan lapsi Maths kuollut syyskuun 15. päivänä ja haudattu 18. päivänä. Kuollessaan hän oli 14 päivän ikäinen ja kuolinsyyksi on ilmoitettu tuntematon sairaus.

2.1.2 Rippikirjat

Seurakuntien papit pitivät tarpeellisena merkitä muistiin seurakuntalaisten kristinopin tietoja. Rippikirjat muotoutuivat vähitellen kahdesta erilaisesta kirkollisten toimitusten kirjanpidosta: kuulusteltavista ja ehtoolliselle tulijoista. Ensimmäinen maininta rippikirjojen pidosta Suomessa on vuodelta 1666. Vuoden 1673 kirkkojärjestys ja vuoden 1686 kirkkolaki antoivat tarkemmat määräykset rippikirjojen pidosta ja säätivät sen pakolliseksi. Papit tutkivat lukutaitoa

kyläkunnittain lukukinkereillä, joiden yhteydessä merkittiin muistiin kuulusteluista selviytyneet ja heidän taitonsa. Aluksi vaadittiin, että piti osata ulkoa katekismuksen kuusi pääkappaletta, aamu-, ilta- ja pöytärukoukset sekä synnintunnustus. Myöhemmin alettiin vaatia myös Lutherin selitysten sekä raamatunlauseiden taitamista, mitä ei enää ulkoluvulla voitu hallita, vaan niiden osaaminen edellytti sisälukutaitoa. Suoritussarakkeiden taito- eli lukumerkit olivat aluksi kirjavia. 1700-luvulla vakiintui merkiksi rasti, jonka hyvin osaava sai kokonaisena, kun taas huonommin osaava sai vain osan siitä. Lukutaidoton ei päässyt ehtoolliselle eikä voinut solmia avioliittoa. Ripille pääsyn jälkeen ehtoollisella käynti oli pakollista ja papit tarkkailivat sitä merkitsemällä ehtoolliskäynnit rippikirjaan.¹⁷

Rippikirjoihin merkittiin seurakunnan ripillä käyneet asukkaat tietyinä, yleensä 5-10 vuoden pituisena ajanjaksona. Joihinkin rippikirjasidoksiin on kirjattu myös lapset. Merkinät tehtiin kylittäin. Kylän nimen jälkeen seurasivat kylään kuuluvat talot. Talo tarkoitti erillisten ruokakuntien, talouksien muodostamaa kokonaisuutta eli maakirjataloa. Yhden maakirjatalon alueella on yleensä ollut useita erillisiä, itsenäisiä taloja ja torppia. Talon tai torpan nimen alle merkittiin kaikki tilan alueella asuvat ruokakunnat edustajineen. Ensin mainittiin ruokakunnan päämies ja hänen vaimonsa ja lapsensa, sitten muut sukulaiset ja henkilöt, jotka asuivat päämiehen perheen kanssa samassa ruokakunnassa. Samalle sivulle on yleensä merkitty useita perheitä. Liite 3 esittää Korpilahden seurakunnan rippikirjan aukeamaa.¹⁸

Kuvassa 2 on ensimmäinen sivu Korpilahden seurakunnan rippikirjan aukeamasta vuosilta 1841-1851. Kuva esittää Oittilan lukualueen Tyystilammin talon merkintöjä. Kuten rippikirjoissa yleensäkin sarakkeiden nimet ovat valmiiksi painettuja. Ensimmäiseen sarakkeeseen, *By. Hemmans- och Bond-namn* on

¹⁷ Hyppönen, Luttinen 1988, 65-70, Karskela 1987, s. 85-87

¹⁸ Hyppönen, Luttinen 1988, 65-70, Karskela 1987, s. 83-92

kirjattuna talon alueella asuneiden henkilöiden nimet ja ammatti tai sosiaalinen asema, sekä suhde perheen päämieheen kuten vaimo, poika, tytär tai vävy. Sitten seuraavat henkilöiden syntymäajat. Tämän jälkeen tulevat sarakkeet sisältävät kristinopin taitomerkintöjä, joita on tässä rippikirjassa annettu sisälukutaidosta, Lutherin katekismuksen lukujen ja niiden selitysten sekä pyhien kirjoitusten osaamisesta. Seuraavaksi on ilmoitettu vuodet, joina henkilö on laiminlyönyt kuulusteluun osallistumisen. *Tillfällige anmärkningar* -sarake sisältää erilaisia talon asukkaita koskevia merkintöjä: kuolinajan, tietoja henkilökohtaisista ominaisuuksista esimerkiksi sokeudesta, sekä merkinnät, jotka ilmoittavat miltä rippikirjan sivulta tai mille sivulle henkilön tiedot on siirretty. Koska rippikirja

77. Oittila Läslag G.B. N. 12.

By. Hemmans och Bonde-namn.	Födelse dag och år.	ABC Bok Tonsåkrång.	D. Luth. Catech. med Förklaring.						D. S. S. Befripar.	Försam- mat Cat. Förhör.	Tillfälliges
			1	2	3	4	5	6			
<i>Nick. Sylv. Lamski</i>	<i>1847</i>										
<i>Henriq. Henriq. Jon.</i>	<i>7/1853</i>										
<i>Geo. Eric Thomas</i>	<i>9/1890</i>	x	x	x	x	x	x				
<i>J. Carlström Fredrik Carlström</i>	<i>5/1825</i>	18	19	20	21	22	23	24	<i>1847</i>	<i>1860-42</i>	
<i>Magn. Filips Mattson</i>	<i>1874</i>	x	x	x	x	x				<i>1874/1877</i>	
<i>Geo. Heland Matts</i>	<i>1876</i>	x	x	x	x	x	x			<i>1877/1879</i>	
<i>Mattias Filip</i>	<i>11/1829</i>										
<i>Mattias Filipsson</i>	<i>5/1850</i>	16	x	x	x	x	x		<i>1850-55</i>		
<i>Geo. Hedvig Henriq. P.</i>	<i>7/1802</i>	x	x	x	x	x	x		<i>1850-55</i>		

Kuva 2 Korpilahden seurakunnan rippikirjan aukeaman ensimmäinen sivu

märkningar:	Nattvardsgång.													
	18 41.	18 42.	18 43.	18 44.	18 45.	18 46.	18 47.	18 48.	18 49.	18 50.	18 51.	18 52.		
	11/7	22/5	16/10	11/6		26/7	8/6	5/7	24/7	1/6	1/10	30/7	27/7	29/7
	11/7	22/5	16/10	11/6		26/7	8/6	5/7	24/7	1/6	1/10	30/7	27/7	29/7
17 2. 7. 6. 1. 6.						22/7				1/6				
			11/6											
	11/7	22/5	16/10	11/6		26/7	8/6	1/7		1/6	1/10	30/7		
	11/7	22/5	16/10	11/6	12/5	26/7	8/6			1/6	1/10	30/7	27/7	
	11/7	22/5	16/10	11/6	12/5	26/7	8/6			1/6	1/10	30/7	27/7	29/7

Kuva 3 Korpilahden seurakunnan rippikirjan aukeaman toinen sivu

kattaa kymmenen vuotta, on siirtoja, etenkin itsellisten osalta, tapahtunut useita ja silloin henkilön nimi on yliviivattu. Aukeaman toisella sivulla, jota kuva 3 esittää, ovat tiedot henkilöiden vuosittaisista ehtoollisella käynneistä.

Rippikirjojen väestötilastolliset taustatiedot paranivat, kun 1700-luvulta lähtien niihin alettiin merkitä kunkin henkilön syntymäaika ja -paikka, kuolinaika ja -paikka, muutto seurakunnan sisällä, seurakuntaan tai sieltä pois sekä sosiaalinen asema tai ammatti. Myöhemmin rippikirjoihin lisättiin rokotusmerkinnät sekä huomautussarakkeet, joihin merkittiin tietoja seurakuntalaisten tekemisistä, lähinnä erilaisista rikkeistä. Tämän monipuolisen tietosisällön vuoksi rippikirjoista muodostui väestökirjanpitojärjestelmän peruselementti.¹⁹

¹⁹ Hyppönen, Luttinen 1988, s. 66-70, Kälvemärk 1977

Rippikirjojen loppusivuille on joissakin seurakunnissa merkitty vuosittain kaikki rippikoululaiset ja arvosteltu heidän kristinopin- ja lukutaitonsa. Useissa Itä-Suomen seurakunnissa oli käytäntönä, että alle rippikouluikäiset lapset vietiin omiin rippikirjoihinsa eli lastenkirjoihin, jollaisesta on esimerkki liitteessä 4. Niiden rakenne ja sisältö on periaatteessa samanlainen kuin rippikirjojen. Kuu-
lusteluista suoriutuneet siirrettiin varsinaisiin rippikirjoihin. Liite 4 esittää Suurlahden kylän lapsia. Sarakkeet ovat samat kuin rippikirjoissa, mutta ehtoolliskäyntimerkinnät luonnollisesti puuttuvat. Lisäksi lastenkirjoissa on sarake, jossa ilmoitetaan ripille pääsystä. Sivun yläosassa on kirjattuina Ruolahden talon isännän Adam Johanssonin lapsiin liittyviä merkintöjä. Vanhin lapsista Hedda Lisa on syntynyt 10.10.1796, hallitsee erinomaisesti sisäluvun, ABC-kirjan sekä Lutherin katekismuksen selityksineen ja on päässyt ripille 26.6.1812.

2.2 Kirkonkirjojen käyttömenetelmiä

Historialliset väestöluettelot syntyivät palvelemaan oman aikansa tarpeita ja niiden tietosisältö ja rakenne muokkautui sen mukaisesti. Kirkonkirjoihin kirjattiin ihmisten syntymät, avioliitot, kuolemat sekä Ruotsissa ja Suomessa myös ehtoollisella käynnit ja kristinopin tiedot. Eri alojen tutkimus voi hyödyntää kirkonkirjojen tietoja ja soveltaa niihin alalleen ominaisia menetelmiä, joista tässä luvussa esitetään esimerkkejä.²⁰

Tavanomainen sukututkimus kohdistuu yhteen sukuun, jonka jäsenten kohtaloita seurataan niin pitkälle taaksepäin kuin historiallisissa asiakirjoissa heistä löytyy merkintöjä. Sukuun kuuluneista henkilöistä selvitetään paitsi nimi ja elinaika myös ammatti, paikkakunta ja asuinpaikan muutokset. Sukututkimuksen keskeisintä lähdeaineistoa ovat seurakuntien kirkonkirjat, joista monipuolisimpia ja

²⁰ Karskela 1987, s. 26

eniten tietoa sisältäviä ovat rippikirjat. Yleensä sukututkimus aloitetaan käymällä läpi esivanhempien asuinseurakunnan rippikirjoja. Niissä henkilöt on kirjattu talouksittain siten, että ensin mainitaan perheen päämies, sitten hänen vaimonsa ja lapsensa. Näillä tiedoilla on ratkaiseva merkitys ruokakuntia ja perhekokonaisuuksia muodostettaessa. Ensin on käytännöllisintä poimia avioliitot vihki- ja kuulutusluetteloista, verrata niiden ja rippikirjojen tietoja ja alkaa sen jälkeen seurata perheeseen kuuluvien henkilöiden merkintöjä rippikirjoista. Perheen jäsenten syntymä- ja kuolinajat, aviopuolisoiden vihkimisajat ja muiden tapahtumien päivämäärät tarkistetaan historiakirjoista. Vaikkakin syntymäajat annetaan rippikirjoissa, on usein tarpeellista tutkia syntyneiden ja kuolleiden luetteloita, koska niissä saattaa olla tietoja kuolleina syntyneistä tai lapsista, jotka kuolivat ennen kuin heistä tehtiin merkintä rippikirjoihin. Jos seurakunnassa on pidetty kronologisia muuttaneiden luetteloita, ne ovat hyödyllisiä paikallistettaessa rippikirjojen henkilöitä, sillä niistä saadaan selville, mistä muuttanut henkilö saattaisi löytyä. Näin jatketaan ajassa taaksepäin niin kauan kun tietoja seurakunnan kirjoista löytyy. Puuttuva rippikirjoilla käynti voi merkitä tilapäistä poissaoloa, matkaa, vankilaa, muuttoa tai osallistumiskieltoa. Silloin on tutkittava muuttaneiden luetteloita tai muita henkilöhistorian lähteitä. Viimeinen merkintä saattaa myös olla viimeinen elinvuosi. Jos perhe katoaa seurakunnan kirjoista, sitä etsitään lähipitäjien kirkonkirjoista joko syntymäaikoja tai nimiä seuraamalla. Elinolojen ja aikakauden ominaispiirteiden liittäminen tutkimukseen antaa sille tieteellistä arvoa.²¹

Kirkollisten luetteloiden tietoihin voidaan soveltaa myös toisenlaista menetelmää. Aluksi yhdistetään kaikki kirkonkirjojen samaan henkilöön liittyvät tiedot. Tämän jälkeen tutkimustavoitteet määräävät minkälaisia kokoonpanoja, perheitä vai muunlaisia ryhmiä, yksittäisistä henkilöistä muodostetaan. Sosiaalhistoria ja perhedemografia ovat kaksi historian tutkimuksen läheistä alalajia, joita kiinnostavat yhteiskunnalliset yhteisöt, niiden rakenteet ja kehitys. Sosiaalhistorian

²¹ Karskela 1987, s. 83-92, Hyppönen, Luttinen 1988, s. 18-26, 61-90

ja historiallisen perhedemografian tutkimuksissa muodostetaan perherakenteita, joissa aviopuolisot liitetään toisiinsa ja lapset avioliittoon, josta ovat syntyneet. Perherakenteet käsittävät tutkittavan alueen ja ajanjakson koko väestön ja niitä kutsutaankin kollektiivisiksi sukututkimuksiksi tai elämäkertoiksi. Kollektiivisia elämäkertoja täydennetään yhdistämällä jokaiseen yksilöön kaikki merkinnät, joita hänestä löytyy veroluetteloissa, oikeusistuinten asiakirjoissa tai peruluetteloissa. Täydennetyt kollektiiviset elämäkerrat ovat erittäin tärkeitä sosiaalishistoriassa talousrakenteen, sukulaisuusverkostojen ja muuttuvan perhe-elämän kehityssuuntien ja eroavaisuuksien tutkimisessa.²²

Tutkija voi yhdistää henkilöt toisenlaisiksi ryhmiksi, kohorteiksi. Kohortti määritellään joukoksi, joka on kokenut jonkin tapahtuman saman ajanjakson kuluessa. Tällainen tapahtuma voi olla esimerkiksi sama syntymäaika ja -paikka. Kohorttien avulla voidaan pitkittäisanalyysiä käyttäen tutkia erilaisia sosiaalisia, taloudellisia ja väestötieteellisiä ilmiöitä niihin kuuluneiden ihmisten elämän koostumuksen perusteella. Esimerkkinä kohortin käytöstä voidaan mainita ruotsalainen tutkimus, jossa vertailtiin paikkakunnalta muuttaneiden ja siellä pysyneiden henkilöiden avioliittoastetta, lasten lukumäärää ja sosiaalisen aseman muutoksia. Tutkimus käsitti 188 miehestä ja 232 naisesta koostuvan kohortin. Nämä henkilöt asuivat 19-20 -vuotiaina eteläruotsalaisessa pikkukaupungissa ja heitä seurattiin tuosta iästä 34-35 vuoden ikään. Esimerkiksi 34-35 -vuotiaina 94 miestä oli muuttanut paikkakunnalta ja heistä 64 % oli solminut avioliiton, kun taas niistä 94:stä, jotka olivat pysyneet kaupungissa 53 % oli solminut avioliiton. Naisista 144 oli muuttanut ja 72 % heistä oli avioliitossa, kun taas kaupunkiin jääneistä vain 56 % oli naimisissa.²³

Historiakirjoja voidaan käyttää myös tilastoina, sillä ne antavat arvokkaan perustan arvioitaessa syntyvyyden, kuolleisuuden, avioliittojen ja muuttaneiden

²² Thaller 1991, Best 1991

²³ Kälvemark 1977, Kälvemark, Roger 1980, Best 1991

määrää ja astetta. Tässä lähestymistavassa ei ole tarpeen yhdistää kaikkia samaan henkilöön liittyviä tietoja, vaan silloin tarkastellaan pelkästään tapahtumien määrää tutkimuksen kohteena olevalla alueella.²⁴

2.3 Kirkonkirjat tieteellisessä tutkimuksessa

Historian tutkimukselle ovat kirkonkirjat itsestään selvää ja luonnollista lähdeaineistoa. Niiden avulla saadaan tietoa menneisyydessä eläneistä henkilöistä, mutta myös yhteiskunnasta ja sen muuttumisesta, kehittymisestä ja taloudellisista rakenteista. Etenkin historiantutkimuksen alalajit, sukututkimus, perhedemografia ja sosiaalhistoria käyttävät kirkonkirjoja peruslähteenään.²⁵

Sukututkimus, genealogia, on historiatiede, jossa arkistolähteitä käyttäen laaditaan luetteloita sukulaisuuden perusteella samaan ryhmään kuuluvista henkilöistä. Varsinainen elämäkerrallinen tutkimus on Suomessa rajoittunut lähes yksinomaan suurmiehiin ja sen ulkopuolelle ovat jääneet tavalliset ihmiset, joiden elämäkertojen selvittelyssä sukututkijalle tarjoutuu tärkeä tutkimustehtävä. Sukututkimus aloitetaan yleensä kirkonkirjoista ja sitä täydennetään muilla historiallisilla väestöluetteloilla, jotka käsittävät kaikki kansalaisryhmät. Väestöluetteloiden apuna käytetään muita elämäkerrallisia asiakirjoja. Sukututkija erittelee keräämänsä tiedot, minkä jälkeen hänen on mahdollista jatkaa syvällisempää analyysiä kuten henkilön tekojen motiivien selvittämistä ja sijoittamista osaksi ympäristöään. Sukuluetteloiden täydentäminen elämäkerrallisilla tiedoilla antaa tutkimuksellista kantavuutta ja merkitsee siirtymistä kohti henkilöhistoriallisia elämäkertoja. Eri tieteenalat käyttävät sukututkimuksen tuloksia sekä kirkonkirjojen yksityiskohdiltaan runsasta aineistoa tärkeänä lähdemateriaalina, kuitenkin kunkin alan omien tarpeiden ja metodien muokkaamana. Geneettinen

²⁴ Karskela 1987, s. 83-92

²⁵ Danell 1985, s. 181-192, Thaller 1991

lääketiede käyttää kirkonkirjojen kuolinsyytietoja, jotka ulottuvat 1700-luvulle asti, suomalaisille ominaisten perinnöllisten sairauksien jäljittämiseen. Oikeustiede hyötyy sukututkimuksista selvittäessään maanomistus-, perimys-, avioliitto- ja rangaistuskäytännön kehittymistä.²⁶

Historian alalajit, perhedemografia ja sosiaalhistoria, ovat läheisiä tieteenaloja. Erona on, että perhedemografia korostaa sukulaisuutta ja sosiaalhistoria perheiden yhteisasumista. Perhedemografia tutkii perheiden ja sukulaisryhmien kokoa ja rakennetta sekä perherakenteen muutoksia ja niiden kehityssuuntaa ja analysoi demografisten tekijöiden vaikutusta näihin. Demografisia tekijöitä ovat esimerkiksi syntymä, kuolema, avioliiton solmiminen sekä perheeseen ja siitä pois muuttaminen. Ranskalainen Louis Henry kehitti demografisten tekijöiden mittaamisen perustaksi perhelehtimenetelmän. Menetelmässä kirkonkirjojen merkinnät kaste-, avioliitto- ja hautaustapahtumista yhdistetään perhetietueiksi ja siirretään lomakkeille. Perherakenteen pohjalta voidaan arvioida esimerkiksi avioliiton solmimisikää, synnytysvälejä ja perhekokoa. Perhelehtimenetelmä on nykyaikaisen historiallisen demografian perusta.²⁷

Perhedemografia käyttää yksilötason mittaamiseen kehitettyjä välineitä kuten perhelehtiä, elämäntauluja ja rakennemalleja kollektiivisen kohteen, sukulaisuuden ja yhteisasumisen sekä niihin vaikuttavien demografisten tekijöiden, tutkimiseen. Rakennemalleja ovat analyttiset mallit, makro- ja mikrosimulaatiot. Vanhimpia ovat analyttiset mallit, joita käytetään arvioitaessa esimerkiksi odotettavissa olevaa sukulaisten määrää, perhekokoa tai aikaa, joka keskimäärin vietetään kussakin asemassa perheessä esimerkiksi lapsena, avioliitossa ja leskenä. Makrosimulaatiot pohjautuvat kollektiivisiin elämäkertoihin ja niiden pohjalla oleviin elämäntauluihin. Menetelmässä väestö jaetaan alaryhmiin eli kohortteihin, iän, sukupuolen, avioaseman ja muiden perheominaisuuksien

²⁶ Hyppönen, Luttinen 1988, s. 21, Danell 1985, s. 181-192

²⁷ King 1990, Ruggles 1990, Schürer 1987

mukaan ja saadut tulokset heijastetaan koko väestöön. Elämätaulut antavat kuvan siirtymisestä perheasemasta toiseen esimerkiksi avioliitosta avioeroon, lapsuudesta kodin jättämiseen ja auttavat siten arvioimaan perheasemaa, sen kestoja ja eri asemien välisiä keskiarvoja. Historiallisen perhedemografian tulevaisuutta on kolmas lähestymistapa, mikrosimulaatio, joka toimii pikemminkin yksilöiden kuin ryhmien tasolla. Mikrosimulaatiomallit luovat hypoteettisen väestön antamalla yksilöille demografisia tapahtumia eli syntyviä, kuolemia ja avioliittoja, ennustettujen todennäköisyyksien pohjalta. Näin mikrosimulaatiota käytetään luomaan elämähistorioita hypoteettiselle väestölle, joka käyttäytyy samoin kuin todellinen väestö.²⁸

Perhedemografian tavoin sosiaalishistorian alaan kuuluva perhehistoria tutkii ihmisten muodostamia yhteisöjä. Perhehistorian tutkijat ovat kiinnostuneet ensisijassa perheistä ja talouksista. Perhe ymmärretään samaan sukuun kuuluvien henkilöiden yhteisasumiseksi; taloudet taas muodostuvat ihmisistä, jotka muodostavat yhteisen ruokakunnan. Historiantutkijat selvittävät talous- ja kulttuuriolosuhteiden vaikutusta perheiden ja talouksien rakenteeseen ja niiden muutoksiin. He pyrkivät vangitsemaan tavallisen kansan mentaliteetin, jossa korostuvat paikalliset sosiaaliset yhteydet ja sukulaisuuden näkeminen arvokkaana voimavarana. Perhehistorian tutkija etenee asettamalla vanhat väitteet uuteen kehykseen seuraavanlaisilla kysymyksillä: "Onko perhe muuttunut suurperheestä ydinperheeksi teollistumisen myötä?" taikka "Johtuiko ydinperheen vallitsevuus demografisista rajoituksista vai ydinperheen asettamisesta etusijalle?".²⁹

Perhehistoria ja sukututkimus liittyvät läheisesti säätykierron ja sosiaalisen liikkuvuuden tutkimiseen, sillä perinteisissä yhteiskunnissa perheet olivat samanaikaisesti sekä tuotantoyksiköitä että näyttelivät ratkaisevaa osaa henkilöiden sosiaalisessa sijoittumisessa. Sosiaalisen liikkuvuuden tutkimiseksi

²⁸ Ruggles 1990

²⁹ King 1990

joudutaan rakentamaan yksilöiden ura ja etsimään sitä varten tietoja hajanaisista lähteistä. Alueellisen muuttoliikkeen syitä, taustaa ja vaikutusta yhteiskuntaan on tutkittu muuttoluetteloiden ja pappien antamien muuttotodistusten avulla.³⁰

Perhedemografian, sosiaalihistorian ja sukutieteen tutkijat ovat hyötäneet tietotekniikan kehityksestä. Uusi tekniikka on tehnyt mahdolliseksi siirtää ja tallentaa lähteiden tietoja koneella luettavaan muotoon. Runsaasti tapauksia sisältävistä väestöluetteloista alettiin rakentaa laajoja historiallisia tietokantoja, joiden tietoja oli mahdollista analysoida joko tilastojen avulla tai kohdistaa niihin yksityiskohtaisia kyselyjä. Yksityisten henkilöiden ja tapahtumien selvittelystä siirryttiin massoja käsittelevän aineiston käyttöön. Historian tutkimus alkoi käsitellä kollektiivisiä ilmiöitä, mikä edellytti kvantitatiivisten menetelmien käyttöönottoa. Kvantitatiiviset menetelmät ovat täydentäneet historiallista kuvaa, sillä niiden avulla on mahdollista testata tutkijan asettamia teorioita ja hypoteeseja korkealla yleistystasolla ja saada esiin yhteyksiä, jotka yksilötasolla eivät ole ilmeisiä. Aluksi näiden menetelmien katsottiin soveltuvan ensisijassa tavallisen kansan tutkimiseen, mutta vähitellen on poliittista osallistumista ja eliittejä koskeva tutkimus tullut tärkeäksi. Esimerkkinä näiden menetelmien käytöstä voidaan mainita Saksassa tehty vertailututkimus, jossa todettiin, että Euroopan parlamenttien jäsenten poliittiseen suuntaukseen 1800-luvun puolivälissä eivät vaikuttaneet ratkaisevasti taloudelliset edut, vaan alueelliset siteet, lähinnä maantieteelliset ja uskonnolliset tekijät. Toinen esimerkki määrällisistä menetelmistä ovat kollektiiviset elämäkerrat.³¹

Tietokoneella toteutetut historialliset tietokannat ovat rikastuttaneet tutkimusta huomattavasti. Tästä on hyvänä esimerkkinä ruotsalaisen henkilöhistoriallisen tietokannan, Demografiska Databasenin, monipuolinen käyttö. Demografiska Databasen eli DDB perustuu eri puolilta Ruotsia valittujen alueiden kirkonkirja-

³⁰ Best 1991, Norberg, Rolén 1980

³¹ Best 1991, Genet, Zampolli 1992

merkintöihin, joita on täydennetty muiden asiakirjojen samaa väestöä koskevilla tiedoilla. Yhtenäisen ja kattavan lähdemateriaalin puuttuminen, mikä on johtunut käytettävien lähteiden heikosta laadusta ja riittämättömästä tietosisällöstä, on aiheuttanut kansainväliselle historiallis-demografiselle tutkimukselle ongelmia.³² DDB on tarjonnut ulkomaisille tutkijoille mahdollisuuden käyttää ajallisesti ja paikallisesti kattavaa ruotsalaista kirkonkirja-aineistoa ja tehdä sen perusteella johtopäätöksiä oman maansa vastaavista ilmiöistä. Uumajan DDB:n aineistoon perustuvat tutkimukset pyrkivät selvittämään teollistumisen sosiaalisia vaikutuksia perheen ja yksilön olosuhteisiin, perhe- ja avioliittomalleihin ja naisen asemaan sekä perheessä että yhteiskunnassa. DDB on avannut uusia mahdollisuuksia alueellisen ja yhteiskunnallisen liikkuvuuden selvittämiseen. Ekologisen näkökannan korostaminen on lisännyt kiinnostusta yrittää ymmärtää toisaalta yksilön ja yksilöryhmien sekä toisaalta yksilön ja ympäristön välisiä tapahtumia. Kirkonkirjojen yksilötiedot muuhun aineistoon yhdistettynä antavat mahdollisuuden tutkia ihmisen vaikutusta ympäristöön sekä yksilön tai yksilöryhmien ja ympäristön välistä suhdetta. DDB:n alueellisia tilastoja voidaan vertailla yleisvaltakunnalliseen tilastoon ja selvittää alueellisia eroja, jotka muuten jäisivät havaitsematta. Esimerkiksi vaikka pienten lasten kuolleisuus osoitti Ruotsissa 1800-luvun alusta tasaista laskua, niin paikalliset erot olivat huomattavat.³³

³² Kälvemarm 1977, History on data 1989, Danell 1985, s. 193-199

³³ Kälvemarm 1977, Danell 1985, s. 181-192, Schürer 1987

3 KIRKONKIRJA-AINEISTON TALLENTAMISEEN LIITTYVIÄ ERITYISPIIRTEITÄ

Kirkonkirjat ovat historiallisia asiakirjoja. Niiden tehtävä oli merkitä muistiin papiston seurakuntalaisiin kohdistamat toimitukset. Historiallisen alkuperänsä vuoksi niiden muokkaamisessa tutkimuksen käyttöön sopiviksi tietokannoiksi on joitakin erityisesti niille ominaisia piirteitä, mistä johtuu, että suunnittelu edellyttää tarkkaa perehtymistä lähdemateriaaliin.³⁴ Seuraavassa esitellään joitakin seikkoja, joista nämä erityispiirteet johtuvat.

Kirkollisissa luetteloissa kuten myös muissa historiallisissa väestöluetteloissa mainitaan nimeltä henkilöt, joihin kohdistetaan erilaisia toimenpiteitä. Samoista henkilöistä saattaa olla hajanaisia tietoja useissa lähteissä. Kun eri lähteiden tiedot yhdistetään on mahdollista luoda elämäkuvat laajan ihmisjoukon yksittäisistä jäsenistä. Tietokantojen tehokkaan hyödyntämisen edellytyksenä on näiden yksittäiseen henkilöön ja hänen läheisiinsä liittyvien tietojen yhdistäminen. Tietojen yhdistäminen eli linkitys on vaikea ongelma historiallisia tietokantoja rakennettaessa. Linkityksen hoitamiseen on kehitetty edistyneitä tietokoneohjelmia. Tietokoneohjelmia on käytetty myös nimien normalisoinnissa eli nimien kirjoitusasujen yhtenäistämässä. Oikeinkirjoitus on kaikissa kielissä vaihdellut eri aikakausina ja alueilla. Kirkonkirjat ovat käsinkirjoitettuja ja siksi tutkijalta vaaditaan vanhojen käsialojen tuntemusta. Linkitys edellyttää, että sekä henkilö- että paikannimiä voidaan vertailla ja nimet on normalisoitava, jotta samannimiset ihmiset voitaisiin löytää eri lähteistä.³⁵

Historialliset tietokannat on tarkoitettu palvelemaan tutkimusta. Koska eri tutkimusten tietotarpeet vaihtelevat, on lähteiden kaikki relevantti tieto tallen-

³⁴ Thaller 1991

³⁵ Winchester 1992, Atack, Bateman, Gegson 1992

nettava mahdollisimman tarkkaan. Myös harvinaiset tiedot ja erikoistapaukset on otettava huomioon ja siksi tallennettujen tietojen joukkoon jää paljon tyhjää. Lähteissä esiintyy usein puutteellisia ja ristiriitaisia mutta myös loogisesti samanarvoisia tietoja. Henkilöllä saattaa olla kaksi tasaveroista ammattia kuten maanviljelijä ja kutoja. Pappien käyttämiä lukutaitomerkkejä ei sellaisinaan voida siirtää tietokoneelle. Lisäksi tietojen pituudet vaihtelevat huomattavasti. Aikaperspektiivi on myös otettava huomioon, sillä on pystyttävä kuvaamaan muutokset henkilöiden suhteissa ja erilaisissa hierarkioissa eri ajanjaksoina. Historiallista tietokantaa ei luonnin jälkeen päivitetä. Uusi ja vanha aineisto voidaan yhdistää keskenään, mutta varsinaista uusien tapahtumien lisäystä tai tietokannassa olevien tietojen muuttamista ei tapahdu.³⁶

Kirkonkirjojen alkuperäisenä tehtävänä oli pitää lukua seurakunnissa suorite-
tuista kirkollisista toimituksista sekä Ruotsissa ja Suomessa lisäksi seurakunta-
laisten kristinopin tiedoista ja ehtoollisella käynneistä. Myöhemmässä vaiheessa
kirkonkirjat toimivat myös valtionhallinnon vaatimien väestötilastojen pohjana.
Näistä erilaisista tehtävistä johtuen niiden tietosisältö vaihteli eri aikoina, mistä
on seurauksena, ettei tietokantaa suunniteltaessa voida tehdä laajoja yleistyksiä
ja luokitteluja.³⁷ Kirkonkirjat sisältävät sekä tekstitietoa että rakenteellista tietoa.
Tekstitiedon määrä vaihtelee eri kirkonkirjatyypeissä, mutta saman ohjelmiston
on kyettävä käsittelemään molempia tietotyyppisiä. Väestöluetteloissa käsitellään
myös henkilöiden välisiä suhteita, ja siksi tietokannan tulisi sisältää sekä henkilö-
tietokanta, jonka jäsenillä on samanlaisia ominaisuuksia että suhdetietokanta,
joka käsittäisi yhteydet erilaisia sosiaalisia rooleja edustavien ihmisten välillä.³⁸

Historialliselle tiedolle on ominaista kontekstiherkkyys, mikä tarkoittaa sitä, että
tietyin ominaisuuden merkitys saattaa olla riippuvainen muiden ominaisuuksien

³⁶ Thaller 1991, Karjalan väestötieteellisen tietokannan loppuraportti 1989

³⁷ Karskela 1987, s. 26, Kälvemark 1977

³⁸ Karjalan väestötieteellisen tietokannan loppuraportti 1989, Thaller 1991

sisällöstä eikä sitä niitä tuntematta voida tulkita oikein tai että jonkin termin merkitys saattaa muuttua ajan mittaan. Lisäksi tieto on sisällöltään sumeaa ja epätäsmällistä, mistä on seurauksena, että useampi kuin yksi tulkinta on oikea. Esimerkiksi vanhat ihmiset vuonna 1640 ja 1830 ovat aivan erilainen ryhmä. Ikätiedot ovat perinteisesti epätäsmällisiä ja ilmoitettiin vielä 1800-luvun lopulla-kin noin 10 vuoden täsmällisyydellä. Historiallinen tieto on moniulotteista. Yksinkertaisessa asiakirjassa, jossa kerrotaan että "X lahjoitti Y:n Z:lle", piilee taustalla koko joukko erilaista tietoa tekstissä esitettyjen tosiasioiden lisäksi. Tällaisen tiedon vieminen tietokantaan on ongelmallista.³⁹

Seuraavissa luvuissa tarkastellaan lähemmin eräitä menetelmiä, joita on kehitetty nimien normalisointiin sekä samaan henkilöön eri lähteissä liittyvien tietojen yhdistämiseen eli linkitykseen.

3.1 Nimien normalisointi

Suuria väkijoukkoja tutkivat historiantutkijat joutuvat kokoamaan henkilöiden tietoja useista toisistaan riippumattomista lähteistä. Yleensä käytäntö on sellainen, että verrataan luettelon A ja B kahta lähdemerkintää toisiinsa ja jos ne tietyiltä yksityiskohdiltaan ovat samanlaiset, päätetään niiden liittyvän samaan henkilöön. Koska rutiinitieto väestöstä on kerätty nimien perusteella, nimet ovat tärkein yksityiskohta henkilön tietojen yhdistämisessä. Nimien oikeinkirjoitus on kuitenkin vaihdellut kaikissa kielissä eri aikakausina.⁴⁰ Suomalaisissa kirkonkirjoissa nimien kirjoitusasu vaihtelee eri kirkonkirjojen pitäjillä ja jopa samallakin muistiin merkitsijällä. Esimerkiksi nimi Kristina on saatettu kirjoittaa seitsemällä eri tavalla eli Christina, Christine, Christin, Kristiina, Stiina ja Stina. Nimien vaihtelu saattaa johtua paikallisista syistä, muiden kielten vaikutuksesta tai

³⁹ Thaller 1991

⁴⁰ Guth 1976, Winchester 1992, Attack, Bateman, Gregson 1992

kopiointivirheistä. Henkilön tietojen yhdistämien nimien perusteella edellyttää nimien muuttamista yhtenäiseen muotoon eli nimet on normalisoitava.⁴¹

Kun lähteiden tietoja yhdistetään on ensimmäisenä tehtävänä yleensä järjestää nimet aakkosjärjestykseen. Tämän jälkeen on mahdollista liittää toisiinsa lähteiden merkintöjä sukunimen tai sukunimiparin mukaan. Nimien normalisointi takaa sen, että lajiteltaessa samaan yksilöön liittyvät nimet tulevat mahdollisimman lähelle toisiaan. Nimien normalisointiin on kehitetty sekä kielen fonetiikkaan että kirjainten aseman tutkimiseen perustuvia menetelmiä.⁴² Seuraavassa tarkastellaan ensin eräitä kielen fonetiikkaan perustuvia nimien normalisointimenetelmiä. Tämän jälkeen kuvataan edellisistä poikkeavaa menetelmää, jossa tutkitaan kirjainten asemaa nimessä. Myöhemmin, luvussa 5, yksityiskohtaisesti kuvatuista esimerkkietokannoista, ruotsalainen Demografiska Databasen ja Karjala-tietokanta käyttävät kielen fonetiikkaan perustuvia menetelmiä. Amerikkalaista Gillespie-järjestelmää käsittelevissä lähteissä ei nimien normalisointia ole käsitelty.

3.1.1 Kielen fonetiikkaan perustuvat menetelmät

Historiantutkijat ovat tuottaneet nimivariaatioiden aiheuttamien ongelmien ratkaisemiseksi erilaisia kielen fonetiikkaan perustuvia automaattisia järjestelmiä.⁴³ Englannin kieltä voidaan muokata poistamalla vokaalit niin, että jäljelle jäävät alkukirjaimet ja jono konsonanteja, eivätkä nimet silti menetä ominaislaatuaan. Tätä tapaa on käytetty Russell Soundex -menetelmässä, joka on suunniteltu angloamerikkalaisten nimien vertailuun. Menetelmä muuttaa nimen koodiksi, joka muodostuu kirjaimesta ja kolmesta numerosta (0-6). Kirjain tulee

⁴¹ Karjalan väestötieteellisen tietokannan loppuraportti 1989

⁴² Winchester 1992

⁴³ De Brou, Olsen 1986

nimen ensimmäisestä kirjaimesta, numerot määräytyvät kolmen seuraavan konsonantin mukaan. Esimerkiksi nimet Smith, Smyth ja Smythe kuuluvat kaikki samaan ryhmään, S530. Russell Soundex -menetelmän heikkoutena on, että saman nimen eri muodot eivät aina joudu samaan ryhmään, jolloin niitä ei voi vertailla. Esimerkiksi nimet, jotka alkavat yhdeltä kirjaimelta kuulostavalla konsonanttiryhmällä kuten nimissä Shmidt ja Shmitz, päätyvät eri ryhmiin. Myös muunkieliset nimet aiheuttavat ongelmia, jos niitä käsitellään kuin ne olisivat englanninkielisiä kuten on laita esimerkiksi nimessä Le Blanc, joka voi saada muotoja De Blanc, White, Whit-son, Whit-sen, Writ-san, Werit-sar. Tutkijat kehittivät Russell Soundex -menetelmästä omaan käyttöönsä paremmin soveltuvia versioita. Eräs näistä on NYSIIS eli New York State Identification and Intelligence System. NYSIIS -järjestelmässä nimet tehdään vertailukelpoisiksi yhdenmukaistamalla niiden kirjoitusasu. Nimien alussa ja keskellä sijaitsevat vokaalit muutetaan a-kirjaimiksi ja nimien lopussa olevat vokaaliyhdistelmät y-kirjaimiksi. Konsonanttien esittäminen yksinkertaistetaan tarkoitusta varten määriteltyjen sääntöjen perusteella. Esimerkiksi nimestä Bush tulee Bas. Tässä järjestelmässä suurin mahdollinen määrä nimen eri variaatioita tulee samaan ryhmään ja parantaa siten oikean linkityksen mahdollisuuksia, koska vertailuja tehdään sekä ryhmien sisällä että niiden välillä. Englanninkielisiä nimiä käsitellessä nämä konsonanttien vertailuihin perustuvat menetelmät ovat osoittautuneet menestykselliseksi.⁴⁴

Kielen fonetiikkaan sidottujen menetelmien tehokas soveltaminen rajoittuu kuitenkin kieleen, jonka tutkimiseen ne on alunperin luotu. Kanadan ranskankielisen väestön nimien käsittelyyn Russell Soundex -menetelmä sopi huonosti. Osa tutkijoista käytti Henry-menetelmää, joka on ranskankielinen versio Russell Soundex -menetelmästä. Menetelmä lyhensi sukunimet alkukirjaimiksi ja kahdeksi merkittäväksi konsonantiksi ja tuntui paremmin sopivan ranskankielisten nimien linkitykseen. Menetelmän arvostelijat esittivät kuitenkin,

⁴⁴ Winchester 1992, Atack, Bateman, Gregson 1992

että se yksinkertaisti ranskankielen foneettista rakennetta. Siksi Henry kehitti toisen version, joka otti paremmin huomioon ranskan kielen foneettiset ominaisuudet. Kuitenkaan esimerkiksi kanadalainen, historiallisen tietokannan rakentanut, Bouchardin ryhmä ei sitä käyttänyt, sillä ryhmän mielestä menetelmä muutti liikaa nimien foneettista rakennetta eikä huomioinut etunimiä, jotka olivat olennainen elementti avioparien linkityksessä. Siksi Bouchardin ryhmä kehitti Fonem-ohjelmiston, joka sisältää 64 muutossääntöön perustuvan taidokkaan foneettisen koodin. Fonem on tehokas keino linkittää automaattisesti ranskankielistä väestöä, mutta jos tutkimus käsittää etnisesti sekalaista väestöä, automaattinen linkitys ei onnistu, vaan nimet on linkitettävä manuaalisesti.⁴⁵

Apulaisprofessori Seppo Sippu on Jyväskylän yliopiston tietojenkäsittelyopin laitoksella Hiski-projektin yhteydessä kehittänyt nimien normalisointiin sopivan ohjelmiston, jonka käyttöönottoa harkittiin Karjala-järjestelmässä. Hiski-järjestelmä kehitettiin Jyväskylän yliopiston tietojenkäsittelyopin laitoksella opinnäytetyönä vuonna 1985 ja sen päämääränä oli tallentaa historiakirjamerkintöjä tietokoneelle. Tietoja on mahdollista lajitella eri perustein esimerkiksi lapsen, isän tai kotipaikan nimen perusteella. Järjestetyt merkinnät voidaan tulostaa paperille luetteloiksi, joista tietoja on huomattavasti helpompi tutkia kuin epäselvistä alkuperäisistä historiakirjoista. Historiakirjamerkintöjä siirretään alkuperäisessä muodossaan magneettinauhalle ja luodaan nimihakemistot, joissa ovat kaikki merkinnöissä esiintyneet etu-, lisä- ja paikannimet, niiden normalisoidut muodot sekä normalisoitujen muotojen tunnusjonot.⁴⁶

Järjestelmä normalisoi nimet syötön yhteydessä. Nimen normalisointi tapahtuu kolmen algoritmin avulla. Algoritmit poistavat, lisäävät ja vaihtavat nimen kirjaimia tiettyjen sääntöjen mukaan. Ensimmäinen ja toinen algoritmi muodostavat ehdotuksen nimen normalisoiduksi asuksi, kolmas algoritmi muodostaa

⁴⁵ De Brou, Olsen 1986

⁴⁶HISKI-järjestelmä 1985

nimelle tunnusjonon. Kahta nimiasua pidetään synonyymeinä, jos niillä on sama tunnusjono. Lähteissä esiintyneet alkuperäiset nimet, niiden normalisoidut muodot ja tunnusjonot on liitetty linkeillä toisiinsa ja tallennettu nimihakemistoon. Käyttäjä voi valita järjestelmän tuottamista vaihtoehtoista tai kirjoittaa itse haluamansa normalisoidun muodon. Mikäli mikään vaihtoehto ei ole oikea, normalisoitu asu voidaan muodostaa korjailemalla jotakin esitetyistä vaihtoehtoista tai kirjoittamalla kokonaan uusi nimiasu. Kaikki kirjoitusasut viedään nimihakemistoon ja tällä tavalla järjestelmä oppii käytön aikana. Syötön edetessä järjestelmä tuntee suurimman osan vastaan tulevista nimiasuista. Kolmas algoritmi muodostaa nimelle tunnusjonon poistamalla vokaalit lukuun ottamatta sanan ensimmäisenä kirjaimena olevaa vokaalia. Tunnusjonoissa kaksinkertaiset konsonantit muutetaan yksinkertaisiksi, kirjaimet c, g ja q muutetaan k:ksi, d:stä tulee t, b:stä p. Sanan alussa ä muutetaan e:ksi ja å o:ksi. Tunnusjonoista tulee konsonanttiyhdistelmiä, esimerkkeinä nimistä Pehr, Pär -> pr, Peder -> ptr ja Arvid, Arffuidh -> arvt.⁴⁷

Patronyymit eroavat etunimistä vain päätteensä osalta ja niiden alkuosat ovat tavallisia etunimiä. Patronyymit normalisoidaan siten, että järjestelmä katkaisee päätteen pois ja käsittelee alkuosan samoin kuin tavalliset nimet. Esimerkiksi Arffuidhsson-nimen alkuosa muutetaan normalisoituun muotoon Arvid ja patronyymimuodoksi saadaan Arvidsson. Valintavaiheessa päätte liitetään vaihtoehtoihin ja merkitään alleviivauksella päätteen alkukohta. Valinnan jälkeen viedään patronyymin ja sen normalisoidun muodon alkuosa sekä tunnusjono nimihakemistoon.⁴⁸

⁴⁷ HISKI-järjestelmä 1985

⁴⁸ HISKI-järjestelmä 1985

Taulukko 1 Etunimihakemiston sisältö nimen Kristina osalta

Alkuperäisasu	Tunnusasu	Normaalisoitu asu
Christina	krstn	Kristina
Christine	krstn	Kristina
Christin	krstn	Kristina
Kristiina	krstn	Kristina
Stiina	stn	Kristina
Stina	stn	Kristina

Nimihakemistoja on kaksi: etunimihakemisto, josta on esimerkki taulukossa 1 sekä lisä- ja paikannimihakemisto. Etunimihakemisto kattaa koko maan ja sitä voidaan käyttää kaikkien seurakuntien etunimien talletuksessa, koska etunimistö on hyvin samankaltaista eri puolilla Suomea. Lisä- ja paikannimihakemisto on seurakuntakohtainen. Apulaisprofessori Sipun kehittämässä ohjelmistossa jokaiselle nimelle annetaan normalisoitu muoto, mutta myös alkuperäinen kirjoitusasu tallennetaan.⁴⁹

Ruotsalaisessa kirkonkirjojen tietoja tallentavassa, Demografiska Databasen -järjestelmässä, nimien normalisointiin käytetään sääntöjä, jotka perustuvat Uumajan yliopiston skandinaavisten kielten laitoksen tekemään tutkimukseen. Kaikista lähteiden merkinnöistä kerätään henkilönimet yhteen seurakunta-kohtaiseen tiedostoon, jossa on osoittimet lähteisiin. Nimet koodataan automaattisesti. Koodausohjelma poistaa ensin kielletyt erikoismerkit ja -jonot, minkä jälkeen nimiin sovelletaan oikeinkirjoitus- ja korvaussääntöjä. Etunimiin sovelletaan sääntöjä, joiden vaikutuksesta tapahtuu esimerkiksi seuraavia muutoksia:

- Oluf -> Olof
- Hin -> Hen
- Jon -> Joh
- Hele -> Lena
- Stina -> Kris
- Kajsa -> Kat
- Greta -> Marg

Sukunimissä korvataan esimerkiksi Ers -> Erik ja Ols -> Olof. Kaikki sukunimet, jotka sisältävät eri muodoissa päätteet -son ja -dotter standardoidaan. Stan-

⁴⁹ HISKI-järjestelmä 1985

dardoinnin jälkeen nimet korvataan viisinumeroisella koodilla, joka löytyy kooditaulukosta. Jos ohjelma ei löydä nimen koodia, nimitietue viedään erilliseen tiedostoon, joka edellyttää manuaalisia operaatioita koodin antamisessa. Nimikoodit ovat peräkkäisiä numeroita ilman, että niillä on ryhmittelyvariantteja samalle tai samanlaisille nimille.⁵⁰

Yksi menetelmä nimien normalisoimiseksi on standardoida nimien oikeinkirjoitus tietoa tallennettaessa eli käyttää nimien perusmuotoja, ja säilyttää sen lisäksi myös nimien alkuperäinen muoto.⁵¹ Tämä ei ole automatisoitu prosessi, mutta saattaa johtaa parempiin tuloksiin joissakin olosuhteissa. Suomalaisessa, luovutetun Karjalan kirkonkirjoja tallentavassa Karjala-tietokannassa nimet normalisoidaan tallennuksen yhteydessä. Nimien normalisoinnissa käytetään nimien perusmuotoja, jotka on esitetty teoksessa, Uusi suomalainen nimikirja. Tässä teoksessa on tiedot Suomessa historiallisena aikana käytetyistä ja käytössä olevista etunimistä sekä niiden vaiheista. Siihen on otettu mahdollisimman tarkoin etunimien murteenmukaiset paikallismuodot, kansanomaiset etunimien puhuttelu-, kutsuma- ja hellittelymuodot sekä niiden viralliset vastineet. Esimerkiksi Frans -nimestä tavataan eri alueilla muodot Pransu, Ransa, Ranse, Ransi, Ranssi ja Ransu, aikaisemmin myös Rantsi ja Rantti, jotka ovat säilyneet paikannimissä. Myös sukunimet normalisoidaan tuossa teoksessa esitettyjen perusmuotojen mukaan. Kirjassa on selvitetty suomalaisten sukunimien ikää, alkuperää, levikkiä ja lähtösijoja ja hahmotettu kuvaa sukunimistömme syntyta-voista ja eri tyypeistä.⁵²

⁵⁰ Demografiska Databasen 1990

⁵¹ Winchester 1992

⁵² Karjalan väestötieteellisen tietokannan loppuraportti 1989, Vilkuna, Huitu, Mikkonen 1988

3.1.2 Kirjainten aseman tutkimiseen perustuvat menetelmät

Amerikkalainen tutkija Guth on kehittänyt algoritmin, joka tunnistaa sukunimien oikeinkirjoitusvaihtelut useista riippumattomista lähteistä. Algoritmi kuuluu ohjelmistoon, jonka perustehtävänä on merkintöjä yhdistämällä etsiä kaikki tieto, mikä liittyy tiettyyn yksilöön useissa väestöluetteloissa. Nimien tunnistus tapahtuu kirjainten aseman perusteella.⁵³

Nimentunnistusalgoritmia voidaan soveltaa erilaiseen lähdemateriaaliin. Guthin käyttämissä väestöluetteloissa suku- ja etunimi olivat ainoat yhteiset muuttujat, joiden avulla henkilöt voitiin tunnistaa ja se pakotti kehittämään menetelmän, joka mahdollisimman tehokkaasti käyttää hyväkseen sukunimien erottelevuutta. Menetelmä käsittää useita ohjelmia. Täydellinen yhdenmukaisuus kahden tutkittavan nimen jokaisen kirjaimen välillä viittaa siihen, että kysymyksessä todennäköisesti on sama henkilö. Saman sukunimen eri kirjoitusasujen vertailu-algoritmi on sen tähden ensimmäinen tehtävä jokaisessa ohjelmassa. Aluksi tutkitaan, ovatko kaksi sukunimeä täysin samat. Jos näin ei ole, algoritmi joutuu silmukkaan, jossa sukunimiä vertaillaan kirjain kirjaimelta kaksi kerrallaan. Tässä kahden sukunimen kirjainten yhteen sovittamisessa käytetään yhdeksää testiä, joissa annettujen sääntöjen mukaan tarkastellaan kirjainten suhteellista asemaa, niiden paikkaa ja järjestystä. Ohjelma suorittaa ensimmäistä testiä niin kauan kuin nimet ovat samoja. Kun ohjelma kohtaa eri kirjaimet samassa asemassa, se etsii samoja kirjaimia muista asemista ja alkaa käydä läpi loppuja kahdeksaa testiä, joita havainnollistetaan taulukossa 2. Jos näissä testeissä löydetään yhteen sopiva kirjain, palataan silmukan alkuun, laskuria lisätään yhdellä ja edetään tutkimaan nimen loppuosaa, jossa verrataan nimien seuraavia kirjaimia. Jos yhteisiä kirjaimia ei löydy, nimet katsotaan erilaisiksi.⁵⁴

⁵³ Guth 1976

⁵⁴ Guth 1977, De Brou, Olsen 1986

Taulukko 2 Sukunimien yhteen sopivuuden testaaminen

	Asema nimessä 1	Asema nimessä 2	Esimerkinimet	
Testi 1	x	x	¹ GLAVIN	GLAWYN
Testi 2	x	x + 1	GLAVIN	GLAWYN
Testi 3	x	x + 2	GLAVIN	GLAWYN
Testi 4	x	x - 1	GLAVIN	GLAWYN
Testi 5	x - 1	x	GLAVIN	GLAWYN
Testi 6	x + 1	x	GLAVIN	GLAWYN
Testi 7	x + 2	x	GLAVIN	GLAWYN
Testi 8	x + 1	x + 1	GLAVIN	GLAWYN
Testi 9	x + 2	x + 2	GLAVIN	GLAWYN

¹ Alleviivaus osoittaa kussakin testissä vertailtavat kirjaimet.

Esimerkkinä tämän algoritmin toiminnasta tutkitaan nimen "Glavin" variaatiota "Glawyn". Nimet ovat samanlaisia kolmen ensimmäisen kirjaimen osalta, mutta neljännet kirjaimet eroavat ja algoritmi alkaa vertailla poikkeavaa kirjainta kummankin nimen lähinnä seuraaviin ja edeltäviin kirjaimiin. Kun tämä ei tuota tulosta, eikä minkäänlaista yhteen sopivuutta ilmene, algoritmi yrittää kahta viimeistä vaihtoehtoa. Viimeinen testi todellakin onnistuu: kirjaimet asemassa "x+2" kummassakin nimessä ovat samoja. Algoritmi kasvattaa asemalaskuria yhdellä ja aloittaa silmukan uudelleen. Tässä esimerkissä nimet loppuvat yhteiseen kirjaimeseen. Nimen loppuosa läpäisee yhteen sovittamisprosessin; näitä kahta nimeä pidetään samoina. Algoritmi hylkää nimet, joiden pituuserot ovat suurempia kuin kaksi kirjainta.⁵⁵

Guthin algoritmin etu on, että se ei ole kieleen sidottu, vaan nimien identifiointi tapahtuu kirjainten aseman ja järjestyksen eikä foneettisen yhteen sopivuuden perusteella. Tutkimus, johon Guth sovelsi tätä menetelmää, käsitti kahdeksan äänestysluettelon äänestäjien nimien vertaamista väestölaskenta- ja kutsuntaluetteloiden nimiin. Luetteloihin kuului sekä ranskan- että englanninkielistä väestöä vuosilta 1792-1842, mutta algoritmi kykeni yhdistämään monikielisen väestön

⁵⁵ Guth 1977

tietoja. Guthin algoritmia kokeiltaessa tiedot lajiteltiin ensin sukunimen, etunimen ja tapahtumavuoden mukaan ja algoritmi piti samoina 78.2 % nimistä. Kuudessa seuraavassa läpiajossa tiedot lajiteltiin muiden tietojen kuten ammatin ja asuinpaikan perusteella. Tämä sekoitti nimien mukaista lajittelua ja järjesti sukunimet eri tavalla, mikä lisäsi tehtyjen vertailujen määrää, mutta muodosti loput yhteen sopivuudet.⁵⁶

Algoritmi yhdistää oikein yksilöiden nimet, joissa on pientä eroavuutta oikeinkirjoituksessa, jos ero aiheutuu korvaavista vokaaleista kuten Foulton ja Fulton tai konsonanteista esimerkiksi Cardbell ja Cardwell sekä konsonanttien kahdentumisesta kuten Morrison ja Morrisson. Se tuottaa tuloksia hyvin erilaisista oikeinkirjoituksista sekä nimistä, jotka eivät ala samalla kirjaimella tai joiden käsiala on epäselvää esimerkiksi Sincere - St.-Cyre, Gignac - Iicnac ja Boutier - Routier. Koska algoritmi ei ole riippuvainen foneettisesta rakenteesta, se kykenee yhdistämään nimet, jotka eivät ole ranskalaisia eivätkä englantilaisia kuten Stoppleben, Stoppleban ja Stoppleben.⁵⁷

Algoritmi tuottaa joitakin virheellisiä yhdistämissä, jotka pitää korjata käsin. Esimerkki nimien Forton ja Garon väärästä yhteen sopivuudesta on opettavainen. Testi tuottaa yhteen sopivuuden kolmansissa kirjaimissa. Kaksi seuraavaa vaihetta tuottavat yhteen sopivuuden, kun vertaillaan kohdekirjainta nimessä Garon:ssa yhtä sijaa oikealle olevaan kirjaimen nimessä Forton. Testi keskeytyy, kun päästään jommankumman sanan loppuun. Lyhyehköt sukunimet ovat erittäin alttiita virheille, koska ohjelma yhdistää nimet, joissa esiintyy vain yksi tai kaksi tavallista vokaalia. Guthin algoritmin täsmällisyyttä ja luotettavuutta vertailtiin käsin saatuihin tuloksiin. Vertailu osoitti, että Guthin algoritmi oli lähes yhtä tehokas kuin tutkijan manuaalisesti suorittama nimien yhdistäminen.⁵⁸

⁵⁶ De Brou, Olsen 1986

⁵⁷ De Brou, Olsen 1986

⁵⁸ De Brou, Olsen 1986

3.2 Linkitys

Tietue sisältää kohteen erilaisia ominaisuuksia. Henkilötietue voi koostua esimerkiksi henkilön nimestä, syntymä- ja kuolinajasta. Tietuelinkityksellä ymmärretään eri historiallisten asiakirjojen samaa henkilöä koskevien tietojen yhdistämistä kronologiseen järjestykseen.⁵⁹ Yksilöiden välillä vallitsee erilaisia suhteita. Linkitys tarkoittaa myös näiden suhteiden yhdistämistä siten, että lapset liitetään vanhempiinsa ja aviopuolisot toisiinsa. Linkiksi kutsutaan yhteyttä, jonka on osoitettu vallitsevan kahden tietueen välillä. Lähteiden tietojen yhdistäminen on edellytyksenä henkilön mahdollisimman täydellisen elämäkerran rakentamiselle. Historiantutkijat alkoivat käyttää tietokonetta linkityksessä 1960-luvun lopulta lähtien. Päämääränä oli aikaansaada asiantuntijajärjestelmä, joka identifioi saman henkilön useista eri lähteistä. Winchester pitää merkittävänä kolmea 1970-luvun alusta käynnissä ollutta historiallista lähde- materiaalia käyttävää linkitysprojektia, jotka ovat Cambridgen ryhmä Englannissa, Demografiska Databasen eli DDB Uumajassa ja Bouchardin ryhmä Kanadassa. Näissä projekteissa yhdistettiin kirkollisten luetteloiden kuten kaste-, avioliitto- ja kuolleiden luetteloiden ja DDB:ssä lisäksi rippikirjojen tietoja. Projektit on toteutettu suur- tai minitietokoneella, mutta tietotekniikan kehittymisen seurauksena myös mikroja on alettu käyttää linkitysprojekteissa. Suuren yksilöjoukon identifioiminen historiallisista väestöluetteloista on monimutkainen tehtävä. Jokaisella projektilla on oma käsityksensä siitä, miten tämä tavoite saavutetaan: mitkä ovat linkityksen päävaikeudet, mitkä tiedon yksityiskohdat parhaiten soveltuvat yksilön tunnistamiseen, mikä olisi tehokkain linkitysstrategia ja miten parhaiten organisoida linkitetyt tiedostot.⁶⁰

Laajoista tietuelinkitysprojekteista saadut kokemukset ovat osoittaneet, että

⁵⁹ Schofield 1992

⁶⁰ Winchester 1992, Atack, Bateman, Gregson 1992

suunnittelussa on tarpeen ottaa huomioon lähdeaineiston luonteen ja laadun läheinen yhteys tutkimuspäämääriin. On osoittautunut tarpeelliseksi suunnitella tutkimuksen päämäärien kanssa yhdenmukainen linkitysjärjestelmä ja varmistaa samalla, että lähdeaineisto on laadultaan sellaista, että se kykenee tukemaan järjestelmää. Historiallinen lähdeaineisto on vaihtelevaa, mikä selittää käytettyjen menetelmien moninaisuutta. Joissakin tapauksissa on yhdistämisen kohteena kaksi samaa väestöä koskevaa tiedostoa, jotka sisältävät yhden merkinnän jokaisesta henkilöstä ja tällöin pyritään yhdistämään tietueet kaksi kerrallaan. Toisaalta lähteet voivat koostua yhdestä ainoasta tiedostosta, joka sisältää useita samaan henkilöön tai aviopariin liittyviä merkintöjä. Tässä tapauksessa joudutaan käyttämään useita erilaisia menetelmiä, koska tietueiden välillä täytyy tehdä vertailuja kaikkien sopivien yhteyksien muodostamiseksi. Vertailut täytyy tehdä pareittain niin, että jokaista tietuetta verrataan peräkkäin useisiin muihin. Vertailu voi päättyä joko positiiviseen tai negatiiviseen tulokseen, mutta se voi johtaa myös vertailuparien muodostamiin ryhmiin, klustereihin, silloin kun päätös on seurausta todennäköisyydestä eikä siihen voida vastata joko kyllä tai ei. Tässä tilanteessa tietue A on linkitettävissä ei ainoastaan B:hen, vaan myös C:hen ja D:hen, ja D taas on linkitettävissä E:hen. Kaikki yhdistelmät ovat siis mahdollisia, mikä johtuu tietojen epävarmuudesta ja ristiriitaisuudesta.⁶¹

Seuraavissa luvuissa tarkastellaan linkityksessä käytettyjä muuttujia ja lähestymistapoja sekä kuvataan lopuksi kanadalaisen Bouchardin tutkimuksessa että ruotsalaisessa DDB:ssä käytettyjä linkitysmenetelmiä. DDB kuuluu myöhemmin, luvussa 5, yksityiskohtaisesti kuvattaviin esimerkkietokantoihin. Muiden esimerkkietokantojen, suomalaisen Karjala-tietokannan ja amerikkalaisen Gillespie-järjestelmän, linkitysmenetelmiä ei kuvata, koska Karjala-tietokannassa ei tietojen linkitystä ole suoritettu ja Gillespie-järjestelmää käsittelevissä lähteissä ei ole selvitystä linkitysmenetelmästä. Koska linkitys on henkilöhistorian tietokantojen rakentamisen keskeisimpiä tehtäviä, on tässä tutkimuksessa katsottu

⁶¹ Bouchard 1992, Bouchard 1986

tarpeelliseksi esitellä DDB:n linkitysmenetelmän lisäksi toinen linkitysmenetelmä, joksi on valittu kanadalaisen Bouchardin käyttämä tapa yhdistää samaan henkilöön liittyviä tietoja.

3.2.1 Linkitysmuuttujat

Kirkonkirjoihin ja muihin historiallisiin väestöluetteloihin kirjattiin käyttötarkoituksen perusteella erilaisia, mutta osittain myös päällekkäisiä tietoja tapahtumista ja henkilöistä. Kun ryhdytään linkittämään henkilöiden tietoja eri lähteistä, on ensin päätettävä mitä tietoja yhdistämisessä käytetään. Kyseisten yksityiskohtien, linkitysmuuttujien, on oltava mukana useammassa kuin yhdessä lähteessä.⁶² Neutraaleimmat ja pysyvimmät ominaisuudet ovat sopivimpia, mutta lajitteluun käytettyjen kriteerien on sisällettävä riittävästi myös erottelevia ominaisuuksia. Nimet muodostavat historiallisen tietuelinkityksen perustan. Niiden lisäksi yleisimmin käytettyjä ovat syntymäaika- ja paikka, sukupuoli, asuinpaikka, ammatti ja sukulaisiin liittyvät tiedot.⁶³ Nimiä voidaan helposti lajitella ja useimmiten ensimmäinen tietuelinkitysaskel onkin järjestää nimitieto täydelliseen tai puolittaiseen aakkosjärjestykseen. Mutta sitä ennen tutkija joutuu selvittämään nimien oikeinkirjoituksessa esiintyvät vaihtelut. Toinen ongelma liittyy samannimisiin henkilöihin, mikä saattaa johtaa siihen, että yhdistetäänkin kaksi eri henkilöä, joilla on sama nimi. Esimerkiksi saattaa löytyä kaksi kastemerkintää samalle nimelle, mutta ainoastaan yksi tuohon nimeen liittyvä avioliittomerkintä. On selvittävä, kumpi lapsista on solminut avioliiton. Jos väestöluettelot mainitsevat iän, ammatin, asuinpaikan tai sukulaisten nimiä, on helppo erottaa samannimiset henkilöt, vaikka samannimisiä ihmisiä löytyykin useampia. Siten jo tietueiden yksinkertainenkin vertailu saattaa paljastaa monia ristiriitoja, jotka täytyy ratkaista ennen kuin voidaan muodostaa oikeita elämä-

⁶² Winchester 1992

⁶³ Bouchard 1992, Atack, Bateman, Gregson 1992

kertoja. Lisäksi saattaa löytyä tapahtumia, joita ei ole mainittu väestöluetteloissa. Lapset, jotka kuolivat ennen kuin heidät ehdittiin kastaa, puuttuvat kaste-merkinnöistä. Ihmisiä pystytään erottelemaan nimien perusteella, mutta samalla edellytetään, että tutkittavien henkilöiden tietojen noudattavan loogisia, biologisia ja historiallisia rajoitteita. Esimerkiksi etunimien täytyy olla yhteen sopivat, kastepäivien täytyy edeltää hautauspäiviä eikä lapsia saa yhdistää yli 50 vuotiaisiin naisiin.⁶⁴

Ruotsalais-suomalaiset rippikirjat ovat tutkimuksen kannalta erittäin hyvänlaatuisia. Niiden perhe- ja taloustieto on tiukkaan pakattu yhteen yhdeksi luetteloksi, jonka sivut muodostuvat toisiaan seuraavista riveistä. Milloin tahansa henkilön elämässä tapahtuu muutos ja hänen tietonsa siirretään toiseen rippikirjaan, tehdään riville merkintä, mistä etsiä seuraavaa mainintaa kuten lehtiartikkelissa sanotaan "jatkuu sivulla xx". Samoin rippikirjaan, jonne siirto on tapahtunut, tehdään viittaus, mistä henkilö tai perhe aikaisemmin on löydettävissä, kuten lehdessä sanotaan "jatkuu sivulta yy". Tällainen yksinkertaisten osoittimien järjestelmä helpottaa linkitystä.⁶⁵

Uumajan DDB:ssä käytetään rippikirjatietoja yhdistettäessä seuraavia linkitysmuuttujia:

- sukunimiä
- etunimiä (saadaan selville myös henkilön sukupuoli)
- syntymävuotta, -kuukautta, -päivää ja -seurakuntaa
- siirtomerkintää jostakin (seurakunta, rippikirja, sivu)
- siirtomerkintää jonnekin (seurakunta, rippikirja, sivu)
- asuinpaikkaa.

Yhdistettäessä rippikirjojen tietoja syntyneiden luetteloihin käytetään syntymän päivämäärää, vihittyjen luetteloihin vihkimisen päivämäärää, kuolleiden luetteloihin kuolinpäivämäärää ja muuttaneiden luetteloihin muuttotodistusta,

⁶⁴ Schofield 1992, Winchester 1992

⁶⁵ Schofield 1992, Winchester 1992

muuton päivämäärää sekä muuttopaikkakuntaa. Päinvastaisessa tapauksessa liitettäessä historiakirjojen ja rippikirjojen tietoja käytetään muuttujia taulukon 3 esittämällä tavalla.⁶⁶

Taulukko 3 Historiakirjojen ja rippikirjojen linkitysmuuttujat

Muuttuja	Syntyneet	Vihityt	Kuolleet	Muuttaneet
sukunimiä		x	x	(x)
etunimiä	x	x	x	(x)
synnyinvuosi		x	(x)	
syntymäkuukausi, -päivä	x	(x)		
viittaus rippikirjaan, sivu	x	x	x	x
viittaus asuinpaikkaan	x	x	x	x
ikä			x	
päämiehen nimi				(x)
aviopuolison nimi		x	x	
isän/äidin nimi	x		x	
muuttopäivämäärä				x
muuttanut jonnekin				x
muuttanut jostakin				x
kuolinpäivämäärä			x	
vihkimispäivämäärä		x		

(x):llä merkityt muuttujat saattavat puuttua

3.2.2 Linkityksen lähestymistavat

Tutkijat ovat käyneet innokkaasti keskustelua linkityksen filosofiasta: pyrkiäkö optimaaliseen täsmällisyyteen vai täydellisyyteen eli siihen, että päästään mahdollisimman lähelle lähdeaineiston linkkien kokonaismäärää. Cambridgen ryhmän edustama lähestymistapa vaatii, että vain vahvimmat ja varmimmat linkit valitaan, vaikka joitakin mahdollisesti tärkeitä tietueita jäisi yhdistämättä. Englantilainen tutkija Skolnick kannattaa toista lähestymistapaa, jonka mukaan pitäisi varmistaa, että muodostettujen linkkien lukumäärä on mahdollisimman lähellä lähdeaineiston linkkien kokonaismäärää, vaikka se saavutettaisiinkin

⁶⁶ Danell 1985, s. 122-124

hyväksymällä virheellisiä linkkejä. Ensimmäisessä lähestymistavassa pitää varautua vinoutumiin silloin, kun halutaan tutkia muuttoliikettä tai demograafista käyttäytymistä. Jokainen linkittämätön tietue johtaa samanaikaisesti virheellisen perhetietueen muodostamiseen ja toiseen epätäydelliseen perhetietueeseen. Toisessa lähestymistavassa on vaarana, että yksilöllisten elämäkertojen, sukulaisuuden ja geneettisten tosiasioiden analysointi estyy, kun perhetietueisiin tuodaan syntymä-, kaste- ja hautaustapahtumia, jotka eivät niihin kuulu. Etusija on pantava joko linkkien luotettavuudelle ja täsmällisyydelle tai kokonaismäärälle ja kattavuudelle. Tietuelinkityksen määrittelyä kahdella erilaisella, ilmeisen vastakohtaisella tavalla ei ole mahdollista ratkaista teoreettisin perustein, vaan määrääviä tekijöitä ovat tietokannan rakentamisen päämäärät ja lähdemateriaalin ominaisuudet. Ristiriita rajoittaa tietokantojen soveltuvuutta hyvin valikoiviin tarkoituksiin.⁶⁷

Tutkijat ovat yhtä mieltä siitä, että kaikkien tietueiden vertailu on mahdotonta silloin, kun tiedon määrä on suuri.⁶⁸ Linkitysprosessin tehokkuuden parantamiseksi on välttämätöntä vähentää tiukan vertailun läpikäyvien parien määrää. Periaatteena pidetään, että ensin muodostetaan helpoimmat ja luotettavimmat linkit, joissa tunnistuskriteerit vastaavat toisiaan täysin, ja sitten epävarmemmat linkit väljentämällä kriteerejä.⁶⁹ Vaikeimpien tapausten selvittäminen edellyttää menettelytapaa, joka takaa, että kaikki linkitykseen ehdolla olevat tietueet sovitetaan yhteen ja vertaillaan. Näitä tapauksia vertailtaessa tietueet lajitellaan linkitysmuuttujien erottelevia ominaisuuksia käyttäen pienempiin osajoukkoihin siten, että kaikki tietueet, jotka teoriassa ovat ehdokkaina yhdistettäväksi, löytyvät samasta osajoukosta. Tämän osajoukon sisällä arvioidaan tietueiden kaikkien ominaisuuksien yhteen sopivuutta. Arvioinnissa voidaan apuna käyttää painokertoimia ja todennäköisyyslaskelmia tai muita normi-, kriteeri- ja hierarkia-

⁶⁷ Bouchard 1986

⁶⁸ Bouchard 1986, Winchester 1992

⁶⁹ Guth 1976, Schofield 1992

joukkoja.⁷⁰

Tarvitaan myös päätöksentekomenetelmä, joka valitsee parhaat linkit ottaen huomioon tutkimuksen kontekstin ja tavoitteet.⁷¹ Päätöksenteon luotettavuutta voidaan arvioida erilaisten pisteitysmenetelmien avulla. Korkeimmat pistemäärät johtavat linkkien muodostamiseen. Maininnat henkilökohtaisista ominaisuuksista toimivat luottamuksen mittaajina. Jokaiselle tietueparille annetaan pisteitä esimerkiksi iän, ammatin, asuinpaikan, sukulaisten nimien ja niiden ristiriidattomuuden perusteella. Pisteitys voi tapahtua automaattisesti tai sen voi suorittaa tutkija ottaen huomioon erityisen historiallisen kontekstin tuntemuksen.⁷² Suurimmat linkkien muodostamat verkot eli klusterit ovat vaikeimpia. Skolnick pyrkii säilyttämään klusterissa kaikki maksimipisteet saaneet linkit. Hän esittää, että matalimmat luottamuspisteet omaavat linkit olisi tarpeen tuhota verkon selvittämiseksi. Toiset tutkijat, esimerkiksi Cambridgen ryhmään kuuluva Schofield, eivät tätä lähestymistapaa kannattaneet. He vaativat, että muodostetaan myös heikommät linkit, jotka ovat loogisia vahvimman, hyväksytyin linkin kanssa. Painokertoimia käyttäessään Cambridgen ryhmä antaa etusijan varmemmille linkeille ja linkkiverkko hajoitetaan hitaasti tuhoamalla vahvempien linkkien kanssa yhteen sopimattomat linkit. Menetelmän valinta riippuu yleensä tutkimuksesta. Järkevin tapa ratkaista ongelma on lykätä linkityspäätöstä kaikissa epäselvissä tapauksissa ja yrittää hankkia lisää informaatiota.⁷³

Käytännössä useilla linkeillä on sama pistemäärä. Tällaisissa tapauksissa annetaan etusija tietueille, joilla on vähemmän kilpailevia linkejä. Lopulta, jos tietueilla on sama pistemäärä ja sama määrä kilpailevia linkejä, silloin ensisijaisen linkin valinta pitää tehdä umpimähkään. Kun yksi linkki on hyväksytty,

⁷⁰ Bouchard 1986, Schofield 1992

⁷¹ Bouchard 1992, Schofield 1992

⁷² Winchester 1992

⁷³ Schofield 1992, Winchester 1992

edetään käsittelemään seuraavaksi luotettavinta linkkiä ja niin edelleen. Jos jokin ominaisuus järjestelmällisesti sisältää enemmän tietoa kuin muut, se saa suhteellisesti enemmän pisteitä.⁷⁴

Tutkijan on ratkaistava käyttääkö osittain vai täysin automaattista linkitystä. Vastaus riippuu esimerkiksi tiedostojen koosta, linkitettävien tietueiden määrästä ja identifiointiin käytettyjen muuttujien puuttumisesta tai niiden sisältämistä ristiriitaisuuksista. On kaksi tapaa järjestää tietueet: kortti-indeksimenetelmä ja osoittimilla varustettujen listojen käyttö. Automaattisessa linkityksessä voidaan käyttää osoittimilla varustettuja listoja. Tietueet yhdistetään loogiseen järjestykseen käyttämällä osoittimia, jotka viittaavat seuraavaan tietueeseen. Tietueen lisääminen keskelle tiedostoa vaatii kopioimista ja väliaikaista varastointia. On kuitenkin joitakin listojen prosessointiin soveltuvia relaatiotietokantaohjelmia. Ruotsalaiset ja suomalaiset rippikirjat ovat näiden kahden yhdistelmä; niissä on käytetty yksinkertaista osoittimien järjestelmää, joka helpottaa linkitystä.⁷⁵

3.2.3 Bouchardin linkitysmenetelmä

Bouchardin ryhmä Kanadassa on kehittänyt ranskankielisten historiallisten lähteiden käsittelyyn SOREP-järjestelmän. Bouchardin menetelmä käsittää neljä perättäistä päävaihetta, joissa linkitys etenee johdonmukaisesti. Taulukko 4 antaa yleiskuvan SOREP-järjestelmästä, sen eri vaiheiden tehtävistä, niissä käytetyistä ohjelmista sekä ohjelmien suorittamista toimenpiteistä. Ensimmäisessä vaiheessa tiedostoille suoritetaan pintapuolinen lajittelu. Lajitteluun käytetään Bouchardin suunnittelemaa Fonem-ohjelmaa, joka normalisoi nimet ja poistaa siten nimivariaatiot. Toisessa ja kolmannessa vaiheessa vertaillaan vihittyjen luetteloiden merkintöjä nimien perusteella. Nimitieto on riittävän erottelevaa,

⁷⁴ Schofield 1992, Winchester 1992

⁷⁵ Winchester 1992

sillä vertailtavia yksityiskohtia on kahdeksan, kuten nähdään taulukosta 5.⁷⁶

Toisessa vaiheessa suoritetaan seitsemän lajittelua. Jokaisen lajittelun aikana Couple-ohjelma muodostaa merkinnöistä ryhmiä, joihin kuuluvat henkilöt nimien perusteella mahdollisesti sopisivat aviopareiksi. Ensin ohjelma yhdistää vihittyjen luetteloiden perustella aviopareiksi sellaiset lähteiden merkinnät, joissa nimet ovat täysin samoja. Seuraavaksi sovitetaan yhteen kirjauksia, joiden nimissä on kaksi tai kolme yhteen sopivaa yksityiskohtaa neljästä. Samalla selvitetään tapaukset, joissa tavataan ristiriitaisia yksityiskohtia. Tässä vaiheessa muodostetaan kahdesta merkinnästä A ja B aviopari välittömästi, jos ne tyydyttävät linkitykselle asetetut vaatimukset. Jokaista merkintää ei siis välttämättä

Taulukko 4 SOREP-järjestelmän linkitysprosessin kulku

	Alijärjestelmät	Ohjelmat	Ohjelmien käyttö
	Kirkonkirjat		
Vaihe 1	Alustava prosessi	NETTOIE FONEM INT	1 Lajittelu 2 Yhteen- sovittaminen
	Vihittyjen luettelot		
Vaihe 2	Perherakenteen muodostaminen	COUPLE	3 Vertailu
	Perhetietueet (a)		4 Päätöksenteko (Linkitys)
Vaihe 3	Perherakenteen muodostaminen	SUB RETRI	5 Eheys
	Perhetietueet (b)		6 Manuaalinen puuttuminen
Vaihe 4	Yksilöelämän- kertojen muodostaminen	MATRICE INVER	7 Tilastot

⁷⁶ Bouchard 1986, Bouchard 1992

Taulukko 5 Avioparien merkintöjen vertailu

	Mies		Vaimo	
	1 Sukunimi	2 Etunimi	3 Sukunimi	4 Etunimi
Merkintä A	Tremblay	Alfred	Simard	Marie
Merkintä B	Tremblay	Maurice	Levesque	Marie

verrata kaikkiin muihin. Kolmannessa vaiheessa suoritetaan kaksitoista täydentävää lajittelua. Niissä muodostetaan henkilöryhmiä, joiden nimitiedoille asetetaan väljemmät kriteerit. Riittää, kun neljästä nimestä yksi on sama sekä ensimmäiset kolme kirjainta toisesta. Tällöin päätös, että merkinnät koskevat samaa avioparia ei voi perustua ainoastaan nimien samanlaisuuteen, vaan on turvauduttava muihin yksityiskohtiin, esimerkiksi tietoihin asuinpaikasta ja ammatista. Nämä vaiheet kokoavat samaan tietueeseen tietyn avioparin kaikki merkinnät, joita yhdellä avioparilla saattaa olla jopa 30-35. Tähän perhetietueeseen liitetään myös näihin merkintöihin sisältyvät lapsia koskevat tiedot. Neljännessä vaiheessa muodostetaan yksilöiden elämäkertoja; siinä yhdistetään kasteluetteloiden tietoja avioliitto- ja hautausluetteloiden tietoihin. Matrice-ohjelma tekee kaikki mahdolliset linkit vertailemalla järjestelmällisesti jokaista kastemainintaa avioliitto- ja hautausmerkintöihin. Tässä vaiheessa muodostetaan linkkiverkkoja eli klustereita ja päätöksenteossa otetaan huomioon kaikki keskenään kilpailevat vaihtoehdot, joita arvioidaan nimien lisäksi muidenkin ominaisuuksien kuten iän ja sukupuolen samanlaisuuden perusteella. Käytännöllä hierarkkista taulukkoa, joka perustuu vertailtujen etunimien samankaltaisuuteen, on muodostettu järjestelmä, joka antaa painokertoimen jokaiselle mahdolliselle linkille (0-6). Linkeille annetut pisteet viedään matriisitaulukkoon, jonka perusteella päätökset tehdään. Joissakin tapauksissa ohjelma ei pysty ratkaisemaan epäselvyyksiä, jotka silloin ohjataan tutkijalle manuaalista päätöksentekoa varten.⁷⁷

⁷⁷ Bouchard 1986, Bouchard 1992

Järjestelmän periaatteena on, että ensin käsitellään varhimmat tilanteet, joihin liittyvät helpoimmat ja luotettavimmat linkit kuten avioparimerkinnät. Monimutkaisimmat tilanteet lykätään viimeiseen vaiheeseen, jossa käsitellään klustereita. Siten linkityssäännöt muuttuvat vaiheesta toiseen mentäessä lähdeaineiston laadun ja luonteen mukaan. Avioparimerkinnöissä verrataan neljää eri nimiparia toisiinsa, muissa kirkonkirjoissa vain kahta ja siksi edellisten suhteen linkitys voidaan tehdä paljon varmemmalta pohjalta kuin jälkimmäisissä. Automaatio on viety niin pitkälle kuin mahdollista, mutta siitä huolimatta on pakko jättää tutkijalle paljon valtaa etenkin loppuvaiheen päätöksiä tehtäessä.⁷⁸

3.2.4 Demografiska Databasen linkitysmenetelmä

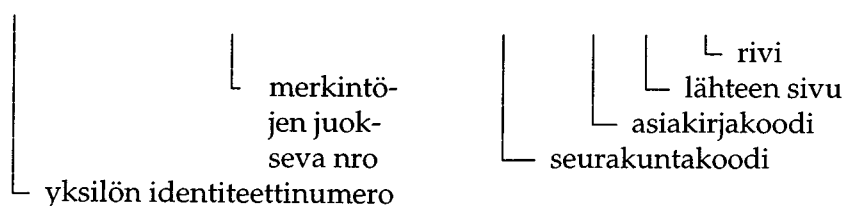
DDB:ssä kaikki samaan henkilöön liittyvät merkinnät etsitään linkitysmuuttujien avulla vuoronperään jokaisesta tiedostosta ja tallennetaan omana tietueenaan linkkitiedostoon. Henkilön merkinnät löytyvät linkkitiedoston riveiltä henkilötunnuksen avulla. Henkilöiden välinen linkitys toteutetaan vasta sen jälkeen, kun yhtä henkilöä koskevat merkinnät on linkitetty. Periaatteena on, että linkitys tapahtuu linkkimuuttujien avulla ja tiedot linkityksestä viedään suhdetiedostoon, joka sisältää henkilöiden väliset yhteydet.⁷⁹

DDB:ssä yksilön merkintöjen yhdistäminen aloitetaan antamalla henkilölle ainutkertainen henkilötunnus kuten nähdään kuvasta 4. Tällä tunnusnumerolla ei ole yhteyttä henkilön muihin tietoihin, millä taataan, ettei numeroa tarvitse korjata, vaikka tiedoissa havaitaankin virheitä. Merkintöjen saaminen kronologiseen järjestykseen varmistetaan käyttämällä tapahtumien juoksevilla numeroilla runsaasti väljyyttä, jolloin uuden merkinnän lisääminen ei aiheuta aikaisempien

⁷⁸ Bouchard 1986, Bouchard 1992

⁷⁹ Demografiska Databasen 1990

Henkilö- tunnus	Merkintöjen järjestysnro	Lähdeviittaus	EHM	LÄNKANF
0000231	001000000000	83520 301 0023 02		110110
0000231	002000000000	83520 010 0425 04		000011
0000231	002500000000	83520 201 0123 11	1	010010
0000231	003000000000	83520 014 0008 22		010100
0000231	003500000000	83520 503 0077 23	1	010101
0000231	004000000000	83520 203 0023 07		111000



Asiakirjakoodit:

0xx = rippikirjat

2xx = muuttaneiden luettelot

3xx = syntyneiden luettelot

5xx = vihittyjen luettelot

EHM -sarake ilmoittaa henkilön suhteen tapahtumaan. Ykkönen (1) tässä sarakkeessa merkitsee, ettei yksilö ole tämän tapahtuman päämies. Kirkonkirjojen merkintä voidaan liittää useisiin yksilöihin, kuten tiedot muutoista tai kuulutuksista ja vihkimisestä.

LÄNKANF -sarake ilmoittaa millä perusteella merkintä liitetään yksilöön. Jokaisella positiolla on oma merkityksensä; se voi esimerkiksi ilmoittaa onko kyseessä koneellinen vai manuaalinen linkitys.

Kuva 4 Linkitystiedoston käyttö DDB:ssä yksilön tietoja yhdistettäessä.

juoksevien numeroiden muuttamista. Linkitystiedostojen lähdeviittaustietueet on varustettu hakemistolla, jotta väärät linkitykset voitaisiin helposti purkaa.⁸⁰

Henkilöiden välisten suhteiden linkittämisessä käytetään apuna suhdetiedostoja kuvan 5 osoittamalla tavalla. Kahden henkilön välinen suhde esiintyy suhdetiedostossa vain kerran. Henkilötunnuksia (henkilötunnus1 ja henkilötunnus2) sisältävät tietueet on indeksoitu eli varustettu hakemistolla, joka sisältää osoitteet tietueisiin. Suhdetietojen päivitys tehdään vain yhteen paikkaan tiedostossa.⁸¹

⁸⁰ Demografiska Databasen 1990

⁸¹ Demografiska Databasen 1990

Henkilötunnus1	Suhde	Henkilötunnus2
00000010	301	00000014
00000010	101	00000231
00000014	203	00000231

suhdekoodi		
100	isä, luultavasti biologinen	suhde- koodeja
101	biologinen isä	
102	kasvatusisä	
.		
.		
200	äiti, luultavasti biologinen	
.		
.		
300	yhteisiä lapsia	
301	vihityt	
302	yhdessä asuvia	
303	vain kuulutettuja	
.		
.		
501	sisaruksia	

Kuva 5 Suhdetiedoston käyttö henkilöiden välisiä suhteita linkitettäessä

Aputiedostot, joista on esimerkki kuvassa 6, sisältävät linkityksessä tarpeellista tietoa, esimerkiksi päivämääriä ja standardoituja nimiä. Kun merkintä on viety linkkitiedostoon, päivitetään aputiedostoja tiedoilla, joita siellä ei vielä ole. Mitä useampia merkintöjä on linkitetty sitä käyttökelpoisemmiksi aputiedostot tulevat. Ne tukevat sekä koneellista että käsin tehtyä linkitystä.⁸²

Henkilötunnus	Koodi	Koodin selitys
00000231	302 755 12 01	synnyinpvm
00000231	801 PERNSDOTTER	sukunimi
00000231	802 KARIN	etunimi
00000231	502 778 06 23	vihkimisen pvm
00000231	202 83520 83540 780	muuttoseurakunnat, vuosi

Kuva 6 Aputiedoston käyttö linkityksen yhteydessä.

⁸² Demografiska Databasen 1990

Alunperin linkitys tapahtui DDB:ssä manuaalisesti, siten että yhdistettiin saman henkilön tietoja sisältävät reikäkortit yhdeksi nipuksi. Linkitys on monimutkainen toimenpide eikä DDB:ssä ole päästy täysin koneelliseen linkitysjärjestelmään.

Karjala-tietokannassa harkittiin samantapaista linkitysmenetelmää kuin DDB:ssä, mutta tietosuojalautakunnan kielteinen kanta on toistaiseksi estänyt Karjala-tietokannassa yhtenäisen väestötietokannan toteuttamisen sekä henkilöihin ja henkilöiden välisiin suhteisiin liittyvien tietojen linkittämisen.⁸³ Karjala-tietokannassa on kehitetty lähdetietokantojen hyväksikäyttöjärjestelmä, jossa kunkin lähdetyyppin merkintöjen perusteella on muodostettu kuusi rekisteriä. Lisäksi on muodostettu henkilöhakemisto-rekisteri, joka sisältää seurakunnittain perustiedot kaikista tallennetuista tapauksista. Henkilöhakemistoon on tallennettu etunimi, patronyyminimi, sukunimi, os. sukunimi, sukupuoli, syntymäaika, kirkonkirjasivu, signum, sidoskoodi ja henkilönumero. Henkilöhakemiston tietojen, dBase IV:n ohjauskeskuksen ja kyselykielen avulla voi käyttäjä tehdä henkilön ja henkilöiden välisiin suhteisiin liittyvien tietojen linkitystä esimerkiksi samoja linkitysmuuttujia käyttäen kuin DDB:ssä.⁸⁴

⁸³ Karjala-tietokanta 1993

⁸⁴ Karjalan väestötieteellinen tietokanta 1992

4 HISTORIALLIISIIN VÄESTÖLUETTELOIHIN PERUSTUVIA TIETOKANTOJA

Tässä luvussa luodaan lyhyt katsaus henkilöhistorian tietokantoihin, joihin on tallennettu erilaisten henkilöhistoriallisten väestöluetteloiden merkintöjä. Seuraavissa luvuissa keskitytään tutkimaan tarkemmin kolmea esimerkkiä, jotka on toteutettu relaatiotietokantoina ja joiden pääasiallisen lähdeaineiston muodostavat kirkonkirjat. Historiallisia väestöluetteleita alettiin tallentaa tietokoneille 1970-luvulla.⁸⁵ Henkilöhistorian tietokantoja rakentavien projektien toteutuksessa on käytetty sekä itse kehitettyjä ohjelmistoja että valmiita kaupallisia järjestelmiä, joita saatettiin muokata omaan käyttöön soveltuviksi. Varhaisimmissa projekteissa käytettiin suur- tai pientietokoneita, mutta mikrotietokoneiden nopea kehitys teki niiden käytön mahdolliseksi henkilöhistoriallisten tietokantojen toteutuksessa. Oman ohjelmiston kehittivät esimerkiksi Cambridge Group Englannissa ja Max Planck -Instituutti Saksassa. Kaupallista ohjelmistoa osittain omalla ohjelmistolla täydennettynä käyttivät kanadalainen SOREP sekä ruotsalainen DDB. Mikrokoneille soveltuvilla kaupallisilla ohjelmistoilla toteutettiin amerikkalaiset Vetterin, Gonzalesin ja Gutmannin, Atackin, Batemanin ja Gregsonin sekä Gillespien projektit ja suomalainen Karjala-tietokanta.

Englantilainen Cambridge Group rakensi täysin automaattisen ohjelmiston, jossa kaikki, myös vaikeat päätökset tietueiden linkityksestä, tehdään tietokoneen avulla. Lähdemateriaali käsittää kirkollisten luetteloiden merkintöjä, jotka yhdistettiin yhtenäiseksi tietokannaksi.⁸⁶ Cambridge Groupin tavoitteena oli muuttaa kirkollisten luetteloiden merkinnät kastetuista, vihityistä ja haudatuista todellisiksi, koko maan käsittäviksi syntyneiden, vihittyjen ja kuolleiden kokonaisuudeksi, minkä jälkeen näiden lukujen perusteella olisi mahdollista arvioida maan väkilukua sekä syntyvyys- ja kuolleisuustasoa eri ajanjaksoina. Parikym-

⁸⁵ Winchester 1992

⁸⁶ Schofield 1992

mentä vuotta kestänyt kehitystyö tuotti joukon tietuelinkitysohjelmia, joiden avulla tietokanta rakennettiin. Ryhmä tallensi kolmentoista Oxfordshiren seurakunnan ja yhden lontoolaisen seurakunnan tietoja. Lontoolaisen seurakunnan tietoihin kuului 40 000 kaste-, 59 000 hautaus- ja 20 000 avioliittomerkintää vuosilta 1550-1753. Kaiken kaikkiaan tehtiin noin miljoona linkkiä ja noin 40 000 perherakennetta. Tietokannan tietoja voidaan yhdistää muihin paikallisiin tiedostoihin ja käyttää niiden käsittelyssä joko ryhmän kehittämiä tai muita ohjelmia.⁸⁷

Max Planck -Instituutissa Saksassa kehitettiin Kleio-ohjelmisto, joka on tarkoitettu historiallisen lähdeaineiston käsittelyyn. Ohjelmiston tavoitteena on tallentaa lähteet muodossa, joka muistuttaa alkuperäisten lähteiden muotoa ja ottaa huomioon historiallisen tiedon ominaislaadun. Kleio-ohjelmisto käsittelee tietokantaa kokonaisuutena ja sen tietomalli muistuttaa semanttista verkkoa. Yksittäiset tiedot on linkitetty taustatietoon, jossa niitä voidaan käsitellä kontekstiin liittyvinä. Tietoja käsitellään tekstipalasten kokoelmana ilman, että tehdään oletuksia niiden merkityksestä. Komentokieli on olio-orientoitunut⁸⁸. Ohjelmistolla voidaan hakea tietoa, jolla on erityinen symbolinen nimi, rakenteessa määrätty sijainti, tietty sisältö tai yhteys yhteisen sisällön kautta. Tavoitteena oli myös, että tieto on tallennettava niin tarkoin alkuperäisen lähteen mukaisena kuin mahdollista.⁸⁹ Esimerkiksi alkuperäinen kirjoitusasu oli säilytettävä silloin, kun on epäilyjä sanan tai tekstin merkityksestä.⁹⁰ Kleio-ohjelmistoa on sovellettu useisiin saksalaisiin, hollantilaisiin ja ranskalaisiin lähteisiin. Esimerkkinä voidaan esittää Ranskassa Montperlierissä Kleio-ohjelmistolla rakennettu historiallinen tietokanta, johon tallennettiin erään etelä-ranskalaisen alueen kastettujen, vihittyjen ja kuolleiden luetteloiden merkintöjä vuosilta 1750-1850.

⁸⁷ Schofield 1987

⁸⁸ Genet, Zampolli 1991

⁸⁹ Thaller 1992

⁹⁰ Thaller 1989

Materiaali käsittää kaiken kaikkiaan 13 miljoonaa merkintää. Kleio-ohjelmisto suorittaa tietueiden linkityksen nimien perusteella. Lähdemateriaalista on tarkoitus muodostaa tietokanta, josta saadaan selville yksilöiden ja perheiden tiedot ja yhdistetään nämä alueelliseksi perhehistoriaksi.⁹¹

SOREP, joka on kanadalaisten yliopistojen välinen väestöntutkimuskeskus, on kehittänyt tietokonepohjaisen järjestelmän, joka tallentaa kirkollisten luetteloiden merkintöjä ja muodostaa perherakenteita. SOREP:n tallentama tietokanta käsittää 1.5 miljoonaa kaste-, avioliitto- ja kuolinmerkintää Quibecin maakunnan alueelta. Tarkoituksena on lisätä tietokantaan vielä koko Quibecin alueen aviopari-merkinnät 1800-1900-luvuilta. Tietokanta on toteutettu Vax-8350 tietokoneella ja Ingres-tietokannan-hallintajärjestelmää käyttäen.⁹²

Ruotsissa Uumajan yliopistossa 1972 käynnistetty Demografiska Databasen on Pohjoismaiden laajin kirkonkirjoja tallentava projekti. Kirkonkirjoja tallennetaan vuosilta 1820-1900. Ensimmäiseen tallennusvaiheeseen tutkijat valitsivat seitsemän seurakuntaa eri puolilta Ruotsia oman kiinnostuksensa perusteella. Myöhemmin valittiin lisää kolme yhtenäistä aluetta, sillä muuttaminen seurakunnasta toiseen oli estänyt luomasta täydellisiä elämäkertoja.⁹³ Vuoden 1989 loppuun mennessä oli tallennettu noin 250 000 yksilön tiedot (mikä käsittää noin 1,2 miljoonaa kirkonkirjamerkintää) ja vuoden 1991 loppuun mennessä noin 365 000 yksilön tiedot.⁹⁴ Ruotsissa on toinenkin merkittävä henkilöhistoriallinen tietokanta, joka ei kuitenkaan ole aivan yhtä laaja kuin DDB. Siihen on tallennettu Tukholman kaupungin manttaalikirjat, jotka sisältävät tietoja kaupungin kiinteistöistä ja niiden asukkaista 1800-luvun loppupuolella ja 1900-luvun alussa. Manttaalikirjoihin merkittiin kiinteistöjen asukkaiden nimet, ammatit, syntymä-

⁹¹ Smets 1987

⁹² Bouchard 1992

⁹³ Könberg 1989

⁹⁴ Preliminary Draft 1989

paikat ja -ajat, perheet palvelijoihin, muutot ja kuolemat. Manttaalikirjoissa on myös yhteiskunnallista tietoa, kuten merkintöjä varallisuudesta ja mahdollisista rangaistuksista. Ne muistuttavat sisällöltään suuresti kirkonkirjoja, mutta niissä on tietoja myös kiinteistöistä ja maasta. Kaikki manttaalikirjojen merkinnät on tallennettu lähdetietokantaan, joka on relaatiotietokanta. Jotta haut olisivat nopeampia ja tehokkaampia, on luotu joukko rekistereitä, joista on osoittimet lähdetietokantaan.⁹⁵

Edellä mainitut projektit on toteutettu suurtietokoneella. Mikrotietokoneet ovat viimeisten 20 vuoden aikana kehittyneet huomattavasti ja ne on otettu käyttöön myös historiallisten väestöluetteloiden tallennuksessa. Lähteessä, Winchester 1992, esitellään amerikkalaisten tutkijoiden Vetterin, Gonzalesin ja Gutmannin projekti. He käyttivät mikrolla toimivaa kaupallista relaatiotietokannan hallintajärjestelmää, johon yhdistettiin jonkin verran omaa suunnittelua.⁹⁶ Heidän esimerkkiään seurasivat Atack, Bateman ja Gregson, joiden tutkimuksessa yhdistettiin 140 pohjoisamerikkalaisen kunnan väestö- ja maanviljelyluetteloiden tietoja vuosilta 1860 ja 1880. Päämääränä oli saada aikaan yhtenäinen, koneella luettavissa oleva tietokanta, jonka avulla voidaan selvittää maanviljelyalueiden kehitystä 1800-luvun loppupuolella. Projektissa käytettiin PC Matchmaker-ohjelmistoa ja mikrotietokonetta.⁹⁷

Leboutten, Alterin ja Gutmannin projektissa tallennetaan Louis Henryn perhelomakemenetelmää käyttäen kerättyä belgialaista aineistoa, joka käsitti 3.600 perherakennetta. Tutkimuksessa käytettiin SAS-ohjelmistoa. Sen avulla oli mahdollista tuottaa kaikki perustaulukot, joita historiallisen demografian elämäkerroissa käytetään, ja säilyttää yleisen tilastollisen ohjelmiston joustavuus ja voima. Tämän ohjelmiston edistyneitä piirteitä tarvittiin perherakennetiedon

⁹⁵ Lemming 1983, Fogelvik, Sperlings 1982

⁹⁶ Winchester 1992

⁹⁷ Atack, Bateman, Gregson 1992

monimutkaisimpien yksityiskohtien käsittelyyn. SAS-ohjelmisto tuotti tilastollisen analyysin kaikesta tiedosta, joka yleensä oli löydettävissä Louis Henryn perhelomakkeelta. Tiedot syötettiin Henryn perhelomakkeessa esitetyssä järjestyksessä. Perhelomake on jaettu seitsemään lohkokon, joista jokainen käsittää useita kenttiä. Kenttien tietosisältö on määritelty.⁹⁸

Amerikkalainen Gillespien projekti tallensi vuosien 1860, 1870, 1880, 1900 ja 1910 väestölaskentaluetteloita, veroluetteloita sekä kirkon kaste-, avioliitto- ja hautausluetteloita. Erillisinä luetteloiden tiedoilla on rajoituksensa, mutta yhdessä niistä saadaan monipuolinen kuva väestön ominaisuuksista. Projekti toteutettiin mikro-tietokoneella Foxbase+ -relaatiotietokannan hallintajärjestelmää käyttäen. Yhtenäinen järjestelmä tarjosi tutkijoille monia mahdollisuuksia. Voitiin esimerkiksi tutkia syntyvyyttä luku- ja kirjoitustaidon, ammatin tai muiden muuttujien suhteen. Tietojen yhdistäminen teki mahdolliseksi muodostaa perheitä Louis Henryn perhelomakkeen mallin mukaan.⁹⁹

Koska suomalaiset kirkonkirjat ovat samanlaisia kuin ruotsalaiset, ryhdyttiin täälläkin DDB:n esimerkin innoittamana selvittämään mahdollisuuksia luoda samankaltainen suomalaiseen kirkonkirjamateriaaliin perustuva tietokanta.¹⁰⁰ Suomen henkilörekisterilaki aiheutti kuitenkin ongelmia Karjala-tietokannan rakentamiselle. Koska henkilörekisterilaki säätelee Karjala-tietokannan rakentamista varten perustetun tietokantasäätötoimintaa, anoi tämä tietosuojalautakunnalta poikkeuslupaa tutkimukseen tarkoitetun, kirkonkirja-aineistoon perustuvan henkilörekisterin perustamiseen. Tietosuojavaltuutettu hylkäsi anomuksen, koska ei mielestään nähnyt painavaa syytä rekisterin perustamiseksi. Tietokantasäätö haki kuitenkin muutosta päätökseen korkeimmasta hallinto-

⁹⁸ Leboutte, Alter, Gutmann 1987

⁹⁹ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

¹⁰⁰ Karjalan väestötieteellisen tietokannan loppuraportti 1989

oikeudesta.¹⁰¹ Päätöksessään 9.11.1992 KHO katsoi, että rekisterin perustamiselle on olemassa henkilörekisterilain mukainen painava syy ja palautti asian tietosuojalautakunnalle uudelleen käsiteltäväksi.¹⁰² Tietosuojalautakunta myönsikin 30.11.1992 poikkeusluvan luovutetun Karjalan evankelis-luterilaisten seurakuntien tietojen keräämiseen ja tallentamiseen Karjala-tietokantaan. Tietosuojalautakunnan päätöksen mukaan Karjala-tietokantaan voidaan tallentaa kirkonkirjojen tietoja 1700-luvulta vuoteen 1949 saakka lukuunottamatta arkaluonteisina pidettäviä tietoja. Tietojen käyttö tutkimuksen lähdeaineistona on mahdollista ja sallittua. Henkilörekisterilakiin ollaan suunnittelemassa muutoksia ja tutkijoiden piirissä toivotaan lain antavan mahdollisuuden kerätä tutkimuksen tarkoituksiin tavanomaisia henkilötietoja tarvitsematta anoa poikkeuslupaa tietosuojalautakunnalta. Tämä merkitsisi näiden toimintojen yleishyödyllisen merkityksen tunnustamista.¹⁰³ Tietokannan rakentamisessa oli määrä käyttää esikuvana DDB:n kehittämää järjestelmää ja sille asetettuja tavoitteita. Ruotsin DDB:n mallin mukaisesta väestötietokannasta kuitenkin luovuttiin ja ryhdyttiin kehittämään lähdetietokannan käyttöä. Vuonna 1992 valmistui lähdetietokannan käyttäjäjärjestelmä, joka mahdollistaa sekä alkuperäisen että standardoidun aineiston käytön tutkimukseen. Karjala-tietokanta toteutettiin mikrotietokoneelle ja tietokantaohjelmisto dBase IV:lla.

¹⁰¹ Karjala-tietokanta 1993

¹⁰² Autio1993a

¹⁰³ Autio 1993b, Hiltunen 1993

5 KIRKONKIRJOIHIN PERUSTUVIA RELAATIOTIETOKANTOJA

Seuraavaksi tässä tutkimuksessa tarkastellaan ja kuvataan yhdenmukaisesti kolmea historiallista tietokantaa, joiden aineiston muodostavat pääasiassa kirkonkirjojen merkinnät. Demografiska Databasen on toteutettu pien-tietokoneella¹⁰⁴, Karjala-tietokanta¹⁰⁵ ja Gillespie -järjestelmä mikrotietokoneella¹⁰⁶. Tietokannat on pääasiallisesti toteutettu kaupallisilla relaatiotietokannan hallintajärjestelmillä; DDB:ssä käytetään lähdemerkintojen tallennukseen itse suunniteltuja ohjelmia. DDB käyttää SGL/DS -, Karjala-tietokanta dBase IV - ja Gillespie-järjestelmä Foxbase+ -ohjelmistoa. Aluksi esitetään järjestelmien tavoitteet ja tietokannan kehitystyön vaiheet sekä kuvataan rakenteet, joihin lähteiden merkinnät tallennetaan. Samoin selvitetään toimenpiteitä, jotka ovat tarpeen siirrettäessä historiallisen alkuperäislähteen merkintöjä tietokoneella luettavaan muotoon. Viimeksi mainittu osuus jää tosin Gillespie-järjestelmän osalta niukaksi, koska siitä ei lähteissä löydy tietoja. Seuraavaksi esitetään järjestelmien tietosisältö ja tietosisältöjen vertailu. Tallennukseen käytetyt rakenteet eivät hyvin sovellu tietojen hakuun, vaan niitä on muokattava edelleen. Seuraavissa luvuissa kuvataankin rakenteita ja hakujärjestelmiä, joita tietojen hakua varten on kehitetty. Tutkimuksessa pyritään osoittamaan, että henkilöhistorian tietokantoihin on mahdollista tallentaa kaikki alkuperäisten lähteiden tiedot ja että tallennukseen käytettyjä rakenteita edelleen muokkaamalla voidaan kehittää rakenteet ja hakujärjestelmät, jotka tukevat tehokasta ja joustavaa tiedon hakua. Tietokantojen kuvauksella pyritään päätelmään, että tietokannat tukevat eri tieteenalojen tutkimusta ja että kaikki tieto, mikä alkuperäisistä lähteistä löytyy, on tutkijoiden käytettävissä, mutta vaivattomammin ja tehokkaammin kuin alkuperäisiä lähteitä käyttäen. Tietokantojen käyttäjille asettavat rajoituksia

¹⁰⁴ Könberg 1989

¹⁰⁵ Karjalan väestötieteellisen tietokannan loppurapotti 1989

¹⁰⁶ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

lähdeaineiston tietosisältö sekä se, että tietokannan rakentamisvaiheessa eivät kaikki tutkijoiden tarpeet ja kysymyksenasettelut olleet tiedossa. Ensiksi esitetään kuitenkin tutkimuksessa käytettyjä relaatiotietokantaan liittyviä käsitteitä.

5.1 Relaatiotietokantamäärittelyt

Tässä tutkimuksessa *entiteetillä* ymmärretään rajattavissa olevaa ilmiötä, joka kuuluu muiden samankaltaisten ilmiöiden joukkoon. Entiteeteillä on *ominaisuuksia*. Ominaisuuden *nimeä* ja siihen liittyvää *arvojoukkoa* nimitetään *attribuutiksi*. Samankaltaiset entiteetit muodostavat *entiteettijoukkoja*. Niiden yhteisten ominaisuuksien määritykset muodostavat *entiteettityypin*. Entiteettityypin kaikkiin entiteetteihin liitetään samat attribuutit. Entiteetit täytyy pystyä erottamaan toisistaan ja siksi entiteettityypin liittyvät attribuutit on määriteltävä niin, että niiden arvoilla pystytään erottamaan entiteetit toisistaan. Entiteettityypin *avaimella* ymmärretään sellaista attribuuttien joukkoa, joiden arvoilla entiteetit pystytään erottamaan toisistaan. Entiteettityyppien E_1, \dots, E_k välillä voi vallita *suhde*. Suhteiden yhteisten ominaisuuksien määritykset muodostavat *suhdetyypin*. Suhdetyyppejä luokitellaan sen mukaan montako entiteettityypin entiteettiä osallistuu suhteeseen. Jos suhdetyypin $1:1$ on määritelty entiteettityyppien E_1 ja E_2 välille, niin silloin yhtä $E_1:n$ tai $E_2:n$ entiteettiä kohti suhde sisältää korkeintaan yhden parin, jossa entiteetti on mukana. Kun suhdetyypin $N:1$ eli *monesta yhteen* on määritelty $E_1:n$ ja $E_2:n$ välille, niin $E_2:n$ entiteettiä e_2 kohti suhde voi sisältää useita pareja $(e^1_1, e_2), \dots, (e^n_1, e_2)$, mutta kutakin $E_1:n$ entiteettiä e_1 kohti voi suhde sisältää korkeintaan yhden parin (e_1, e_2) . Suhdetyypissä $N:M$ eli *monesta moneen* ei aseteta mitään rajoituksia suhteessa mukana olevien entiteettien määrille.¹⁰⁷

Tietokantaan tallennetaan tietoja entiteeteistä ja niiden suhteista. Tiedot

¹⁰⁷ Ullman 1988, 34-53

muodostuvat entiteettien attribuuttien arvoista. Entiteetit ja niiden väliset suhteet voidaan esittää *relaatiomallia* käyttäen. Relaatiotietomallissa voidaan entiteetit ja niiden väliset suhteet esittää *taulukoina*. *Taulukkotietokanta* on joukko taulukoita.¹⁰⁸ Taulukoiden yhteisten ominaisuuksien määritykset muodostavat *taulukkokaaavion*. Yleensä taulukkokaaviot nimetään. Taulukkokaavio R on äärellinen joukko $\{A_1, \dots, A_n\}$ *sarakkeiden* nimiä. Taulukoiden sarakkeet nimetään entiteettityypin attribuuttien mukaan ja niille annetaan alkuarvot. Kuhunkin sarakkeen nimeen A_i liittyy ei-tyhjä joukko D_i , $1 \leq i \leq n$, jota kutsutaan *arvojoukoksi*. Taulukon *rivi* sisältää yhden entiteetin tiedot. Suhteita voidaan taulukoissa esittää eri tavoin. Entiteettityyppien E_1, \dots, E_k välillä voidaan suhde esittää taulukkona, joka koostuu entiteettityyppien E_1, \dots, E_k avaimien arvoista ja tällöin $E_1:n, \dots, E_k:n$ avainattribuuttien nimien on oltava eri nimiä. Suhde on mahdollista esittää myös liittämällä toisen entiteettityypin avainattribuutti sarakkeena toiseen taulukkoon. Tätä avainattribuuttia nimitetään silloin *viiteavaimeksi*.¹⁰⁹

Entiteeteillä, attribuuttien arvoilla ja entiteettien välisillä suhteilla saattaa olla lukuisia erilaisia *rajoitteita*.¹¹⁰ Rajoite on lause, jolla ilmoitetaan, että vain osa määriteltyjen taulukkotyyppien taulukoista on sallittuja. Tietokanta määritellään *tiedon-määrittelykielellä*. Monissa relaatiomalliin pohjautuvissa tietokannan hallintajärjestelmissä käytetään *SQL*-kieltä, joka sisältää sekä tiedonmäärittely - että *kyselykielen*. Relatiomalliin pohjautuvissa järjestelmissä kyselykieltä voidaan käyttää määrittelykielen osana eräiden rajoitteiden määrittelemiseksi.¹¹¹ Avaimen määrittely on yksi tapa esittää rajoite. Tyypin R taulukon r avain on sellainen $R:n$ (taulukkokaaavion) osajoukko $K = \{B_1, \dots, B_n\}$, jolle ei ole kahtia riviä, joissa kaikki $K:ssa$ mainittujen sarakkeiden sarake-arvot olisivat samoja. Avain on *minimaalinen*, jos se ei sisällä aitona osajoukkona toista avainta. Usein taulukkokaaviossa

¹⁰⁸ Ullman 1988, 7-53

¹⁰⁹ Ullman 1988, s. 34-53

¹¹⁰ Ullman 1988, s. 376-382

¹¹¹ Ullman 1988, s. 12-13, 210-234

määritellään yksi tai useampia avaimia ja näistä yhtä sanotaan *perusavaimeksi*. Kun taulukkokaavio kirjoitetaan sarakenimien joukkona ja osa nimistä on alleviivattu, tarkoittaa se sitä, että alleviivatut sarakenimet muodostavat avaimen.¹¹² Tietokannan suunnittelun perinteinen lähestymistapa pyrkii mahdollisimman pitkälle *normaalimuotoisiin* relaatioihin. Tätä lähestymistapaa käytettäessä kaikki relaatiot ovat vähintään *ensimmäisessä normaalimuodossa*. *Normalisoimattoman* relaation rivien sarakearvot voivat olla rakenteellisia: joukkoja, jonoja tai taulukoita. Ensimmäisessä normaalimuodossa taulukko on silloin, kun sen rivien sarakearvot ovat *atomisia* eli jokaisella (rivi, sarake)-positiolla on tarkalleen yksi arvo ei koskaan arvojoukkoa.¹¹³

Indeksointi eli *hakemiston käyttö* tarkoittaa sitä, että tiedosto on avaimen mukaan lajiteltu. Avaimelta vaaditaan normaalisti yksikäsitteisyyttä. Esimerkiksi haku nimen ja iän perusteella saattaisi joissakin sovelluksissa tulla kyseeseen. Tällaista avainta kutsutaan *toissijaiseksi* avaimeksi ja se voidaan muodostaa mistä tahansa kentästä.¹¹⁴

5.2 Tietokantojen tavoitteet

Henkilöhistorian tietokantoja rakensivat tutkijat joko omaa työtään varten tai ne olivat laajempia projekteja, joiden tulokset oli tarkoitettu eri alojen tutkimukseen ja monipuoliseen käyttöön. Esimerkkijärjestelmissä yksilö muodostaa pienimmän analysointiyksikön, jota on voitava seurata lähteestä toiseen sekä perheensä että taloutensa yhteydessä.¹¹⁵

¹¹² Ullman 1988, s. 35-50

¹¹³ Codd 1970

¹¹⁴ Ullman 1988, s. 312-323

¹¹⁵ Gutmann, Fliess, Holmes, Fairchild, Teas 1989, Danell, 1985, Karjalan väestötieteellisen tietokannan loppuraportti 1989

Demografiska Databasen ja Karjala-tietokanta rakennettiin tutkijoiden yhteiskäyttöön ja niiden toivottiin lisäävän kirkonkirja-aineiston hyödyntämistä ja vapauttavan tutkijat tietojen aikaa vievästä kokoamisesta. Eri alojen tutkijat haluttiin yhteisen materiaalin pariin, minkä uskottiin rikastuttavan tutkimusta ja johtavan poikkitieteellisen tutkimuksen lisääntymiseen. Kansainväliselle tutkimukselle haluttiin tarjota mahdollisuus käyttää aineistoa, jota vastaavaa ei ole muualla kuin Ruotsissa ja Suomessa.¹¹⁶ DDB:ssä kehitystyön peruseriaatteenä olikin, että eri tutkimusalojen kiinnostuksen kohteet ohjasivat suunnittelua. Tutkijat valitsivat ajankohdan ja alueet, joilta kirkonkirjamateriaalia tallennetaan. Rippikirjat valittiin tiedonkeruun päälähteeksi, koska niiden tietosisältö on monipuolisin ja niiden avulla on mahdollista seurata yksityisiä henkilöitä, perheitä ja talouksia vuosi vuodelta. Ainutlaatuisia ne ovat siksi, että papisto on niihin kirjannut valmiiksi perherakenteet, jotka muissa maissa oli kerättävä hajanaisista lähteistä. Muita kirkonkirjoja käytettiin rippikirjojen tietojen täydentäjinä.¹¹⁷

Sekä DDB:n että Karjala-tietokannan toteutukselle asettivat tutkijat seuraavat yleiset vaatimukset:

1. Järjestelmän on kuvattava lähteet niitä muuttamatta. On voitava verrata ja kontrolloida tietokannan tietoja alkuperäisiin lähteisiin nähden. Kutakin lähdetä on pystyttävä analysoimaan erikseen.
2. Järjestelmän tulee sisältää kaikki olennainen ja relevantti tieto.
3. Järjestelmän on oltava avoin eli tietokokoelmaa on voitava laajentaa sekä alueellisesti että sisällöltään samantyyppisillä lähdesarjoilla.
4. Järjestelmän on oltava taloudellinen.
5. Järjestelmän on oltava tutkimukseen suuntautunut ja palveltava sekä demografista tutkimusta että sukututkimusta. On voitava analysoida koko väestöä joko tietyssä vuonna tai tietyllä aikavälillä. Lisäksi on voitava tehdä sekä poikittaistutkimusta koko väestöstä että pitkittäistutkimusta demografi-

¹¹⁶ Danell 1985, s. 11-14, Stenflo, Sundin 1987, Karjalan väestötieteellisen tietokannan loppuraportti 1989

¹¹⁷ Danell 1983, Danell 1985, s. 11-14

sista muutoksista. Jokaisesta ihmisestä on oltava mahdollisuus tehdä sekä elämän- että perhekuvaus.¹¹⁸

DDB:sta tutkimukselle saatavan hyödyn lisäämiseksi on ryhdytty tallentamaan muutakin aineistoa, jossa käsitellään saman alueen ja aikakauden väestöön liittyviä tietoja kuin kirkonkirjoissakin. Tällaista aineistoa ovat esimerkiksi koulu- ja vankilalaitosta, murrसानoja, tuomioistuimia, köyhäinhoitoa, teollistamista ja väestönlaskentaa koskevat asiakirjat sekä manttaaliluettelot, tilikirjat, sairaskertomukset, rakennuspiirustukset ja perukirjat. Tämä merkitsee DDB:n tietosisällön huomattavaa laajentamista.¹¹⁹

Gillespien tietokanta on suunniteltu kehittäjiensä omaa demograafista tutkimusta varten, mutta siihen tallennettiin kaikki lähteiden tiedot alkuperäisessä muodossa niin, että sitä on mahdollista jatkossa käyttää muuhun tutkimukseen. Gillespiejärjestelmän tutkijoiden yhtenä tavoitteena oli selvittää, miten relaatiotietokannan hallintajärjestelmää voidaan käyttää monimutkaisten historiallisten yhteyksien esittämiseen.¹²⁰

5.3 Järjestelmän kehitystyön vaiheet ja lähdeaineiston tietojen tallennus

Järjestelmän kehitystyö noudatti sekä DDB:ssä että Karjala-tietokannassa seuraavaa vaihejakoa:

1. *Lähdeaineiston kuvaaminen ja vaatimusten muotoilu*
Suunnittelun ensimmäisen vaiheen tulokset, lähdekuvaukset, selvittävät lähteiden rakennetta ja tehtävää, jota varten asiakirja aikanaan oli syntynyt.
2. *Informaatio- ja vaatimusanalyysi*
Informaatio- ja vaatimuserittely puolestaan on tavoitellun järjestelmän kuvaus ja selvitys siitä, mihin tarkoituksiin järjestelmän sisältämää materiaalia käytetään.

¹¹⁸ Danell 1985, s. 14, 78, Karjalan väestötieteellisen tietokannan loppuraportti

¹¹⁹ History on data 1989

¹²⁰ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

tetään ja kysymyksenasetteluja, joita sen avulla on mahdollisuus tutkia. Ensimmäisen vaiheen kuvauksia analysoimalla täsmennettiin vaatimuksia ja tehdään tarkat muuttujakuvaukset. Tuotokuvaus selvittää vaatimuksia vastaavien tulosteiden rakenteen yksityiskohtaisesti.

3. *Järjestelmän suunnittelu*

Kahden ensimmäisen vaiheen kuvausten pohjalta kiinnitetään systeemimalli eli karkea järjestelmärakenne.

4. *Järjestelmän toteuttaminen*

Muotoillaan järjestelmän yksityiskohdat.¹²¹

Järjestelmän rakentamistehtävät ovat: tallennus, varastointi ja käyttö. Tallennusvaiheessa kirkonkirjojen tieto siirretään lähteistä koneella luettavaan muotoon. Varastointi tarkoittaa tallennettujen tietojen säilyttämistä työstämisen eri vaiheissa joko raaka- ja lähdetiedostoina tai lopullisessa muodossa relaatiotietokantana. Käyttö on lähdetiedostojen ja tietokannan hyödyntämistä ja tiedon hakua eri tavoin. Seuraavassa käsitellään tarkemmin tallennusta.¹²²

Tallennuksen tehtäviä ovat lähdevalmistelu, tiedonrekisteröinti ja muokkaus. Lähdeaineiston historiallinen alkuperä edellyttää lähdevalmistelua eli huolellista perehtymistä lähteisiin. Lähdevalmistelu aloitetaan selvittämällä, mitä arkistomateriaalia tallennuksen kohteena olevissa seurakunnissa on säilynyt ja tutkimalla sen laatu. Työn tuloksena syntyvät erilaiset yleiskuvaukset alueen seurakunnista ja niiden lähdeaineistosta sekä yksityiskohtaiset selvitykset kustakin kirkonkirjasidoksesta. Seurakuntien luettelot on sidottu arkistokokonaisuuksiksi eli sidoksiksi, jotka erotetaan toisistaan arkistotunnusten avulla. Lähdeaineiston yleiskuvauksissa arvioidaan kirkonkirjojen sisältöä, laatua, lukukelpoisuutta ja puutteita sekä niiden työstämiseen tarvittavaa työpanosta. Yksityiskohtaiset kuvaukset ovat tarkkoja selontekoja arkisto-sidoksista. Ne sisältävät sidoksen arkistotunnuksen, tiedon siitä, miltä vuosilta sidoksen merkinnät ovat, mitä tietoja sidos sisältää, sivujen ja kirjattujen tapausten lukumäärän, missä muodossa tiedot esiintyvät, luettavuuden, kielen, sidoksen rakenteen ja mahdollisesti lähde-

¹²¹ Danell 1985, 81-88

¹²² Danell 1985, 93-96

tyypin yleisestä käytännöstä poikkeavan tietosisällön sekä erikoisominaisuudet. Jos sidos on jaettu sarakkeisiin, luetellaan niiden nimet ja järjestys sekä niiden sisältämä tieto.¹²³

Lähdevalmistelu antaa pohjan aineiston jatkokäsittelylle. DDB:ssä tietojen siirtoa tietokoneella luettavaan muotoon kutsutaan tiedonrekisteröimiseksi. Sen jälkeen kun tiedot on siirretty tietokoneelle suoritetaan tallennettujen tietojen oikeellisuuden tarkistus ja mahdollisten virheiden korjaus.¹²⁴ Tallennuksen virheettömyyttä valvotaan tarkistuksilla. Karjala-tietokannassa osa tarkistuksista on samoja kaikille lähdetypeille. Sen lisäksi on joillekin lähdetypeille määriteltyjä tarkistuksia. Tarkistuksilla pyritään varmistamaan, että muuttujien arvot ovat sallituilla arvoalueilla ja loogisesti oikeita. Niillä pyritään myös eliminoimaan tallentajien tekemät virheet. Yksinkertaisuustarkistuksissa eli tallennuksen aikaisten virheiden tarkistuksilla pyritään varmistamaan, että jokaiseen henkilöön tai tapahtumaan liittyvät tiedot on tallennettu vain kerran. Siinä tutkitaan, että kunkin merkinnän yksilöivä sivu- ja rivitieto on yhdessä sidoksessa ainoastaan yhden kerran käytössä. Jos löytyy useita samoja sivu- ja rivitietoja, se viittaa siihen, että sama henkilö tai tapahtuma on vahingossa tallennettu kahteen kertaan. Lisäksi tarkistetaan, etteivät peräkkäisten rivien sarakearvot ole samoja, jolloin näyttäisi todennäköiseltä, että sama henkilö tai tapahtuma olisi vahingossa tallennettu kahteen kertaan. Haamutarkistuksissa varmistetaan, ettei tietokantaan ei ole tallennettu tyhjiä tietueita eli ettei riville ole tallennettu pelkästään oletusarvoja, jolloin olisi todennäköisesti kyse tyhjän merkinnän tallennuksesta. Tarkistukset eivät kuitenkaan takaa täysin virheetöntä tulosta, sillä niiden avulla ei pystytä huomaamaan esimerkiksi kaikkia tallentajien tekemiä kirjoitus- ja poimintavirheitä. Tiukan alkuperäisyysvaatimuksen johdosta tallennetaan lähteissä esiintyvät virheellisetkin tiedot, jotka myöhemmin tehtävien tarkistusten

¹²³ Danell 1985, s. 82-85, Karjalan väestötieteellisen tietokannan loppuraportti 1989

¹²⁴ Danell 1985, s. 82-84, 112-114

yhteydessä on korjattava.¹²⁵

Siirrettäessä tietoja lähteistä tietokoneelle tehdään sekä DDB:ssä että Karjala-tietokannassa tallennettavien tietojen muokkausta eli muuttujien koodausta, standardointia ja generointia. Muokkaustoimenpiteillä pyritään parantamaan tallennuksen virheettömyyttä ja yhtenäisyyttä. DDB:ssä väestötietokanta pyrkii mahdollisimman tehokkaaseen käyttöön samoin kuin Karjala-tietokannan käyttäjille kehitetty Karhu-järjestelmä, ja siksi muokkaustoimenpiteitä jatketaan niitä rakennettaessa.¹²⁶

Tutkijat asettivat sekä DDB:lle että Karjala-tietokannalle tavoitteen, että niiden tuli sisältää kaikki lähteissä esiintyvä relevantti tieto alkuperäisessä muodossaan. Myös Gillespie-järjestelmään on tallennettu kaikki tutkimukseen kuuluvissa lähteissä esiintyvä tieto.¹²⁷ Vaatimus kaiken tiedon tallentamisesta alkuperäisessä muodossaan on ristiriitainen. Alkuperäistä lähdettä ei paraskaan jäljenne pysty korvaamaan. Tehokkaan käytön edellytys on, että tiedot jatkokäsittelyssä muokataan vertailukelpoisiksi keskenään. On siis ratkaistava, onko tietosisältö vai tiedon esitysmuoto tärkeämpi. Nimistöntutkimuksen kannalta tietojen esitysalla on ehkä merkitystä, mutta muuten tietokannan käyttäjille tietosisältö on todennäköisesti tärkeämpi. Vaikka muuttujien koodaus säästää muistitilaa ja helpottaa hakuja, ovat DDB:ssä saadut kokemukset osoittaneet, että on parempi mitä vähemmän tietoja esitetään koodeina. Koodien on myös oltava yksinkertaisia ja helposti muistettavia. DDB:ssä koodattuina tallennetaan henkilö- ja paikannimet, ammattinimikkeet, sukupuoli, sukulaisuus ja taitomerkinnot. Karjala-tietokannassa on koodauksessa eroja siten, että seurakunta- ja paikannimet sekä ammattinimikkeet tallennetaan lähdemerkinnän muodossa. Myös rokotusta ja

¹²⁵ Danell 1985, 134-139, Karjalan väestötieteellisen tietokannan loppuraportti

¹²⁶ Danell 1985, s. 110, 141-144, Karjalan väestötieteellisen tietokannan loppuraportti

¹²⁷ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

aviosyntyä koskevat tiedot tallennetaan lähdemerkinnän muodossa.¹²⁸ Gillespie-järjestelmässä koodeja käytetään seurakunta- ja paikannimistä sekä muuttujissa, jotka arvioivat tietojen luotettavuutta.¹²⁹ Joitakin tietoja, esimerkiksi taitomerkinntöjä, on mahdoton tallentaa sellaisessa muodossa kuin ne esiintyvät lähteissä. Taitomerkinntöjen tietosisältö pystytään kuitenkin siirtämään tietokantaan käyttämällä lähteessä esiintyvää merkintää vastaavaa koodiarvoa. DDB:ssä lukutaitomerkinntät muutetaan numerokodeiksi, jotka on muodostettu niin, että niistä voidaan päätellä rippikirjassa käytetty merkkijärjestelmä.¹³⁰

Standardointi eli esitetyn tiedon yhtenäistäminen tapahtuu määriteltyjen ohjeiden mukaan. Sen yhteydessä käydään aineisto läpi, jotta voitaisiin taata muuttujakuvausten ja tallennettujen tietojen yhdenmukainen esittäminen. Standardointi merkitsee esimerkiksi epäloogisten muuttuja-arvojen siirtämistä huomautuskenttään. Nimien ja ammattinimikkeiden normalisointi on välttämätöntä. Normalisoituja muotoja tarvitaan tietojen automaattisessa järjestämisessä ja yhdistämisessä.¹³¹

Generointi tarkoittaa sellaisten muuttujien ja niiden arvojen määrittelyä, jotka voidaan johtaa lähteissä annetuista tiedoista, vaikka ne eivät varsinaisesti esiinny henkilön kohdalla kirkonkirjamerkinntöinä. Generoituja sarakkeita ja niille saatuja arvoja ovat esimerkiksi lasten lukumäärä, siviilisääty ja sen muutokset eri aikoina, suhteen kesto, muutto- ja asumistiedot sekä syntymään, kasteeseen, kuolemaan ja hautaukseen liittyviä vuosilukuja. Jos ammattitiedot puuttuvat, ne on mahdollista päätellä yksilöstä annettujen muiden tietojen perusteella. Etenkin sellaisia tietoja generoidaan, joita kokemuksesta tiedetään tutkijoiden usein

¹²⁸ Danell 1985, s. 142, Karjalan väestötieteellisen tietokannan loppuraportti

¹²⁹ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

¹³⁰ Berggren, Danielsson 1989, Karjalan väestötieteellisen tietokannan loppuraportti

¹³¹ Danell 1985, s. 141, Karjalan väestötieteellisen tietokannan loppuraportti

kysyvän. Vasta sen jälkeen kun koko alueen tiedot on liitetty yhteen, on DDB:ssä mahdollista tuottaa esimerkiksi väestötietokannan päätaulukon tiettyjen generoitujen sarakkeiden arvot. Generointi edellyttää, että muut tiedot on jo aikaisemmin tarkistettu ja standardoitu.¹³²

DDB:ssä ja Karjala-tietokannassa suurin osa tiedoista tallennetaan sellaisessa muodossa kuin ne esiintyvät lähteissä, osa tallennetaan sekä lähdeuskollisesti että koodina, osa pelkästään koodimuodossa ja pieni osa generoituna eli pääteltynä tietona.

5.4 Lähdeaineiston tietosisältö

Kaikissa esimerkkijärjestelmissä tietosisältö muodostuu eri lähdetyyppien sisältämistä merkinnöistä, jotka yhdessä antavat kuvan lähdeaineistoon kuuluvien ihmisten elämäntilanteista. Taulukossa 6 esitetään ja verrataan DDB:n, Karjala-tietokannan ja Gillespie-järjestelmän tietosisältöä. Vertailussa on otettava huomioon se, että DDB ja Karjala-tietokanta perustuvat tietosisällöltään samanlaisiin lähteisiin ja siksi niiden tietosisältö on sama. Poikkeuksen muodostaa henkilönumero, joka DDB:ssä identifioi yksilön ja joka puuttuu Karjala-tietokannasta. Gillespie-järjestelmän kirkonkirjat käsittävät kaste-, avioliitto- ja hautausluettelot, joista saadaan selville näihin tapahtumiin liittyvää tietoa. Ikä, sukupuoli, ammatti, lukutaito, varallisuus, taloudet ja asuinpaikka selviävät väestölaskenta- ja veroluetteloista. Väestöluettelot kirjaavat henkilöiden asuinpaikat 20 vuoden välein, mutta niistä puuttuu ruotsalaisten ja suomalaisten rippikirjojen ja muuttoluetteloiden jatkuva seuranta, joka kirjaa ihmisten elämäntilanteissa tapahtuneet muutokset ja muutosten ajankohdan niiden tapahtumahetkellä. Gillespie-järjestelmästä puuttuvat myös seurakuntalaisten ehtoollis- ja kinkerikäyntejä koskevat tiedot.

¹³² Danell 1985, s. 142-143, Karjalan väestötieteellisen tietokannan loppuraportti

Taulukko 6 Esimerkkietokantojen tietosisältö ja vertailu

Tieto	DDB	Karjala	Gillespie
Talletettava seurakunta	X	X	X
Sidoksen tiedot	X	X	X
Henkilön perustiedot			
- henkilönnumero	X		X
- nimitiedot	X	X	X
- sukupuoli	X	X	X
- syntymän päivämäärä ja paikka	X	X	X
- kasteen päivämäärä ja paikka	X	X	X
- kuulutusten päivämäärä ja paikka	X	X	
- avioliiton solmimisen päivämäärä ja paikka	X	X	X
- kuolinpäivämäärä ja paikka	X	X	X
- kuolinsyy	X	X	
- ikä kuollessa	X	X	
- henkilön siviilisäätö eri ajankohtina	X		
- ammatti/ammattit			
- asuinpaikka	X	X	X
- henkilöön liittyvät muut tiedot (esim. rampa)	X	X	X
	X	X	
Henkilön suhteet			
- isä			
- äiti	X	X	X
- aviopuoliso	X	X	X
- lapsi	X	X	X
- kummin nimi	X	X	X
- kummin ammatti	X	X	X
- kummin asuinpaikka	X	X	
	X	X	
Henkilön perustapahtumatiedot			
- syntymään liittyvä tieto (kaksonen ym.)			
- kasteeseen liittyvä tieto (kuka kastoi ym.)	X	X	
- vihkimiseen liittyvä tieto (kuka vihki, missä ym.)	X	X	
- vihkimisen todistajien tiedot	X	X	
- hautaukseen liittyvä tieto (kellojen soitto ym.)	X	X	
- tapahtumiin liittyvät muut tiedot	X	X	
Henkilön muuttotiedot			
- muuton päivämäärä	X	X	
- mistä muuttaa (kylä, talo)	X	X	
- lähtöseurakunta	X	X	
- minne muuttaa (kylä, talo)	X	X	
- muuttaneiden miesten lukumäärä	X	X	
- muuttaneiden naisten lukumäärä	X	X	
- kokonaislukumäärä	X	X	
Henkilön ehtoollis- ja kinkerikäynnit			
- konfirmaation päivämäärä	X	X	
- ehtoollisella käynnit	X	X	
- kinkereillä käynnit	X	X	
- lukutaito	X	X	X
- kristinopin taito	X	X	
- rokotus	X	X	
- rippikirjojen siirtomerkinnot	X	X	
- äänioikeuteen liittyvät tiedot	X	X	X
Henkilön varallisuus			X

Kaikissa esimerkkijärjestelmissä tallennetaan lähteiden merkinnät kullekin lähdetyypille ominaisiin rakenteisiin. DDB:ssä eri kirkonkirjatyyppit kuten rippikirjat sekä syntyneiden, vihittyjen, kuolleiden ja muuttaneiden luettelot muodostavat jokainen oman lähdetiedostonsa. Lähdetietojen tallennus tapahtuu sitä varten kehitetyllä ohjelmistolla.¹³³ Tallennusyksikkönä ovat lähteiden merkinnät, joilla tarkoitetaan lähteen rivillä olevaa, yhtä henkilöä koskevaa tietoa. Tallennus aloitetaan syntymä- ja kasteluetteloista. Syntymätiedot kuuluvat henkilötunnuksen muodostaviin tietoihin ja ovat keskeisiä linkityksen oikeellisuuden varmistamiseksi. Syntymätiedot tallennetaan ID-rekisteriin, joka käsittää kaikki käsittelyn kohteena olevalla sekä sitä välittömästi ympäröivällä alueella syntyneiden henkilöiden syntymätiedot ja henkilötunnukset. Henkilötunnus muodostetaan syntymävuodesta, -kuukaudesta, -päivästä, -seurakuntakoodista, sukupuolesta, samana päivänä syntyneiden järjestys- ja kontrollinumerosta, joka lasketaan henkilön numeron ensimmäisistä numeroista ja on jokin luku 0-9.¹³⁴

Lähdetiedostot on järjestetty samalla tavoin kuin alkuperäiset lähteet ja henkilöstä voi niissä esiintyä useita merkintöjä. Yksilölle annettu tunniste, henkilö-numero, on yhteinen eri lähdetyyppien samaan yksilöön liittyvissä tiedoissa. Huomautukset ja kuolinsyyt sekä yksilöiden keskinäiset suhteet erotetaan muusta tiedosta ja tallennetaan omiin tiedostoihinsa. Yksilöiden väliset linkit tallennetaan suhdetiedostoon. Lähteiden riveille tehtyjen merkintöjen lisäksi lähdetiedostoiksi tallennetaan myös lähteiden otsikkotiedot, sidostiedot, seurakuntarekisterit ja kaikkien koodattujen muuttujien tieto kirjoitettuna.¹³⁵

DDB muodostuu kahdesta osajärjestelmästä: kirkonkirja- ja väestöosajärjestelmästä. Kirkonkirjaosajärjestelmä käsittää lähdetiedostot eli rippikirja-, syntymä-, kuolin-, vihkimis- ja avioliittotiedostot. Lisäksi siihen kuuluu aputiedostoja kuten

¹³³ Danell 1985, s. 96-97

¹³⁴ Preliminary draft 1989

¹³⁵ Danell 1985, s. 96-97

suhde-, tunniste-, huomautus- ja kuolinsyytiedostot. Väestötietokantaa rakennettaessa seurakuntien lähdetiedostot yhdistetään lähdetyyppeittäin omaksi itsenäiseksi tietokannaksi. Viimeisessä vaiheessa seurakuntien lähdetyypitietokannat yhdistetään alueen yhteiseksi väestötietokannaksi. Väestötietokanta yhdessä hyväksikäyttöjärjestelmän kanssa muodostaa väestöosajärjestelmän.¹³⁶

DDB:n kirkonkirjamateriaali varastoidaan tallennusprosessin eri vaiheissa neljässä erilaisessa muodossa:

Raakatiedosto 1	rekisteröidyt, muokkaamattomat tiedot
Raakatiedosto 2	rekisteröidyt, osittain muokatut tiedot
Lähdetiedostot	kirkonkirjaosajärjestelmä
Väestötietokanta	väestöosajärjestelmä, muodostaa väestötietokannan.

Aineiston erilaisilla varastointitavoilla pyritään takaamaan, että työprosessit voidaan tarvittaessa aloittaa uudelleen. Raakamateriaalitiedostot ovat tallennuksen välituotteita, jotka ovat syntyneet muokkausprosessin eri vaiheissa. Niitä säilytetään turvallisuussyistä eikä niitä, toisin kuin lähdetiedostoja ja väestötietokantaa, ole tarkoitettu hauissa käytettäväksi. Ne tallennetaan ennen kuin muokkausvaiheen kaikki toiminnot on suoritettu. Jos virheitä ilmenee muokkauksessa, pitää työ tehdä uudestaan materiaalin lähinnä edeltävästä oikeasta versiosta lähtien.¹³⁷

Karjala-tietokannassa tallennusohjelmat on luotu dBase IV:n ohjelmointikielellä. Rippikirjojen ja lastenkirjojen sekä syntyneiden ja kastettujen, vihittyjen ja kuulutettujen, kuolleiden ja haudattujen sekä muuttaneiden luetteloiden merkinnät tallennetaan taulukoihin. Tallennusohjelmisto rakentuu kahdeksasta ohjelmatiedostosta. Kullakin lähdetyypillä on omat lomakepohjat, joita täyttämällä tiedot viedään tietokantaan. Jokainen sidos muodostaa oman tietokannan. Kaikki saman lähdetyypin sidokset tallennetaan rakenteeltaan samanlaisiin taulukoihin. Taulukko 7 esittää eri lähdetyyppien taulukot ja niiden tietosisällön. Kunkin lähdetyypin taulukkorakenne on määritelty lähdekuvauksissa ja on

¹³⁶ Danell 1985, s. 86-88

¹³⁷ Danell 1985, s. 98-99

Taulukko 7 Karjala-tietokannan lähdeaineiston taulukot

Taulukko	Taulukon tietosisältö
Syntyneiden ja kastettujen luettelot:	
Syka	syntyneiden ja kastettujen luetteloiden perustiedot
Kummit	kummien tiedot
Huom	huomautustiedot
Kumhuom	kummien huomautustiedot
Muuttaneiden luettelot:	
Mulu	muuttaneiden luetteloiden perustiedot
Huom	huomautustiedot
Kuolleiden ja haudattujen luettelot:	
Kuha	kuolleiden ja haudattujen luetteloiden perustiedot
Omat	omaisten tiedot
Huom	huomautustiedot
Vihittyjen luettelot:	
Vilu	vihittyjen luetteloiden vihkitapahtumaa ja kuulutuksia koskevat tiedot
Sumo	sulhasen ja morsiamen tiedot
Todi	vihkimisen todistajat
Vhuom	vihkimistapahtumaan liittyvät huomautukset
Vmhuom	sulhaseen tai morsiameseen liittyvät huomautukset
Lastenkirjat:	
Laki	lastenkirjojen perustiedot
Vanh	lasten vanhempien tiedot
Huom	lasten vanhempiin liittyvät huomautukset
Taid	taito- ja suoritusmerkinnät
Kink	kinkerikerrat
Kinv	kinkerivuodet
Huom	lapsiin liittyvät huomautukset
Rippikirjat:	
Riki	rippikirjojen perustiedot
Taid	taitomerkinnot
Kink	kinkerikerrat
Kinv	kinkerivuodet
Aani	äänioikeusmerkinnät
Avio	avioliittomerkinnot
Puol	puolison rivitiedot
Suhde	suhteiden rivitiedot
Ehto	ehtoollismerkinnät
Huom	huomautustiedot
Asuinpaikat:	
Asuinp	rippikirjojen asuinpaikat
Asuinp	lastenkirjojen asuinpaikat

muotoutunut sidosten tietosisällön ja merkintöjen muodostamien asiakokonaisuuksien mukaan.¹³⁸

Sidosten merkinnät tallennetaan lähteen kirjaamisjärjestyksessä. Taulukon rivi muodostuu sidoksen rivillä olevasta yhtä henkilöä koskeva merkinnästä. Rivit erotetaan toisistaan ykkösestä alkavan juoksevan järjestysnumeron avulla, jonka sarakenimi on *Hnro*, henkilönnumero. *Hnro* ei yksilöi henkilöä, vaan häneen liittyvän tapahtuman ja samasta henkilöstä saattaa ollakin useita rivejä yhden lähteen muodostamassa tietokannassa, kuitenkin eri henkilönnumerolla. Liitteessä 5 on esitetty syntyneiden ja kastettujen luetteloiden lapsen syntymään ja kasteeseen liittyviä tietoja sisältävä *Syka*-taulukko, sen sarakkeiden nimet ja niiden tietosisältö.¹³⁹

Gillespie-järjestelmässä on kunkin lähdetyyppin tiedoista muodostettu tietokanta. Kastetapahtumat muodostavat kastetietokannan, ja esimerkiksi kaikki vuoden 1880 väestöluetteloissa mainitut yksilöt muodostavat yhden väestöluettelotietokannan. Tallennettu raakamateriaali haluttiin säilyttää koskemattomana, jotta tutkimukseen käytettyä lähdeaineistoa voitaisiin hyödyntää myöhemmin. Tallennusjärjestys vaihtelee kussakin lähdetyyppissä. Kirkonkirjat on tallennettu kronologisessa järjestyksessä, veroluettelot on ryhmitelty sukunimen ensimmäisen kirjaimen mukaan. Väestöluetteloiden tallennusjärjestys noudattaa reittiä, jota verovirkailija kulki merkintöjä tehdessään.¹⁴⁰

5.5 Tietokantojen käytön tehostaminen

Tallennuksen yhteydessä pyritään lähdeaineisto mahdollisimman tarkoin ja alkuperäisessä muodossaan siirtämään tietokoneella luettavaan muotoon.

¹³⁸ Karjalan väestötieteellisen tietokannan loppuraportti 1989

¹³⁹ Karjalan väestötieteellisen tietokannan loppuraportti 1989

¹⁴⁰ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

Tietojen haku DDB:n lähdetiedostoista sekä Karjala-tietokannan ja Gillespie-järjestelmän lähdetaulukoista on hankalaa ja siksi niitä on muutettava käytön kannalta tehokkaammiksi ja joustavammiksi. Tiedon hakua varten DDB:ssä on rakennettu väestötietokanta, Karjala-tietokannassa Karhu-käyttäjäjärjestelmä ja Gillespie-järjestelmään hakemistot.

5.5.1 Demografiska Databasenin väestötietokannan taulukot

Vaatus aineiston käytön nopeuttamisesta toteutettiin DDB:ssä kehittämällä kirkonkirjaosajärjestelmä, joka koostuu seurakuntien lähdetyypeittäin tallennetuista tiedoista ja niiden käyttöön tarkoitettuista ohjelmista. Se sisältää kaikki lähteiden tiedot. Joihinkin sovelluksiin riittää tällainen lähdetyypeittäin tallennettu, yhtä seurakuntaa koskeva tieto. Yksilön elämäkuva ei kuitenkaan ole täydellinen, ennen kuin alueen muodostavien seurakuntien tiedot on yhdistetty tietokannaksi. Väestötietokanta muodostaa DDB:n lopputuotteen ja toteuttaa järjestelmälle asetetut vaatimukset. Sen avulla voidaan tutkia lähteisiin merkittäviä yksityisiä henkilöitä, perheitä ja koko väestöä yhdessä seurakunnassa tai useampien seurakuntien muodostamalla alueella. Tutkimus voi keskittyä joko määrättyyn vuoteen tai ajanjaksoon. Lisäksi väestöosajärjestelmä toteuttaa tietojen joustavan saannin. Väestötietokanta rakennetaan kahdessa vaiheessa. Ensin rakennetaan seurakuntakohtainen tietokanta, joka seuraavaksi yhdistetään aluekohtaiseksi tietokannaksi.¹⁴¹

Sekä kirkonkirja- että väestöosajärjestelmä käsittävät kaikki lähteiden tiedot. Lähteiden merkinnät saattavat olla ristiriitaisia, epätäydellisiä tai yhteydeltään epäloogisia. Tällaiset tiedot tallennetaan kirkonkirjaosajärjestelmään, mutta väestötietokannassa ne oikaistaan. Moninkertaista tietoa vähennetään. Jos kirkonkirjoissa esiintyy samasta tapahtumasta useita merkintöjä, tiedon valinta noudattaa periaatetta, jonka mukaan tapahtuman tyypillisen lähteen tiedoille

¹⁴¹ Danell 1985, s. 87-88, 93

annetaan etusija ja niitä täydennetään rippikirjojen tiedoilla. Ero lähdetiedostojen ja väestötietokannan välillä muodostuu siis osittain eroista muuttujien sisällössä, sillä väestötietokannan muuttujia määriteltäessä otetaan jokaisen muuttujan kohdalla huomioon kaikki saatavissa oleva ja relevantti tieto.¹⁴²

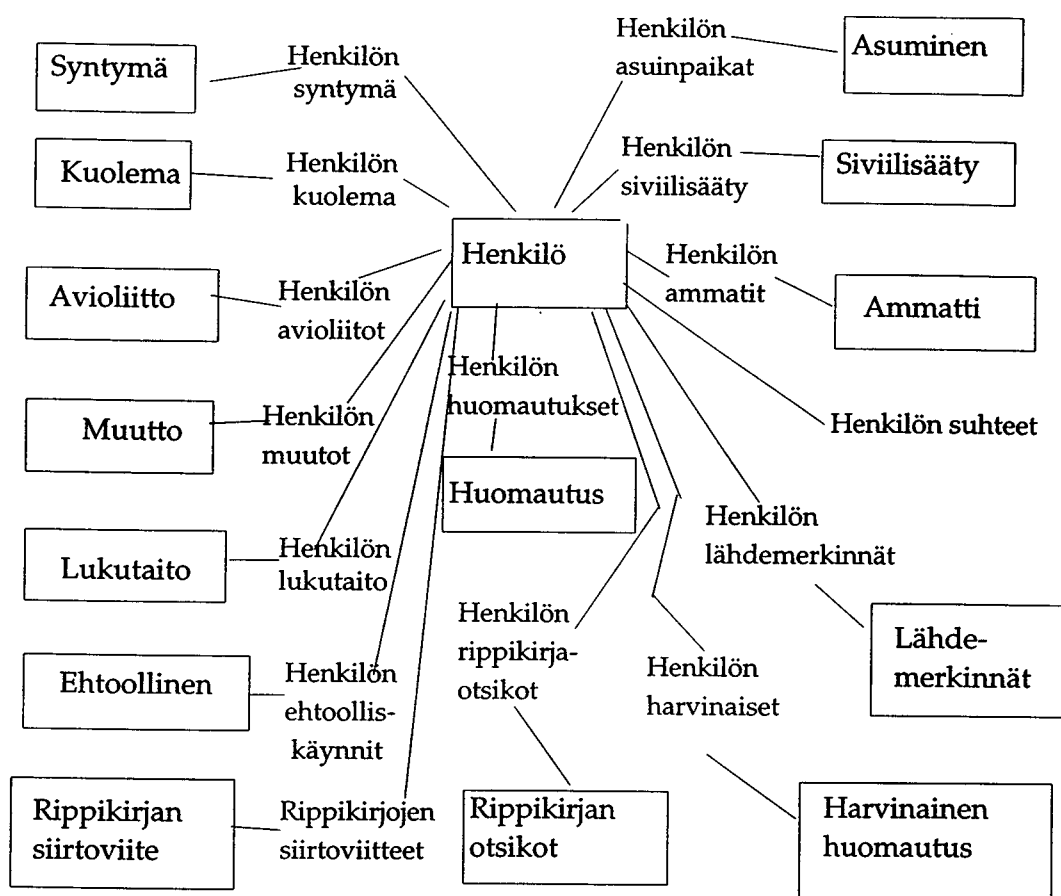
Väestötietokannalle annettujen vaatimusten toteuttamiseksi lähdetiedostojen tietoa käsitellään edelleen niin, että jokaisesta yksilöstä voidaan luoda looginen ja uskottava elämänkuva. Yli 99 prosentissa kaikista tapauksista on mahdollista esittää yksilöstä seuraavia tietoja: nimi, syntymäaika ja syntyperä, maininta syntymistavasta kuten esimerkiksi kaksonen, kuolin- ja vihkimisaika, muutot seurakuntaan ja sen sisällä tai sieltä pois sekä muuttojen päämäärä, ammatti, sosiaalinen asema, siviilisääty, asuinpaikka, perheasema, kuuluminen perheeseen ja talouteen sekä suhteet muihin yksilöihin. Henkilön suhdetiedoista selviää, kuka on kenenkin lapsi, aviomies tai -vaimo, ja ketkä ovat lapsen vanhemmat. Henkilökuvaa voidaan täydentää tiedoilla ehtoollisella ja kuulusteluissa käyneistä sekä annetuista todistuksista ja rokotuksista. Lisäksi on mahdollista ilmoittaa rippikirjojen nidokset ja niiden rivit, joissa henkilö on mainittu sekä viittaukset muihin lähteisiin. Täydelliseen elämänkuvaan sisältyvä muuttujien valikoima sisältää siis tunnistustiedon lisäksi myös demograafisia tapahtumia ja yksilön asemaa eri ajankohtina käsitteleviä tietoja.¹⁴³

Väestötietokantaa luotaessa lähdetiedostojen rakenne on purettu. Kuva 7 esittää väestötietokannan käsitekaaviota. Keskeinen entiteettityyppi on *Henkilö*, jonka attribuutit käsittävät sekä yksilön identifioimiseen tarvittavaa perustietoa että koko alueen aineiston läpikäynnin jälkeen muodostettua generoitua tietoa. Entiteettityypin avaimena on *henkilönumero*. Muita entiteettityyppejä on muodostettu ihmisen elämänkaaren erilaisiin tapahtumiin liittyvistä lähdemerkinnöistä: syntymästä ja kasteesta, kuulutuksista ja avioliitosta, kristinopin osaamisesta, ehtoollisella ja kinkereillä käynneistä, kuolemasta ja hautauksesta

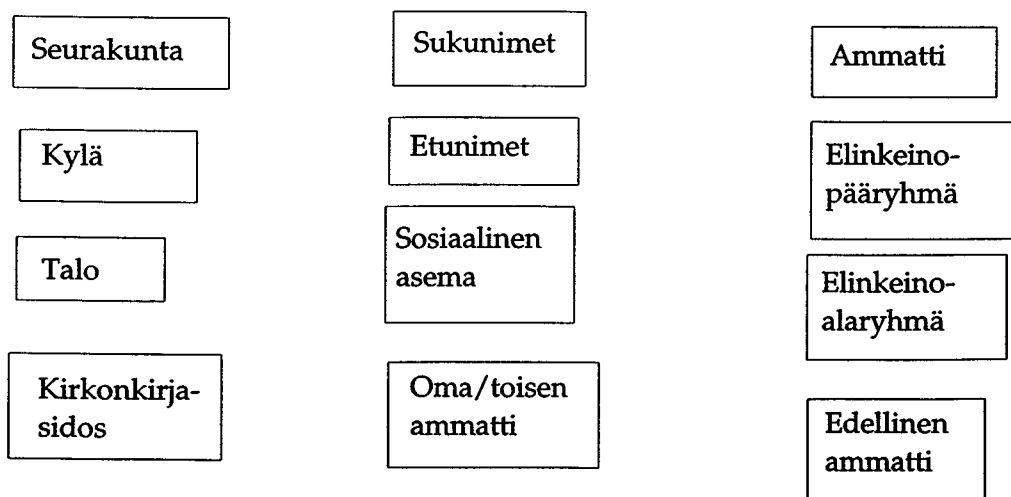
¹⁴² Danell 1985, s. 96-99

¹⁴³ Danell 1985, s. 146-147

Lähteiden tietojen perusteella muodostetut entiteettityypit



Koodeista ja niiden selityksistä muodostetut entiteettityypit



Kuvassa nelikulmainen laatikko esittää entiteettityyppejä ja entiteettityyppien välissä oleva teksti niiden välillä vallitsevaa suhdetyyppeä.

Kuva 7 DDB:n väestötietokanta

sekä henkilön siviilisäädystä ja asuinpaikasta sekä niissä ajan kuluessa tapahtuneista muutoksista.¹⁴⁴

Entiteettityypin Henkilö ja muiden henkilön tietoja sisältävien entiteettityyppien välillä on muodostettavissa suhdetyyppejä. Entiteettityyppiä Henkilö olevat entiteetit muodostavat entiteettijoukon, joka sisältää yhden entiteetin kutakin yksilöä kohti, josta alueen kirkonkirjalähteissä löytyy merkintöjä. Tapahtumamerkintöjä sisältävissä entiteettijoukoissa voi yhdestä henkilöstä olla nolla, yksi tai useampia entiteettejä. Jos ihminen elämänsä aikana on asunut vaikkapa kolmessa paikassa, on hänestä kolme entiteettiä, jotka ovat tyyppiä *Asuminen*. Tässä tapauksessa entiteettityyppien *Asuminen* ja Henkilö välillä vallitsee suhdetyyppi N:1. Jokaista entiteettityyppiin Henkilö kuuluvaa entiteettiä kohti on ainakin yksi, tavallisesti useita, entiteettejä tyyppiä *Lähdemerkinnät*. Lähdemerkinnät sisältävät kaikki viittaukset lähteisiin, joista löytyy merkintöjä entiteetin esittämästä henkilöstä. Entiteettityyppien *Lähdemerkinnät* ja Henkilö suhdetyyppi on N:M.¹⁴⁵

Paikannimien, henkilönimien, sukulaisuussuhteiden, ammattien, kirkonkirjojen arkistotunnusten sekä erilaisten vaihtelevien muuttujien tallennuksessa on käytetty koodeja. Koodien ja henkilön tietoja sisältävien entiteettityyppien välille voidaan muodostaa suhdetyyppi esimerkiksi *Henkilön syntymäseurakunta* (*Syntymä, Seurakunta*), jolloin saadaan henkilön syntymäseurakunnan nimi.

Relaatiomallissa entiteetit ja niiden väliset suhteet esitetään taulukoina. DDB:n taulukot, jotka esitetään taulukossa 8, muodostetaan lähdetiedostoista C-kielisen ohjelman avulla. Taulukoita muodostettaessa on erityistä huomiota kiinnitetty hakujen tehokkuuteen ja yksinkertaisuuteen.¹⁴⁶

¹⁴⁴ Berggren, Danielsson 1989

¹⁴⁵ Berggren, Danielsson 1989

¹⁴⁶ Berggren, Danielsson 1989

Taulukko 8 DDB:n tärkeimmät taulukot ja niiden tietosisältö vuoden 1989 tammikuussa tapahtuneen uudelleenorganisoinnin jälkeen

<u>Lähdemerkintöjä sisältävät taulukot:</u>	
Henkilö/DDB.xxxPERSONV	Tietokannan päätaulukko
Henkilön syntymä/xxx.FDV	Syntymään liittyvät merkinnät
Henkilön kuolema/DDB.xxxDBV	Kuolemaan liittyvät merkinnät
Henkilön avioliitot/DDB.xxxLVV	Avioliiton solmimiseen ja kuulutuksiin liittyvät merkinnät
Henkilön muutot/DDB.xxxFLV	Muuttoihin liittyvät merkinnät
Henkilön lukutaitot/* DDB.xxxBETUV	Kristinopin osaamiseen liittyvät merkinnät
Henkilön ehtoolliskäynnit/ DDB.xxxHL1V	Ehtoolliskäynteihin liittyvät merkinnät
Rippikirjojen siirtoviitteet/ DDB.xxxHL2V	Rippikirjojen viitteet henkilön tietojen siirtämisestä sivulta toiselle
Henkilön huomautukset/* DDB.xxxANMV	Henkilöön liittyvät huomautukset
Henkilön harvinaiset huomautukset/* DDB.xxxANMVARV	Henkilöön liittyvät harvoin esiintyvät tiedot
Henkilön asuinpaikat/ DDB.xxxBOENDE	Asuinpaikkatiedot
Henkilön siviilisääty/ DDB.xxxCIVV	Siviilisääty ja siinä ajan kuluessa ajan kuluessa tapahtuneet muutokset
Henkilön ammatit/ DDB.xxxYRKEV	Ammattiin liittyvät tiedot
Henkilön lähdemerkinnät/ DDB.xxxLANKV	Sisältää tiedot kaikista lähteistä, joista löstä löytyy merkintöjä
Henkilön suhteet/DDB.xxxRELV	Henkilön suhteet eli vanhemmat, kumppanit
Rippikirjapaikannimet/ DDB.xxxRUBORT	Rippikirjojen otsikkoina esiintyvät paikannimet
<u>Aputaulukot (käytetyt koodit ja niiden selitykset):</u>	
Seurakunta/DEDIK	Seurakuntien koodit selityksineen
Kylä/KODORT	Kylien koodit selityksineen
Talo/KODSUB	Talojen koodit selityksineen
Standardoidut sukunimet/KODNME	Sukunimikoodit selityksineen
Standardoidut etunimet/KODNMF	Etunimikoodit selityksineen
Sukulaisuus/KODSLK	Perheasema/sukulaisuuskoodit selityksineen
Muuttujat/KODTEX	Erilaisten muuttujien koodeja selityksineen (esim. AB=aviosynty)
Ammatit/KODYRK	Ammattien koodit selityksineen
Edellinenammatti/KYFD	
Elinkeinopääryhmä/KYNHU	
Elinkeinoalaryhmä/KYNUG	
Oma/toisen ammatti/ KYREL	Ammatti-informaatio ja sen selitykset
Sosiaalinen asema/KYSOC	
Sidos/VOLYM	Arkistosidosten koodit, selitykset ja kuvaukset

(xxx taulukon nimessä tarkoittaa seurakuntamerkintää, esimerkiksi Skellefteå lsf = SKE, Sundsvalls stad = SST jne.).

* Nämä taulukot otetaan käyttöön erikoistilauksesta.

DDB:n tietokannan muodostavat yhden alueen maantieteellisesti toisiinsa rajoituvien seurakuntien lähdemerkinnät. Kuvassa 8 on esitetty DDB:n taulukot, sarakkeet ja sarakkeiden asiasisältö. Alueet erotetaan toisistaan aluekoodien avulla. Taulukoiden sarakkeet muodostuvat entiteettityypin attribuuttien arvojoukoista. Taulukon rivi käsittää yhden entiteetin tiedot. Ihmisen elinkaaren tapahtumia kuvaavien taulukoiden rivit liitetään tiettyyn henkilöön käyttämällä viiteavaimena henkilönumeroa, joka on *Henkilö*-taulukon avaimen osa. Ihmisen elinkaareen liittyvien entiteettien pohjalta muodostetuissa taulukoissa yhteen henkilöön liittyville merkinnöille on annettu numero, joka kertoo tuohon henkilöön liittyvien merkintöjen lukumäärän ja järjestyksen. Useimmissa taulukoissa tämä numero yhdessä aluekoodin tai seurakuntakoodin ja henkilönumeron kanssa muodostaa taulukon avaimen.¹⁴⁷

Tietokannan päätaulukko, *Henkilö/DDB.xxxPersonv*, on sijoitettu kaikki alueen lähdeaineistossa esiintyvät henkilöt. Jokaisesta henkilöstä on yksi rivi tässä taulukossa. Taulukko sisältää sekä yksilöä identifioivaa tietoa kuten henkilönumeron, nimivariaatiot sekä kuolinajan ja syntymän ajan ja paikan. Lasten lukumäärä, henkilöä koskevien kirkonkirjamerkintöjen alkamis- ja loppumisajankohtaa sekä koko eliniän huomioimista merkitsevien sarakkeiden arvot saadaan vasta sitten, kun lähdeaineisto on kokonaisuudessaan käsitelty. Henkilönumero on ainutkertainen ja muodostuu yhdeksästä numerosta, joista kolme ensimmäistä saadaan syntymävuodesta, muut kuusi ovat juoksevia numeroita. Taulukon avaimen muodostavat aluekoodi ja henkilönumero. Hakujen nopeuttamiseksi useimmat sarakkeiden arvot ovat koodimuodossa.¹⁴⁸

DDB:ssä henkilöiden nimet on tallennettu päätaulukkoon. Henkilöstä tallennetaan useita nimivariaatioita. Tavallisimmin käytetty etu- ja sukunimi on tuloksen helpottamiseksi tallennettu kirjoitetussa muodossa, muut variaatiot ovat koodatussa muodossa. Haut tapahtuvat koodattujen nimien avulla, jotta

¹⁴⁷ Berggren, Danielsson 1989

¹⁴⁸ Berggren, Danielsson 1989

HENKILÖ/DDB.xxxPERSONV (REGKOD - aluekoodi PRN - henkilönro KON - sukupuoli FODFRS - syntymäsrk NAMNKODER - nimitiedot FODAT - syntymän pvm FODAR - synnyinvuosi DODAT - kuoleman pvm DODAR - kuolinvuosi BORTYP - 1. merkinnän tyyppi srk/alue BORDAT - 1. merkinnän pvm srk/alue BORAR - 1. merkintävuosi SLUTYP - viimeisen merkinnän tyyppi srk/alue SLUDAT - viimeisen merkinnän pvm srk/alue SLUAR - viimeinen merkintävuosi AB - aviosynty ANTBB - biologisten lasten lkm ABTBT - lasten kokonaislkm OBSPER - koko elinikä huomioitu/ei)	HENKILÖN LUKUTAITO/ DDB.xxxBETUV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro VOL - sidos SID - sivu SIDTYP - sivun tyyppi RAD - rivi NRRAD - numero rivillä BET - lukutaitomerkit)	HENKILÖN ASUINPAIKAT/ DDB.xxxBOENDE (REGKOD - aluekoodi NOFRS - merkintäsrk PNR - henkilönro ORTKOD - kyläkoodi SUBKOD - paikkakuntakoodi ROTE - ruotu BSTT - talo BOBDAT - asumisen 1. pvm BOSDAT - asumisen viimeinen pvm BINTYP - tulotyyppi BUTTYP - lähtötyyppi)
HENKILÖN SYNTYMÄ/DDB.xxxFDB (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi FODFRS - syntymäsrk FODAT - syntymän pvm FDNR - syntymän järjro/srk DOPFRS - kastesrk DOPDAT - kastepvm FB - monisynty (kaksoset jne.) AB - aviolapsi BM - kättilösynty BNR - lapsen järjro/perhe KT - kirkkoonottaminen KTDAT - kirkkoonottamis(pvm)	HENKILÖN EHTOOLLISKÄYNNIT/ DDB.xxxHL1V (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi VACC - rokotus FAMNR - järjro perheessä SLKOD - asema perheessä/sukulaisuus FNVDAT - 1. ehtoollisen pvm NVO - aikaisempia ehtoollisia NV - ehtoollisten pvm:t HFO - aikaisempia lukusia HF - lukuset)	HENKILÖN SIVIILISÄÄTY/DDB.xxxCIVV (REGKOD - aluekoodi PNR - henkilönro CIVKAL - lähde CIVNR - siviilisäädyn nro VIGNR - vihkimisen nro CIV - siviilisäätty CIBDAT - siviilisäädyn alkamis(pvm) CISDAT - siviilisäädyn loppumis(pvm)
HENKILÖN KUOLEMA/DDB.xxxDBV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi DODFRS - kuolinsrk DODAT - kuolinpvm LAKUTL - lääkärintlausunto DODORS - kuolinsyy DBNR - kuolemantapausten järjro/srk BEGFRS - hautausrk DEGDAT - hautauspvm)	RIPPIKIRJOJEN SIIRTOVIITTEET/ DDB.xxxHL1V (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi FODAT - syntymän pvm FODFRS - syntymäsrk AB - aviolapsi INTYP - merkinnän alkutyyppi UTFK - ulosmuutosrk:n tyyppi FLUFRS - ulosmuutosrk INK - sisäänkirjoittautumis(pvm:n) tyyppi FLIDAT - saapumis(pvm) FLINR - saapumisten järjro/srk UTSK - lähtöviittauksen tyyppi FRNVOL - edellinen sidos FRNSID - edellinen sivu UTTYP - merkinnän lopputyyppi INFK - lähtösrk:n tyyppi FLIFRS - lähtösrk UTK - uloskirjoittautumis(pvm:n) tyyppi FLUDAT - poistumis(pvm) FLUNR - poistumisten järjro/srk INSK - saapumisviittauksen tyyppi TILVOL - seuraava sidos TILSID - seuraava sivu)	HENKILÖN AMMATIT/DDB.xxxYRKEV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro RTYP - suhteen tyyppi TYRNR - ammatin järjro merkinnässä TYRPAR - sulut/ei sulkuja TYRFD - edellinen ammatti TYRREL - oma/toisen ammatti TYRSOC - sosiaalinen asema TYRNHU - elinkeinon pääryhmä TYRNUG - elinkeinon alaryhmä TYRKOD - ammattinro)
HENKILÖN AVIOLIITOT/DDB.xxxLVV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi LYSFRS - kuulutusrk LYSDAT - kuulutus(pvm) LYSNR - kuulutusten järjro/hlö LVLNR - kuulutusten järjro/srk VIGFRS - vihkimissrk VIGDAT - vihkimisen pvm VIGNR - vihkimisen järjro/hlö LVVNR - vihkimisten järjro/srk)	HENKILÖN HUOMAUTUKSET/DDB.xxxANMV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro ANR - huomautuksen osanro ANMKOD - huomautuskoodi ANM - huomautus)	HENKILÖN LÄHDEMERKINNÄT/ DDB.xxxLANKV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi NOKAL - merkinnän lähde VOL - sidos SID - sivunro SIDTYP - sivun tyyppi RAD - rivi NRRAD - nro rivillä INTYP - merkinnän alkutyyppi NOBDAT - merkinnän alkupvm UTTYP - merkinnän lopputyyppi NOSDAT - merkinnän loppupvm)
HENKILÖN MUUTOT/DDB.xxxFLV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro NOTYP - merkinnän tyyppi FLTTY - tulo srk:aan/lähtö srk:sta FLTFRS - muutosrk FLTDAT - muuttopvm FLTNR - muuton järjro/srk FLTHM - muuttoryhmän päämies NFLTM - muuttavien miesten lkm NFLTK - muuttavien naisten lkm)	HENKILÖN HARVINAISET HUOMAUTUKSET/ DDB.xxxANMVARV (NOFRS - merkintäsrk PNR - henkilönro NONR - henkilön merkintöjen järjro ANMKOD - huomautuskoodi ANMNMN - muuttujanimi ANMVRD - muuttujanarvo ANMTXT - huomautusteksti)	HENKILÖN SUHTEET/DDB.xxxRELV (REGKOD - aluenro PNR - henkilönro RTYP - suhteen tyyppi RBES - suhteen määrittäminen RPNR - toisen osapuolen henkilönro RNR - suhteen nro REBDAT - suhteen alkupvm RESDAT - suhteen loppupvm)
		RIPPIKIRJAN OTSIKOT/ DDB.xxxRUBORT (NOFRS - merkintäsrk VOL - sidos SID - sivu SIDTYP - sivutyyppi RAD - rivi ROTE - ruotunro ORTKOD - kyläkoodi)

Kuva 8 DDB:n lähde-merkintöihin perustuvat taulukot, niiden sarakkeet ja sarakkeiden sisältö; alleviivatut sarakkeet ilmoittavat taulukon avaimen.

vastausajat pysyisivät kohtuullisina. Koodit ja niiden selitykset kirjoitetussa muodossa on tallennettu aputaulukoihin, joista selitykset saadaan kirjoitetussa muodossa, kun liitetään hakuun koodeja selittävät taulukot. Kaikki lähteissä esiintyvät erilaiset nimiversiot on alkuperäisessä kirjoitusasussaan tallennettu omaan taulukkoonsa, joka voidaan yhdistää tietokantaan tarvittaessa. Henkilön nimitiedot muodostavat oman kokonaisuutensa ja on suunniteltu niiden siirtämistä omaan taulukkoon.¹⁴⁹

Henkilön syntymästä, kuolemasta, avioliitoista ja muutoista kertovien taulukoiden sarakkeet on muodostettu lähinnä yhden lähdetyypin perusteella. Näissä taulukoissa on sarakkeet tapahtuman paikkaa, aikaa ja kunkin tapahtuman erikoispiirteitä kuvaaville tiedoille. Esimerkiksi henkilön syntymään liittyvä taulukko sisältää sarakkeet, joihin on kirjattu merkintäseurakunta, syntymäseurakunta ja -ajankohta, seurakunnan syntyneiden vuotuinen järjestysnumero, kasteseurakunta ja -ajankohta ja lapsen järjestysnumero perheessä. Lisäksi taulukko sisältää sarakkeet seuraaville tiedoille: syntyikö lapsia yksi vai useampia eli oliko kysymyksessä monisynty, oliko lapsi aviolapsi, auttoiko kättilö synnytyksessä sekä kirkkoon ottamisen ajankohta. Rippikirjojen runsaasta tietoaineksestä on muodostettu useita taulukoita, joissa on tietoja lukutaidosta, ehtoolliskäynneistä sekä asuinpaikoista ja niiden muutoksista. Osaan taulukoista on yhdistetty eri lähdetyyppien tietoja. Asumisen ja suhteen kesto on huomioitu lisäämällä sarakkeet, joissa ilmoitetaan asumisen ja suhteen alkamis- ja loppumispäivämäärä.¹⁵⁰

Eri lähdetyyppien henkilöön liittyvät huomautukset on kerätty kahteen taulukkoon, joissa on tietoja esimerkiksi vihkimisten suorittajista ja henkilön maineesta ja erityisominaisuuksista. Pitkät huomautukset jaetaan ja osat erotetaan toisistaan huomautuksen järjestysnumeron avulla. Useimmissa taulukoissa on sarake,

¹⁴⁹ Berggren, Danielsson 1989

¹⁵⁰ Berggren, Danielsson 1989

jonka arvo ilmoittaa samaan yksilöön liittyvien merkintöjen järjestyksen. Kun tätä järjestysnumeroa käytetään yhdessä henkilönumeron kanssa, saadaan selville, mistä lähdeasiakirjasta merkintä löytyy. Taulukossa on tieto siitä, mistä kaikista lähteistä, miltä lähteen sivulta ja riviltä löytyy henkilöä koskevia merkintöjä sekä lisäksi ajanjakso, jota merkintä koskee.

Henkilön sukulaisuussuhteet sekä ammatin, siviilisäädyn ja asumisen vaihtelut on johdettu eri lähdetyyppien merkinnöistä. Henkilön mahdollisesti useat avioliitot erotetaan järjestysnumerolla. Sukulaisuussuhteet tuotetaan rippikirjoista ja aikaisemman järjestelmän manuaalisesta linkitysprosessista. Jos yksilöltä, esimerkiksi lapselta, puuttuu oma ammatti, voi ammattikoodi tarkoittaa isän tai äidin ammattia. Yksilöllä voi olla samaan lähteen merkintään liitettyinä useita ammatteja, jotka erotetaan järjestysnumeron avulla. Koodia käyttäen ilmoitetaan, onko kysymyksessä esimerkiksi isän nykyinen vai entinen ammatti tai viittaako ammatti sosiaaliseen asemaan. Jokaista ammattikoodia vastaa ainutkertainen ammatti tai titteli. Koodi on hierarkkinen siten, että ensimmäinen numero ilmoittaa sosiaalisen aseman, toinen elinkeinon pääryhmän, seuraavat kaksi numeroa ilmoittavat elinkeinon aliryhmän ja viimeiset kaksi numeroa muodostavat juoksevan numeron. Koodien kirjalliset vastineet löytyvät aputaulukoista. Koodeja käytettiin esimerkiksi paikan- ja henkilönimistä, arkistosidoksista, perheasemasta, sukulaisuussuhteista ja ammasteista. Henkilön sisä- ja ulkolukutaito sekä katekismuksen ja muun kristinopin osaaminen on tallennettu numerokoodina, jotka on muokattu pappien merkintäjärjestelmästä. Nämä koodit on laadittu niin, että niiden perusteella pystytään päättämään, minkälaista merkintäjärjestelmää alkuperäisessä lähteessä on käytetty. Aputaulukot ovat ennen kaikkea tarkoitettut helpottamaan hakuja ja linkitystä.¹⁵¹

¹⁵¹ Berggren, Danielsson 1989

5.5.2 Karjala-tietokannan käyttäjäjärjestelmän taulukot

DDB:n tavoitteet painottavat väestötietokannan luomista ja hyväksikäyttöä, mutta Karjala-tietokannan suunnittelussa päädyttiin toisenlaiseen ratkaisuun: lähdetietokantojen käyttöön. Sen jälkeen kun tiedot on tallennettu, niiden virheetömyys tutkittu tarkistusohjelmien avulla ja havaitut virheet korjattu, pyritään standardoinnilla ja generoituja attribuutteja lisäämällä saamaan tallennetut tiedot käyttökelpoisempaan muotoon ja siten parantamaan niiden käyttöä tutkimuksen lähdeaineistona. Kun tietosuojalautakunta vuonna 1992 oli myöntänyt poikkeusluvan Karjala-tietokannan muodostamiseksi, tuli tallennettujen tietojen käyttö erilaisten tutkimusten lähdeaineistona mahdolliseksi. Samana vuonna valmistui Karjala-tietokannan käyttäjäjärjestelmä nimeltään Karhu. Karhu-järjestelmässä on kunkin kirkonkirjatyypin sidokset, jotka tallennusjärjestelmässä muodostavat oman tietokannan, yhdistetty yhdeksi tietokannaksi, jota kutsutaan rekisteriksi. Tallennuksessa käytetyt taulukot on purettu ja uudet taulukot on muodostettu asiasisällön perusteella. Muutoksella on pyritty rakenteeseen, joka tukee käytön joustavuutta ja tehokkuutta. Karjala-tietokannassa ei henkilöä koskevien tietojen linkitystä ole suoritettu, eikä näin ollen ole myöskään kehitetty ainutkertaista henkilötunnusta. Rekistereissä taulukon rivi sisältää alkuperäislähteen merkintärivillä yhteen henkilöön liittyvää tietoa. Tunnusta Henkilönumero (*Hnro*) käytetään yksilöimään tämä tieto ja samalla se on sidoksen tallennettujen tapausten juokseva numero. Henkilönumeroon liitetään lisäksi sidoksen tunnus. Yhdessä ne muodostavat sarakkeen *Henkilo*, jonka avulla voidaan yksilöidä eri sidosten rekisteriin tallennetut tapaukset. Rekisterissä on samalle henkilölle useita henkilönumeroita Henkilo-tunnuksella. Tutkija voi halutessaan käyttää joko yhden kirkonkirjasidoksen tai samanaikaisesti yhden lähdeyyppin kaikkia tietoja.¹⁵²

Karjala-tietokannan käyttäjäjärjestelmässä rekisterin taulukoiden lukumäärä

¹⁵² Karjalan väestötieteellisen tietokannan loppuraportti 1989

määräytyy lähdeyyppin tietomäärän perusteella. Esimerkiksi syntyneiden ja kastettujen luetteloista on muodostettu kuusi taulukkoa, joiden nimet ja tietosisältö on esitetty taulukossa 9. Taulukossa 10 on esitetty lapsen tietoja sisältävän SHE-taulukon sarakkeet ja niiden tietosisältö.¹⁵³

Taulukko 9 Syntyneiden ja kastettujenluetteloiden taulukot Karhu-hyväksikäyttöjärjestelmässä

Taulukko	Tietosisältö
SHE	lapsen henkilötiedot
SVH	lapsen vanhempien henkilötiedot
SAS	asuinpaikka- ja sidostiedot
SKV	lapsen kastetiedot
SKT	lapsen kummien tiedot
SHM	huomautustiedot

Taulukko 10 SHE-taulukon sarakkeet ja niiden tietosisältö

Sarake	Tietosisältö (lapsen henkilötiedot)
Hnro	henkilönumero (yksi sidos)
Henkilo	henkilönumero (kaikki saman lähdeyyppin sidokset)
SPP	syntymäpäivä
SKK	syntymäkuukausi
SVV	syntymävuosi
SVU	syntymävuosi (lähteeseen merkitty tai generoitu)
Enimi	lapsen 1. etunimi (kirjoitusasu)
Enimi2	lapsen muut etunimet (kirjoitusasu)
N_enimi	normalisoitu lapsen 1. etunimi
N_enimi2	normalisoitu lapsen 2. etunimi
Ssrk	syntymäseurakunta (koodi)
Syntsrk	syntymäseurakunta (normalisoitu asu, lähteeseen merkitty tai generoitu)
Sp	sukupuoli (koodi)
Avios	aviosynty (koodi)
Monis	monisynty (koodi)
Reknro	syntymän rekisteröintinumero /srk
Sjnro	järjestys sisarussarjassa
Kuol	kuolemaan liittyvät merkinnät (koodi)
Yliv	tietojen yliviivaus (koodi)

¹⁵³ Karjalan väestötieteellinen tietokanta 1992

Eri rekisterien taulukoiden lukumäärä vaihtelee neljästä kahteentoista. Kaikille yhteisiä taulukoita ovat tapahtumaan liittyvän henkilön *Henkilötiedot, Tapahtuman tiedot, Asuinpaikka- ja sidostiedot* sekä *Huomautustiedot*. Kuhunkin rekisteriin kuuluu taulukoita, joihin on sijoitettu lähdetyypeille ominaisia tietoja. Esimerkiksi rippikirjojen perusteella on muodostettu kaksitoista taulukkoa, joista niille tyypillisiä tietoja sisältävät *Rippikirjojen siirtotiedot, Kinkeritiedot, Taito-merkinnät, Kirkollinen äänioikeus, Ehtoollismerkinnät, Suhdetiedot* ja *Huomautustiedot*.¹⁵⁴

Koska eri alojen tutkijoiden tietotarpeet ovat moninaisia ja ennalta-arvaamattomia, on Karhu-käyttäjäjärjestelmän suunnittelussa lähdetty periaatteesta, että käyttäjät itse määrittelevät haluamansa tiedot, hakuperusteet ja tulosteet.

5.5.3 Gillespie-järjestelmän käyttäjille kehitetyt taulukot

Gillespie-järjestelmässä käytetään relaatiotietokantaa, sillä sen avulla on mahdollista yhdistää samaan henkilöön tai perheeseen viittaavat tietueet. Järjestelmän kehittämisessä pyrittiin joustavuuteen analysoitavien yksiköiden valinnassa. Ensisijaiseksi kohteeksi on valittu yksilö, mutta tietokanta on organisoitu niin, että on mahdollista analysoida myös perheitä, talouksia tai jopa maapalasta. Taulukoiden rivit sisältävät tietoa yhdestä seurakuntaluettelon tapahtumasta tai väestöluettelossa mainitusta yksilöstä. Jokainen yksilö ja avioliitto saavat ainutkertaiset tunnusnumerot. Nämä tunnusnumerot muodostuvat yleensä sukunimen ensimmäisestä kirjaimesta ja nelinumeroisesta luvusta. Kuitenkin kun nainen avioliiton solmittuaan muutti nimensä, hänelle annetaan lisäksi toinen henkilönumero, joka liittyy uuteen avioliittoon ja sen kautta saatuun nimeen. Näin on menetelty siksi, että on haluttu pitää erillään henkilön asema eri perheissä. Yhteys, jolla ilmoitetaan, että kysymyksessä on sama henkilö,

¹⁵⁴ Karjalan väestötieteellinen tietokanta 1992

hoidetaan ristikkäisillä viittauksilla. Sekä henkilön aikaisemmat, ensimmäiseen perheeseen kuuluvat henkilö- ja perhenumerot että avioliiton kautta saadut uuteen perheeseen liittyvät perhe- ja henkilönumerot tallennetaan samaan taulukkoon. Ristikkäisistä viittauksista saa käsityksen taulukosta 11.¹⁵⁵

Gillespie-järjestelmän käytön kannalta keskeisiä ovat kolme ristiinviittaustaulukkoa. Kaksi näistä, *Person* ja *Family*, antavat mahdollisuuden yhdistää yksilöt ja perheet niin, että saadaan selville henkilöiden väliset suhteet: mihin perheisiin henkilö kuuluu ja keitä kuuluu perheeseen. Taulukossa 11 esitetään *Person*-taulukon tärkeimmät sarakkeet sekä sisältöä kuvaava esimerkkitapaus. Siinä on esitetty kahden henkilön tiedot: toinen on isä, toinen tytär. Kolmas taulukko, *Event*, yhdistää yksilön lähdetietokantoihin, joissa hänestä löytyy mainintoja. Se sisältää henkilön elämän tapahtumaluettelon ja sitä käytetään hakuja suoritettaessa. *Event*-taulukosta on esimerkki Gillespien hakujärjestelmää esittelevässä luvussa, taulukossa 15.¹⁵⁶

Taulukko 11 *Person*-taulukon tärkeimmät sarakkeet sekä sisältöä kuvaava esimerkki

Yksilöno	A 49	A1086	A 30
Sukunimi	Alberthal	Alberthal	Ahrends
Etunimet	Peter	Clara	Clara
Sukupuoli	M	F	F
Syntymäaika	1830.58	1869.35	
Syntymäajan luotettavuus	1 (korkea)	1 (korkea)	0 (puuttuu)
Syntymäpaikka	Prussia	Fred.	Texas
Vanhempien perheno		A 31	
Avioliiton perheno 1	A 31	A 18	A 18
Avioliiton perheno 2	A 37		
Avioliiton perheno 3			
Yksilöno (taaksepäin)		A1086	
Yksilöno (eteenpäin)		A 30	

¹⁵⁵ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

¹⁵⁶ Gutmann, Fleiss, Holmes, Fairchild, Teas 1989

Person-taulukko sisältää sarakkeita, jotka identifioivat yksilön ja kuvaavat hänen pysyviä ominaisuuksiaan. Sarakkeet *Vanhempien perheno*, *Avioliiton perheno* ja *Yksilöno* liittävät hänet muihin tietokannan yksilöihin. Esimerkkitapauksen Peter Alberthal on solminut kaksi avioliittoa, joiden numerot ovat A 31 ja A 37. Hänen tyttärensä Clara on syntynyt ensimmäisestä avioliitosta, minkä kertoo sarake *Vanhempien perheno*, joka on A 31. Hän on itse solminut yhden avioliiton numeroltaan A 18. Syntymäperheen ja avioperheen tiedot yhdistetään Claraan yksilönumeroiden ja perhenumeroiden avulla.

Family-taulukko kertoo yksilöiden perhesuhteet kuten taulukosta 12 nähdään. Event-taulukossa on selvitetty, mistä kaikista lähteistä tiettyyn henkilöön liittyviä merkintöjä löytyy sekä merkintöihin liittyvien tapahtumien päivämäärät ja laatu. On mahdollista, että yksilöstä löytyy enemmän kuin yksi rivi kussakin näistä taulukoista. Event-taulukon kohdalla syy on ilmeinen: jokainen henkilö liittyy todennäköisesti useampaan tapahtumaan. On myöskin mahdollista, että Family-taulukossa henkilö esiintyy useammin kuin kerran. Henkilö liitetään sekä siihen avioliittoon, josta hän on syntynyt, että niihin, jotka hän itse solmii, ja niinpä hänestä pitäisi olla rivi sekä syntymäperheen että avioliiton kautta saatujen perheiden tunnusnumerolla. On myös mahdollista, että yksilöstä on useampia rivejä Person-taulukossa, koska avioliiton solmiminen tuottaa järjestelmässä henkilölle uuden, tuohon avioliittoon liittyvän henkilönumeron. Ristikkäiset viiteavaimet takaavat henkilön tietojen yhdistämisen.¹⁵⁷

Taulukko 12 Family-taulukon tärkeimmäsarakkeet

Perheno	A 31	A 31	A 18
Yksilöno	A 49	A1086	A 30
Suhde	aviomies	lapsi	aviovaimo
Sukunimi	Alberthal	Alberthal	Ahrends
Etunimet	Peter	Clara	Clara

¹⁵⁷ Gutmann, Fleiss, Holmes, Fairchild, Teas 1989

5.6 Hakujärjestelmät

Tiedonsaanti on kokonaan riippuvainen tehokkaasta hakujärjestelmästä, joka mahdollistaa erityyppiset haut ja antaa välineet aineiston muokkaamiseksi. Hauilla ymmärretään tietokannan aineistosta eri perustein muodostettuja tietovalikoimia. Kaikissa esimerkkietokannoissa hakujärjestelmät eroavat toisistaan. DDB:ssä hakuihin käytetään ensi sijassa väestötietokantaa, mutta jossain määrin myös lähdetiedostoja. Karjala-tietokannassa haut tehdään lähdetietokannan käyttäjäjärjestelmästä. Gillespie-järjestelmään on kehitetty ohjelmat, jotka ristikkäistaulukoiden avulla hakevat tietoja lähdetietokannoista.

5.6.1 Demografiska Databasenin lähdetiedostojen hakujärjestelmä

DDB:ssä hakujärjestelmä käsittää työprosessin loppuvaiheen ja muodostuu hakuja suorittavien ohjelmien kehittämisestä ja tuotannosta. DDB sisältää välineet yksilön tietojen saannille ja antaa sen lisäksi mahdollisuuden suorittaa tiettyjä erikoishakuja sekä seurakunta- että aluetasolla. Taulukko 13 kuvaa DDB:n eri osajärjestelmien hakumahdollisuuksia ja niiden välisiä eroja. DDB:ssä kirkonkirjamateriaali on ensiksi saatavana lähdetiedostoina. Lähdetiedostot ovat valmiita käyttöön sitä mukaa, kun tiedot on saatu tallennetuksi. Ne sopivat parhaiten alkuperäisten lähteiden merkintöjen tutkimiseen. Järjestelmässä on valmiita välineitä lähdetiedostojen tietojen hakemiseen. Lähdetiedostot ovat käytettävissä osittain peräkkäistiedostoina, osittain tietokantakehitin *Mimerin* avulla. Lähdetiedostoista on mahdollista saada asiakkaan erittelyn mukaisia, yksinkertaisia valikoimia eri lähdetiedostojen materiaalista. Alkuperäisiin asiakirjoihin verrattuna lähdetiedostojen käytettävyys on huomattavasti parempi, koska tietoja on muokattu ja ne ovat helposti luettavassa muodossa. Lähdetiedostoja on täydennetty lisäämällä niihin erilaista tietoa. Esimerkiksi muuttaneiden luetteloita on kirjattu. Täydennetystä muuttaneiden luettelosta on esi-

Taulukko 13 DDB:n eri osajärjestelmien väliset hakumahdollisuudet ja niiden erot

Osajärjestelmä	Hakumahdollisuudet
Kirkonkirjat	Lähdesidoksen tieto: hakuja, jotka koskevat yksittäisten lähteiden merkintöjä eli antaa seurakunnittain pääsyn lähteisiin niiden alkuperäisessä muodossa
Väestö/seurakunta	Seurakunnan kirkonkirjojen tieto: hakuja, jotka koskevat yksilöitä yksittäisessä seurakunnassa eli antaa pääsyn väestötietoihin seurakunnittain
Väestö/alue	Koko alueen kirkonkirjojen tieto: hakuja, jotka koskevat yksilöitä koko alueella eli antaa lopultakin pääsyn koko alueen väestön informaatioon

merkki liitteessä 6. Lisätyt tiedot on merkitty sulkuihin. Joistakin lähdesarjoista on tehty henkilörekistereitä, joihin on liitetty erilaisia lajittelumahdollisuuksia. Liitteen 7 esimerkissä kuolleidenluettelon tietoja on lajiteltu kuolinsyyn mukaan. Lähdetiedostoista voidaan hakea materiaalia mikrotietokoneella muokattavaksi, mitä varten on kehitetty standardiohjelmat *Surf* ja *Resan*. Surf-ohjelman avulla voidaan henkilörekistereistä hakea yksittäisen henkilön tietoja. Etsittävien joukkoa on mahdollista rajoittaa esimerkiksi syntymäajan ja sukunimen mukaan. Joukosta voidaan sitten valita haluttu henkilö lähempää tarkastelua varten. Ohjelma tuottaa luettelon lähteistä, joissa tutkittavasta henkilöstä löytyy merkintöjä. Resan-ohjelmalla on mahdollista saada selville henkilön sukulaiset ja esittää hänen sukutaulunsa, mistä on esimerkki liitteessä 8.¹⁵⁸

¹⁵⁸ Danell 1985, s. 149-160

5.6.2 Demografiska Databasenin väestötietokannan hakujärjestelmä

DDB:n väestötietokannan tietokannanhallintajärjestelmänä on SQL/DS, jonka kyselykieli on SQL. Tiedon saanti väestötietokannasta tapahtuu joko eräajoina tai vuorovaikutteisen kyselyohjelman avulla. Eräajo-ohjelmia käytetään pääasiassa standardihakujen tekemiseen. Vuorovaikutteiset ohjelmat antavat tutkijalle mahdollisuuden itse suorittaa hakuja tietokannasta ja analysoida aineistoa erilaisten ohjelmien avulla. SQL-kyselykielen avulla on helppo hakea tietoa, jos käytetään vain muutamia taulukoita. Usein on kuitenkin välttämätöntä liittää useita taulukoita oikean tiedon saamiseksi. Hakujen helpottamiseksi on kehitetty vuorovaikutteinen hakujärjestelmä, Indiko, jonka avulla tutkijat voivat luoda kohortteja haluttuja yhteisiä ominaisuuksia omaavista ihmisistä. Tutkija voi myös valita vain yhden ihmisen. Kaikki elämäntapahtumat voidaan esittää graafisten elämänlinjojen avulla. Kuva 11 tarjoaa esimerkin graafisesta elämänkuvasta. Se esittää perherakennetta, jossa näkyvät perheenjäsenten syntymät, kuolemat, avioliitot sekä muutot seurakunnassa, seurakuntaan ja sieltä pois. Kaavioon on liitetty henkilönnumerot, perheasema sekä henkilöiden syntymävuosi.¹⁵⁹

DDB:stä valitut tiedot voidaan muuttaa erilaisiksi tiedostotyypeiksi lisäanalyysia varten, esimerkiksi SAS-tilasto-ohjelman käsittelyyn sopiviksi tiedostoiksi. DDB tuottaa standarditilastoja rekisteröityjen seurakuntien tiedoista. Lähes joka vuodelta tehdään tilastoja syntyneiden, avioliittojen, seurakuntaan ja sieltä pois muuttaneiden ja kuolleiden määrästä iän ja sukupuolen mukaan jaoteltuna. Joka viidenneltä vuodelta tehdään tilastoja väestön jakaantumisesta eri lähteissä ja demograafisina mittareina käytetään tällöin esimerkiksi ikää ja sukupuolta, syntyvyyttä, avioliittojen määrää ja kuolleisuutta. Jonkin verran yleistä tilastoa on saatavilla myös elämäkertoista jaoteltuna eri syntymäkohorttien mukaan. Tämän lisäksi on saatavilla vuosittaisia tilastoja väestöstä, avioliitoista, muutto-
liikkeestä, asumisesta, kuolleista ja syntyneistä siviiliaseman, sukupuolen ja iän

¹⁵⁹ Preliminary draft 1989

Henkilö numero	Perhe- asema	Henkilön nimi
01924	WF	Sonesdotter, Pernilla (Pernilla Sonentytär)
01553	HB	Påhlsson, Hans (Hans Pählinpoika)
02730	HB	Jonasson, Johannes (Johannes, Jonaksen poika)
03958		Dödfödda d. (kuolleena syntynyt tytär)
04027		Petter (Petter, Jonaksen poika)
04172		Hansdotter Maria (Maria, Hansin tytär)
04353		Hansdr, Sophia (Sophia, Hansin tytär)
04522		Hansdotter, Johanna (Johanna, Hansin tytär)
04658		Carolina (Carolina, Hansin tytär)
04781		Hansdotter, Anna Stina (Anna Stina, Hansin tytär)
05238		Jonasson, Emma Karolina (Emma Karolina, Jonaksen tytär)
05671		Hans Petter (Hans Petter, Jonaksen poika)

Hnro	Sytymä- vuosi	1820	1830	1840	1850	1860	1870	1880	1890
01924	1822	B			M	W	M		D
01553	1817				M		D		
02730	1833		B			I	C	CM	W
04027	1846				B		D		
04172	1848				B		MC		C
04353	1850				B		D		
04522	1851				B			M	
04658	1853				D				
04781	1854				B			O	I
05338	1859					B			
05671	1864						D		

Perheasema: WF=Vaimo, HB= Miäs

Elämänlinja:

—	Henkilö asuu seurakunnassa
B	Syntymä (birth)
D	Kuolema (death)
I	Muutto seurakuntaan (in-migration)
O	Muutto seurakunnasta (out-migration)
C	Muutto seurakunnan sisällä (circular migration)
M	Vihkiminen (marriage)
W	Avioliiton purkautuminen (dissolution of marriage)

Kuva 9 Esimerkki Demogra-ohjelmalla tapahtuneesta hausta. Pernilla Sonentyttären, hänen puolisoidensa ja perheen lasten kirkonkirjoista löytyneiden merkintöjen perusteella on laadittu yllä esitetyn kaltainen elämänkaarikuvaus.

mukaan tilastoituna.¹⁶⁰

Toistaiseksi ovat yksilötiedot ja tilastotiedot yleisin DDB:stä toimitetun tiedon muoto. Tutkija voi kuitenkin tilata muuttujista haluamansa valikoiman, joka toimitetaan magneettinauhalla tai levykkeellä. DDB:ssä on kehitteillä tehokkaampia välineitä kaiken tiedon saamiseksi vuorovaikutteisten hakujen avulla.¹⁶¹

Väestötietokanta sisältää kaikki lähdetiedostojen tiedot ja toteuttaa sille asetetut tehokkuus- ja joustavuusvaatimukset. Taulukot on pyritty rakentamaan niin, että yleisimmin haettujen tietojen hakuajat ovat mahdollisimman lyhyitä ja joustavat haut käyttäjän antamien kriteerien mukaan ovat mahdollisia. Prio-
risointeja tekemällä on saatu yksiselitteiset tiedot esimerkiksi henkilön syntymän päivämäärästä, joka voi vaihdella eri lähteiden välillä. Hakujen helpottamiseksi on taulukoihin lisätty sarakkeet tutkijoiden usein tarvitsemille, lähdemerkintöjen pohjalta generoiduille tiedoille, joita ovat esimerkiksi siviilisääty ja sukupuoli.¹⁶²

5.6.3 Karjala-tietokannan hakujärjestelmä

Karjala-tietokannan käyttäjäjärjestelmässä, Karhussa liikutaan valikon avulla. Päävalikosta valitaan haluttu tietolähde eli tietokanta- ja henkilöhakemistot, syntyneiden ja kastettujen luettelot, muuttaneiden luettelot, vihittyjen luettelot, kuolleiden ja haudattujen luettelot, lastenkirjat tai rippikirjat. Tietokanta- ja henkilöhakemistossa on tietoa tallennetuista seurakunnista, seurakuntien arkistoista, tietokannan taulukoista ja taulukoiden sisällöstä. Henkilöhakemisto sisältää seurakunnittain perustiedot kaikista kirkonkirjoihin merkityistä henkilöistä. Henkilöhakemistossa heistä kerrotaan etunimi, patronyyminimi,

¹⁶⁰ Preliminary draft 1989

¹⁶¹ Preliminary draft 1989

¹⁶² Danell 1985, s. 88-91

sukunimi, os. sukunimi, sukupuoli, syntymäaika, kirkonkirjasivu, kirkonkirjan arkistotunnus eli signum, sidoskoodi ja henkilönnumero. Hakemistossa olevat nimet on normalisoitu yhtenäiseen muotoon. Haku voidaan suorittaa minkä tahansa perustiedon perusteella, mutta on nopeampi, jos tiedetään henkilön syntymävuosi ja etu- tai sukunimi. Jos henkilöhakemistosta haetaan esimerkiksi hakuehdolla "Mannonen" ja "1887" saadaan taulukon 14 mukainen luettelo. Taulukosta 14 havaitaan esimerkiksi, että Anni Mikontytär Mannonen esiintyy seurakunnan kirkonkirjoissa seitsemän kertaa. Arkistotunnuksen eli signumin perusteella päätellään, että merkinnät ovat rippikirjoissa. Nyt kun tiedetään, mistä lähdeyyppistä merkintöjä etsitään, voidaan henkilönumeroiden avulla tutkia rippikirjalähdetietokannan taulukoita ja löytää tietoja Anni Mikontytärrestä, hänen puolisoistaan, lapsistaan, asuinpaikastaan ja vanhemmistaan. Sukututkimus Karjala-tietokantaa käyttäen muistuttaa paljon alkuperäisten kirkonkirjojen perusteella tehtävää tutkimusta ja edellyttää tarkempaa perehtymistä asia-kirjoihin kuin DDB-tietokannan käyttäminen, jossa jokaisella henkilöllä on ainutkertainen henkilönnumero, jonka avulla hänestä esiintyvää tietoa voidaan hakea. Henkilöhakemisto onkin tarkoitettu lähinnä sukututkijoiden käyttöön.¹⁶³

Taulukko 14 Ote Karhu-käyttäjäjärjestelmän henkilöhakemistosta etunimen mukaan jaoteltuna

Etunimi ja patron.	Sukunimi	Synt.aika	Sivu	Signum	Henkilönro	
Anni	Mannonen	2/ 8/1887	87	I Bab	1201	1700
Anni Matintytär	Mannonen	2/ 8/1887	4	I Aa 12	00120	29
Anni Matintytär	Mannonen	2/ 8/1887	5	I Aa 14	00140	24
Anni Matintytär	Mannonen	2/ 8/1887	5	I Aa 15	00150	23
Anni Mikontytär	Mannonen	17/ 7/1887	4	I Aa 9	00090	75
Anni Mikontytär	Mannonen	17/ 7/1887	6	Aa 10	00101	88
Anni Mikontytär	Mannonen	17/ 7/1887	67	I Aa 1	00101	968
Anni Mikontytär	Mannonen	17/ 7/1887	27	I Aa 10	00102	124
Anni Mikontytär	Mannonen	17/ 7/1887	114	I Aa 11	00110	147
Anni Mikontytär	Mannonen	17/ 7/1887	151	I Aa 14	00140	1195
Anni Mikontytär	Mannonen	17/ 7/1887	151	I Aa 15	00150	851
Edla	Mannonen	21/ 1/1887	1	I Aa 9	00090	4

¹⁶³ Karjalan väestötieteellinen tietokanta 1992

Tietokannan hallintajärjestelmä dBase IV ja kyselykieli SQL antavat käyttäjille mahdollisuuden hakea lähdetietokannoista sarakkeiden arvoja ja myös itse määrittellä sarakkeet, joihin haku kohdistuu. Haettuja tietoja voidaan lajitella ja tulostaa niistä erilaisia raportteja. Käytön periaate on siis, että käyttäjä voi hyödyntää valmiita tulostusohjelmia, mutta hänellä on mahdollisuus myös omiin hakuihin ja tulostuskomentoihin. Tallennetuista lähdetietokannan tiedoista haetaan ja tulostetaan lähdesarjoittain eri perustein määriteltyjä tilastoja, hakemistoja ja taulukoita, joista painatetaan Karjala-tietokantajulkaisu. Karjala-tietokannan käyttäjäjärjestelmään kuuluvat sidoskohtaiset vakioraportit, joita voidaan havainnollistaa Applause II -grafiikkaohjelmiston avulla.¹⁶⁴

Karjala-tietokannan tietokannan hallintajärjestelmän kyselykielenä on SQL. SQL on erinomainen väline sitoa joustavalla tavalla yhteen valinnaisia tietokannan taulukoita. Monimutkaisia hakuja tehtäessä tarvitaan kuitenkin tietoa siitä, miten sitominen tapahtuu, että varmasti saataisiin haluttu valikoima. Tiedon haku on yksinkertaista silloin, kun tietoja haetaan vain muutamista taulukoista.¹⁶⁵

SQL-kyselykielen lisäksi voidaan Karjala-tietokannassa käyttää tietojen hakuun dBase IV:n ohjauskeskusta. Ohjauskeskus tukee *query by example* -lähestymistapaa, jossa näytöllä olevaan taulukkoon voidaan suoraan kirjoittaa ehtoja, joita hauissa noudatetaan. Ohjauskeskus mahdollistaa vuorovaikutteisten kyselyjen tekemisen eikä edellytä SQL-kielen osaamista. Kiinnostavat relaatiot on mahdollista valita relaatioluettelosta ja poimia halutut sarakkeet tulosrelaatioon käyttäen yhdistävänä linkkinä jotain yhteistä saraketta. Tietyt rivit sarakkeista saa, jos käyttää ehtoruutua. Näennäistaulukkoa voidaan muokata kysely-, kooste- ja vertailuoperaattoreita sekä aritmeettisiä funktioita käyttäen. Ohjauskeskusta käyttäen voi myös laatia ja muokata erilaisia raportteja ja lo-

¹⁶⁴ Karjalan väestötieteellinen tietokanta 1992

¹⁶⁵ Karjalan väestötieteellinen tietokanta 1992

makkeita. Koska Karjala-tietokannassa ei henkilön eikä häneen suhteessa olevien henkilöiden tietoja ole linkitetty, ei yksikäsitteistä henkilönnumeroa käyttämällä voi selvittää henkilön elinkaaren tapahtumia, mutta henkilöhakemiston avulla on mahdollista etsiä yksittäisen henkilön tiedot. Jos Karjala-tietokannassa halutaan etsiä rippikirjoista tietoja henkilöistä, joiden syntymäaika on vuosien 1890 ja 1900 välillä näyttäisi kysely seuraavalta:

```
select n_enimi, sukun, svv
from rn285
where svv between 1890 and 1900.166
```

5.6.4 Gillespie-järjestelmän hakujärjestelmä

Gillespie-järjestelmässä on tiedon hakua varten muokattu ohjelmat *Family* ja *Person*. *Person*- ja *Family* -taulukoiden ristikkäiset viiteavaimet tekevät mahdolliseksi etsiä yksilön ja hänen perheensä tiedot sekä henkilöiden väliset suhteet. *Family*-ohjelma on suunniteltu tunnistamaan ydinperheen jäsenet ja näiden yhteydet sukulaisperheisiin ajassa eteen- ja taaksepäin. Jos haetaan joko perheen isän nimellä tai perhenumerolla, *Family*-ohjelma tuottaa luettelon ydinperheen jäsenistä, heidän suhteistaan toisiinsa ja heidän henkilönumeronsa. *Person*-ohjelma esittää joko nimellä tai henkilötunnuksella haetun henkilön elämän kulun sellaisena kuin se on lähteisiin kirjattu. Elämätkuvalomakkeessa ovat yksilön henkilötunnukset, nimet, syntymäpaikka ja -aika, vanhempien ja avio-puolison nimet sekä perhenumerot. Lisäksi siinä on numeroitu luettelo lähteistä, jotka sisältävät merkintöjä henkilöstä.¹⁶⁷

Gillespie-järjestelmässä tiedon saanti perustuu indeksin eli hakemiston käyttöön, mikä mahdollistaa tietojen nopean haun. Fyysinen järjestys on sama, jossa

¹⁶⁶ Ullman 1988, s. 195-197, Karjalan väestötieteellinen tietokanta 1992

¹⁶⁷ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

tietueet on tallennettu. Hakemistoavaimen avulla tietoja voidaan hakea tutkimuksen kannalta tarkoituksenmukaisesti, esimerkiksi sukunimen tai henkilötunnuksen perusteella, ja käsitellä niitä halutulla tavalla. Kastetietokannan hakemistoavaimina voidaan käyttää vanhempien tai lasten nimiä. Jos haku tapahtuu vanhempien nimen perusteella, voidaan tutkia syntyvyystasoa, lasten nimillä haetuista tiedoista voidaan tutkia esimerkiksi avioitumisikää. Haut yhteen tai useampaan tietokantaan voidaan ryhmitellä ja muodostaa vastausten perusteella raportti esimerkiksi perherakenteesta. Perherakennelomakkeita voidaan saada jonkin väestöluettelon kaikista perheistä tai halutun listan mukaan. Tuloksia voidaan analysoida joko välittömästi tietokannan hallintajärjestelmä Foxbase+:lla tai tutkia syntyvyyttä, avioliittoastetta ja kuolleisuutta tilasto-ohjelmia käyttäen.¹⁶⁸

Prosessit, joilla perheiden ja yksilöiden elämän kulkua kuvaavia tietoja kerätään, perustuvat relaatiotietokannan hallintajärjestelmän toimintoihin. Person- ja Family -taulukot on molemmat varustettu hakemistolla. Taulukon Person hakemistoavaimina ovat nimi ja henkilönnumero, taulukon Family nimi ja perhenumero ja kummastakin on lisäksi indeksoimaton muoto. Jos tunnetaan perheen aviomiehen ja vaimon nimi, saadaan Family-taulukosta perheen rakenne ja sen jäsenet perhenumerolla. Person-ohjelma käyttää Person-taulukkoa. Henkilön nimellä saadaan aakkosellinen nimiluettelo, ja sen jälkeen henkilönnumerolla henkilön elämän kulku. Henkilön elämehistoriassa on ilmoitettu vanhemmat sekä aviopuoliso, mikä on mahdollista siksi, että Person-taulukko sisältää kaikkien niiden perheiden perhenumerot, joissa henkilö on jäsenenä joko syntymän tai avioliiton kautta.¹⁶⁹

Gillespie-järjestelmän tärkein ominaisuus on kyky analysoida henkilöiden elämäkertoja tehokkaasti. Event-taulukko on tapahtumaluettelo, jonka sisällöstä

¹⁶⁸ Gutmann, Fliess, Holmes, Fairchild, Teas 1989, Ullman 1988, s. 3, 312-323

¹⁶⁹ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

saa kuvan taulukosta 15. Event-tilukko käsittää viittaukset lähdetaulukoiden riveille, joilta voidaan rakentaa yksilön, perheen tai koko yhteisön elämän kulku. Analyysissä on otettava huomioon, että toiset merkinnät osoittavat muutoksia yksilön tai perheen olosuhteissa tietyssä aikana ja toiset vain merkitsevät muistiin yksilön tai perheen ominaisuuksia. Analyysissä on olennaista tietää muutoksen aika. Event-tilukko sisältää henkilö- ja perhenumeron lisäksi lähdetyyppiin, tapahtuman aikaan ja paikkaan liittyvät sarakkeet sekä henkilön aseman tapahtumassa ilmoittavan sarakkeet. Relatiotietokanta antaa mahdollisuuden liikkua nopeasti Event-tilukosta todelliseen lähdeaineistotilukkoon. Paikka, lähde, lähteen tyyppi ja tapahtuman järjestysnumero tarvitaan rivin identifioimiseen. Tietueet tallennetaan kirjaamisjärjestyksessä ja kullekin lähdetypille annetaan tallennusjärjestyksen mukaiset numerot. Niinpä on vain yksi kasteluetteloiden numero, jonka arvo on 1067. Event-tilukko on indeksoitu kirkon nimen ja järjestysnumeron mukaan.¹⁷⁰

Gillespie-järjestelmä on muodostettu kahdesta linkitetystä modulistista relatiotietokanta-järjestelmässä.. Person- ja Family-tilukot muodostavat identifiointimodulin, jonka avulla yksilöidään ihmiset ja esitetään yhteydet heidän välillään. Identifiointimoduli on yhteydessä tapahtumamoduliin, joka

Taulukko 15 Event-tilukon tärkeimmät sarakkeet

Henkilöno	A1086
Perheno	A 31
(Tapahtuman) paikka	Fredriksburg
Lähde	Vereinskirche (kirkko)
(Lähteen) tyyppi	Kasteluettelo
Tapahtuman järjestysno lähteessä	1067
Merkinnän päivämäärä	1869.35
Päivämäärän luotettavuus	1 = korkea
Henkilön asema tapahtumassa	henkilö kastettu , koodi = 1
Tapahtuman luonne	syntymä, koodi = 2

¹⁷⁰ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

sisältää linkitetyt lähdeaineistotaulukot. Event-taulukko liittää yksilöt ja perheet lähdetaulukoiden merkintöihin, joissa heidät on mainittu. Gillespie-järjestelmän vahvuus on kyky liikkua eri modulien välillä. Henkilö voidaan nimen avulla etsiä Person-taulukosta ja löytää hänen vanhempansa tai aviopuolisonsa Family-taulukosta ja sitten siirtyä Event-taulukkoon ja löytää sieltä demograafisesti tärkeitä tapahtumia.¹⁷¹

Gillespie-järjestelmän analysointimenetelmillä voidaan arvioida väestön kokoa annettuna ajankohtana. Väestö voidaan jakaa erilaisten ominaisuuksien kuten ikä, sukupuoli, aviollinen sääty, varakkuus, poliittinen suuntaus, luku- ja kirjoitustaito mukaan ryhmiin ja saada siitä tulosteita ominaisuuksien mukaan lajiteltuna. Väestöä koskevat tiedot ovat kahdenlaisia: joko muuttumattomia ominaisuuksia kuten sukupuoli ja etunimet tai tapahtumia, esimerkiksi syntymää ja avioliiton solmimisia. Tietojen analyysi vaatii määrittämään, mitkä tapahtumat ovat todennäköisiä erilaisia ominaisuuksia omaaville ryhmille. Esimerkiksi kuolleisuuden tutkimista varten ihmiset jaetaan iän perusteella ikäryhmiin ja riippuen siitä, mihin ikäryhmään ihminen kuuluu, hänelle arvioidaan todennäköinen jäljellä oleva elinaika niiden ihmisten tietojen perusteella, joista tutkimuksessa on muodostettu elämätkuva.¹⁷²

¹⁷¹ Gutmann, Fliess, Holmes, Fairchild, Teas 1989

¹⁷² Gutmann, Fliess, Holmes, Fairchild, Teas 1989

6 YHTEENVETO

Tämän tutkimuksen tarkoituksena on ollut antaa yleiskuva henkilöhistorian tietokannoista ja niiden lähdeaineistosta. Henkilöhistorian tietokannat eroavat liiketaloudellisista tietokannoista paitsi lähdeaineiston historiallisen alkuperän myös käytön suhteen. Sen jälkeen kun lähteiden tiedot on siirretty tietokoneelle ja suoritettu tarvittavat tarkistukset ja korjaukset, tietokantoja ei enää tarvitse päivittää, vaan niiden käyttö rajoittuu hakuihin tietokannoista. Lisäksi tutkimuksessa on yhdenmukaisesti kuvattu ja verrattu kolmea henkilöhistoriallista relaatiotietokantaa, jotka ovat ruotsalainen Demografiska Databasen, suomalainen Karjala-tietokanta ja amerikkalainen Gillespie-järjestelmä. Tavoitteena on ollut muodostaa kehys, jota voitaisiin käyttää henkilöhistorian tietokantojen arvioinnissa.

Toisessa luvussa on käsitelty kirkonkirjojen rakennetta ja asiasisältöä. Päämääränä on ollut esitellä, minkälaisien tietojen tallentamiseen henkilöhistorian tietokantoja rakennettaessa on varauduttava. Yhden väestöluettelolajin rakenteen ja asiasisällön tuntemus antaa suuntaviivoja myös muiden käsittelyssä, sillä vaikka eri historialliset väestöluettelot eroavat tietosisällöltään alkuperäisen tehtävänsä mukaan, niin ne rakenteeltaan muistuttavat toisiaan. Toisessa luvussa on myös selvitetty eräitä kirkonkirjojen käyttömenetelmiä sekä kirkonkirjoja lähdemateriaalinaan käyttävää tutkimusta. Käyttömenetelmien valinta riippuu tieteenalan kysymyksenasettelusta ja tavoitteista. Historiallisten väestöluetteloiden tietosisältö määrää, minkä tieteenalan tutkimukseen ne soveltuvat. Kirkonkirjojen tiedot syntyneistä, avioliiton solmineista ja kuolleista ja haudatuista, ehtoollisella ja kinkereillä käymisestä ja lukutaidosta ovat tärkeää lähdeaineistoa sukututkimuksen, perhehistorian ja historiallisen perhedemografian tutkijoille, ja siksi niiden käytön tehostaminen ja helpottaminen automaattisen tietojenkäsittelyn avulla on perusteltua.

Kolmannessa luvussa on selvitetty historiallisten väestöluetteloiden erityis-

piirteitä, jotka aiheuttavat ongelmia tallennuksessa. Kirkonkirjat ovat sisällöltään heterogeenisia. Niitä on pidetty vuosisatoja ja aikojen kuluessa niihin merkityt tiedot ovat vaihdelleet. Tieto, joka jonakin aikana on merkitty yleisesti, puuttuu toisena aikana kokonaan. Koska ei voida tietää, mikä tieto on merkittävää tutkijalle, on kaikki tieto, myös harvinaiset ja erikoistapaukset, tallennettava. Tämä johtaa taulukoina toteutetuissa relaatiotietokannoissa tyhjiin kenttiin. Kirkonkirjat sisältävät rakenteisen tiedon ohella vaihtelevan mittaista tekstitietoa. Tekstitiedon tallennus kiinteämittäisiin taulukoihin on osittain ratkaistu käyttämällä tekstiä korvaavia koodeja, joiden pituus on suunnittelijan määräämä. Saman aikakauden kirkonkirjat kuten muutkin väestöluettelot sisältävät tietoja samoista ihmisistä. Tutkijalle on tärkeää yhdistää nämä erilliset tiedot ja saada kokonaiskuva kunkin henkilön elämästä. Linkityksen automaattiseen käsittelyyn onkin kehitetty edistyneitä ohjelmistoja. Koska vertailtavien tietojen lukumäärä on suuri, linkitys sopii hyvin tietokoneella suoritettavaksi. Henkilön tietojen yhdistämiseen liittyy läheisesti nimien kirjoitusasujen yhtenäistäminen, joka usein muodostaa linkitysohjelmistojen ensimmäisen vaiheen. Nimien normalisointiin on kehitetty sekä kielen fonetiikkaan että kirjainten aseman tarkasteluun perustuvia algoritmeja. Nimien normalisointiohjelmiä pitäisi kehittää siten, että samalla ohjelmistolla voitaisiin vertailla eri kielisiä nimiä. Nimien normalisointi ei ruotsalaisten ja suomalaisten kirkonkirjojen tietojen linkityksessä ole yhtä ongelmallista kuin muualla, sillä rippikirjat sisältävät valmiit perherakenteet sekä muita tunnistamista helpottavia yksityiskohtia. Suomessa ja Ruotsissa eivät myöskään sukunimet olleet vanhimpina aikoina kovin yleisiä, mikä vaikeuttaa niiden käyttöä linkityksessä.

Neljännessä luvussa on lyhyt selonteko projekteista, joiden päämääränä on ollut suunnitella ja toteuttaa henkilöhistoriallisia tietokantoja. Näissä projekteissa tietokannan lähdeaineiston muodostavat erityyppiset historialliset väestöluettelot, eikä niitä aina ole toteutettu relaatiotietokantoina. Selonteon tavoitteena on ollut sitoa myöhemmin luvussa viisi kuvatut esimerkkietokannat yleiseen henkilöhistorian tietokantojen kehukseen. Tietotekniikan ja ohjelmistojen

kehittyminen näkyy myös henkilöhistorian tietokantojen toteutuksessa. Kaupalliset ohjelmistot ovat monipuolistuneet, mikä on lisännyt niiden käyttöä, vaikkakin usein itse suunnitelluilla ohjelmilla täydennettynä. Aluksi henkilöhistorian tietokannat toteutettiin suur- tai pienkoneilla, mutta myöhemmin, 1980-luvulta lähtien, mikrotietokoneita on alettu käyttää myös suuria tietomääriä sisältävien historiallisten väestöluetteloiden tallentamiseen. Monet projektit esimerkiksi Demografiska Databasen, Cambridge Group, SOREP eli kanadalaisien Yliopistojen välinen väestön-tutkimuskeskus, Tukholman historiallinen tietokanta, Karjala-tietokanta ja Gillespie-järjestelmä on toteutettu relaatiotietokantoina. Saksassa Max Planck -Instituutissa kehitetty Kleio-ohjelmisto perustuu tietomalliin, joka muistuttaa semanttista verkkoa ja tarjoaa olio-orientoituneen komentokielen.

Viidennessä luvussa on yhdenmukaisesti kuvattu ja verrattu kolmea relaatiotietokantaa, jotka ovat ruotsalainen Demografiska Databasen eli DDB, suomalainen Karjala-tietokanta sekä amerikkalainen Gillespie-järjestelmä. DDB:n ja Karjala-tietokannan dokumentoinnissa on kiinnitetty runsaasti huomiota lähdevalmisteluun eli lähteiden rakenteen ja sisällön selvittämiseen ennen tallennukseen ryhtymistä, mistä työstä on myös käyttäjälle hyötyä. Tutkielman käyttöön saaduissa lähteissä ei Gillespie-järjestelmän tätä puolta ole huomioitu. Esimerkkietokantojen tavoitteena on, että kaikki lähdeaineiston sisältämä tieto on lähdeuskollisesti tallennettava. Tietosisältöjen yhdenmukainen vertaaminen on mahdollista, mutta siinä on kuitenkin otettava huomioon, että DDB ja Karjala-tietokanta perustuvat rakenteeltaan ja sisällöltään samanlaisiin lähteisiin. Gillespie-järjestelmän lähteistä puuttuvat rippikirjojen monipuoliset tiedot, mutta kirkonkirjoja on täydennetty vero- ja väestölaskentaluetteloiden merkinnöillä. Lähteiden perusteella on pääteltävissä, että lähdemerkinnöistä tallennetut tiedostot ovat kaikissa hyvin samankaltaisia.

Tietosisällön ohella muita keskeisiä tarkastelun kohteita tietokantojen arvioinnissa ovat käytön tehostamista varten muokatut rakenteet sekä hakujär-

jestelmät. Hakujärjestelmien yhdenmukainen vertailu on ongelmallista. Tietokantojen käytölle asetetut tavoitteet poikkeavat ja lisäksi toteutuksessa on käytetty erilaisia ohjelmistoja ja ratkaisuja. Siksi tietokantojen käytön tehostamiseen kehitettyjen rakenteiden ja hakujärjestelmien kohdalla on eroja esimerkkietokannoissa. DDB:ssä jokaisella tietokantaan kuuluvalla henkilöllä on ainutkertainen henkilönnumero, jonka avulla on mahdollista selvittää henkilön koko elämäntietä sekä hänen perheenjäsenensä. Karjala-tietokannassa ei tietosuojavaltuutetun kielteisen kannan takia ole ollut mahdollista suorittaa samaan henkilöön liittyvien tietojen linkitystä, mikä tekee hakujärjestelmästä kömpelön. Koska henkilön tietoja ei voida hakea yksikäsitteisen henkilönnumeron avulla, on käytettävä käyttäjäjärjestelmän tarjoamia mahdollisuuksia eli SQL-kielisiä kyselyjä, ohjauksrakennetta ja henkilöhakemistoa. Gillespie-järjestelmässä on käytössä sekä henkilönnumero että perhenumero. Kumpikaan näistä ei ole ainutkertainen, sillä henkilö saa avioliiton solmiessaan uuden henkilö- ja perhenumeron. Henkilön tietojen hakua varten tarvitaan taulukot, joihin on tallennettu sekä henkilön edelliset että tulevat tunnistusnumerot.

Tämän tutkimuksen lähteet muodostuvat lehtiartikkeleista ja esimerkkietokantoja käsittelevistä raporteista. Lähteissä käsitellään ja painotetaan henkilöhistorian tietokantojen rakentamisen eri puolia: linkitystä, nimien normalisointia, lähdevalmistelua, tallennusta, hakurakenteita ja -järjestelmiä. Tutkimuksessa on kuitenkin pyritty luomaan monipuolinen kokonaiskuva tietokantojen rakentamiseen liittyvistä tehtävistä ja ongelmista. Siksi esimerkiksi lähdevalmistelu on kuvattu DDB:n ja Karjala-tietokannan raporttien pohjalta, koska Gillespie-järjestelmästä ei vastaavaa tietoa ole ollut saatavilla. Karjala-tietokannassa ei ole suoritettu linkitystä. Gillespie-järjestelmän osalta tieto puuttuu, vaikka linkitys ilmeisesti on tehty. Tämän vuoksi on linkitystä käsittelevässä luvussa esitelty muissa tietokannoissa käytettyjä ratkaisuja. Nimien normalisointia ei Karjala-tietokannassa suoriteta automaattisen tietojenkäsittelyn avulla. Gillespie-järjestelmää käsittelevissä lähteissä ei nimien normalisointia ole selvitetty. Tutkimuksessa on kuitenkin haluttu tarkastella erilaisia, automaattiseen tietojenkäsitte-

lyyn perustuvia nimien normalisointimenetelmiä, joita henkilöhistorian tietokannoissa on käytetty. Siksi on kuvattu sekä kieleen perustuvia sekä kirjainten aseman vertailua koskevia menetelmiä. Koska esimerkkietokantoja jatkuvasti parannetaan, osa lähteiden tiedoista saattaa joiltakin yksityiskohdiltaan olla vanhentunutta, mutta ei olennaisilta osiltaan.

Tietokantojen käytöstä ei tutkimuksessa ole tehty empiiristä tutkimusta, vaan on tyydytty tietokantojen kuvaamiseen kirjallisuuden avulla. Empiirinen tutkimus jollakin tai kaikilla esimerkkietokannoilla on mahdollinen jatkotutkimuksen aihe.

Automaattisen tietojenkäsittelyn omaksuminen historiallisten väestöluetteloiden tallentamiseen on lisännyt niiden käyttöä tutkimuksessa. Se on siirtänyt historiallisen tutkimuksen painopistettä yksityisten henkilöiden ja tapahtumien selvittelystä kollektiivisten ilmiöiden käsittelyyn. Kollektiiviset ilmiöt edellyttävät kvantitatiivisten menetelmien käyttöä. Kvantitatiivisten menetelmien avulla on mahdollista testata tutkijan asettamia teorioita ja saada esiin yhteyksiä, jotka eivät ole ilmeisiä yksilötasolla. Automaattiseen tietojenkäsittelyyn perustuvat tietokannat tekevät mahdolliseksi tutkia suuria tietomääriä nopeasti ja tehokkaasti. Tutkija voi saada tutkimuksensa vaatimat tiedot ilman, että hänen on tarpeen käydä läpi käsinkirjoitettua, vaikeaselkoista lähdeaineistoa. Automaattisesti käsiteltävät tietokannat antavat myös mahdollisuuden suorittaa vertailevaa tutkimusta kansainvälisellä aineistolla. Joissakin tietokannoissa (DDB, Karjala-tietokanta) tutkijapalvelu suorittaa tutkijan määrittelemiä hakuja ja toimittaa tiedot levykkeillä. Tulevaisuudessa on mahdollista, että tietokannat ovat tutkijoiden vapaasti käytettävissä tietoverkkojen välityksellä. Koska tietokannat sisältävät henkilötietoja, on tietosuojaa varmistettava.

Tutkielman tieteellisenä antina voidaan pitää tietokantojen yhdenmukaisen kuvauksen ja vertaamisen kautta luotua kehystä, jonka avulla voidaan tarkastella vastaavia järjestelmiä. Lisäksi tutkielma nostaa esille hyvin spesifin aiheen,

historialliset väestöluetteloiden tallentamisen tietokantoihin, joiden tarkastelu ei tietojärjestelmätieteissä ole saanut paljoa huomiota. Tutkielma pyrkii hyödyttämään sekä tietokantojen suunnittelijoita että niiden käyttäjiä. Suunnittelijat saavat tietoa henkilöhistorian lähdeaineistosta ja sen erityisominaisuuksista sekä lisäksi historiallisten väestöluetteloiden käyttömenetelmistä. Heille on apua myös linkitykseen ja nimien normalisointiin käytettyjen menetelmien sekä esimerkkitietokantojen erilaisten ratkaisujen tarkastelusta. Käyttäjän on tärkeää arvioida onko tietokannasta hyötyä hänen tutkimukselleen. Tietosisältöjen ja hakujärjestelmien kuvaaminen ja vertailu auttaa selvittämään, soveltuuko tietokanta hänen tutkimukseensa ja mitkä ovat hakujärjestelmän mahdollisuudet ja rajoitukset.

Monet henkilöhistorian tietokannat on toteutettu relaatiotietokantoina. Kleio-ohjelmiston kehittäjä, saksalainen tutkija Thaller arvostelee relaatiomallia, joka ei hänen käsityksensä mukaan parhaalla mahdollisella tavalla sovellu tietoon liittyvän kontekstin välittämiseen. Thallerin käsityksen mukaan historiallinen tieto tarvitsee tietomallin, joka on herkkä ottamaan huomioon ympäröivän kontekstin. Abstraktin tietomallin ominaisuuksia tärkeämpi on tutkijalle komentokielen helppous ja käyttökelpoisuus. Thallerin kehittämässä Kleio-ohjelmistossa on käytetty mallia, joka muistuttaa semanttista verkkoa. Jatkossa esitetään tutkittavaksi ja vertailtavaksi, miten relaatiomallin lisäksi muut tietomallit, kuten esimerkiksi semanttinen verkko, oliosuuntautunut tietomalli ja hyperteksti, soveltuvat historiallisten väestöluetteloiden tallentamiseen.¹⁷³

¹⁷³ Thaller 1987, Thaller 1991

LÄHTEET

Painetut lähteet

- ATAACK, BATEMAN, GREGSON 1991. Atack, J., Bateman, F., Gregson, M.
 "Matchmaker, matchmaker, make me a match." A general personal
 computer-based matching program for historical research. *Historical
 Methods*, 25, 2, Spring 1992, s. 53-65.
- AUTIO 1993a. Autio, V.-M., Tietosuoja puhuttaa - vaihteeksi myönteistäkin
 kerrottavaa. *Sukutieto*, 1, 1993.
- AUTIO 1993b. Autio, V.-M., Henkilörekisterilain tarkistusta odotellaan.
Sukutieto, 3, 1993.
- BEST 1991. Best, H., Technology or methodology? Quantitative historical social
 research in Germany. *Computers and the Humanities*, 25, 1991, 163-171.
- BOUCHARD 1986. Bouchard, G., The processing of ambiguous links in
 computerized family reconstruction. *Historical Methods*, 19, 1, Winter
 1986, 9-19.
- BOUCHARD 1992. Bouchard, G., Current issues and new prospects for com
 puterized record linkage in the province of Quebec. *Historical Methods*,
 25, 2, Spring 1992, 67-73.
- BROU, OLSEN, 1986. De Brou, D., Olsen, M., The Guth algorithm and the
 nominal record linkage of multi-ethnic populations. *Historical Methods*,
 19, 1, Winter 1986, 20-24.
- CODD 1970. Codd, E., A relational model for large shared data banks.
Communications of ACM, Vol. 13, No 6, 1970, 377-387.
- DANELL 1983. Danell, C. (toim.), Demografiska databasen. Särtryck ur *Släkt
 och Hävd*, 2, 1983.
- DANELL 1985. Danell, C. (toim.), Demografiska databasen vid Umeå
 universitet. En presentation av verksamheten. Umeå, 1985.
- FOGELVIK, SPERLINGS 1982. Fogelvik, S., Sperlings, S., Stockholms
 historiska databas - en presentation. *Släkt och hävd*, 1982, 199-211.

- GENET, ZAMPOLLI 1992. Genet, J.-Ph., Zampolli, A. (toim.), *Computers and the Humanities*. Billing and Sons Ltd, Worcester, 1992.
- GUTMANN, FLIESS, HOLMES, FAIRCHILD, TEAS, 1989. Gutmann, M., Fliess, K., Holmes, A., Fairchild, A., Teas, W., *Keeping track of our treasures: Managing historical data with relational database software*. *Historical Methods*, 22, 4, Fall 1989, 128-143.
- GUTH 1976. Guth, G. J. A., *Surname spellings and computerized record linkage*. *Historical Methods Newsletter*, 10, 1, December 1976, 10-19.
- HILTUNEN 1993. Hiltunen, P., *KHO kumosi tietosuojalautakunnan päätöksen*. *Sukutieto*, 2, 1993.
- HYPPÖNEN, LUTTINEN 1988. Hyppönen, M., Luttinen, R., *Sukututkimuksen käsikirja*. WSOY, Porvoo, 1988.
- KARSKELA 1987. Karskela, S., *Sukututkijan tietokirja*. S. Karskela, Turku, 1987.
- KING 1990. King, M., *All in the family? The incompatibility and reconciliation of family demography and family history*. *Historical Methods*, 23, 1, Winter 1990, 32-40.
- KÄLVEMARK 1977. Kälvemmark, A.-S., *The country that kept track of its population*. *Scandinavian Journal of History*, 2, 1977, 211-230.
- KÄLVEMARK, ROGER 1980. Kälvemmark, A.-S., Roger, J., *Methodiska aspekter på svenskt historiskt-demografiskt källmaterial*. *Teoksessa: Meddelande från familjehistoriska projektet*, Historiska Institutionen, Uppsala Universitet, 2, 1980, 37-42.
- LEBOUTTE, ALTER, GUTMANN 1987. Leboutte, R., Alter, G., Gutmann, M., *Analysis of reconstituted families. A package of SAS programs*. *Historical Methods*, 20, 1, Winter 1987, 29-34.
- LEMMING 1983. Lemming, P., *Stockholms historiska databas*. *Nordisk arkivnyt*, 3, 1983, 78- 83.
- NORBERG, ROLÉN 1980. Norberg, A., Rolén, M., *Industrialiseringen, familjen och individen, Sundsvallsdistriktet 1850 - 1930*. *Teoksessa: Meddelande från Familje-historiska Projektet*, Historiska Institutionen,

- Uppsala Universitet, 2, 1980, 17-22..
- PRICE, GRAY 1994. Price, G., Gray, A., Object oriented databases and their application to historical data. *History and Computing*, 6, 1, 1994, 44-52.
- RUGGLES 1990. Ruggles, S., Family demography and family history: problems and prospects. *Historical Methods*, 23, 1, Winter 1990, 22-30.
- SCHÜRER 1987. Schürer, K., Historical demography, social structure and the computer. *Teoksessa: History and Computing* (toim. Denley, P., Hopkin, D.), Manchester University Press, Manchester, 1987, 33- 45.
- SCHOFIELD 1992. Schofield, R., Automatic family reconstitution. The Cambridge experience. *Historical Methods*, 25, 2, Spring, 1992, 75-79.
- SMETS 1987. Smets, J., South French society and the French revolution. The creation of a large database with Clio. *Teoksessa: History and Computing* (toim. Denley, P., Hopkin, D.), Manchester University Press, Manchester, 1987, 49-58.
- STENFLO, SUNDIN 1987. Stenflo, G., Sundin, J., Using a large historical database. An example from the demographic database in Umeå. *Teoksessa: History and Computing* (toim. Denley, P., Hopkin, D.), Manchester University Press, Manchester, 1987, 58-62.
- THALLER 1987. Thaller, M., Methods and techniques of historical computation. *Teoksessa: History and computing* (toim. Denley, P., Hopkin, D.) Manchester University Press, Manchester, 1987, 147-155.
- THALLER 1991. Thaller, M., The historical workstation project. *Computers and the Humanities*, 25, 1991, 149-162.
- ULLMAN 1988. Ullman, J., Principles of database and knowledge-base systems. Principles of computer science series, 14. Computer Science Press, Rockville 1988
- VILKUNA, HUITU, MIKKONEN 1988. Vilkuna, K., Huitu, M., Mikkonen, P., *Uusi suomalainen nimikirja*. Otava, Helsinki, 1988.
- WINCHESTER 1992. Winchester, I., What every historian needs to know about record linkage for the microcomputer era. *Historical methods*, 25, 4, Fall 1992, 149-165.

Painamattomat lähteet

- BERGGREN, DANIELSSON 1989. Berggren, M., Danielsson, M., Tabeller, kortfattad översikt av innehållet i databasen. Reviderad version (14.2.19889). Umeå Universitet, 1989.
- DEMOGRAFISKA DATABASEN 1990. Demografiska Databasen. Kokoelma järjestelmän dokumentteja. Uumaja, 1990.
- HISKI-JÄRJESTELMÄ 1985. HISKI-järjestelmä, Suttuprojekti. Jyväskylän yliopisto, Tietojenkäsittelyopin laitos, 1985.
- HISTORY ON DATA 1989. History on data. The Demographic Data Base. Umeå university, Tryckeri City, 1989.
- KARJALAN VÄESTÖTIETEELLISEN TIETOKANNAN LOPPURAPORTTI 1989. Karjalan väestötieteellisen tietokannan loppuraportti liitteineen (13.12.1989), 1989.
- KARJALAN VÄESTÖTIETEELLINEN TIETOKANTA 1992. Karjalan väestötieteellinen tietokanta, Karjala-tietokannan hyväksikäyttö (3.11.1992), 1992.
- KARJALA-TIETOKANTA 1993. Karjala-tietokanta, tietokannan suunnittelijan Satu Soivasen haastattelu heinäkuussa 1993, Mikkeli.
- KÖNBERG 1989. Könberg, E., Kyrkoboksmaterial i en relationsdatabas (26.7.1989), 1989.
- PRELIMINARY DRAFT 1989. Access to the material. The Demografic Database. Preliminary draft (29.11.1989), 1989.
- ROUHIAINEN 1983. Rouhiainen, J., Esitutkimusraportti demografisen tietokannan toteutumismahdollisuuksista. Jyväskylän yliopisto, Historian laitos, 1983.
- SIPPU 1985. Sippu, S., Nimien normalisointi. HISKI-järjestelmä, Suttuprojekti. Jyväskylän yliopisto, Tietojenkäsittelyopin laitos, 1985.

LIITTEET

- LIITE 1. Syntyneiden luettelot vuonna 1791, nide 1C1, s. 189.
- LIITE 2. Kuolleiden luettelot vuonna 1791, nide 1C1, s. 312.
- LIITE 3. Rippikirja 1841 - 1851, nide 1Aa11, s. 77.
Korpilahden kirkonarkisto.
- LIITE 4. Lastenkirja 1800 - 1817, nide 1Ab4, ss. 140 - 141.
Kuhmoisten kirkonarkisto.
- LIITE 5. Karjala-tietokannan syntyneiden ja kastettujen luetteloita
tallentava taulukko SYKA, sen sarakkeiden nimet ja tietosisältö
- LIITE 6. Esimerkki muokatusta muuttoluettelosta.
- LIITE 7. Esimerkki henkilörekisteristä, kuolleiden ja haudattujen
luettelo.
- LIITE 8. Esimerkki Resan-ohjelmalla tuotetusta sukutaulusta.

Nati 1791.

54.	28. 29.	Ewa Augustti	F. Soldat Henric Klar Maths Hennric	M. Caisa Forans Dr Sophia Samuels Dr Eena Maths Dr	Vittilar
55.	6. 7.	Ewa	F. Torpn Michel Johans Eric Ersjon Johan Fona'son.	M. Sophia Johans Dr Ewa Sigfrids Dr Hedwig Adams Dr	Epomäki
56.	6. 7.	Maths	F. Inhus Alexander Mats Maths Hennric Johan Elias	M. Stina Abraham Dr Eisa Philips Dr	Pytänmieni, to
57.	8. 10.	Ewa	F. Soldat Adam Tits, Henric Hennric Jonas Hennric	M. Maria Thoma Dr Sophia Johans Dr Ewa Hennric Dr	Kommo, Sold. tor,
58.	13. 14.	Jonas	F. Torparen Alex. der Ersjon Johan Daniels Elias Simons	M. Stina Simons Dr Eena Forans Dr Regina Ers Dr.	Pihlaja mäki
59.	17. 20.	Anna Eisa.	F. Bondeson Johan Kilsjon Elias Hennric Maths Vik	M. Eena Thoma Dr Sophia Pehrs Dr	Murame Eilli
60.	16. 20.	Carl	F. Soldat Maths wih Johan Hennric Johan Michels	M. Sophia Thoma Dr Ewa Michels Dr Susanna Simons Dr	ibid.
61.	18. 21.	David	F. Thomas Hennric Johan Jacob	M. Brita Simons Dr Eisa Thoma Dr Ewa Thoma Dr.	Puthilaa Eppä
62.	16. 21.	Maths	F. Torparen Johan Sigfrids Eric maths Maths Mathsson.	M. Hedwig Anders Dr Ewa Forans Dr Eisa Forans Dr	Varajaawi.
September					
63.	3. 4.	Maths	F. Inhus Anders Pahl Maths Forans Johan Johans	M. Maria Anders Dr Eustrosina Mats Dr Hedwig Adams Dr	Pamaja.
64.	2. 4.	Jonas	F. Dronger Olof Simons Philip Bertels Johan Mathes	M. Valborg Elias Dr Maria Sigfrids Dr Ewa Bertels Dr.	Peuha.
65.	10. 11.	Marander	F. Torparen,	M. Stina Michels Dr	Markula.

312

Denati 1791.

			Julii					
46.	Tiais	Drängen Petter Olofssons Bartholena	13.24.	1.				Siffjuka
47.	Kopijesa	Gårdvärden Enkilingen Johan Thomason	9.24.	81.				af Alder v.
48.	Murame Tervala	Soldatn Jonas Wters Elias	16.24.	1				Oungifjuka
49.	Tikkala	Soldatn Mathis Troijens de Sena	3.7.		1	14.		Oungifjuka
50.	Tikkala Sipilä	Bonden Philiphenricsons Barn Alexander	5.7.			14.		Oungifjuka
51.	Flurtia Kako.	Bonden Johan Johanssons Barn Alexander.	15.21.	6.				Vattenfot.
52.	Ulimurame Myllytiuhora	Enhyses Enkan. Mariami- Nills dotter	3.14.	66	2.			Häll och flyng.
53.	Sutola	Bonden Eric anderssons dotter Eva			12	6.		Siffjuka.
54.	Murame Kestinen.	Torparen Esaias Johans- sons barn Jonas.	September	1.4.	1.	6.		Siffjuka.
55.	Sahnämäki.	Bonden mathis Gustafssons barn Evaastina.	2.4.		2.	6.		Okändjuka.
56.	Murame Sipilä.	Bonden Förän Johanssons Barn, Alexander.	1.4.	2.	3.			Siffjuka.
57.	Murame. Kietala.	Bonden Michel Ericssons Barn Eric	9.11.	5.	4.			Rödfot.
58.	Mändymäki	Torparen Jonas Philipssons Barn Eric.	9.11.	7.				Rötfeber.
59.	Ulimurame Pooka.	Bonden Mathis Mikelssons Barn Elias.	10.18.	10.		14.		Rödfot.
60.	Tiffola.	Kyrkovärden Alexander Johansson barn Johan.	12.18.	1.	4.	14.		Siffjuka.
61.	Murame. Taipale.	Torp ⁿ Eric Johanssons Barn, Johan	12.18.	4.	4.			Siffjuka.
62.	Painaja	Enhyses Anders Pähls Barn Mathis	15.18.			14.		Okändjuka.
63.	Päiväkinda Kasioja.	Torparen Mathis Föranssons Barn Lisa.	8.18.	12.	10.			Rikhofta.
64.	Murame. Kauppila.	Bonden Johan Mathissons Barn Tobias	15.18.	4.	2.			Siffjuka.

77.

Oittila Läsag G.B. n. 72

By. Heddman. och Bonde-namn.	Födelse dag och år.	ABC Bok Inskrivning.	D. Luth. Catech. med Förklaring.						D. S. S. Begrifv.	Försam- mat Cat. Förhör.	Tillfallige
			1	2	3	4	5	6			
Nob. Tyttlarabi	1811								41/47		
Herrig Henrikson	1783										
Thomae	1790	x	x	x	x	x	x				
Henrikson	1825	18	11	4	1	1	1	1	1	1	1840-42 1844
Filipson	1774	x	x	x	x	x	x			1844	
Helena Matts	1776	x	x	x	x	x	x			1849	
Matts Filipson	1800	x	x	x	x	x	x			1850	
Hedvig Henrikson	1802	x	x	x	x	x	x			1849	
Henrikson	1815	x	x	x	x	x	x	50		1849	
Maria Jacobi	1820										
Carl Gustafson	1812	y	y	y	y	y	y			1849	
Johanna	1826	x	x	x	x	x	x				
Carl Eriksson	1821	1	1	1	1	1	1			1849	
Lars Eriksson	1819	y	y	y	y	y	y			1849	
Henrikson	1826	x	x	x	x	x	x			1849	
Johanna	1850	y	y	y	y	y	y			1849	
Jonas Mattsson	1834	x	x	x	x	x	x	50		1849	
Alexander Eriksson	1826	1	x	x	x	x	x			1849	
Carl Eriksson	1846									1849	
Carl Eriksson	1828	x	x	x	x	x	x			1849	
Carl Eriksson	1817	x	x	x	x	x	x			1849	
Carl Eriksson	1826	y	y	y	y	y	y			1849	
Carl Eriksson	1825	1	1	1	1	1	1			1849	
Carl Eriksson	1798									1849	
Carl Eriksson	1821									1849	
Carl Eriksson	1778	x	x	x	x	x	x			1849	
Carl Eriksson	1814	x	x	x	x	x	x	50		1849	

Muurijärvenkadun

141

By- Hemmans- Föräl- Födel- krats och Barns Namn	Födel- dag & år	Innanför	D. Luth. Cat. med Förkl.						Antmärkingar när de blifvit (Admit. m. m.)	
			A	B	C	1	2	3		4
<i>Puolake Matti Adamsonson & Fru</i>										
<i>Hedda Sofia</i>	<i>10/9 96</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>sv. ad. Admit 1814 26</i>
<i>Suflor Fred.</i>	<i>15/3 99</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>admit: 1814</i>
<i>Jonas Adam</i>	<i>2/3 80</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	
<i>Johannes J.</i>	<i>7/7 808</i>									
<i>Wendla Magdalena</i>	<i>1810</i>									
<i>Puolake Matti Adamsonson & Fru</i>										
<i>Johan Joh. Joh.</i>	<i>16/8 804</i>									<i>sv. ad. den 26 April 1809</i>
<i>Eva Johanna</i>	<i>808</i>									<i>sv. ad. den 27 April 1809</i>
<i>Johannes Johanna</i>	<i>809</i>									<i>sv. ad. den 29 Nov: 1814</i>
<i>Maria Sofia</i>	<i>21/11 810</i>									
<i>Johanna Sofia</i>	<i>1812</i>									
<i>Adem Sofia</i>	<i>1814</i>									
<i>Matti Sofia</i>	<i>2/7 1814</i>									<i>sv. ad. 106-1814-adj</i>

LIITE 5 Karjala-tietokannan syntyneiden ja kastettujen luetteloita
tallentava taulukko SYKA, sen sarakkeiden nimet ja tietosisältö

<u>SARAKE</u>	<u>SELVITYS</u>
HNRO	henkilönumero, juokseva numero
TSRK	tallennuseurakunta
SIDOS	tallennetun sidoksen sidostunnus
SIVU	tallennetun sivun numero
ABSIVU	samannumeroiset sivut erotteleva tunnus
RIVI	tallennetun rivin numero
NRO	samalla rivillä olevien järjestysnumero
SPP	syntymäpäivä
SKK	-kuukausi
SVV	-vuosi
SSRK	syntymäseurakunta
SP	sukupuoli
REKNRO	syntymän rekisteröimisnumero
AVIOS	aviosynty
MONIS	monisynty
SJNRO	järjestys sisarussarjassa
KUOL	kuolemaan liittyvät merkinnät
ENIMI	ensimmäinen etunimi
ENIMI2	muut etunimet
ISAMM	isän ammatti
ISENIMI	isän ensimmäinen etunimi
ISPNIMI	isän patro-/matronyyminimi
ISSNIMI	isän lisänimi (sukunimi)
ISIKA	isän ikä
ISSIV	isän siviilisäätty
AIAMM	äidin ammatti
AIENIMI	äidin ensimmäinen etunimi
AIPNIMI	äidin patro/matronyyminimi
AISNIMI	äidin lisänimi (sukunimi)
AIOSN	äidin aikaisempi lisänimi (tyttönimi)
AIIKA	äidin ikä
AISIV	äidin siviilisäätty
VIHPP	vanhempien vihkipäivä
VIHKK	" -kuukausi
VIHVV	" -vuosi
KASMUU	kasteeseen liittyvät muut tiedot
KASPP	kasteen päivä
KASKK	" kuukausi
KASVV	" vuosi
KASRK	kasteseurakunta

KASAMM	kastajan ammatti
KASNIMI	kastajan nimi
VAHPP	kasteen vahvistamisen päivä
VAHKK	" " kuukausi
VAHVV	" " vuosi
VAHAMM	kasteen vahvistajan ammatti
VAHNIMI	kasteen vahvistajan nimi
VAHMUU	kasteen vahvistamiseen liittyvät muut tiedot
KKSIVU	kirkonkirjan sivu, jolle siirretty
ASP	asuinpaikka
YLIV	tiedot yliviivattu
N_ENIMI	normalisoitu etunimi
N2_ENIMI	" etunimi, 2. asu
N_ENIMI2	" toinen etunimi
N2_ENIMI2	" toinen etunimi, 2. asu
N_KYNIMI	" kylän nimi
N_TANIMI	" talon nimi
N_ISAMM	" isän ammatti
N2_ISAMM	" isän toinen ammatti
N3_ISAMM	isän kolmas ammatti
N_ISEN	isän etunimi
N2-ISEN	isän etunimi, 2. asu
N-ISEN2	isän toinen etunimi
N2_ISEN2	isän toinen etunimi, 2. asu
N_ISPN	isän patronyyminimi
N2_ISPN	isän patronyyminimi, 2. asu
N_ISSN	isän sukunimi
N_AIAMM	äidin ammatti
N2_AIAMM	äidin toinen ammatti
N3_AIAMM	äidin kolmas ammatti
N_AIEN	äidin etunimi
N2_AIEN	äidin etunimi, 2. asu
N_AIEN2	äidin toinen etunimi
N2_AIEN2	äidin toinen etunimi, 2. asu
N_AIPN	äidin patronyyminimi
N2_AIPN	äidin patronyyminimi, 2. asu
N_AISN	äidin sukunimi
N_AIOSN	äidin tyttönimi
G_SVV	generoitu syntymävuosi
G_ENIMI	etunimi
G_SNIMI	sukunimi
G_AVIOS	aviosynty
G_SSRK	syntymäseurakunta
G_KKSIDOS	kirkonkirjasidos
G2_KKSIDOS	toinen kirkonkirjasidos
G_KKSRK	kirkonkirjaseurakunta

G_KASVV	kastevuosi
G_KASRK	kasteseurakunta
G_VAHVV	kasteen vahvistuvuosi
G_VAHSRK	kasteen vahvistusseurakunta
G_USKO	uskonto
G_ISUSKO	isän uskonto
G_ISKSRK	isän kotiseurakunta
G_AIUSKO	äidin uskonto
G_AIKSRK	äidin kotiseurakunta

LIITE 6 Muokattu muuttoluettelo Attmarin seurakunnasta vuosilta 1869 - 1886. Muuttoluettelosta puuttuu alkuperäisessä muodossa sekä muuttajien syntymäaika että päämiehen mukaan muuttaneiden henkilöiden nimet. Tämä tieto on lisätty muista lähteistä ja merkitty sulkeisiin.

Vuosittainen järjestysno	Muuttopvm	Nimi	(Syntymäaika)	Ammatti	Muuttanut mistä	mihin	(Vol/Sidos)	rivi
021	1884-11-16	Larsdr, Sigrid Kajsa	(1864-07-22)	piika Hassela	Östansjö		(AI:10B 0396)	01
022	1884-11-16	Andersdr, Sigrid Lisa	(1863-05-26)	piika Tuna	Attmarby		(AI:10A 0001)	02
023	1884-11-21	Ström, Karl Johan	(1858-07-13)	renki	Stora Kil	Sörlingsjö	(AI:10B 0338)	03
024	1854-11-22	Homlström, Erika Nikolina	(1865-01-30)	piika	Bygdeå	Karläng	(AI:10A 0169)	04
025	1854-11-26	Skoglund, Karin	(1857-11-05)	piika	Hassela	Sörfors	(AI:10B 0337)	05
026	1884-11-26	Lypen, Nils	(1853-03-16)	sulattaja	Stora Tuna	Sörfors	(AI:10b 0846)	06
026	1884-11-26	(Spångberg, Anna Mari)	(1847-04-21)		Stora Tuna	Sörfors	(AI:10B 0486)	06
026	1884-11-26	(Anna Theresia)	(1881-06-10)		Stora Tuna	Sörfors	(AI:10B 0486)	06
026	1884-11-26	(Beda Eleonora)	(1883-10-10)		Stora Tuna	Sörfors	(AI:10B 0486)	06
026	1884-11-26	(Hulda Viktoria)	(1878-10-09)		Stora Tuna	Sörfors	(AI:10B 0486)	06
026	1884-11-26	(Karl Viktor)	(1873-07-06)		Stora Tuna	Sörfors	(AI:10B 0486)	06
026	1884-11-26	(Oskar)	(1876-08-12)		Stora Tuna	Sörfors	(AI:10B 0486)	06

LIITE 7 Esimerkki kuolleiden ja haudattujen luettelon henkilörekisteristä, Attmarin pitäjä, 1814 - 1893

Attmarin srk	Kuolleiden ja haudattujen luettelo	Vuosilta 1814 - 1893	Lajittelu kuolinsyyn perusteella	Ikä	Vol/Sid/Rivi
Nimi	Syntynyt	Kotipaikka	Kuollut	Kuolinsyvy	
Märta Olsdr	1810-06-15	Attmar	1830-11-14	Blodstörtning (Verensyöksy)	20 (C:3) 0196 15
Erik Sörlin	1814-12-17	Attmar	1884-10-05	Blodstörtning (Verensyöksy)	70 (C:6) 0179 16
Segrin Henricsdr	1797-08-07	Fjolsta	1842-03-22	Blodstörtning (Verensyöksy)	44 (C:3) 0206 16
Nil Peter Forssberg	1827-07-26	Attmar	1867-01-01	Bloduppspottning	56 (C:6) 0189 01
Jonaa Olof Bergström	1823-05-22	Njurunda	1892-05-25	Blaskatarr	69 (C:7) 0158 !
Johan Peter Sörgren	1865-12-25	Attmar	1866-12-06	Bräck	20 (C:6) 0157 16
Nils Ericsson	1765-10-24	Rude	1838-04-21	Bräck	71 (C:3) 0203 12
Olof Ersson	1838-04-26	Östansjö	1879-05-23	Bräck	41 (C:6) 0162 12
Engla Gustfva Berg	1801-07-30	Rude	1876-05-23	Bräck	74 (C:6) 0154 09
Eric Persson	1754-01-27		1840-06-04	Bräck	86 (C:3) 0205 10
Juliana Maria Hammarström	1826-09-30	Hapf	1893-05-30	Bräck	64 (C:7) 0163 07
Brita Ersdr	1868-09-25	Backe	1827-12-29	Bräck	59 (C:3) 0193 36
Sven Svensson	1745-07-12	Söderang	1828-04-04	Bräck	82 (C:3) 0184 05
Nils Pålsson	1763-03-04	Westerböle	1827-12-05	Bräck	64 (C:3) 0193 35
Nils Ersson	1759-03-22	Söderlindsjö	1822-11-25	Bräck	63 (C:3) 0189 35
Eric Dahlbom	1793-01-05	Njurunda	1832-06-21	Bräck inflammation	59 (C:4) 0105 26

Liite 8 Resan-ohjelmalla tuotettu sukutaulu. Se on ensimmäinen osa sukutaulusta, jonka päämies on syntynyt Skellefteån seurakunnassa v. 1795.

1	19864	2	12661	4	6559	8	⇒	
				9				
		2	12351	5	7037	10		
				11				
				6	6044	12		3415
	7	6480	13	3240	14	3884		⇒
			15	3711				

Numero	Etunimi	Sukunimi	Syntynyt
1 19864	Jonas	Johaninpoika Söderberg	1795-06-12
2 12661	Jonas Johan	Söderberg	1766-01-22
3 12351	Greta Fredriikka	Jonintytär	1764-10-04
4 6559	Johan	Andersinpoika Storman	1726
5 7037	Liisa	Olofintytär	1729
6 6044	Jon Jonas	Andersinpoika Lundström	1723
7 6480	Katariina Karoliina	Jonintytär	1726-07-24
12 3415	Anders	Joninpoika Lundström	1692
13 3240	Maarit		1685
14 3884	Jon	Kristierinpoika Varg	1700
15 3711	Maarit	Nilsintytär	1698

**II PIENTEN LASTEN KUOLLEISUUDESTA
LAVANSAARELLA VUOSINA 1751-1899**

TIIVISTELMÄ

Lahti, Eeva Marjatta

Pienten lasten kuolleisuudesta Lavansaarella vuosina 1751-1899 / Eeva Lahti

Jyväskylä: Jyväskylän yliopisto, 2001

42 s.

Tutkielma

Tutkielman tässä osassa tarkastellaan pienten lasten kuolleisuuden kehitystä Lavansaarella vuosina 1751-1899. Sen lisäksi, että kuolleisuus vaikuttaa väestön kokoon ja muovaa sen rakennetta, se voidaan tulkita myös väestön terveydentilan osoittimeksi. Suomessa 1700-luvun puolivälistä alkavat tilastot tekevät mahdolliseksi kuolleisuuden muutosten seuraamisen varsin pitkältä ajalta.

Tutkimuksessa on lähteinä käytetty pääasiassa Lavansaaren kuolleiden ja haudattujen sekä syntyneiden ja kastettujen luetteloita. Kun näissä havaittiin puutteita, käytettiin lasten- ja rippikirjojen tietoja. Lavansaaren kirkonkirjat on tallennettu Karjala-tietokantaan, jonka yksityis-kohtainen esittely on tutkielman ensimmäisessä osassa. Tutkimuksen lähdettiin oletuksesta, että alle neljä vuoden ikäisinä kuolleisiin lapsiin kuului ikäryhmiä, joiden kuolleisuus tutkimusajan kuluessa kehittyi eri tavoin. Tutkimus vahvisti tämän oletuksen ja osoitti sen lisäksi, että kuolleisuuden aleneminen tapahtui eri aikoina eri ikäryhmissä.

Pienten lasten kuolleisuuden tutkiminen antaa paitsi tietoa määrällisistä muutoksista myös viitteitä ihmisten asenteista ja niiden muuttumisesta, esimerkiksi suhtautumisesta lapsiin, pienten lasten kuolemaan ja imettämiseen. Lasten kuolleisuuden vaihtelut kertovat taloudellisen hyvinvoinnin ja hygieenisten olojen tasosta ja niiden muutoksista.

AVAINSANAT: kuolleisuus, syntyneisyys, tartuntatauti, epidemia, kuolleiden ja haudattujen luettelot, syntyneiden ja kastettujen luettelot, rippikirjat, lastenkirjat

Sisällys

1. Johdanto	1
1.1. Tutkimusalue	5
1.2. Tutkimustehtävä ja lähteet	7
2. Pienten lasten kuolleisuuden kehitys	11
2.1. Lavansaaren yleinen kuolleisuus	11
2.2. Alle neljän vuoden ikäisten lasten kuolleisuus	16
2.3. Yhdestä neljään vuoden ikäisten kuolleisuus	18
2.4. Imeväiskuolleisuus	24
3. Pohdinta	34
Lähteet	40
Liitteet	42

1 JOHDANTO

Käsite kuolleisuus ilmoittaa kuolleiden osuuden väestöstä. Kuolleisuudella on historiantutkimuksessa merkitystä ilmiönä, joka vaikuttaa väestön kokoon ja muovaa sen rakennetta. Kuolleisuus voidaan myös tulkita väestön terveydentilan osoittimeksi, jolloin sitä käytetään hyvinvoinnin mittarina. Kuolleisuuden ja sen syiden selvittäminen on osa väestö- ja perhehistorian tutkimusta. 1700-luvun puolivälistä alkavat tilastot tekevät Suomessa mahdolliseksi kuolleisuuden muutosten seuraamisen varsin pitkältä ajalta.¹

Vaikka Suomen väestö kasvoi 1700-luvun jälkipuoliskolla ja seuraavalla vuosisadalla, kasvu ei kuitenkaan johtunut siitä, että kuolleisuus olisi Suomessa ollut erityisen alhainen. Nykytasoon (noin 1 %)² verrattuna kuolleisuus oli moninkertainen. 1700-luvulle oli ominaista korkea peruskuolleisuus, se osa kuolleisuudesta, jossa ei esiintynyt kovin suuria ja äkkinäisiä ajallisia vaihteluita. Tähän kuului ensinnäkin suuri osa imeväiskuolleisuudesta, jolla tarkoitetaan alle yhden vuoden ikäisten kuolleisuutta. 1700-luvulla syntyneistä lapsista useampi kuin joka viides kuoli ensimmäisen ikävuotensa aikana. Kaiken kaikkiaan hädin tuskin puolet syntyneistä lapsista eli aikuisikään. Myös verraten korkea aikuiskuolleisuus oli aikakaudelle ominaista. Korkea kuolleisuus kieli myös merkittävästä sairastuvuudesta. Peruskuolleisuuden lisäksi toistuvat tartuntatautiepidemiat nostivat ajoittain kuolleiden määrän jopa syntyneiden määrää korkeammalle. Kuolonvuosien uhrit olivat suurelta osin alle 15 vuoden ikäisiä lapsia, jotka näinä epidemiovuosina läpikävivät aikakauden yleiset lastentaudit, iso-, tuhka- ja tulirokon sekä hinkuyskän - kaikki tuohon aikaan verrattain

¹ Koskinen, Martelin 1994, s. 150-152

² Pitkänen 1994, s. 46

tappavia tauteja. Myös osa aikuiskuolleisuudesta johtui infektioitaudeista. 1700-luvun loppupuolella Suomessa oli erityisen korkean kuolleisuuden aikoja 1760-luvun alussa ja Kustaa III:n sodan aikana (1788-90).³

Vaikka kokonaiskuolleisuus ei alentunut huomattavasti ennen 1800-luvun loppua, niin kuolleisuuden rakenne muuttui. Suomessa kuten muuallakin Euroopassa imeväiskuolleisuus ja kuolleisuus isorokkoon pienentyi. Kun 1700-luvun jälkipuoliskolla runsaat 20 prosenttia elävänä syntyneistä lapsista kuoli ensimmäisen ikävuotensa kuluessa, niin 1820-luvulla oli imeväiskuolleisuuden taso pudonnut alle 20 prosentin, ja vuosisadan viimeisinä vuosikymmeninä vastaava luku oli enää noin 15 prosenttia.⁴ Imeväiskuolleisuuden pienentyminen ei johtanut mainittavaan kokonaiskuolleisuuden alenemiseen, sillä kuolleisuus ei pienentynyt muissa ikäryhmissä. Isorokon aiheuttama kuolleisuus väheni rokotuksen tehostuttua 1800-luvun alkuvuosikymmenien jälkeen, mutta sen vaikutus jäi tilapäiseksi. Vanhemmalla aikuisväestöllä kuolleisuus jopa lisääntyi. 1800-luvulla Suomessa oli korkea kuolleisuus 1830- ja 1850-luvuilla, mutta sen lisäksi sattui kaksi dramaattista kuolleisuuskriisiä, jotka suuressa määrin vaikuttivat myös aikuisväestöön. Ensimmäinen osui Suomen sodan vuosiin 1808-1809 ja toinen ajoittui suurten nälkävuosien aikaan 1866-1868.⁵ Kuolleisuus pysyi 3 prosentin tuntumassa aina 1860-luvulle asti ja aleni selvästi vasta vuosina 1866-68 koettujen, terävän kuolleisuushuipun aiheuttaneiden suurten nälkävuosien jälkeen.⁶

1700- ja 1800-luvuilla Euroopan väestönkasvu, joka siihen asti oli ollut vaatimatonta, ylsi melko korkeisiin kasvulukuihin (1750-1850 vuotta kohden laskettu väestönkasvu oli Suomessa 1,2 prosenttia, Englannissa 1,1 prosenttia).

³ Pitkänen 1988, s. 19-23, Jutikkala 1968

⁴ Turpeinen 1979a

⁵ Koskinen, Martelin 1994, s. 152, Pitkänen 1988, s. 28, Jutikkala 1993

⁶ Jutikkala 1968, s. 29, Pitkänen 1994 s. 42-44

Väestönkasvun nopeutumisen syynä on pidetty kuolleisuuden alentumista. Väestöhistorioitsijat ovat esittäneet erilaisia selitysmalleja siitä, mitkä tekijät ovat aikaansaaneet kuolleisuuden pienentymisen. Osa tutkijoista, etenkin englantilaiset, katsoi, että kuolleisuuden pienentyminen johtui elintason kasvusta ja isorokkorokotuksen käyttöönotosta. Englannista on olemassa kohtuullisen hyvät arviot hedelmällisyyden ja kuolleisuuden kehityksestä 1600-luvun alusta alkaen. Englantilaiset tutkijat Wrigley ja Schofield, jotka ovat väestöhistorian tutkimuksessa käsitelleet kuolleisuuskriisejä ja kuolleisuuden paikallisia eroja myös imeväis- ja lapsikuolleisuuden osalta, ovat havainneet, että nopean väestönkasvun ajanjaksoille on tyypillistä rajujen kuolleisuushuippujen puuttuminen ja korkea syntyneisyyden taso.⁷ Ranskalaiset tutkijat esittivät, että maatalouden kehitys vaikutti ravitsemusta parantamalla kuolleisuuden alenemiseen.⁸ Ruotsalaiset ovat tutkineet lasten, etenkin imeväisikäisten, kuolleisuuden muutoksia ja kuolemansyitä 1700-1800-luvuilla. Kuolemansyitä selvittämällä saatiin tietoja hygieenisistä oloista ja rintaruokintatavoista ja sitä kautta myös viitteitä ihmisten asenteiden muutoksista. He etsivät myös vastausta kysymykseen, osallistuivatko ihmiset tietoisesti muutoksiin, jotka johtivat imeväiskuolleisuuden alenemiseen vai oliko aleneminen suunnittelemattoman yhteiskunnallistaloudellisen muutoksen seurausta.⁹

Suomessa Oiva Turpeinen on tutkinut taloudellisten tekijöiden, epidemioiden ja liikkuvuuden vaikutusta kuolleisuuteen. Turpeinen katsoi, että liikkuvuus oli kuolleisuuden kannalta merkittävä tekijä, koska se lisäsi tartuntatautien vaaraa.¹⁰ Sekä Oiva Turpeinen että Kari Pitkänen ovat tutkineet kriisiaikojen kuolleisuuden

⁷ Pitkänen 1994, s. 40-45, Wrigley, Schofield 1981

⁸ Pitkänen 1994, s. 40-45

⁹ Sundin, s. 93-96

¹⁰ Turpeinen 1979b

syitä.¹¹ Tutkijoita on kiinnostanut erityisesti Suomen imeväiskuolleisuuden aleneminen, joka alkoi 1700-luvun lopulta lähtien. Suomalaisessa yhteiskunnassa vallitsevissa olosuhteissa olisi odottanut imeväiskuolleisuuden pikemminkin nousevan kuin laskevan. Maa pysyi pitkään maanviljelysvaltaisena ja maattomien sekä muiden alempiin sosiaaliluokkiin kuuluvien määrä kasvoi nopeasti. Samanaikaisesti lisääntyi aviottomien lasten osuus syntyneistä. Näissä ryhmissä imeväis-kuolleisuus oli keskimääräistä korkeampi. Imeväiskuolleisuus pysyi alenemisesta huolimatta edelleen korkeana 1800-luvulla.¹² Eri tutkijat, esimerkiksi Jutikkala, Turpeinen, Pitkänen, Moring ja Sundin, ovat yhtä mieltä siitä, että korkean imeväiskuolleisuuden ja rintaruokinnan välillä vallitsee yhteys. Muiden tekijöiden kuten epidemioiden, elintason, viljasatojen, hygienian, synnytyksen aiheuttamien vaurioiden ja paikallisten olojen vaikutuksesta on erimielisyyttä.

Tämän tutkimuksen kohderyhmäksi valittiin alle neljän vuoden ikäiset lapset, joiden kuolleisuuden kehitystä vuosina 1751-1899 tarkastellaan. Koska tutkimus käsittää pitkän ajanjakson, sata viisikymmentä vuotta, oli välttämätöntä käyttää koneelliseen muotoon tallennettuja seurakuntatietoja. Lavansaaren kirkonkirjojen tiedot olivat kokonaisuudessaan jo tallennettuja Karjala-tietokantaan ja siksi Lavansaari valittiin tutkimusalueeksi. Lavansaari osoittautuikin mielenkiintoiseksi tutkimusalueeksi. Se kuului Suomenlahden ulkosaariin ja oli maantieteellisen sijaintinsa takia edullisessa asemassa, koska saariston väki pystyi nälänhädän aikana hankkimaan elannon täydennystä merestä. Toisaalta voisi ajatella, että Lavansaari olisi ollut syrjäinen ja eristäytynyt alue. Lavansaaren asukkailla oli kuitenkin vilkkaat yhteydet Suomenlahden rannikkoalueille.

¹¹ Turpeinen 1986, Pitkänen 1993

¹² Pitkänen 1988, s. 25-27

1.1 Tutkimusalue

Lavansaari (Lövsjär) sijaitsee itäisellä Suomenlahdella. Alue käsittää varsinaisen pääsaaren (Lavansaari) ja siitä kahdeksan kilometriä itään olevasta Peninsaaren. Pinta-alaltaan saaret ovat yhteensä 15,9 neliö-kilometriä. Lavansaari sijaitsee noin 30 kilometrin päässä Venäjän ja 70 kilometrin päässä Suomen mannermaalta.¹³

Hallinnollisesti Lavansaari kuului vuoden 1539 veroluetteloissa yhdessä muiden Suomenlahden ulkosaarten kanssa Äyräpään kihlakuntaan ja Viipurin linnan veroalaisuuteen. Myöhemmin Lavansaari kuului Koiviston pitäjään. Suuren Pohjan sodan jälkeen (1721) saaret olivat jonkin aikaa lahjoitusmaana. Lahjoitusmaakauden päätyttyä Lavansaari liitettiin Venäjän kruunun alaisen Viipurin maakomissariaatin hallintoon. Myöhemmin 1700-luvulla Lavansaari kuului muiden ulkosaarten kanssa Haminan piiriin. 1800-luvulla alkoi uudelleen hallinnollinen yhteys Koiviston kanssa, mikä jatkui Lavansaaren kunnan itsenäistymiseen, vuoteen 1905, asti.¹⁴

Lavansaari kuului 1500- ja 1600-luvuilla Vehkalahden seurakuntaan. Vuonna 1721 Uuden-kaupungin rauhanteon yhteydessä se yhdistettiin Viipurin seurakuntaan Koiviston seurakunnan ja Viipuriin kuuluvan Kakin eli Johanneksen kappelin papistojen yhteisesti hoidettavaksi. Vuodesta 1763 lähtien Lavansaari tuli pelkästään Kakin kappalaisen alaisuuteen ja ainoastaan tämä sai käydä Lavansaarella. Vuonna 1810 Lavansaari siirrettiin Viipurin emäseurakunnan hoitoon. Vuonna 1817 Lavansaari liitettiin jälleen Koiviston seurakuntaan, jonka yhteydessä se pysyi aina siihen asti, kun se 1896 sai itsenäisen seurakunnan aseman. Seurakunnalliset olot olivat vielä 1700-luvun puolivälissä sangen sekavat kaikissa Suomen ulkosaarissa. Niiden hallinnolliset rajat olivat muuttuneet tiheään, eivätkä saarelaiset aina tienneet mihin piiriin kuuluivat. Papit kävivät

¹³ Karjala-tietokanta, Lavansaaren seurakuntatiedot

¹⁴ Karjala-tietokanta, Lavansaaren seurakuntatiedot

saarissa hyvin epäsäännöllisesti, sillä pitkät ja vaaralliset matkat haittasivat heidän käyntejään.¹⁵

Lavansaarella oli kiinteää asutusta jo 1500-luvulla, mutta asutus kärsi pahoin Venäjän ja Ruotsin välille 1570 puhjenneen sodan takia. 1600-luvulla asutus elpyi ja vakiintui ja siitä lähtien saarilla oli pysyvää asutusta. Lavansaaren pääsaarella oli 1800-luvulla kaksi kylää: Lounatkylä (Eteläkylä) ja Pohjakylä. Peninsaari oli 1700-luvun puolivälistä lähtien asumaton. Alunperin Lavansaaren asukkaat tulivat pääasiassa Suomen-lahden pohjoisrannalta: Vanhalta Vehkalahdelta, Vironlahdelta, Kymistä ja Sippolasta. Asukkaita tuli myös Viipurista ja Koivistolta sekä jonkin verran Virosta. Saaren asukkaat olivat suomenkielisiä. Saaren väkiluku kasvoi 1700- ja 1800-luvuilla nopeasti ja oli 975 henkeä vuonna 1900. Pinta-alaan nähden saarten asukasmäärä oli siis varsin suuri.¹⁶

Lavansaarelaiset saivat toimeentulonsa lähinnä kalastuksesta, hylkeenpyynnistä ja kauppamerenkulusta. Tärkein elinkeino oli silakanpyynti, jota harjoitettiin nuotilla ja verkoilla. Maatalouden harjoittamiseen saarella ei ollut edellytyksiä, sillä maaperä oli karua. Puuttuva vilja hankittiin kalastuksen tuotolla. Perunaa kasvatettiin, mutta ei kuitenkaan edes kotitarpeiksi. Lehmiä oli jonkin verran, jopa pari kolme lehmää joissakin taloissa. Pietarin kaupungin perustaminen lisäsi saarten elämisen mahdollisuuksia, kun sen markkinat avautuivat Lavansaaren tuotteille. Kauppa-merenkulku oli vilkkaimmillaan 1800-luvun loppupuolella ja 1900-luvun alussa. Saarelaiset omistivat tällöin noin sata alusta, jotka enimmäkseen olivat pieniä. Niillä he harjoittivat kaupankäyntiä Venäjän, Suomen ja Ruotsin rannikkokaupunkien välillä.¹⁷

¹⁵ Hannula 1947, s. 24-25

¹⁶ Korhonen 1996, s. 113-114, Talve 1996, s. 121

¹⁷ Korhonen 1996, s. 115-116, Hannula 1947, s. 185-188, 55-57

1.2 Tutkimustehtävä ja lähteet

Kuolleisuuden aleneminen Suomessa tapahtui vaiheittain siten, että muutokset keskittyivät eri aikoina eri ikäryhmiin.¹⁸ Lasten kuolleisuus vaihteli alueittain ja ikäryhmittäin. Tässä yhteydessä tarkastellaan Lavansaaren iältään alle 4-vuotiaiden lasten kuolleisuuden kehitystä vuosina 1751-1899. Tutkimuksessa on lähdetty olettamuksesta, että jos kohderyhmä iän perusteella jaetaan pienempiin ryhmiin, niin näiden ryhmien kuolleisuus tutkimusajan kuluessa kehittyi eri tavoin. Siksi kohderyhmää tarkastellaan aluksi yhtenä kokonaisuutena, minkä jälkeen selvitetään erikseen 1-4-vuotiaiden sekä alle vuoden ikäisten kuolleisuuden muutoksia. Lopuksi selvitetään vielä eri ikäisten imeväisten kuolleisuuden muutoksia. Miten kuolleisuus vaihteli eri ikäryhmissä? Milloin minkin ryhmän kuolleisuus kääntyi laskuun? Oliko eri ryhmien kuolleisuuden alenemisessä ajallisia eroja? Jossain määrin pyritään myös vastaamaan kysymykseen, mitkä syyt aiheuttivat lasten kuolleisuutta ja olivatko eri ikäisten lasten kuolemansyyt erilaisia. Koska tutkimus kattaa 150 vuotta, oli välttämätöntä käyttää tietokantoihin tallennettuja kirkonkirjojen merkintöjä. Tavoitteena on myös ollut saada kokemuksia siitä, miten tietokantaan tallennetut kirkonkirjat soveltuvat tieteelliseen historialliseen tutkimukseen. Tähän mennessä Karjala-tietokanta on ollut lähinnä sukututkijoiden käytössä.

Tutkittavalle ajanjaksolle osui imeväisikäisten kuolleisuuden aleneminen Suomessa. Myös pienten lasten kuolleisuuden lasku alkoi. Tutkimuksen kannalta on mielenkiintoista tarkastella, miten erityisen korkean kuolleisuuden ajat, 1760-luku, Kustaa III:n sodan aika, Suomen sodan aika, 1830-, 1850- ja 1860-luvut, heijastuvat Lavansaaren kuolleisuuteen. 1800-luvulle osui yleisen kuolleisuuden tason aleneminen, jota on nimitetty epidemiologiseksi transitioksi. Sille on ominaista muun muassa, että kuolleisuuden vuosittainen vaihtelu vähenee, kuolleisuuden painopiste siirtyy vanhimpiin ikäluokkiin ja kuolemansyy rakenne

¹⁸ Pitkänen 1988, s. 40-56

muuttuu oleellisesti.¹⁹

Tämän tutkimuksen pääasiallisina lähteinä ovat olleet Karjala-tietokantaan tallennetut Lavansaaren kuolleiden ja haudattujen luettelot sekä syntyneiden ja kastettujen luettelot. Silloin kun syntyneiden ja kuolleiden luetteloista jonain vuosina puuttui merkintöjä, on myös lasten- ja rippikirjoja käytetty tietojen saamiseksi. Mikäli lasten- ja rippikirjoissa oli lapsen syntymä- tai kuolinpäivämäärä erilainen kuin syntyneiden tai kuolleiden luetteloissa, on etusija annettu luetteloiden tiedoille. Lavansaaren syntyneiden ja kastettujen luettelot sekä kuolleiden ja haudattujen luettelot alkavat vuodesta 1750 ja jatkuvat vuoteen 1949. Rippikirjat alkavat vuodesta 1744 ja päättyvät 1950. Lastenkirjat kattavat vuodet 1800-1884 ja 1905-1921. Syntyneiden ja kastettujen luettelot sisältävät lapsen henkilötietoja, vanhempien tietoja, asuinpaikkatietoja, kastetapahtumaan liittyviä tietoja ja kummien tietoja. Syntyneiden luetteloista on saatu vuosittain syntyneiden lasten määrä, lapsen syntymän päivämäärä sekä lapsen ja vanhempien nimet. Vuosien 1818, 1819 ja 1821 syntyneiden luetteloista puuttui merkintöjä useiden kuukausien ajalta. Näinä kuukausina syntyneiden tiedot löytyivät lastenkirjoista. Vuoden 1880 syntyneiden luetteloissa oli vain tammi-kesäkuussa syntyneet ja vuoden 1881 luetteloissa elo-joulukuussa syntyneet. Puuttuvien kuukausien tiedot löytyivät lastenkirjoista. Vuoden 1885 syntyneiden luetteloista puuttuivat tammi-toukokuussa ja syys-joulukuussa syntyneet. Vuoden 1886 syntyneistä puuttui myös muutama. Nämä puuttuvat tiedot löytyivät rippikirjoista.

Kuolleiden ja haudattujen luettelot sisältävät vainajan henkilötietoja, kuolemaan ja hautaukseen liittyviä tietoja, asuinpaikkatietoja ja vainajan omaisten tietoja. Kuolleiden luetteloista saatiin vuosittaiset kuolleiden määrät, kuolleen henkilön nimi ja useimmiten myös vainajan ikä. Aluksi merkinnöissä oli mainittu kuolinvuosi ja hautauksen päivämäärä, mutta kuolinpäivämäärä puuttui.

¹⁹ Koskinen, Martelin 1994, s. 151, Jutikkala 1968

Kuolinpäivämäärät ilmestyivät merkintöihin 1780-luvulla. Myös kuolemansyyt puuttuivat aluksi kokonaan. Niitä alkoi olla merkinnöissä vuodesta 1770 lähtien, mutta senkin jälkeen ne saattoivat toisinaan puuttua joko kokonaan tai joidenkin kuolleiden kohdalta. Pienten lasten iät, joita ei ollut merkitty Ikä-sarakkeeseen, saattoivat olla merkittyinä huomautustietoihin. Vuoden 1880 kuolleiden luettelossa oli vain aikuisina kuolleita ja lapset oli kirjattu lastenkirjaan. Myös vuoden 1881 luetteloista puuttui kuolleista joitakin merkintöjä, jotka löytyivät lastenkirjoista. Vuosien 1886 ja 1887 kuolleiden luetteloista puuttuvat tiedot löytyivät rippikirjoista. Verrattaessa syntyneiden ja kuolleiden luetteloja havaittiin, ettei kaikkia pian syntymänsä jälkeen kuolleita lapsia ollut lainkaan merkitty syntyneiden luetteloihin. Nämä lapset on tutkimuksessa lisätty syntyneihin, jos kuolleiden luetteloissa on mainittu heidän ikänsä, eikä selvästikään vastaavaa lasta esiinny syntyneiden luettelossa. Tämän varmistettiin siten, että muodostettiin syntyneiden luetteloista lasten vanhempien nimien perusteella perhekokonaisuuksia. Jos kuolleiden luetteloissa mainittua lasta ei esiintynyt perheen tiedoissa ja lapsi ikänsä puolesta sopi perhekokonaisuuteen, lisättiin lapsi myös syntyneihin asianomaisen vuoden kohdalle. Kuolleiden luetteloista jäi lisäämättä syntyneihin joitakin sellaisia tapauksia, joiden ikää eikä vanhempia ollut mainittu ja joista näin ollen ei voinut tietää, minä vuonna he olivat syntyneet.

Syntyneiden ja kastettujen sekä kuolleiden ja haudattujen luetteloiden pohjalta on laadittu taulukko (LIITE 1), jossa on syntyneiden ja kuolleiden lukumäärät vuosittain. Taulukon Huomautus-sarakkeeseen on vuosittain merkitty myös kuolleina syntyneet sekä sellaiset tapaukset, joita ei ole lisätty syntyneihin, koska kuolleen henkilön iästä ja syntymävuodesta ei ole varmuutta. Lapset on kuoliniän perusteella jaoteltu ryhmiin. Lisäksi taulukossa on mainittu kymmenvuosittain kunkin ikäryhmän lukumäärä sekä osuus syntyneistä prosentteina.

Yksityiskohtainen kuvaus Karjala-tietokannasta on tämän tutkimuksen teoreettisessa osassa. Tässä on syytä mainita vain, että vuoden 1994 syksyyn

mennessä oli kaikki Lavansaaren seurakunnan kirkonkirjojen merkinnät Karjala-tietokannan toimesta tallennettu ja siirretty tietokantoihin. Tässä tutkimuksessa on tietokannan hakuohjelmana käytetty Visual Database-ohjelmistoa.

2 PIENTEN LASTEN KUOLLEISUUDEN KEHITYS

2.1 Lavansaaren yleinen kuolleisuus

Lavansaaren asukasluvusta ei ole saatavissa vuosittaisia tietoja. Väkilukutiedot perustuvat arvioihin. Vuoden 1778 asukasluvuksi Korhonen arvioi 313 ja Talve 348 henkeä. Vuoden 1812 asukasluvuksi Korhonen arvioi 351 henkeä. Talven mukaan asukkaita oli 1800-luvun puolivälissä noin 600 henkeä. Vuonna 1900 asukasluku oli 975 henkeä. 1820-luvulta lähtien asukasmäärä kasvoi koko 1800-luvun ajan aina vuoteen 1921 saakka, jolloin se oli korkeimmillaan noin 1300 henkeä. Saaren väkiluku ei kuitenkaan kasvanut tasaisesti. Korhonen on päätellyt, että asukasluvun kasvu perustui lähes tyystin väestön luonnolliseen lisäykseen ja että muuttoliike saarelle päin pysyi vähäisenä.²⁰

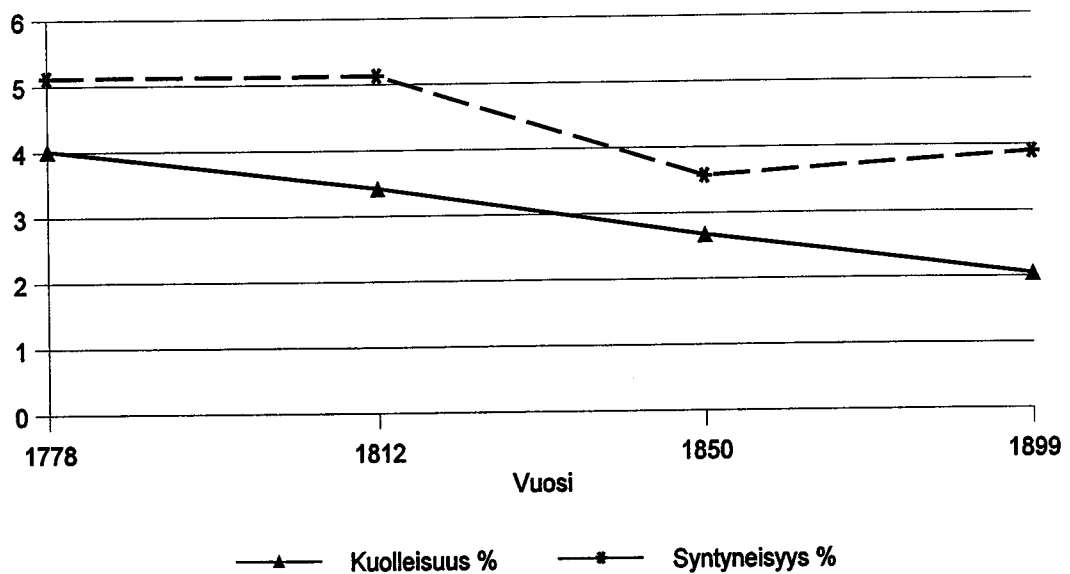
Asukaslukuarvioiden perusteella on laskettu Lavansaaren kuolleisuus vuosina 1778, 1812, 1850 ja 1899 ja esitetty sen vaihtelua Kuvassa 1. Koska väkilukutiedot perustuvat arvioihin, on kuolleisuuden laskemisessa käytetty vastaavan kymmenluvun kuolleiden keskiarvoa. Kuolleisuus laski Lavansaarella vuoden 1778 yli neljästä prosentista lähes puoleen 1800-luvun loppuun mennessä. Koska väkiluvusta on tietoja niin harvoin, eivät kuolleisuuden vaihtelut tule näkyviin Lavansaaren kuolleisuuden kehitys näyttäisi poikkeavan koko Suomen kehityksestä. Suomessa kokonaiskuolleisuus alentui huomattavasti vasta 1800-luvun lopulla, mutta Lavansaarella kuolleisuuden aleneminen näyttää alkaneen jo 1700-luvulla ja jatkuneen vielä jonkin verran kiihtyen 1800-luvun puolella.²¹

Kuvassa 1 on esitetty myös Lavansaaren syntyneisyyden muutoksia 1778-1900.²²

²⁰ Talve 1996, s. 121, Korhonen 1996, s. 113

²¹ Koskinen, Martelin 1994, s. 152

²² Syntyneisyys ilmoittaa syntyneiden osuuden asukasluvusta.



Kuva 1 Lavansaaren kuolleisuuden ja syntyneisyyden kehitys vuosina 1778-1899.

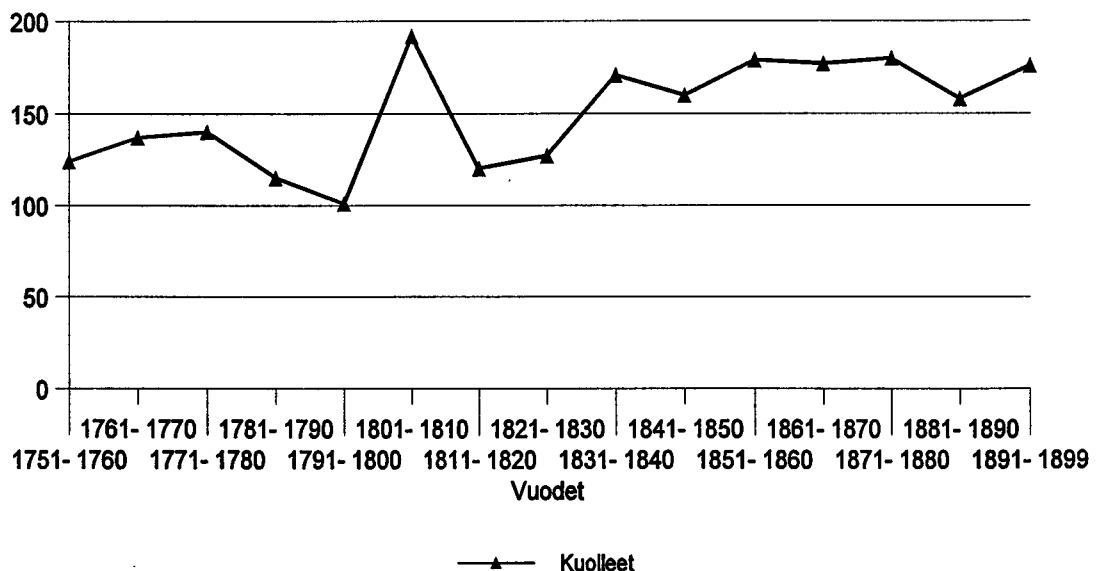
Suomen syntyvyydessä ei tapahtunut tasomuutoksia 1700-luvulta 1800-luvun alkuvuosikymmeneihin mentäessä. Myös Lavansaarella vuosina 1778-1812 syntyneisyys pysyi hyvin tasaisena, mutta laski sitten 1800-luvun puoliväliin mennessä. Koska väkiluvusta on seuraava arvio vasta 1800-luvun puolivälistä, ei varmasti voi sanoa, milloin syntyvyyden lasku alkoi. Vuosisadan puolivälin jälkeen syntyneisyys kääntyi jälleen nousuun. Pitkäsen käsitys on, että syntyneisyyden pientyminen 1800-luvulla liittyi alentuneeseen avioliittoisuuteen, mikä oli seurausta kiristyneistä taloudellisista olosuhteista.²³ Lavansaarella syntyneisyyden ja kuolleisuuden välinen ero kasvoi 1800-luvun alkupuolelle asti, minkä jälkeen se kapeni vuosisadan puoliväliin asti. Syntyvyyden ja kuolleisuuden välinen ero kasvoi jälleen 1800-luvun loppua kohti mentäessä ja 1800-luvun lopulla syntyneisyys oli Lavansaarella jonkin verran korkeampi kuin koko Suomessa yleensä.²⁴ Tämä näyttäisi vahvistavan Korhosen päätelmää, että

²³ Pitkänen 1994, s. 43-44

²⁴ Pitkänen 1994, s. 42

väkiluku lisääntyi voimakkaasti 1700-luvun lopulla ja ettei lisääntyminen johtunut muuttoliikkeestä. 1800-luvun ensimmäisellä puoliskolla väestönkasvu näyttäisi hiljentyneen, mutta sen jälkeen kohonnut syntyvyys heijastui myös asukasluvun kasvussa; väkiluku oli 1920-luvulla yli 1300 henkeä. Se, että asukkaiden määrä saattoi lisääntyä, viittasi siihen, että Lavansaari pystyi tarjoamaan elinmahdollisuuksia kasvavalle asukasmäärälle.²⁵

Kuvassa 2 on esitetty graafi, joka kuvaa kuolleiden määrän vaihteluita Lavansaarella 1751-1899. Koska väkiluku kasvoi 1800-luvulla, niin myös kuolleiden määrä luonnollisesti lisääntyi, eikä graafi anna kuvaa kuolleisuuden tasosta, vaan vain lukumääräiset vaihtelut. 1760- ja 1770-luvuilla kuolleiden määrä oli nousussa, mutta laski sitten 1780-luvulta alkaen ja oli matalimmillaan 1790-luvulla. Tutkimuskauden korkein huippu Lavansaarella osui vuosikymmenelle 1801-1810. 1751-1899 välisenä aikana kuolleiden määrä ylitti syntyneiden määrän vain kahtena kymmenvuotiskautena: 1801-1810 ja 1831-1840.



Kuva 2 Muutokset Lavansaaren kuolleiden lukumäärissä vuosina 1751-1899.

²⁵ Korhonen 1994, s. 114

1801-1810 kuolleiden määrä oli yli kolmekymmentä henkeä suurempi kuin syntyneiden määrä. 1830-luvulla kuolleita oli kahdeksan henkeä enemmän kuin syntyneitä. 1830-luvulta lähtien kuolleiden määrä pysyi suunnilleen saman suuruisena vuosisadan loppuun asti. Kun väkiluku kasvoi, tämä tarkoitti sitä, että kuolleisuus laski, mitä osoittaa myös kuvan 1 kuolleisuutta esittävä graafi.

1760- ja -70-luvuille oli ominaista ilmaston huononeminen ja siitä johtuvat heikot sadot useina vuosina. Tilanteen teki vielä vaikeammaksi se, että myös Venäjällä, josta viljaa oli totuttu ostamaan, oli huono viljasato.²⁶ 1800-luvun ensimmäisellä vuosikymmenellä, toisin kuin myöhemmin 1860-luvulla ei alueen merellinen sijainti eikä merestä saatu ruoka auttanut ihmisiä selviämään katovuosista. Silloin Venäjä oli sodassa Ranskaa vastaan, mikä todennäköisesti vaikeutti viljan saantia sieltä. Pahimmat kuolonvuodet osuivat vuosikymmenen alkuun, ennen Suomen sotaa. Korhonen toteaa, ettei vuosien 1808-1809 sota suuremmin vaikuttanut saarten ympäristöön.²⁷ 1860-luvulla kuolleisuus ei Lavansaarella noussut samoihin mittoihin kuin Suomessa keskimäärin. Näyttää siltä, että 1860-luvun nälkävuosien aikana Lavansaarella, kuten myös Ahvenanmaalla, saarten elinkeinorakenne suojaasi kaikkein tuhoisimmalta nälänhädältä. Kun saaristolaiset voivat hankkia ruokansa merestä myös talvisaikaan, niin heillä oli mahdollisuus säilyä hengissä ja terveenä toisin kuin sisämaan asukkailla, joiden oli pakko turvautua hätäravintoon.²⁸ Kuolleisuudessa oli tosin mantereen puolellakin suuria eroja 1860-luvulla eri paikkakuntien välillä.²⁹

Lavansaarella kuolleiden määrä oli nousussa 1700-luvun puolivälistä vuoteen 1780 asti, mutta laski sitten vuosisadan loppua kohti mentäessä ja kun syntyneisyys pysyi korkeana, on todennäköistä, että väkiluku nousi 1700-luvun

²⁶ Sundin 1996, Alanen 1963

²⁷ Korhonen 1996, s. 97

²⁸ Moring 1994, s. 133-134

²⁹ Pitkänen 1993, s. 83-89

lopulla. Kuitenkin asukkaiden määrä oli lähes sama vuosina 1778 ja 1812. Tämä on mahdollisesti johtunut vuosien 1803-1806 ja 1808 korkeista kuolleisuusluvuista, joista väestö ei vielä vuoteen 1812 mennessä ollut ehtinyt toipua.

Lavansaarella korkeat kuolleisuusluvut olivat seurausta erilaisista kriiseistä, joita huonot sadot ja sotavuodet aiheuttivat. Huonoja satoja seurasivat erilaiset tauti-epidemiat. Ihmisten vastustus-kyvyille ei ainoastaan ravinnon määrä ollut tärkeä, vaan myös ruuan laatu. Ruuan puute sekä jokainen erillinen tartunta heikensivät yksilön vastustuskykyä ja lopullinen kuolemansyy oli moninaisten tautien yhteistulosta.³⁰ Aikuisten ihmisten tavallisiksi kuolemansyiksi oli kuolleiden luette-loihin merkitty vanhuuden heikkous, keuhko-, vatsa-, kuume- ja polttotauti sekä halvaukset. Ympäristöstä johtuvana erityispiirteenä ovat olleet erilaiset onnetto-muuksista johtuvat kuolemat ja hukkumistapaukset. Kuolleiden luetteloiden merkintöjen mukaan hukkui vuonna 1831 viisi ja vuonna 1842 seitsemän henkeä yhdellä kertaa.

Suomessa korkean kuolleisuuden aikoja olivat 1760-70-lukujen lisäksi, vuosikymmen 1801-1810, 1830- sekä 1850-1870-luvut. Lavansaarella oli kuolleisuus vuosina 1778, 1812 ja 1850 jonkin verran korkeampi kuin Suomessa keskimäärin. Vuoteen 1899 mennessä erot olivat tasoittuneet ja kuolleisuus oli Lavansaarella suunnilleen samaa tasoa kuin Suomessa keskimäärin.³¹

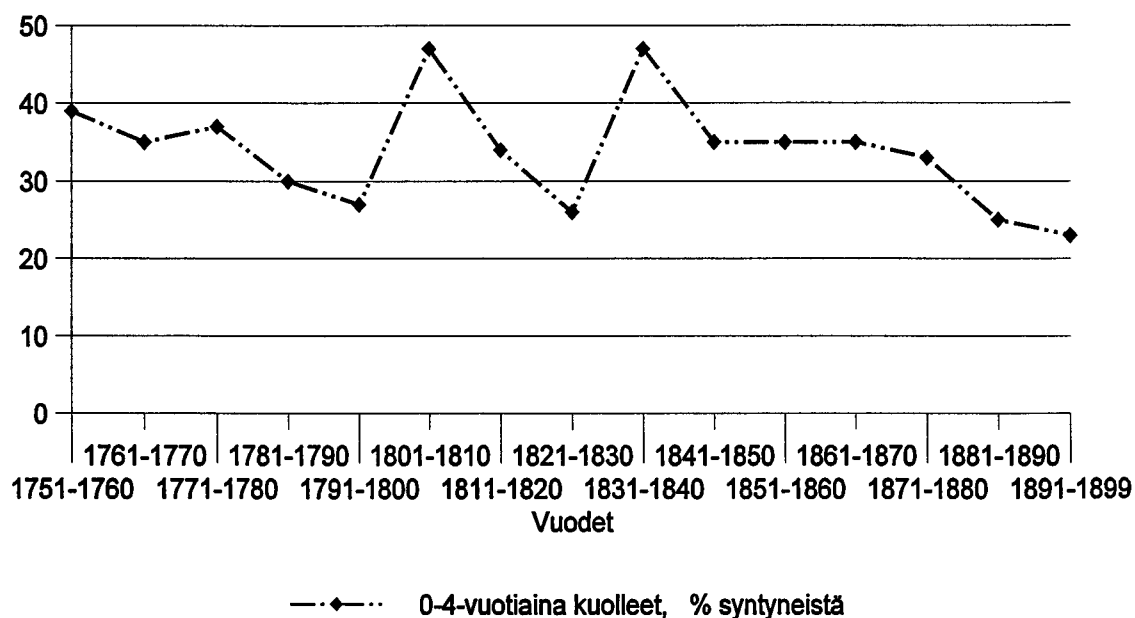
Lasten kuolleisuuden muutoksia tarkastellaan vertaamalla tutkittavan ikäryhmän kuolleiden osuutta syntyneistä, mihin tutkimukseen kuolleiden ja syntyneiden luettelot tarjoavat huomattavasti täydellisemmät tiedot kuin mitä oli käytettävissä Lavansaaren yleisen kuolleisuuden selvittämisessä.

³⁰ Sundin 199 Suomen virallinen tilasto 1871, s. 14-15

³¹ Pitkänen 1994, s. 42,46, Jutikkala 1968

2.2 Alle neljän vuoden ikäisten lasten kuolleisuus

Kuvassa 3 on esitetty Lavansaaren alle 4-vuotiaiden kuolleisuuden vaihteluita vuosina 1751-1899 kuvaava graafi. Syntyneistä lapsista kuoli 1700-luvun puolivälissä lähes 40 prosenttia alle neljän vuoden ikäisinä. 1700-luvun lopulla, jolloin kuolleisuus kokonaisuudessaan oli laskussa, tämän ikäisten kuolleisuus laski alle 30 prosentin. Kun etsitään selitystä 1790-luvun aikaisempaa matalampaan lasten kuolleisuuteen, niin havaitaan, että tuolla vuosikymmenellä ainoastaan yhtenä vuonna oli lievähkö rokkoepidemia. Myöskään 1820-luvulla, jolloin kuolleisuus oli matala, ei esiintynyt laajoja tartuntatautiepidemioita. Vuosikymmenien 1801-1810 ja 1831-1840 korkeat kuolleisuushuiput nostivat alle 4-vuotiaiden kuolleisuuden lähes 50 prosenttiin. Noina vuosikymmeninä oli melkein puolet kaikista Lavansaaren kuolleista alle neljän vuoden ikäisiä lapsia. 1801-1810 oli useita korkean kuolleisuuden vuosia. Niistä ankarin oli vuosi 1803. Silloin saarella oli rokko-epidemia ja kaikkiaan 39:stä kuolleesta 22 oli alle 4 vuoden ikäisiä. 1830-luvulla kuolleisuus oli tavallista korkeampi useina vuosina ja lapsia



Kuva 3 Lavansaaren 0-4-vuotiaiden kuolleisuuden muutokset vuosina 1751-1899.

kuoli erilaisiin tauteihin. Erityisen paljon lapsia kuoli vuonna 1832, jolloin Lavansaarella raivosi ankara isorokkoepidemia. Alle neljä vuotiaiden kuolleisuus laski 1840-luvulla noin 35 prosentin tuntumaan ja pysyi lähes neljän vuosikymmenen ajan suunnilleen samoissa lukemissa ennen kuin kääntyi selkeään laskuun 1880-luvulla ja oli sitten vuosisadan lopussa alle 25 prosenttia. Alle vuoden ikäisten lasten kuolleisuuden taso ei noussut 1860-luvun nälkävuosina kuten ei aikuistenkaan. Lavansaaren 0-4-vuotiaiden kuolleisuuden huiput osuivat samaan aikaan korkean yleisen kuolleisuuden kanssa. Kuolleisuuden lasku oli Lavansaarella samansuuntainen kuin yleensä Suomessa. Koko Suomen alueella 1-2 vuoden ikäisten lasten kuolleisuus alkoi pudota nopeasti 1870-luvulla, 3-4 vuoden ikäisten vuosisadan lopulla ja 5-9-vuotiaiden seuraavan vuosisadan alussa.³²

Kun tarkastellaan 0-4-vuotiaiden kuolleisuuden pitkäaikaista kehitystä, niin havaitaan, että kuolleisuus laski selvästi tarkastelujakson aikana. Karkeasti arvioiden laski alle neljän vuoden ikäisten kuolleisuus noin viidenneksellä 1700-luvun puolivälistä 1800-luvun loppuun mennessä. Korkean kuolleisuuden vuosina lasten kuolleisuus näytti johtuvan erilaisista tartuntatautiepidemioista. Voisi ajatella, että Lavansaari sijaintinsa takia oli syrjäinen ja eristynyt paikka, jossa ei ollut kovinkaan suurta pelkoa tartuntataudeista. Kuitenkin Lavansaaren elinkeinorakenteelle oli ominaista liikkuvuus, joka vei ihmiset eri puolille Suomenlahden rannikkoa ja kaupungeja. Tämä edisti tartuntatautien leviämistä ja sitä kautta kuolleisuutta. Lisäksi asutus oli keskittynyt kyliin, joissa talot sijaitsivat tiiviisti lähekkäin, mikä antoi taudeille oivallisen lähtöaseman.³³

Korkeaa lasten kuolleisuutta pidettiin esiteollisena aikana annettuna; se liittyi läheisesti uskonnollisuuteen. Koska lasten kuolleisuus vaihteli alueellisesti, on

³² Jutikkala 1993

³³ Hannula 1947, s. 35

syytä otaksua, että oli eroja siinä, miten ihmiset tarkastelivat elämää ja kuolemaa. Ja kun lasten kuolleisuus laski ja kuolleiden ikärakenne muuttui, se vaikutti myös asenteisiin. Asenteet maallistuivat, mikä tarkoitti sitä, että ihmiset eivät enää pitäneet aiempaa tilannetta, jossa vallitsi korkea lasten kuolleisuus itsestään selvyytensä ja annettuna, Jumalan tekona.³⁴ Asenteiden maallistuminen Lavansaarella näkyi esimerkiksi lasten nimien muuttumisena ja lukumäärän nopeana lisääntymisenä, mikä alkoi 1870-luvulla. Aiemmin käytettyjen, lähinnä uskonnollista perua olevien nimien lukumäärä oli ollut suhteellisen suppea. 1860-luvulla Lavansaarella oli käytössä 14 tytön- ja 23 poikien nimeä. 1870-luvulla oli jo 30 tytön- ja 39 pojannimeä.³⁵

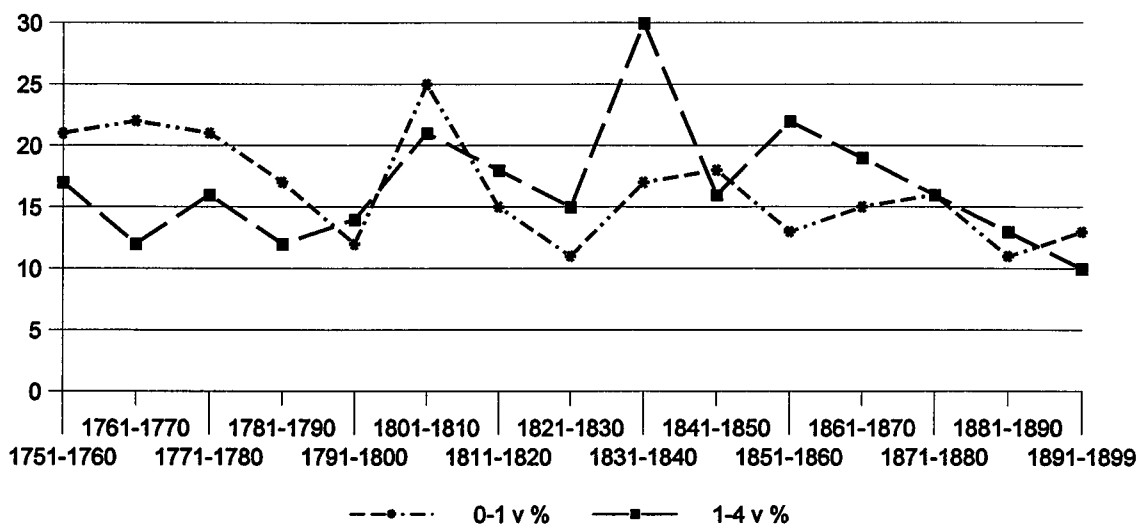
Tutkimuksen oletuksena oli, että pienten lasten kuolleisuudessa on eroja eri ikäryhmien välillä. Seuraavaksi tarkastellaankin erikseen 1-4-vuotiaiden ja imeväis-ikäisten kuolleisuuden vaihteluita vuosina 1751-1899.

2.3 Yhdestä neljään vuoden ikäisten kuolleisuus

Kuvassa 4 on esitetty Lavansaaren imeväisikäisten ja 1-4-vuoden ikäisten kuolleisuuden vaihtelua vuosina 1751-1899 kuvaavat graafit. Imeväisikäisten kuolleisuus oli korkeampi lähes koko 1700-luvun loppupuolen ja ero kasvoi 1760-luvulla yli seitsemään prosenttiyksikköön. Vähitellen ryhmien väliset erot pieneivät ja 1790-luvulla imeväisikäisten kuolleisuus oli jo matalampi kuin 1-4-vuotiaiden. Molempien ikäryhmien kuolleisuus nousi korkeaksi kymmenluvulla 1801-1810. Imeväisikäisten kuolleisuus oli tuolloin korkeimmillaan koko tarkastelujakson aikana, yli 25 prosentissa. 1820-luvulta alkaen imeväisten kuolleisuuden vaihtelut tasaantuivat ja lasku jatkui entistä selvempänä. Vaikka imeväisikäisten kuolleisuus nousikin jyrkästi 1801-1810, näyttäisi kuolleisuuden lasku alkaneen jo 1780-luvulla. 1-4-vuotiaiden kuolleisuus oli korkeimmillaan,

³⁴ Turpeinen 1984

³⁵ Talve 1996, s. 310-311



Kuva 4 Lavansaaren imeväisikäisten ja 1-4-vuotiaiden kuolleisuuden vaihtelut vuosina 1751-1899.

noin 30 prosentissa, 1830-luvulla. Vielä 1850-luvulla tämän ikäryhmän kuolleisuus kohosi yli 20 prosentin, mutta kääntyi sen jälkeen selkeään laskuun ja oli 1800-luvun viimeisellä vuosikymmenellä vähän päälle 10 prosenttia. Imeväisikäisten kuolleisuus kohosi 1800-luvun viimeisellä vuosikymmenellä ja oli vuosisadan lopussa hieman yli 13 prosenttia.

Varsinkin vuotta vanhemmat lapset sairastuivat 1700- ja 1800-luvuilla lapsille yleisiin tarttuviin tauteihin, jotka tuohon aikaan olivat usein kohtalokkaita ja johtivat kuolemaan. Suurin uhka heille oli isorokko. Isorokko vei hautaan enemmän ihmisiä kuin mikään muu sairaus. Se oli 1700-luvulla niin merkittävä tappaja, että sitä on toisinaan kutsuttu tämän aikakauden kansantaudiksi. Koska isorokko oli tarttuva tauti, jonka sairastaminen käytännössä aikaansai immuni- teetin, se oli etupäässä lastentauti. Isorokkoepidemioita esiintyi tietyin välein, kun uusi sukupolvi, jolta puuttui immunitetti, oli kyllin suuri. Useimmat lapset saivat taudin epidemioiden aikana ja silloin kuolleisuus kohosi huomattavasti.³⁶

³⁶ Sundin 1996, s. 98, Pitkänen 1994 s. 44-45, Jutikkala 1968, s. 29-30

Isorokkoepidemioiden välit näkyivät hyvin myös Lavansaarella. Ensimmäisen kerran tarkastelu- jakson aikana isorokko esiintyi Lavansaaren kuolleiden ja haudattujen luetteloissa epidemiana vuonna 1770, jolloin 33 lasta iältään viidestä viikosta 17 vuoteen kuoli tähän tautiin. Seuraava epidemia oli kolmentoista vuoden päästä vuonna 1783 ja silloin tautiin kuoli 30 lasta. Ja seuraava rokkoepidemia oli jälleen kolmentoista vuoden kuluttua 1796, jolloin 12 lasta iältään 4 kuukaudesta 6 vuoteen kuoli. Vuonna 1815 oli lievä rokkoepidemia. Seuraavan kerran isorokko esiintyi rajuna epidemiana vuonna 1832. Silloin siihen kuoli 27 lasta, joista vanhin oli 15 vuotta vanha ja nuorin kuuden kuukauden ikäinen. Tämän jälkeen kuolleiden luetteloissa ei ole mainittu kuolemansyynä isorokkoa, vaan on käytetty nimitystä rokko. Vuoden 1832 rajun epidemian jälkeen Lavansaarella esiintyi rokkoepidemioita esimerkiksi vuosina 1854, 1867 ja viimeisen kerran jakson aikana isorokko esiintyi Lavansaaren kuolleiden ja haudattujen luetteloissa epidemiana tarkastelujakson aikana vuonna 1892. Sen lisäksi, että tauti ilmeni epidemioina, sattui vuosittain joitakin kuolemantapauksia. Kun otetaan huomioon väkiluvun kasvu, niin 1800-luvun lopun epidemioiden aikana kuoli suhteellisesti vähemmän lapsia kuin aikaisemmin. Isorokon väheneminen vaikutti positiivisesti lasten vastustuskykyyn tartuntatautien suhteen, sillä jokainen infektiio vähensi yksilön vastustuskykyä.

Vuoden 1825 määräys ei vielä tehnyt isorokkorokotusta pakolliseksi, vaan se tuli pakolliseksi vasta 1883.³⁷ Lavansaarella ensimmäiset rokotusmerkinnät ilmes- tyivät 1800-luvun alkupuolella lastenkirjoihin ja rippikirjoihin. Suomessa keski- määräinen isorokkokuolleisuus, jossa paikalliset erot eivät näy, alkoi 1800-luvun ensimmäisen neljänneksen jälkeen huomattavasti pienentyä. Suuren määrän uhreja vaatineita epidemioita ilmeni tosin aina vuosisadan lopulle saakka, mutta niitä esiintyi aiempaa paljon harvemmin. Lavansaarella isorokkokuolleisuuden pienentyminen näyttäisi alkaneen jonkin verran myöhemmin kuin Suomessa yleensä. Suomalaisen aineiston perusteella voi tehdä johtopäätöksen, että 1800- luvulla tapahtunut isorokkokuolleisuuden pienentyminen johtui nimenomaan

³⁷ Forsius 1986

vuoden 1825 jälkeen tehostuneesta lasten rokotuskampanjasta.³⁸

Isorokon lisäksi myös tuhkarokko oli hyvin tuhoisa pienille lapsille. Siihen kuoli jokunen lapsi vuosittain, mutta se saattoi yltyä myös rajuksi epidemiaksi. Tarkastelujakson aikana tuhkarokko esiintyi ensimmäisen kerran epidemiana 1803. Silloin kuoli 13 lasta, joiden ikä vaihteli kahdesta kuukaudesta 8 vuoteen. Tuhkarokon lisäksi on tuolta vuodelta maininta 9 rokoon kuolleesta lapsesta. On mahdollista, että myös tämä tarkoitti tuhkarokkoa. Vuoden 1848 epidemia oli erittäin raju, silloin 27 lapsen kuolemansyyksi on merkitty tuhkarokko. Lasten ikä vaihteli 6 kuukaudesta 9 vuoteen. 1800-luvun toisella puoliskolla oli useita tuhkarokkoepidemioita. Vaikka rokkoepidemioita esiintyi koko vuosisadan ajan, niiden aikana kuolleiden lasten lukumäärä väheni 1800-luvun loppupuolella. Rokkotartuntoihin kuolleiden väheneminen oli osoitus lasten vastustuskyvyn parantumisesta.

Lavansaarella esiintyi myös muita hengityselinten sairauksia kuten hinkuyskää, kuristustautia (kurkkumätä) ja kuumetta. Joskus nekin saattoivat yltyä epidemioiksi. 1864-1865 esiintyi kuristustautiepidemia, johon kuoli 17 lasta, joiden ikä vaihteli kuudesta kuukaudesta 4 vuoteen. Hinkuyskää esiintyi myös epidemiana esimerkiksi vuonna 1897, jolloin siihen menehtyi kuusi lasta, joiden ikä vaihteli kuukaudesta 6 vuoteen. Hinkuyskä tarttui varsin pieniinkin lapsiin.³⁹

Suolistoperäiset taudit, joiden syynä oli saastunut vesi tai ravinto, olivat yleisiä koko tarkastelujakson ajan. Yleisin näistä oli polttotauti, jota esiintyi Lavansaarella koko tarkastelukauden ajan.⁴⁰ Se ei ollut yksinomaan lasten tauti, vaan

³⁸ Pitkänen 1994, s.44-45, Jutikkala 1968, s. 29-30

³⁹ Forsius 1986

⁴⁰ Polttotauti-nimitystä on käytetty useista eri tautioireista, joita olivat kova kuume, velttous ja tiedottomuus. Se saattoi tarkoittaa useita eri tauteja kuten lavantautia, pikkulavantautia, pilkkuumetta, toisintokuumetta ja mahdollisesti myös influenssaa.

siihen kuoli sekä lapsia että aikuisia. Polttotautiin kuoli lähes joka vuosi jokunen, mutta se saattoi myös kehittyä epidemiaksi, jolloin siihen menehtyi kymmeniä saarelaisia. Näin tapahtui vuosina 1804 ja 1805. Jo vuoden 1803 tuhkarokko-epidemian aikana esiintyi runsaasti polttotautia, mutta seuraavina vuosina tauti laajeni epidemiaksi. Vuonna 1804 siihen kuoli 18 henkeä ja 1805 23 henkeä. Polttotauti esiintyi yleisenä kuolemansyynä aina vuosisadan loppuun asti. Punatauti esiintyi usein samoihin aikoihin kuin polttotautiakin oli liikkeellä. Esimerkiksi vuosina 1776-1777 ja vielä 1779 ilmoitettiin kuolemansyiksi sekä polttotauti, punatauti, vatsatauti ja lisäksi pistos, jota Sundin pitää ripuliin liittyvänä oireena.⁴¹ Näihin tauteihin kuoli kaiken ikäisiä ihmisiä aina kahden viikon ikäisistä lapsista vanhuksiin ja niiden yhtäaikainen esiintyminen sekä 1770-luvulla että vuosina 1801-1810 viittaa siihen, että huonojen satojen aikana jouduttiin käyttämään ravintoa, joka oli joko saastunutta tai muuten ravinnoksi sopimatonta. Voi hyvin perustein päätellä, että nämä taudit olivat seurausta heikosta ravinto-tilanteesta.⁴² Toimeentulon heikkenemiseen viittaa sekin, että vuosina 1801-1810 riisitautia oli merkitty kuolemansyiksi. Riisitautiin kuolleet olivat aikuisia, eivät lapsia. Vaikka vielä 1808 kuolleisuus oli korkea, pahimmat kuolonvuodet, 1803-1806, osuivat aikaan ennen Suomen sodan syttymistä. Noina vuosina on täytynyt olla aivan erityisiä syitä, jotka vaikeuttivat ravinnon hankkimista merestä ja viljan hankkimista Venäjän puolelta.

Suolistoperäiset sairaudet pysyivät hengityselinsairauksien ohella lasten yleisinä kuolemansyinä koko 1800-luvun ajan. Esimerkiksi 1830- ja 1850-luvuilla 1-4-vuotiaiden korkea kuolleisuus oli paljolti rokkojen aiheuttamaa, mutta kuolleisuutta lisäsi kirjava joukko erilaisia tauteja: näivetystauti, polttotauti, ripuli, punatauti, kuume, keuhkotauti, halvaus, hinkuyskä, vesipöhö, hengenahdistus sekä hammaskivut. Hammaskivut alkoivat 1800-luvun puolivälistä esiintyä kuolemansyynä entistä useammin, esimerkiksi vuonna 1854 kuoli peräti

⁴¹ Sundin 1996, s. 98

⁴² Suomenmaan virallinen tilasto 1871, s. 13-15, Keränen 1994, s. 55

viisi lasta, joiden ikä vaihteli 9 kuukaudesta kahteen vuoteen, näihin vaivoihin. 1-4-vuotiaiden lasten kuolleisuus kääntyi laskuun 1830-luvulla, vaikka se vielä kerran 1850-luvulla kohosi lähes 22,5 prosenttiin. Kuolleisuuden lasku kiihtyi 1860-luvulta lähtien. Hannula mainitsee teoksessaan, että vuonna 1886 Lavansaarella oli vaikea lavantautiepidemia, johon kuoli 38 lasta. Hannula ei kuitenkaan ilmoita, mihin lähteisiin hän tietonsa perustaa. Vuoden 1886 kuolleiden luettelo on epätäydellinen ja apuna on käytetty rippikirjoja. Näissä lähteissä kuolleita on kaiken kaikkiaan 27 henkeä, joista 18 on alle 4 vuoden ikäisiä lapsia. Kuolemansyitä ei rippikirjoissa ole mainittu.⁴³

Kuolemansyiden yksityiskohtainen tarkastelu poikkeuksellisen korkean kuolleisuuden vuosikymmeninä, 1801-1810 ja 1831-1840, osoittaa, että yhdestä neljään vuoden ikäisinä kuolleiden lasten kuolemansyyt olivat pääasiassa moninaisia tarttuvia tauteja: rokkoja, hinkuyskää, kuumetta, vesipöhää, polttotautia, punatautia sekä ripulia ja muita vatsatauteja. Vain vuonna 1832 oli yksi, erittäin selkeä syy: isorokko. Alle kuukauden vanhoina kuolleiden kuolemansyiksi on kuolleiden luetteloihin merkitty useimmiten tuntematon tauti tai heikkous. Vanhemmilla imeväisikäisillä esiintyy joitakin mainintoja rokkoihin ja polttotautiin kuolleista. Äidiltä peritty immuniteetti haihtui nopeasti ja rokkoja esiintyi jo 2-5 kuukauden ikäisillä lapsilla. Kuumeeseen ja polttotautiin kuoli 3-5 kuukauden ikäisiä lapsia.

1-4-vuotiaiden kuolleisuuden pitkäaikainen kehitys Lavansaarella osoittaa hienoista nousua vuosien 1751-1899 välisenä aikana, kun taas alle vuoden ikäisten kuolleisuus lähes puolittuu. Seuraavassa luvussa verrataan aluksi Lavansaaren ja koko Suomen imeväiskuolleisuuden kehitystä, minkä jälkeen tarkastellaan lähemmin eri ikäisten imeväisten kuolleisuuden kehitystä Lavansaarella.

⁴³ Hannula 1947, s. 93

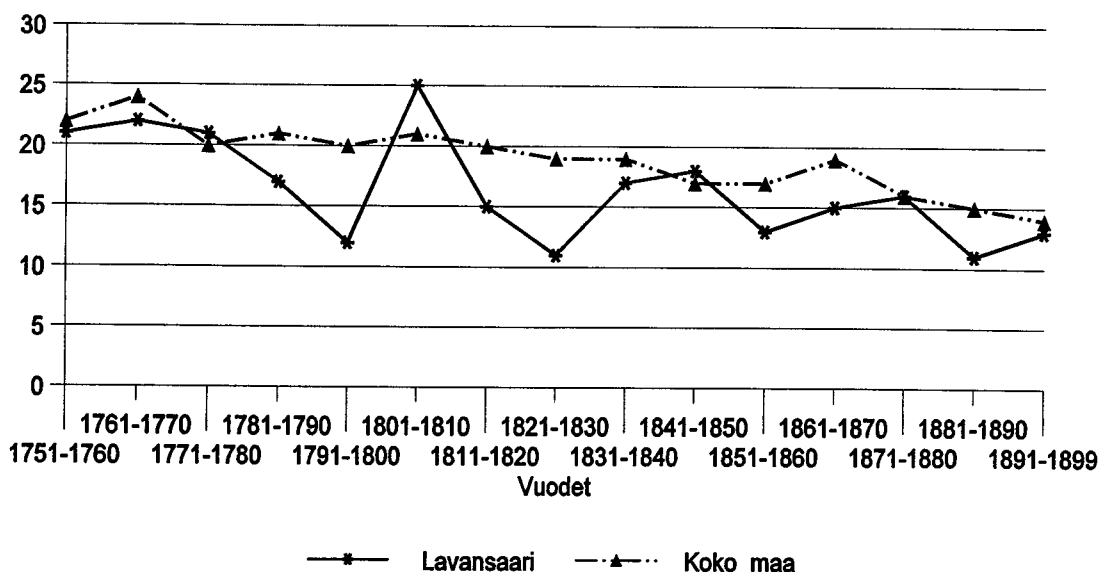
2.4 Imeväiskuolleisuus

1770-luvulta lähtien tapahtui Lavansaarella alle vuoden ikäisten kuolleisuudessa huomattavaa alenemista. Alimmillaan, noin 12 prosenttia, kuolleisuus oli 1790-luvulla. Vuosikymmen 1801-1810 oli kuitenkin ankara varsinkin imeväisikäisten kohdalla. Tuolla vuosikymmenellä joka neljäs kaikista syntyneistä lapsista kuoli ennen ensimmäistä ikävuottaan. Tämän jälkeen kuolleisuudessa ei enää tapahtunut niin jyrkkiä muutoksia, vaan vaihtelut tasaantuivat.

Imeväiskuolleisuus kohosi korkeaksi, noin 19 prosenttiin, 1840-luvulla, mutta jäi selvästi matalammaksi kuin 1801-1810. Vaikka vuosikymmenelle 1801-1810 osui vielä korkea kuolleisuushuippu, näyttäisi imeväiskuolleisuuden lasku Lavansaarella alkaneen jo 1780-luvulla. Samansuuntainen kehitys tapahtui koko Suomen alueella.⁴⁴

Kuvassa 5 on esitetty Lavansaaren ja koko maan imeväiskuolleisuuden vaihteluita kuvaavat graafit. Koska koko maata käsittävä aineisto on erittäin laaja, sen kuolleisuudesta puuttuvat jyrkät vaihtelut, jotka Lavansaaren pienemmän aineiston kohdalla näkyvät selvästi. Sekä koko maan että Lavansaaren imeväiskuolleisuus laski 1700-luvun puolivälin yli 20 prosentista 13-14 prosenttiin 1800-luvun loppuun mennessä. Lavansaaren imeväiskuolleisuus oli koko maan kuolleisuutta korkeampi 1770-luvulla, kymmenluvulla 1801-1810 sekä 1840-luvulla. Lavansaaren imeväiskuolleisuus oli vuosina 1801-1810 lähes neljä prosenttiyksikköä korkeampi kuin Suomessa keskimäärin. Imeväiskuolleisuus oli korkea perättäisinä vuosina 1803-1806 ja vielä vuonna 1808. Tuo vuosikymmenen ei ollut ankara vain imeväisikäisille. Silloin kuoli melkein kaksinkertainen määrä lavansaarelaisia verrattuna viereisiin vuosikymmeniin. Kaikkiaan kuoli 192 henkeä ja syntyi 140 lasta. 1800-luvun ensimmäisellä vuosi-

⁴⁴ Jutikkala 1993



Kuva 5 Lavansaaren ja koko Suomen imeväiskuolleisuuden vaihtelu vuosina 1751-1899.

kymmenellä kuutena vuonna kymmenestä kuoli enemmän ihmisiä kuin syntyi. Yleensä syntyvien lasten lukumäärä Lavansaarella korkeampi kuin kuolleiden määrä.

Imeväiskuolleisuuden lasku Suomessa johti aiemmin vallinneiden huomattavien alueellisten erojen tasoittumiseen. Kuolleisuuden alentuminen tapahtui eri aikoina eri ikäryhmissä. Eri ikäisten kuolemansyyt myös olivat erilaisia ja painottuivat eri tavoin. Kuolleisuuden alenemista on selitetty yleisen elintason, hygienian ja terveydenhuollon parantumisella. Elintason kohentuminen heijastui ensinnäkin lasten hoidon ja ravitsemuksen paranemisena, minkä seurauksena sairastavuus ja kuolleisuus joihinkin tartuntatauteihin on saattanut olennaisesti pienentyä. Toisaalta kohentuneet elinolot ovat tuoneet parannusta asumisahtauteen ja henkilökohtaiseen hygieniaan, joilla myös on ollut vaikutuksia tartuntatauteihin sairastumiseen ja kuolleisuuteen.

Taloudellinen kehitys näkyi Lavansaarella asuntojen koon kasvamisena. 1600-

luvulla yleisiä savutupia alettiin 1700-luvulla muuttaa uloslämpiviiksi, kun savupiiput tulivat käyttöön. Tuvat saivat ensimmäiset lasi-ikkunat. 1800-luvulla asuntona yleistyivät uloslämpivä tupa eteisineen tai tupa-kamari-eteinen- sekä mahdollisesti myös kolmihuoneinen tupa-kyökki-kamari-asuinrakennus. Huomattava taloudellinen nousukausi alkoi 1870-luvun lopulla ja 1880-luvun alussa ja jatkui vuosisadan vaihteen yli. Taloudellinen vaurastuminen näkyi asuntojen, vaatetuksen ja jossain määrin myös ruokatalouden parantumisena. 1860-luvun jälkeenkin oli tupa-kamari-yhdistelmä tosin yleisin, mutta sen rinnalle ilmestyi kolmi- ja nelihuoneisia asuinrakennuksia. Uusia lisärakennuksia olivat erilliset käymälät.⁴⁵ Hygienian tasoa paransi myös se, että luovuttiin vanhasta tavasta naulata hylkeennahat lattiaan. Aikaisemmin ei tästä ollut huomattu aiheutuvan mitään vaaraa, vaikka lapset ryömivät lattialla mätänevien nahkojen seassa.⁴⁶

Ruokatalouden uutuus, peruna, levisi maaseudulle 1800-luvun alkupuolella ja oli tärkeä lisä ruokavalioon. 1800-luvun lopulla toivat lisääntyneet yhteydet ruokatalouteen lisää ostotavaroita: ryynit, riisin, sokerin, kahvin ja teen. Ruokatalous perustui kuitenkin kuten ennenkin omavarais-talouteen ja vaihtokaupalla hankittuihin peruselintarvikkeisiin.⁴⁷

Kuolleisuuden alueellinen esiintyminen ei kuitenkaan tue käsitystä taloudellisten tekijöiden ensisijaisesta merkityksestä. Satovaihtelut eivät suoraan heijastuneet imeväiskuolleisuuteen. Imeväiskuolleisuuden taso näyttää eniten riippuneen lastenhoitotavoista, etenkin imetystavoista. Suurin syy imeväiskuolleisuuden paikallisiin eroihin oli imetyksen laiminlyönti. Tutkijat ovatkin yhtä mieltä siitä, että imettäminen oli tärkein imeväiskuolleisuuteen vaikuttava tekijä. Muiden tekijöiden, kuten epidemioiden, elintason, hygienian, synnytys-

⁴⁵ Talve 1996, 345-348

⁴⁶ Hannula 1947, s. 90

⁴⁷ Talve 1996, s. 346-348

komplikaatioiden ja mahdollisesti pelkästään paikallisten olosuhteiden, vaikutuksesta vallitsee erilaisia näkemyksiä.⁴⁸

Kansa koki imeväiskuolleisuuden olevan kohtalon sallimaa: Jumala korjasi satoaan. Tämän katsantotavan väistyminen ja fatalistisen maailman-katsomuksen otteen heikkeneminen vaikutti myös osaltaan imeväiskuolleisuuden alenemiseen.⁴⁹ Lääkärit ja viranomaiset taas katsoivat korkean imeväiskuolleisuuden suurelta osalta olevan seurausta äitien tavasta syöttää lapsilleen lehmän maitoa, joka valui sarvesta. Sarven käyttö lisäsi infektioiden vaaraa. Myös Lavansaarella oli käytössä lapsensarvi eli naudansarvi, jonka pää oli suljettu lampaannahalla, säämiskälevyillä tai lehmän nisänahalla. Suomen puolella kiinnitettiin jo 1700-luvun puolivälissä huomiota rinta-ruokinnan tehostamiseen. Lavansaari kuului tuolloin Venäjään, eikä tämä kehoitus niin ollen koskenut saaren asukkaita. 1840-luvulla myös tsaari antoi määräyksiä imetyksen tehostamiseksi. Ajan mittaan lääkärit oppivat ymmärtämään, millä tavoin sairauksia voitiin ehkäistä. He alkoivat neuvoa ihmisiä esimerkiksi lastenhoidossa sekä lapsille yleisten ripuli- ja hengityselinsairauksien hoidossa.⁵⁰

Lastenhoitotavat muuttuivat Suomessa 1800-luvun loppupuolella. Pitkät 1-3 vuoden imetysajat lyhenivät. Varhemmin pitkä imettäminen oli ollut tavallista köyhissä perheissä ja niillä alueilla, joilla karjatalous oli huonosti tuottavaa. Äidinmaito oli tällöin ainoa ravinto, jota pikkulapselle voi antaa.⁵¹ Vuonna 1778 Lavansaarella oli 86, 1830-luvulla vain 30 ja 1920-luvulla 161 lehmää. Näistä luvuista voi päätellä, ettei ainakaan 1800-luvun alkupuolella kaikilla perheillä ollut lehmiä. Siksi tuntuisi luontevalta ajatella, että niissä perheissä, joissa

⁴⁸ Moring 1994, s. 113, Turpeinen 1984

⁴⁹ Turpeinen 1979 a, Pitkänen 1993

⁵⁰ Turpeinen 1979b, Talve 1996, s. 311

⁵¹ Helsti 1996

lehmä ei ollut, lapsia imetettiin pidempään ainakin vielä 1800-luvun alkupuolella ja riskialtis vieroitusaika osui lapsen 1-4 vuoden ikään.⁵² Talven mukaan imetysaika Lavansaarella oli puolisen vuotta. Lasta alettiin vieroittaa, kun ensimmäinen hammas oli tulossa. Tieto puolen vuoden imetysajasta koskee 1800-lukua, mutta tarkkaa tietoa siitä, milloin näin lyhytaikaiseen imettämiseen siirryttiin, ei Talven tutkimuksesta löydy. Imetysajan lyheneminen siirsi vaarallista vieroitusaikaa aikaisemmaksi, lasten ensimmäisen elinvuoden toiselle puoliskolle.⁵³

Lapset tarvitsivat äidinmaidon lisäksi muuta ravintoa. Lavansaarella lapsille annettiin lisäravintona velliä tai esimerkiksi kalasta ja perunasta tehtyä kiinteää ruokaa. Äiti hienonsi ruoan suussaan.⁵⁴ Moringin käsityksen mukaan peruna vaikutti merkittävästi siihen, että 1-12-kuukauden ikäisten lasten kuolleisuus aleni Ahvenanmaalla 1700-luvun lopulla ja 1800-luvun alussa. Peruna paransi myös äitien ravintotilannetta. Lavansaarellakin perunasta tuli tärkeä lisä ruokavalioon. Sitä hankittiin vaihtokaupalla, mutta myös sen viljelyä kokeiltiin. Koska imeväisikäiset olivat taipuvaisia vatsasairauksiin, peruna oli tärkeä, sillä se oli ravitseva ja puhdas lisäravinto.⁵⁵

Eri ikäisten kuolleisuuden kehityksen tarkastelemista varten on eri ikäiset imeväiset jaettu iän perusteella neljään ryhmään: 0-7 päivän ikäiset, 8 päivästä - alle kuukauden ikäiset, kuukaudesta -alle 6 kuukauden ikäiset ja 6 kuukaudesta - alle vuoden ikäiset.

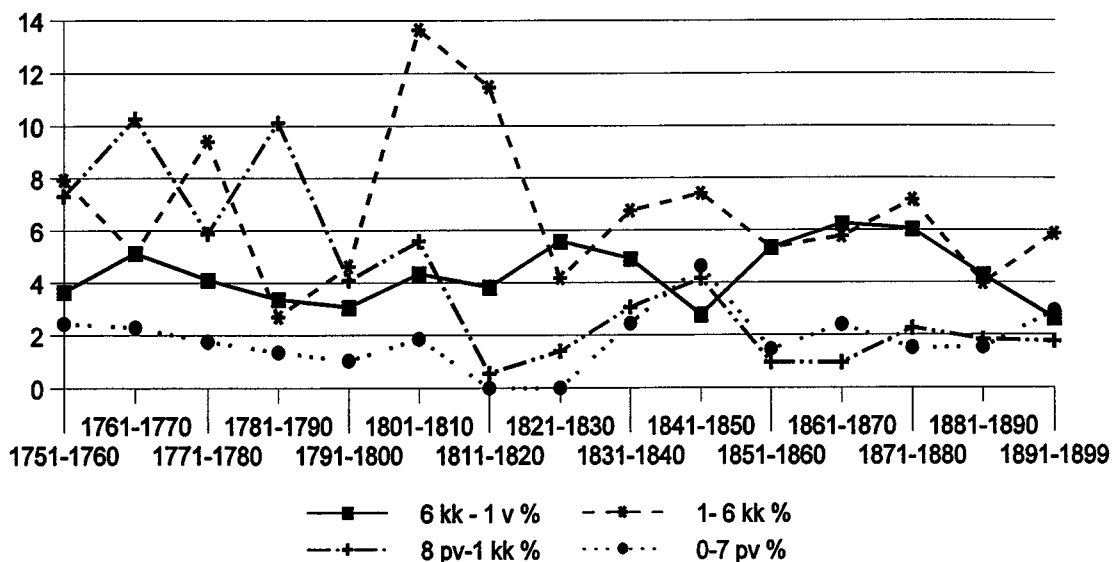
Kuvassa 6 on esitetty Lavansaaren eri imeväisryhmien (0-7 päivää, 8 päivää - alle

⁵² Talve 1996, s. 188

⁵³ Talve 1996, s. 311

⁵⁴ Talve 1996, s. 311

⁵⁵ Moring 1994, s. 123-125



Kuva 6 Lavansaaren eri imeväisryhmien kuolleisuuden vaihtelut vuosina 1751-1899.

kuukausi, kuukausi - alle 6 kuukautta, 6 kuukautta - alle vuosi) kuolleisuuden vaihteluita vuosina 1751-1899. Kaikkein nuorimpien lasten kuolleisuus pysyi jokseenkin tasaisena 1751-1800 välisenä aikana. Kuolleisuus vaihteli tuona aikana noin 1,5-2,5 prosentin välillä. 1800-luvun ensimmäisellä vuosikymmenellä, joka Lavansaarella oli erittäin korkean kuolleisuuden aikaa, tämän ryhmän kuolleisuus kohosi vajaan prosenttiyksikön. Vuosien 1811-1830 tilanne, jolloin kuolleiden ja haudattujen luetteloissa ei ole lainkaan 0-7 päivän ikäisinä kuolleita, ei tunnu uskottavalta, vaan on todennäköistä, että merkinnät jostain syystä puuttuvat. Myös 8 päivää-1 kuukauden ikäisten kuolleisuus 1811-1820 vaikuttaa liian matalalta. Merkintöjä kuolleista ei löydy myöskään lastenkirjoista eikä rippikirjoista.

Huomattavin nousu 0-7 päivän ikäisten kuolleisuudessa oli 1840-luvulla. Silloin lähes viisi prosenttia syntyneistä kuoli syntymää seuraavalla viikolla. Nälkävuodet 1860-luvulla nostivat tämän ikäisten kuolleisuutta vain lievästi.

Kuten Moringin tutkimissa pitäjissä Ahvenanmaalla, myös Lavansaarella tämän ikäryhmän kuolleisuus oli nousussa vuosisadan viimeisellä vuosi-kymmenellä. 0-7 päivän ikäisinä kuolleiden kuolemansyiksi on kirkonkirjoihin merkitty epämääräisiä syitä kuten heikkous, tuntematon tauti, pistos tai synnytys. Tutkijoiden käsityksen mukaan yhtäkkiset kuolemat synnytyksen jälkeen aiheutuivat sisäsyntyisistä tekijöistä, jotka johtuivat raskauden ja synnytyksen aikaisista olosuhteista.⁵⁶ Syntyminen kaksosena oli lapselle riskitekijä. Ongelmallisiin synnytyksiin ei Lavansaarella 1800-luvulla ollut mahdollista saada lääkärinapua, sillä lääkärit olivat kaukana mantereella. Vasta vuoden 1817 jälkeen, jolloin Lavansaari liitettiin Koivistoon, alkoi sen väestö saada jatkuvampaa hoitoa Koiviston lääkäreiltä. Vuosisadan keskivaiheilta lähtien nämä lääkärit kävivät säännöllisesti saarella ja loppupuolella ainakin kerran kuukaudessa, mutta tarpeen vaatiessa useamminkin. Diakonissa, kättilö, sairaanhoitaja ja oma lääkäri saatiin saareen vasta 1900-luvulla.⁵⁷

Merkinnät kuolleena syntyneistä ilmestyivät kuolleiden luetteloihin vuodesta 1785 lähtien. 1820-luvulla kuolleena syntyneiden osuus kaikista syntyneistä oli 4,2 prosenttia, 1840-luvulla 2,3 prosenttia, 1850-luvulla 1,9 prosenttia ja 1860-luvulla 3,8 prosenttia. Koska kuolleena syntymisen syyt olivat samantapaiset kuin 0-7 päivän ikäisinä kuolleilla, nämä tiedot täydentävät kuvaa alle 7 päivän ikäisten lasten kuolleisuudesta. Esimerkiksi 1820-luvulla, jolloin ei ole merkintöjä 0-7 päivän ikäisinä kuolleista, kuolleina syntyneiden osuus oli korkea. 1840-luvulla oli sekä 0-7 päivän ikäisinä kuolleiden että kuolleena syntyneiden määrä korkea. Kuolleena syntyneiden määrä oli taipuvainen olemaan suuri sellaisina vuosina, jolloin syntyneisyys oli matala ja kuolleisuus korkea. Sen perusteella voi päätellä, että odottavat äidit kärsivät puutetta noina vuosikymmeninä.⁵⁸ 1860-luvun korkea kuolleena syntyneiden lasten määrä

⁵⁶ Sundin 1996, s. 102

⁵⁷ Hannula 1947, s. 89-97, Suomenmaan virallinen tilasto 1871, s. 24-25

⁵⁸ Suomenmaan virallinen tilasto 1871, s.23

osoittaa, että vaikka kuolleisuus ei tuolla vuosikymmenellä Lavansaarella ollut kovin korkea muuhun maahan verrattuna, niin katovuosilla oli kuitenkin vaikutusta.

Seuraavaksi vanhimman ikäryhmän, joka käsittää 8 päivää - alle 1 kuukauden ikäiset, kuolleisuudessa on huomattavampia vaihteluita, nousuja ja laskuja, kuin nuorimman ikäryhmän kohdalla. Jyrkimmät kuolleisuuden nousut osuivat 1760- ja 1780-luvuille, jolloin yli kymmenen prosenttia kaikista elävänä syntyneistä kuoli tämän ikäisinä. Kuten 0 - 7 päivän ikäisinä kuolleilla myös tämän ikäisten kohdalla näyttäisi olevan puutteita merkinnöissä 1810-luvulla. Vuosikymmenen 1791-1800 jälkeen nousi tämän ikäisten kuolleisuus yli viiden prosentin vain 1801-1810 välisenä aikana. Sen jälkeen ei 1800-luvulla tämän ryhmän kuolleisuudessa ollut enää jyrkkiä nousuja, ei edes 1860-luvun nälkävuosien aikana. Kirkonkirjoihin merkityt kuolemansyyt olivat koko tarkasteltavana olevan ajanjakson ajan hyvin epämääräisiä kuten nuorimmalla ikäryhmälläkin, eivätkä ne tarkentuneet edes 1800-luvun loppupuolelle tultaessa. Yleensä kuolemansyiksi on merkitty heikkous, vaikea sairaus tai tuntematon tauti. 1700-luvun loppupuolella on kuitenkin joitakin merkintöjä, joissa on tarkempia tietoja tämän ikäryhmän kuolemansyistä. Polttotauti, joka Lavansaarella oli yleinen kuolemansyy koko tarkastelujakson ajan, sekä pistokset on mainittu tämän ikäisten kuolemansyynä muutaman kerran. Nämä molemmat oireet viittaavat suolistoperäisiin sairauksiin. Lisäksi on jokin maininta yskän, keuhkotaudin (14 päivän ikäinen lapsi, 1789) ja tuhkarokon (13 päivän ikäinen lapsi) aiheuttamista kuolemista. Ensimmäisen viikon jälkeen ulkoiset tekijät kuten ripuli ja muut suolistoperäiset taudit sekä hengityselinsairaudet tulivat mukaan kuolemansyihin. Tämän ikäisten lasten suolistosairauksien on katsottu viittaavan siihen, että rintaruokinta ei ole ollut säännöllistä.⁵⁹

⁵⁹ Sundin 1996, s. 102

Kuukaudesta alle kuuden kuukauden ikäisten kuolleisuuden vaihtelut olivat jyrkempiä kuin muilla imeväisryhmillä. Etenkin 1820-luvulle asti oli tämän ikäisten kuolleisuudessa jyrkkiä nousuja ja laskuja. 1770-luvun nousun jälkeen kuolleisuus laski 1790-luvulla noin viiteen prosenttiin, mutta sitten seurasi nousu yli 13 prosenttiin vuosikymmenellä 1801-1810. Kuolleisuus oli vielä 1820-luvullakin yli 11 prosenttia. Kuolemansyitä ei ole merkitty tuolta ajalta, on vain maininta tuntematon tai ilmoittamaton. Muidenkin imeväisryhmien kuolleisuus kohosi tuolla vuosi-kymmenellä noin prosentin verran, mutta tässä ikäryhmässä nousu oli aivan eri tasoa, edelliseen vuosikymmeneen verrattuna lähes yhdeksän prosenttiyksikköä. 1820-luvulla kuolleisuus laski 1790-luvun tasolle, ja kuolleisuuden vaihtelut tasaantuivat, vaikka vielä 1840- ja 1860-luvuilla tämän ryhmän kuolleisuus olikin korkeampi kuin muilla imeväisryhmillä. Yleisimpiä lastentauteja alkoi esiintyä 1-6 kuukauden ikäisten kuolemansyinä. Nuorin rokoon (isorokko) vuonna 1770 kuollut lapsi oli viiden viikon ikäinen. Äidiltä peritty immunitetti haihtui nopeasti. Isorokko ei kuitenkaan ollut Lavan-saarella yleinen lapsilla ensimmäisten elinkuukausien aikana. 1800-luvulla esiintyi kuolemansyinä tuhkarokkoa, tulirokkoa, hinkuyskää, kuumetta, rintatautia ja yskää. Hengityselinsairaudet ja niistä johtuva kuolleisuus lisääntyivät ensimmäisen elinkuukauden jälkeen. Suolistoperäiset oireet kuten polttotauti, pistos ja vatsatauti esiintyivät kautta koko tutkimuskauden kuolemaan johtavina syinä myös tämän ikäryhmän kohdalla. Jopa ensimmäisten elinkuukausien aikana kuolleisuus oli suhteellisen korkea tartuntatauti-epidemioiden aikana.⁶⁰

Pienimmät vaihtelut kuolleisuudessa olivat 6 kuukautta - alle vuoden ikäisillä. Tämän ikäisten kuolleisuuteen vaikuttivat kriisiajat, mutta eivät läheskään niin jyrkästi kuin 1-6 kuukauden ikäisten kohdalla. Kuolleisuus nousi 1760-luvulla ja myöskin 1801-1810, mutta nousu jäi 1-2 prosenttiyksikköön. Korkeimmillaan tämän ikäryhmän kuolleisuus oli 1820-luvulla. 1840-luvun laskun jälkeen tämän

⁶⁰ Sundin 1996, s.98-101

ikäisten kuolleisuus nousi jälleen 1850-luvulla ja oli 1860- ja -70-luvuilla kuuden prosentin tuntumassa, minkä jälkeen se laski selkeästi vuosisadan loppua kohti. Tämän ikäistenkin kuolleiden luetteloihin merkityistä kuolemansyistä suuri osa oli epämääräisiä: heikkous, tuntematon tai vaikea sairaus ja maininta ilmoittamaton. Kuitenkin niistä kuoleman-syistä, joita kuolleiden ja haudattujen luetteloissa on mainittu, selviää, että kuolleisuus tarttuviin lastentauteihin muodosti merkittävän osan tämän ikäryhmän kuolleisuudesta. Äidiltä saadun immuniteetin haihduttua tartunnat isorokko- ja tuhkarokko-epidemioiden aikana muodostuivat tälle imeväisryhmälle tuhoisiksi. Rokkojen lisäksi kuume, yskä, hinkuyskä, rintatauti, vesipöhö ja kuristustauti olivat yleisiä kuolemansyitä. Erityinen kuolemansyy olivat hammaskivut, jotka ilmestyvät merkin-töihin 1800-luvun puolivälin jälkeen. Myös polttotauti oli yleinen kuolemansyy. Muita suolistoperäisiä sairauksia tai niihin viittaavia oireita kuten punatauti, ripuli ja pistos esiintyi.

3 POHDINTA

Tämän tutkimuksen tavoitteena oli selvittää Lavansaaren pienten lasten (ikä alle 4 vuotta) kuolleisuuden vaihtelua 1751-1899 välisenä aikana. Olettamuksena oli, että alle neljän vuoden ikäisiin lapsiin kuului ikäryhmiä, joiden välillä on eroja kuolleisuuden vaihtelussa. Tämän osoittamiseksi tutkimuksessa tarkasteltiin ensin alle neljän vuoden ikäisten lasten kuolleisuutta kokonaisuudessaan ja sen jälkeen lähemmin 1-4-vuotiaiden ja imeväisikäisten kuolleisuuden kehitystä. Koska imeväisikäisten kuolleisuus laski tutkittavan aikajakson kuluessa lähes puoleen, haluttiin tässä tutkimuksessa selvittää eri ikäisten imeväisten kuolleisuuden kehitystä yksityiskohtaisemmin, mitä varten imeväisikäiset jaettiin kuoliniän perusteella pienempiin ryhmiin.

Tarkastelu osoitti, että eri ikäisten lasten kuolleisuus vaihteli eri aikoina eri tavalla. Esimerkiksi vuosina 1801-1810, jolloin kuukaudesta alle kuuden kuukauden ikäisten kuolleisuus kohosi lähes 9 prosenttiyksikköä edellisen vuosikymmenen lukemiin verrattuna, nousi muiden imeväisryhmien kuolleisuus vain lievästi. Tämä osoittaa selvästi, etteivät syyt, jotka aiheuttivat 1-6 kuukauden ikäisten kuolleisuuden nousua, vaikuttaneet merkittävästi muiden ikäryhmien kuolleisuuteen ja vahvistaa aikaisempien tutkimusten päätelmiä, joiden mukaan eri ikäisten lasten kuolemansyyt olivat erilaisia. Tärkein tämän ikäisten kuolleisuuteen vaikuttava tekijä oli imettäminen. Mikä aiheutti imettämisen laiminlyönnin tai mitä muita syitä oli 1-6 kuukauden ikäisten kuolleisuuden jyrkkään nousuun, jää jatkotutkimuksen selvitettäväksi. Kirkonkirjalähteet eivät tähän kysymykseen anna vastausta.

Alle 7 päivän ikäisten kuolleisuuden katsotaan johtuvan raskauden tai synnytyksen aikaisista ongelmista. Lavansaarella 0-7-päivän ikäisinä kuolleiden lukumäärä vuosikymmenessä oli pieni. 1840-luvulla esiintyneen 0-7 päivän ikäisten aikaisempaa korkeamman kuolleisuuden lähempi tarkastelu osoitti, että

kymmenen lapsen joukossa oli neljä kaksosta, kaksi ensimmäistä lasta sekä kaksi, joiden synnyttämisen aikana ilmeni ongelmia. Sekä kaksoset että ensimmäiset lapset kuuluivat riskiryhmään.⁶¹ Jatkotutkimuksen tehtäväksi jää selvittää, oliko kuolleisuuden kohoamisessa kysymys sattumasta vai vaikuttivatko siihen jotkin muut syyt, esimerkiksi odottavien äitien toimeentulon heikkeneminen.

Lavansaarella ikäryhmien, 8 päivää-1 kuukausi ja 1-6 kuukautta, kuolleisuudessa olivat jyrkimmät vaihtelut ja ne osuivat 1700-luvulle. Kuitenkaan näiden ryhmien kuolleisuushuiput eivät osuneet samoille vuosikymmenille. 1-6 kuukauden ikäisten kuolleisuus oli korkea 1770-luvulla sekä erittäin korkea 1800-luvun kahdella ensimmäisellä vuosikymmenellä. 8 päivää - alle kuukauden ikäisten kuolleisuus oli korkea 1760- ja 1780-luvuilla. Molemmissa ryhmissä kuolleisuuden vaihtelut tasaantuivat 1800-luvun alkupuolelta lähtien. Näiden kahden ikäryhmän kuolleisuus-huippujen eriaikaisuus, herättää kysymyksiä, sillä molempien kohdalla imetys oli tärkein kuolleisuuteen vaikuttava tekijä, eikä heidän kohdallaan tartuntataudeilla vielä ollut merkittävää osuutta. Vastausta ei löydy kuolleiden ja haudattujen luetteloista, sillä niiden kuolemansyymerkinnät ovat suurelta osin epämääräisiä. Ruotsalainen tutkija Sundin havaitsi, että tartuntatautien aikana kuolleisuus oli korkea jopa ensimmäisten elin-kuukausien aikana.⁶² Tästä voisi olla kysymys Lavansaarellakin, sillä esimerkiksi vuosina 1801-1810 Lavansaarella esiintyi rokkojen lisäksi epidemiaa polttotautia, joka tarttui myös alle kuuden kuukauden ikäisiin lapsiin. Tämän lisäksi kuolemansyiksi on mainittu kirjava joukko tauteja tai tautioireita. On myös mainintoja riisitautiin kuolleista. Tämä kaikki viittaa siihen, että tuolla vuosikymmenellä jo ennen sotavuosia, toimeentulo on huonontunut, mikä on johtanut väestön vastustuskyvyn heikkenemiseen. Jatkotutkimuksen selvitettäväksi jää, oliko esimerkiksi ravintotilanteen heikkeneminen johtanut imetyksen

⁶¹ Moring 1994, s. 118-119, Lutz 1984

⁶² Sundin 1996, s. 98

vähentymiseen ja sitä kautta lasten sairastavuuden ja kuolleisuuden lisääntymiseen vai onko löydettävissä muita syitä. Imettämisen säännöllisyyttä on mahdollista tutkia selvittämällä kuolleisuuden jakautumista kuukausittain, sillä olosuhteissa, joissa vallitsee huono hygienia ja lapsia ruokitaan lehmän maidolla, kuolleisuus kohoaa kesäkuukausina.

Kuudesta kuukaudesta alle vuoden ikäisten kuolleisuus pysyi suhteellisen tasaisena aina 1820-luvulle asti. 1840-luvun laskun jälkeen kuolleisuus kohosi ja oli aikaisempaa korkeampi kolmen vuosikymmenen ajan, kunnes laski selkeästi 1870-luvulta lähtien. Tämän ikäisten kuolleisuudessa näyttelivät tarttuvat lastentaudit jo tärkeää osaa ja niiden esiintymien epidemioina selittää suurelta osin kuolleisuuden vaihteluja. Samoin on laita 1-4-vuotiaiden kuolleisuuden kohdalla. Tarttuvat lastentaudit olivat heidän kohdallaan merkittävin kuolleisuuteen vaikuttava tekijä. Imetysajat olivat 1860-luvulle asti olleet pitkiä. Olosuhteissa, joissa vallitsi korkea altistuminen tartuntataudeille, imetyksestä vieroittamisen aika johti usein korkeaan kuolleisuuteen. 6 kuukauden - 1 vuoden ikäisten kuolleisuuden tasaisuuden perusteella voisi päätellä, että imettäminen jatkui yli vuoden ikäiseksi ainakin 1850-60-luvuille asti.

Tutkimus osoitti myös, että eri ikäisten lasten kuolleisuuden aleneminen tapahtui eri aikoina. Alle 4-vuotiaiden kuolleisuuden pitkäaikainen kehitys osoittaa karkeasti arvioiden vajaan neljäsosan laskua 1751-1899 välisenä aikana. Kuolleisuus jakautui kolmeen vaiheeseen: 1700-luvun puolivälistä 1790-luvulle, 1800-luvun ensimmäiseltä vuosikymmeneltä 1830-luvun loppuun ja 1840-luvulta 1800-luvun loppuun. Ensimmäistä vaihetta leimasi aleneva kuolleisuus, toista kaksi korkeaa kuolleisuushuippua ja kolmatta ensin kolme vuosikymmentä kestävä tasaisen ja sitten selkeästi laskevan kuolleisuuden vaihe. Alle 4-vuotiaiden kuolleisuuden selkeä aleneminen alkoi Lavansaarella 1870-luvulta lähtien.

Vuodesta alle neljän vuoden ikäisten kuolleisuuden pitkäaikaista kehitystä kuvaava graafi osoitti hienoista kuolleisuuden nousua tarkastelukauden aikana. Tämän ryhmän kuolleisuus jakautuu noin 50 vuoden pituisiin vaiheisiin: 1700-luvun loppupuolisko sekä 1800-luvun alku- ja loppupuolisko. Ensimmäistä vaihetta leimasivat matalahkot nousut ja laskut, mutta kuolleisuuden taso oli matalampi kuin seuraavassa vaiheessa. 1800-luvun alkupuolelle ominaisia olivat korkeat kuolleisuushuiput ja selkeästi korkeampi kuolleisuuden taso kuin 1700-luvun lopulla. 1850-luvulta alkoi hyvin selkeä kuolleisuuden lasku.

Imeväisikäisten pitkäaikaista kuolleisuuden kehitystä kuvaava graafi osoittaa, että kuolleisuus laski karkeasti arvioiden lähes puoleen 1700-luvun puolivälistä 1800-luvun loppuun. 1800-luvun ensimmäisellä vuosikymmenellä oli jyrkkä nousu tämän ryhmän kuolleisuudessa, mutta aleneva suunta jatkui sen jälkeen. Kun sitten tutkitaan imeväisikäisten kuolleisuutta yksityiskohtaisemmin, huomataan että eri ikäryhmien kuolleisuus kehittyi eri tavoin. Sekä 0-7-päivän että 6 kuukaudesta alle vuoden ikäisten kuolleisuus osoittivat molemmat hienoista nousua, kun taas 8 päivästä alle kuukauden ikäisten sekä 1-6-kuukauden ikäisten kuolleisuus laskivat selvästi. Nuorimman ikäryhmän (0-7 päivän ikäiset) kuolleisuus laski tasaisesti 1700-luvun puolivälistä aina vuosisadan loppuun asti. Sen jälkeen tämän ryhmän kuolleisuuden tarkastelua haittaavat vuosien 1811-1830 puuttuvat tiedot, mutta kuolleisuus näyttäisi olevan nousussa 1800-luvun alusta aina 1840-luvulle asti, minkä jälkeen tasaisempien vaihteluiden jakso. Tämän ryhmän kuolleisuus näyttäisi olevan hienoisessa nousussa aina 1790-luvulta alkaen.

Kahdeksasta päivästä alle kuukauden ikäisten kuolleisuudessa voi havaita kolme vaihetta. Ensin 1700-luvun puolivälistä 1790-luvulle asti oli korkean kuolleisuuden ja jyrkkien vaihteluiden jakso. Sen jälkeen kuolleisuus laski matalammalle tasolle, mutta vaihteluiden arviointia haittaavat puutteelliset tiedot. 1840-luvulta lähtien alkoi vähäisten vaihteluiden vaihe, jossa kuolleisuus pysytteli 1-2 prosentin tuntumassa. Tämän ryhmän kuolleisuus alkoi laskea

1790-luvulta lähtien.

Yhdestä kuuden kuukauden ikäisten kuolleisuudessa oli jyrkkiä nousuja ja laskuja aina 1820-luvulle asti. 1700-luvun lopun suotuisan kehityksen jälkeen seurasi jyrkkä kuolleisuushuippu vuosikymmenellä 1801-1810. Vaikka kuolleisuus vielä 1810-luvulla oli korkea, voi tämän ikäisten kuolleisuuden laskun katsoa alkaneen tuolloin. Myös tässä ryhmässä kuolleisuuden vaihtelut tasaantuivat 1820-luvulta alkaen.

Vanhimman imeväisryhmän (6 kk -1 v) kuolleisuuden vaihtelu oli tasaisempaa kuin muilla ikäryhmillä lukuun ottamatta alle 7 päivän ikäisinä kuolleita. Tämän ikäryhmän kuolleisuus laski 1760-luvulta aina vuosisadan loppuun. 1800-luvun alusta kuolleisuus alkoi kohota ja nousua jatkui aina 1860-, 1870-luvuille, minkä jälkeen 1880-luvulla seurasi selkeä kuolleisuuden lasku.

Tämä tutkimus osoitti, että imeväisikäisten kuolleisuuden lasku osui lähinnä yli viikon ikäisistä alle kuuden kuukauden ikäisiin. Tämän tuloksen perusteella voi tehdä johtopäätöksen, että rintaruokinnasta, jota tutkijat yleisesti pitävät tärkeimpänä alle kuuden kuukauden ikäisten lasten kuolleisuuteen vaikuttavana tekijänä, tuli Lavansaarella entistä säännönmukaisempaa ja että sitä jatkettiin ainakin yli kuuden kuukauden ajan lapsen syntymästä.

Karjala-tietokannan tallentamat Lavansaaren kirkonkirjat soveltuvat erinomaisesti numerotiedon hakemiseen ja lajitteluun. Karjala-tietokannassa eri lähdetyypeistä on muodostettu omat tieto-kannat. Tästä on seurauksena, että kaikkia henkilöstä kirjattuja tietoja ei saa yhdestä ja samasta tietokannasta hakemalla esimerkiksi nimen ja syntymävuoden perusteella, vaan on käytävä läpi kaikki eri lähdetyyppien muodostamat tietokannat. Lasten- ja rippikirjojen muodostamissa tietokannoissa henkilöstä saattaa olla useita eri merkintöjä, mutta ne ovat eri henkilönumerolla ja niiden käyttö on sen tähden aikaa vievää. Nimien normalisointi on jäänyt keskeneräiseksi, sillä samasta henkilöstä on

käytetty nimen eri muotoja. Tallennettujen kirkonkirjojen käyttö edellyttää, että tutkija tuntee kirkonkirja-aineiston ja tietää, minkälaisia tietoja niistä voi löytää. Numerotiedon hakuun ei tarvita tietokantaohjelmiston hyvää hallintaa, mutta jos tutkimus edellyttää monipuolisemman tiedon hakua, silloin ohjelmistoon on perehdyttävä perusteellisemmin.

LÄHTEET

Arkistolähteet

Karjala-tietokanta, Lavansaaren kirkonkirjat:

Rippikirjat 1744-1899

Lastenkirjat 1800-1884

Syntyneiden ja kastettujen luettelot 1750-1899

Kuolleiden ja haudattujen luettelot 1750-1899

Karjala-tietokanta, Lavansaaren seurakuntatiedot

Kirjallisuus

ALANEN 1963. Alanen, Aulis, Suomen historia vapaudenajalla.

Suomen historia. IX osa. Helsinki 1963.

FORSIUS 1986. Forsius, Arno, Epidemioista, väestötapahtumista ja kuolinsyistä.

Suomen sukututkimusseuran julkaisuja 40. Suomen sukututkimusseura. Helsinki 1986.

HANNULA 1947. Hannula, Vaalimo, Lavansaari. Historian pohjalle rakentuva kertomus "Suomenlahden selkäsaaresta". Forssa 1947.

HELSTI 1996. Helsti, Hilikka, Tabun rikkojista raatajaäiteihin. Yksinsynnyttäjien tarinat suomalaisten maalaisnaisten kertomusperinteessä. Elektroloristi 1/1996.

JUTIKKALA 1968. Jutikkala, Eino, Suomen talous- ja sosiaalhistoria. Historian korkeakoulu 2. Toimittanut Eino Jutikkala. Porvoo 1968.

JUTIKKALA 1993. Jutikkala, Eino, How many lived to maturity? 1749-1865. Yearbook of population research in Finland, XXXI. Helsinki 1993, s. 16-24.

KERÄNEN 1994. Keränen, Jorma, Nälkämaan laulu. Kainuu esimerkkinä nälkäalueesta. "Pane leipään puolet petäjäistä". Nälkä- ja pulavuodet Suomen historiassa. Jyväskylän yliopisto. Historian laitos. Suomen

- historian julkaisuja. Jyväskylä 1994.
- KORHONEN 1996. Korhonen, Martti, Suomen ulkosaarten historia. Suomenlahden ulkosaaret Lavansaari, Seiskari, Suursaari, Tytärsaari. Toimittaneet Risto Hamari, Martti Korhonen, Timo Miettinen, Ilmar Talve. Suomalaisen Kirjallisuuden Seuran Toimituksia 630. Jyväskylä 1996.
- KOSKINEN, MARTELIN 1994. Koskinen, Seppo ja Martelin, Tuija, Kuolleisuus. Teoksessa Suomen väestö. Toimittaneet Seppo Koskinen, Tuija Martelin, Irma-Leena Notkola, Veijo Notkola, Kari Pitkänen. Hämeenlinna 1994.
- LUTZ 1984. Lutz, W. The changing nature of the link between infant mortality and fertility in Finland 1776-1978. Yearbook of population research in Finland, XXII. Helsinki 1984, s. 26-45.
- MORING 1994. Moring, Beatrice, Skärgårdsbor. Hushåll, familj och demografi i finländsk kustbygd på 1600-, 1700- och 1800-talen. Helsingfors 1994.
- PITKÄNEN 1994. Pitkänen, Kari, Suomen väestön historialliset kehityslinjat. Teoksessa Suomen väestö. Toimittaneet Seppo Koskinen, Tuija Martelin, Irma-Leena Notkola, Veijo Notkola, Kari Pitkänen. Hämeenlinna 1994.
- PITKÄNEN 1993. Pitkänen, Kari, Deprivation and disease. Mortality during the great finnish famine of the 1860s. Publications of the finnish demographic society, 14, Helsinki 1993.
- PITKÄNEN 1988. Pitkänen, Kari, Väestötutkimus ja yhteiskunta. Suomalaisen väestötutkimuksen historia 1700-luvulta noin vuoteen 1950. Suomen väestötieteen yhdistyksen julkaisuja, 11, Helsinki 1988.
- SUOMENMAAN VIRALLINEN TILASTO 1871. VI. Väkiluvun-Tilastoa. Toinen vihko. Syntyneet, Vihityt ja Kuolleet vuosina 1865-68 ynnä katsahdus väkiluvun muutoksiin vuodesta 1812 alkaen.
- SUNDIN 1996. Sundin, Jan, Child mortality and causes of death in a swedish city, 1750-1860. Historical Methods. Summer 1996. Volume 29. Number 3, 93-106.
- TALVE 1996. Talve, Ilmari, Suomenlahden ulkosaarten kansankulttuuri 1800-luvun loppu puolelta talvisotaan. Suomenlahden ulkosaaret Lavansaari,

- Seiskari, Suursaari, Tytärsaari. Toimittaneet Risto Hamari, Martti Korhonen, Timo Miettinen, Ilmar Talve. Suomalaisen Kirjallisuuden Seuran Toimituksia 630. Jyväskylä 1996.
- TURPEINEN 1979a. Turpeinen, O. Infant mortality in Finland 1749-1865. The Scandinavian economic history review, XXVII, 1. Keuruu 1979.
- TURPEINEN 1979b. Turpeinen, O. Monthly mortality in Finland in 1751-1806. Yearbook of population research in Finland, XVII. Helsinki 1979, s. 58-73.
- TURPEINEN 1984. Turpeinen, O. The percentage of deaths under one year of age of all deaths in Finland in 1749-1865. Yearbook of population research in Finland, XXII. Helsinki 1984, s. 46-54.
- TURPEINEN 1986. Turpeinen, O. Näлкä vai tauti tappoi? Kauhunvuodet 1866-1868. Historiallisia tutkimuksia, 136. Helsinki 1986.
- WRIGLEY, SCHOFIELD 1981. Wrigley, E, Schofield, R. The Population History of England 1541-1871. A Reconstruction. London 1981.

LIITTEET

- Liite 1. Lavansaaren kuolleet ja syntyneet vuosittain 1751-1899. Alle neljän vuoden ikäisinä kuolleiden lasten ikäryhmiin jaotellut lukumäärät vuosittain ja kymmenvuosittain sekä ikäryhmien osuus prosentteina.

LIITE I Lavansaaren kuolleet ja syntyneet vuosittain 1751-1899. Alle neljän vuoden ikäisinä kuolleiden lasten ikäryhmiin jaotellut lukumäärät vuosittain ja kymmenvuositain sekä ikäryhmien osuus prosentteina.

* Merkintä kuolemasta kuolleiden luettelossa, mutta ei ole selvyttä henkilön syntymävuodesta eikä iästä, mistä syystä ei ole lisätty syntyneihin.

** Kuolleetna syntyneet

Vuosi	Syntyneet	Kuolleet	Kuolleet 0 pv-7 pv	%	Kuolleet 8 pv-1 kk	%	Kuolleet 1-6 kk	%	Kuolleet 6 kk-1 v	%	Kuolleet 0 pv - 1 v	%	Kuolleet 1-4 v	%	Kuolleet 0 pv- 4 v	%	Huomautukset
1750	12	3											2		2		
1751	12	6	1		1								1		3		
1752	12	20	1		1				1				8		12		1*
1753	21	18			4				1				5		11		
1754	15	8	1						2						4		
1755	13	9			1										5		
1756	20	21			2				1				7		10		
1757	18	12							1				4		6		1*
1758	12	8			2								1		4		
1759	23	11			1								1		4		
1760	18	11	1		1								2		5		
	164	124	4	2.44	12	7.32	13	7.93	6	3.66	35	21.34	29	17.68	64	39.02	

Vuosi	Syntyneet	Kuolleet	Kuolleet <=7 pv	Kuolleet 8 pv-1 kk %	Kuolleet 1-6 kk %	Kuolleet 6 kk-1 v %	Kuolleet 0 pv - 1 v	Kuolleet 1-4 v %	Kuolleet 0 pv- 4 v	Huomautukset						
1761	13	7	1		1				2							
1762	18	5	1	2	1	1			5							
1763	13	21	4		1	1			9							
1764	18	6	3						4							
1765	18	7														
1766	17	12			1				7							
1767	19	13	1		2	2			8							
1768	20	14							4							
1769	20	12		6	1	1			9							
1770	19	40	1	1	2	4			14	1*						
	175	137	4	2.29 %	18	10.29 %	9	5.14 %	9	5.14	40	22.86 %	22	12.57	62	35.43

Vuosi	Syntyneet	Kuolleet	Kuolleet 0-7 pv	%	Kuolleet 8 pv-1 kk	%	Kuolleet 1-6 kk	%	Kuolleet 6 kk-1 v	%	Kuolleet 0 pv - 1 v	%	Kuolleet 1-4 v	%	Kuolleet 0 pv- 4 v	Huomautukset
1771	19	10			1		1					1		2	1*	
1772	9	28			2		1					7		10		
1773	20	8			4		1							5		
1774	11	4	2									2		4	2*	
1775	25	10	1		3		2		1			1		8		
1776	23	14			1				2			4		7		
1777	14	27			1		4		1			6		12		
1778	20	5			1									1	2*	
1779	12	29					3		2			6		11	1*	
1780	17	5					3					1		4		
	170	140	3	1.76	10	5.88	16	9.41	7	4.12	36	21.18	28	16.47	64	37.65

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv- 4 v	Huomautukset				
1781	17	6	1						1					
1782	13	6	1	1				1	3					
1783	16	42	4	1	2			11	18	8* rokko				
1784	20	3	1		1			1	3					
1785	15	9		1				1	2	1**				
1786	15	3			1				1					
1787	13	11	4					1	5					
1788	16	7	2	1				1	6					
1789	12	12	1						1					
1790	11	16	1			1		3	5					
	148	115	2	1.35	4	2.70	5	3.38	26	17.57	19	12.84	45	30.41

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1 - 6 kk	Kuolleet % 6 kk -1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1 v - 4	Kuolleet 0 pv - 4 v	Huomautukset						
1791	21	11	1	1	1	1		2	5							
1792	20	5	1	1	1	1			3	1*						
1793	19	6						3	3							
1794	11	7		1	1			3	4	1**						
1795	22	10		3				3	6	1**						
1796	20	17	1		2			8	11	1**						
1797	24	12		1		3		2	6							
1798	20	14		2	2			1	5							
1799	25	10	1		1			3	5							
1800	13	9			1	1		4	6	1***						
	195	101	2	1.03	8	4.10	9	4.62	6	3.08	25	12.82	29	14.87	54	27.69

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1 v - 4	Kuolleet 0 pv - 4 v	Huomautukset				
1801	21	16	1		1	1		7	10					
1802	16	6							0	1**				
1803	21	39	1	2	7	1		11	22					
1804	20	30	1	3	4			3	11					
1805	15	32		1				2	3					
1806	15	20		1	3	2		5	11					
1807	21	7			1	1		1	3					
1808	12	22		2		2		3	7					
1809	10	6			2				2					
1810	10	14			3			4	7					
	161	192	3	1.86	21	13.04	7	4.35	41	25.47	35	21.74	76	47.20

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet 8 pv-1 kk %	Kuolleet 1-6 kk %	Kuolleet 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet 1-4 v %	Kuolleet 0 pv- 4 v	Huomautukset				
1811	20	10			1	1		4	6	I**				
1812	15	13			2	2		6	10					
1813	16	4			1			1	2					
1814	22	19			1	1		1	3					
1815	13	23			7			12	19					
1816	10	5			1				1					
1817	14	6			1			4	5					
1818	23	5			2	3			5					
1819	29	21			3			3	7	I**				
1820	21	14			2			3	5					
	183	120	0	0.00	21	11.48	7	3.83	29	15.85	34	18.58	63	34.43

Vuosi	Syntyneet	Kuolleet	Kuolleet 0-7 pv	Kuolleet % 8 pv- 1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv- 4 v	Huomaukset						
1821	27	9				2		3	5							
1822	19	13				4		2	6							
1823	24	14				1		4	5	2**						
1824	18	7		1		2		3	6	1**						
1825	29	13		3				1	4	1**						
1826	17	8		1				4	5	1**						
1827	27	12		2		1		1	6	3**						
1828	24	13				2		4	6	1**						
1829	14	21		1				8	9	1*						
1830	16	17			2			4	6							
	215	127	0	0.00	3	1.40	9	4.19	12	5.58	24	11.16	34	15.81	58	26.98

Vuosi	Syntyneet	Kuolleet	Kuolleet 0-7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1 kk-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv - 4 v	Huomautukset						
1831	18	18	2	2				2	6							
1832	16	46		1	1			14	16							
1833	15	24	1	1				7	9							
1834	15	9		2	1			2	5	2**						
1835	16	15	1	2	1			4	9							
1836	17	16		1	1			8	10							
1837	12	18	1		1			7	9							
1838	16	6		1				1	2							
1839	18	11	1		2			4	8							
1840	20	8		1	1				3	1**						
	163	171	4	2.45	5	3.07	11	6.75	8	4.91	28	17.18	49	30.06	77	47.24

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet 8 pv-1 kk	%	Kuolleet 1-6 kk	%	Kuolleet 6 kk-1 v	%	Kuolleet 0 pv - 1 v	%	Kuolleet 1-4 v	%	Kuolleet 0 pv- 4 v	Huomautukset	
1841	12	10		2		1		2				2		5		
1842	21	17		2		1		1				1		4		
1843	16	9	2	1								2		5		
1844	14	13	1			2						2		5		
1845	28	14		1								3		4		
1846	19	7		1		2								3	1**	
1847	28	18	3	1		1						1		6	2**	
1848	25	43	1	1		2		1				22		27		
1849	32	20	2	4		2		1				3		12	2**	
1850	21	9	1			3		2						6		
	216	160	10	9	4.63	16	4.17	6	7.41	2.78	41	18.98	36	16.67	77	35.65

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet 8 pv-1 kk %	Kuolleet 1-6 kk %	Kuolleet 6 kk-1 v %	Kuolleet 0 pv - 1 v	Kuolleet 1-4 v %	Kuolleet 0 pv - 4 v	Huomautukset					
1851	19	10						5	5						
1852	24	15	1		2	1		1	5	1**					
1853	24	8		1					1						
1854	19	45		1	1	3		22	27						
1855	27	10	1		1	2		1	5	1**					
1856	19	12	1		2				3						
1857	14	27				1		4	5	1**					
1858	20	15			1	1		1	3						
1859	21	13				1		5	6						
1860	19	24			4	2		7	13	1**					
	206	179	3	1.46	11	11	5.34	11	5.34	27	13.11	46	22.33	73	35.44

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv- 4 v	Huomautukset	
1861	18	10			2	1		2	5	1**	
1862	20	8			1	1		1	3		
1863	15	13			1			1	2	1**	
1864	29	26	2			3		10	15	1**	
1865	26	22			3	2		6	11	1**	
1866	18	20		1	1	2		3	7		
1867	21	28	1		2			9	12		
1868	23	11	1		2				3	2**	
1869	15	19				2		5	7	1*2**	
1870	23	20	1	1		2		4	8		
	208	177	5	2	12	13	6.25	15.38	41	19.71	35.10

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv - 4 v	Huomautukset						
1871	20	12		1	3			1	5							
1872	19	11			2	1			3	I**						
1873	26	14				1		1	2							
1874	27	23	3	1	3	2		5	14							
1875	34	22	1	1	4			2	8							
1876	26	25			2	2		9	13	I**						
1877	33	13		1	2	1		2	6							
1878	26	7		2	1				3							
1879	28	13			2	2		1	5	I**						
1880	26	40				7		22	30							
	265	180	4	1.51	6	2.26	19	7.17	16	6.04	45	16.98	43	16.23	88	33.21

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv- 4 v	Huomaukset			
1881	31	8						5	5	1*			
1882	34	13						6	6	2*			
1883	29	16	1		2	1		7	11	1*			
1884	24	13			3			1	4				
1885	35	12		2	3	1			6	1**			
1886	24	27	1			5		12	18				
1887	36	25	1	2	2	4		5	14	2**			
1888	40	11		2	2			2	4				
1889	29	17	1		1	3			5				
1890	45	16	1	2				6	9				
	327	158	5	1.53	1.83	14	4.28	38	11.62	44	13.46	82	25.08

Vuosi	Syntyneet	Kuolleet	Kuolleet 0=7 pv	Kuolleet % 8 pv-1 kk	Kuolleet % 1-6 kk	Kuolleet % 6 kk-1 v	Kuolleet 0 pv - 1 v	Kuolleet % 1-4 v	Kuolleet 0 pv - 4 v	Huomautukset						
1891	32	15	1		5	2		2	10							
1892	34	27			2	1		11	14							
1893	43	9		1		1		2	4	1**						
1894	36	16	2		2			1	5	2*						
1895	32	16	1		3	1		2	8							
1896	38	16	2		4				6	2*						
1897	34	29		2	2	2		11	17							
1898	54	31	4	1	2	1		3	11	1*						
1899	38	17		1		1		3	5							
	341	176	10	2.93	6	1.76	20	5.87	9	2.64	45	13.20	35	10.26	80	23.46