

De fødte, de gifte og de døde - demografi i Tromsø

Gunnar Thorvaldsen

Denne historiografiske artikkelen oppsummerer databehandlingen av historiske mikrodata med opprinnelse i historiske kilder av protokolltypen ved Universitetet i Tromsø. Det viktigste kildemateriale er folketellinger og kirkebøker. Hovedtemaer er transkripsjon, koding, postkobling og brukergrensesnitt, samt forskningsresultatene. Med omsyn til konferansetemaet «Lokalhistorie og globalhistorie» er det fristende å eksemplifisere forskningen med henholdsvis artikkelen «Befolkningsutvikling i Tromsø 1866 til 1900» og boka *Censuses and Census Takers – a Global History* (Thorvaldsen, 1984, 2018).

De historiske instituttene i Oslo og Bergen startet databehandlingen av historiske kilder i Norge, men etter opprettelsen av Registreringsentral for historiske data (RHD) ved UiT – Norges arktiske universitet i Tromsø fra 1978 ble dette hovedsentret for forskningsorientert transkribering. Siden 1985 har RHD vært en permanent avdeling innenfor Fakultet for samfunnsvitenskap og humaniora. Opprettelsen kom etter lobbying fra lokale og regionale politikere og administratorer, blant annet overfor Stortinget, og i møte med denne entusiasmen ble også administrasjonen ved UiT gode støttespillere. RHD betjener forskere, lærere, studenter og slektsforskere over hele landet, og hovedmålet er et nasjonalt befolkningsregister for 1800- og 1900-tallet. RHD publiserte trykte versjoner av originalene fra folketellingene 1865, 1875, 1900 og 1910 samt kirkebøker sammen med alfabetiske indekser, supplert med digitale versjoner ifølge en nasjonal standard for dataregistrering og datadistribusjon (Nygaard, 1995). De kodede versjonene av folketellingene ble opprinnelig distribuert på disketter sammen med en valgfri statistisk programpakke. I samarbeid med Riksarkivets Digitalarkiv ble folketellingsavskriftene utvidet fra enkeltkommuner til nasjonale utgave. Transkripsjon av kirkebokslister var en mer arbeidskrevende virksomhet enn folketellingene, siden håndskriften i disse kildene ofte er gotisk, noe som gjorde den geografiske dekingen mer begrenset. Papirkopier av kildematerialet ble gjort tilgjengelig fra Riksarkivet eller Statsarkivene, men nå hovedsakelig som skannede bilder via Internett i Digitalarkivet.

De transkriberte folketellingene er tilgjengelige i en ordrett fulltekstversjon så vel som en kodet versjon. Sistnevnte er standardisert for å produsere statistiske aggregater og inneholder tallkoder for yrke, familiestatus og sogn eller fødekommune. I tillegg til statistikk er kodene

også nyttige i prosessen med lenking på individnivå, fordi informasjonen om familiestatus og fødested kan brukes på en mer enhetlig måte. Et halvautomatisk kodingsprogram gjør det mulig å supplere tekstvariablene med numeriske koder, og lage datafiler for aggregering vha statistiske programvarepakker. Igjen er konsistens på tvers av tid og rom vektlagt slik at resultater fra tverrsnittsanalyser av flere folketellinger kan være sammenlignbare (Thorvaldsen, 1994).

I tillegg standardiseres personnavn og stedsnavn for å hjelpe forskere som ønsker å identifisere personer i flere kilder. Jordbruksdata om husdyr og utsed finnes i folketellingene 1865 og 1875 og kan knyttes til individdataene. Matrikkelskattelistingene fra 1838 og 1886 ble transkribert med optisk (OCR – optical character recognition) eller ved manuell transkripsjon til en database som også er tilgjengelig på Internett. Også annet trykt kildemateriale har blitt «skrevet av» ved å anvende datateknikker for optisk lesing. Det er satt i gang arbeid med indekser til skifteregistrene, som lar forskere finne fram til originalene eller mikrofilmkopiene.

Transkribering

Grunnprinsippet ved transkribering av folketellinger, kirkebøker og andre kilder har vært å kopiere innholdet så ordrett som mulig. For å oppnå dette må man korrekturlese avskriftene. Vi fant at det er bedre enn dobbel transkripsjon av ergonomiske grunner. Vi sorterer materialet for å oppdage sjeldne og muligens feilaktige varianter, og vi programmerer datamaskinen til å lete etter ulogiske kombinasjoner av egenskaper, for eksempel tenåringsenker. Ved lenking på individnivå oppdager vi motstridende informasjon ved å kombinere data fra flere kilder.

Det finnes unntak fra regelen om ordrett transkripsjon. Ett er innføringen av et skille mellom fornavn og etternavn, som ble skrevet inn i samme felt i originalkilda. Folketellerne brukte dessuten et spesialtegn for å indikere gjentakelse av informasjon fra forrige person, og dette erstattes av den relevante informasjonen. Hvis informasjonen om en person er motstridende, for eksempel når en "datter" er merket som mann i kjønnsfeltet, skal avskriveren rette det som åpenbart er en feil i kilda basert på navn og annen informasjon, og flagge hva som er rettet i kommentarfeltet. Slike kontroller er også innebygd i transkripsjonsappene. For å oppsummere: transkripsjonsregler er et kompromiss mellom å lage en ordrett kopi og å forbedre brukervennligheten til den resulterende databasen (Thorvaldsen et al., 2015).

Forskningseksempel 1: Sergei Sergel's field research among the Sami

Mens vi var ved Universitetet i Aberdeen og fokuserte på polarområdene, dukket det opp en interessant publikasjon av den russiske student-etnografen Sergei Sergel, som tilbrakte flere måneder sammen med norske samer i 1907-1908. Han sluttet seg til den nomadiske Sara-familien, og beskrev hvert familiemedlem, deres næringer, relasjoner, livsstil osv. på russisk. Hva som skjedde med denne familien før og etter at Sergei Sergel' møtte dem, fant vi i de transkriberte norske folketellingene og kirkebøkene (Glavatskaya & Thorvaldsen, 2013). Denne forskningen viste igjen viktigheten av nominative folketellinger og kirkebøker og deres potensiale for å studere etniske og andre minoriteter, en forskning som ofte blir mer kvalitativ enn kvantitativ.

Standardisering og koding av kildematerialet

Hovedformålet med kodingen er å lage statistikk på grunnlag av opplysningene i folketellingene. Det er flere grunner til å ikke bruke de publiserte aggregatene fra hver folketelling. Grensene mellom kommunene og systemet for kategorisering av yrker etc. ble ofte endret, noe som gjør historiske sammenligninger over tid vanskelig. En annen grunn er at variablene i publisert statistikk er kombinert på gruppenivå, ikke på individnivå, noe som øker risikoen for å introdusere økologiske feilslutninger (Langholm, 1976). Et tredje motiv er at de standardiserte kodene gjør det lettere å følge personer fra kilde til kilde ved lenking, for eksempel er kode 0724 for fødested mer konsistent enn endrede kommunenavn som Sandar og Sandehherred. Mens enkle felt som kjønn, sivilstand og alder krever lite standardisering og få regler, er koding av yrker, familiestatus, fødesteder og etnisitet komplekse oppgaver (Thorvaldsen, 1994). Relasjoner mellom personer i samme husholdning eller familie kan analyseres fordi den kodede versjonen av folketellingene er utstyrt med spesialkonstruerte variabler. For eksempel gir en relasjonsvariabel informasjon om hvem som var ektefellen fordi den gjensidig inneholder ID-numrene til ektefellene til gifte personer i hver husholdning. Ved å bruke disse ID-numrene blir det eksplisitt hvilke ektemenn og koner som hørte sammen, selv om det var mer enn ett par i en husholdning. Tilsvarende variabler «peker» fra barna til hver av foreldrene. En advarsel: I store, kompliserte husholdninger kan dataprogrammet som lager disse variablene introdusere feilaktige relasjoner, for eksempel knytte barn til feil foreldre.

Forskningseksempel 2: Nedgangen i tjenertallene

Et eksempel på bruk av husholdnings- og yrkesvariablene i forskning er revurderingen av antall hushjelper i tiårene omkring år 1900. Det ble hevdet i Aschehougs norgeshistorie, bind 9 i samsvar med arbeid fra andre historikere, at antall hushjelper gikk betydelig ned mot slutten av 1800-tallet. Dette er ikke forenlig med analyse av det primære kildematerialet, i form av transkriberte og kodede folketellingsskjemaer på individnivå fra 1865, 1875 og 1900. Funnet av en nedgang i antall tjenere bygget på en sammenligning av de numeriske folketellingene fra 1845 og 1855 samt aggregater fra de senere nominative folketellingene. Dette er imidlertid sekundært kildemateriale og yrkeskategoriene som ble brukt i disse folketellingene endret seg mye. Sammensetningen av tjenerne i arbeidsstyrken utviklet seg i samsvar med mer generelle strukturelle endringer. De migrerte fra landlige til urbane steder og jobbet mindre i landbruket og mer for arbeidsgivere i sekundær og tertiær økonomisk sektor. Relasjonene mellom husstandsmedlemmer i de nominative folketellingslistene lot oss skille mellom det synkende antallet tjenere som var gårdsarbeidere, og hushjelpene hvis antall ikke gikk ned før andre verdenskrig (Thorvaldsen, 2008).

Standardisering av navn

Standardisering av navn er nyttig både for kvantitative analyser av navnefrekvenser og for lenking av individdata. Både for- og etternavn er standardisert i samarbeid med professor i nordiske språk Gulbrand Alhaug (Alhaug, 2005), i et prosjekt finansiert av Forskningsrådet. Prosjektet utarbeidet dermed en liste over personnavn der varianter av stavemåter og nærliggende språklige varianter ble standardisert til samme standardiserte markørnavn. For eksempel ble Kristian med K og Ch kodet til markørnavnet Kristian, og Fredrek er standardisert til Fredrik, mens forskjellen mellom Anne og Anna ble bevart siden variantene uttales ulikt. Siden standardiseringen er ganske konservativ, er det nødvendig å bruke et tilleggsprogram som beregner den grafemiske avstanden mellom navneformer, ved hjelp av en spesiell algoritme (Winkler, 1990). Dermed vil personregistre med lignende navneformer lenkes selv om den opprinnelige standardiseringen var konservativ. En sammenligning av navnestandardisering i amerikanske og norske folketellinger viste at antall lenker i de norske kildene økte med 17 % på grunn av standardiseringen og at det var arbeidet med etternavn som bidro mest (Vick & Huynh, 2011). En annen faktor er at navneskikker endrer seg over tid (Fure, 1990). Dette påvirker både hyppigheten av ulike navn i ulike sogn og hvilke navneformer som oppfattes som synonyme. De kodede dataene fra folketellingene i 1801, 1865, 1875, 1900 og 1910 og

forhåpentligvis snart 1920 er tilgjengelig via internett fra vår partner ved University of Minnesota. (Se tabell 1 [Thorvaldsen 2024](#) og [nappdata.org](#).)

Historisk befolkningsregister (HBR)

Historisk befolkningsregister blir et nasjonalt register som skal dekke de 9,5 millioner menneskene, som bodde i Norge i perioden 1801–1964 med anslagsvis 85 millioner registreringer i de viktigste kildene med tverrsnittsdata, vitale hendelser og migrasjonskilder (Holden, Boudko, & Thorvaldsen, 2020). HBR har en åpen del frem til 1920-tallet som kun har offentlige opplysninger om avdøde personer. Deretter er det en lukket del, som er basert på kilder som kun er tilgjengelige for forskere som kan publisere aggregater. Den lukkede delen inneholder personer som fortsatt lever og som er lenkes til Det sentrale folkeregistret, startet i 1964. Automatisk lenking gjøres både ved UiT – Norges arktiske universitet og ved Norsk Regnesentral (NR) som utvikler nettstedet [histreg.no](#), hvor manuelle lenker legges til av frivillige.







Motivasjonen for å bygge HBR er å skape en sentral nasjonal infrastruktur for forskning innen historie, samfunnsvitenskap, medisin og en rekke andre disipliner. HBR har også en viktig kulturell komponent ved å følge slekter og bosetninger i over 200 år, og ved å sammenligne inkonsekvente poster oppfyller den et avgjørende kildekritisk mål. Dataprogrammer skaper lenker mellom forekomster av samme person i ulike kilder og lager pekere til relasjoner mellom familiemedlemmer. I tillegg til de nevnte institusjonene er Folkehelseinstituttet, Statistisk sentralbyrå, Nasjonalbiblioteket og Norges Handelshøyskole prosjektpartnere. Historisk folkeregister mottok nylig sin andre betydelige finansiering fra Infrastrukturprogrammet til Norges forskningsråd.

Sammenkobling av forekomster av personer i ulike kilder har tradisjonelt vært utført i sjangeren gårds- og slektshistorie for distriktskommuner med detaljerte studier av de lokale kirkebøkene og folketellingene supplert med andre skriftlige og muntlige kilder (Fure, 2004). Studiene har kartlagt bosetningen på gårder i større grad enn personer uten eiendom og folk som flyttet. I de fleste tilfeller utføres slikt arbeid av personer med detaljert innsikt i de lokale forhold. Tradisjonelt har arbeidet blitt utført manuelt, men datamaskiner har i økende grad blitt brukt for å effektivisere og systematisere arbeidet med kildematerialet og analysene. De digitale kildeutgavene forenkler formidlingen av den lenkede databasen via Internett (Kjelland, 2018). Det har lenge vært et mål blant faghistorikere i Norge å aktivere lokalhistoriebøkene detaljerte

gårds- og slektsoversikter i historiske analyser (Hovland, 1977), og vi mener at Historisk folkeregister vil fremme dette målet.

Tidslinjer for å følge enkeltpersoner og familier

«Tidslinjene» laget av Trygve Andersen ved RHD, UiT – Norges arktiske universitet, er et system som kan vise longitudinell informasjon fra Historisk befolkningsregister (Thorvaldsen, Sommerseth, & Holden, 2020). Folketellingene 1865, 1875, 1900, 1910 og snart 1920 er tilgjengelige for søk på UiTs nettsider via et enkelt og et avansert brukergrensesnitt, se <http://rhd.uit.no>. Etter å ha funnet en person ved å søke via navn osv, vil et klikk på hussymbolet vise informasjon om husstanden. Lenkede personer er utstyrt med en markør, jmf. venstre marg i figur 1, noe som betyr at ytterligere informasjon er tilgjengelig fra andre kilder via unike kildereferanser generert av Riksarkivet.

Name	Family status	Marital status	Occupation	Birth year	Place of birth
 Mikkel Nilsen	hf	g	Gaardbr. og Selveier	1812	Stokke Pr.
 Elen L. Hansdatter	Hans Kone	g		1823	Stokke Pr.
 Hans Mikkelsen	Deres Søn	ug	Matros	1844	Stokke Pr.
 Mathias Mikkelsen	Deres Søn	ug	Hjelper Fdr. med Grbr.	1847	Stokke Pr.
 Thorvald Mikkelsen	Deres Søn	ug		1849	Stokke Pr.
 Rikard Mikkelsen	Deres Søn			1852	Stokke Pr.

Figur 1: Familie på gården Sjuvestok i Stokke prestegjeld i 1865 med lenkesymboler i venstre marg.

Ved å klikke på lenkemarkøren vil en oversikt over de tilknyttede datapostene vises fra andre folketellinger og kirkebokregistre. Det advares om at lenkene genereres automatisk, og det kan ikke utelukkes at programvaren har introdusert feil og mangler. Thorvald Mikkelsen ble lett identifisert i folketellingene i både 1865, 1900 og 1910, men ble ikke automatisk knyttet til hans innførsel i folketellingen 1875 fordi han på den tiden ble adoptert som fosterbarn av den barnløse bonden som hadde kjøpt gården av foreldrene hans. Ved å klikke på +-tegnet i kolonnen til venstre i figur 2 vises informasjon om hele den husstanden personen tilhørte i det aktuelle tellingsåret.

+	1865	FT1865	0720-001-0027-00-005
+	1875	FT1875	0720-001-0103-00-003
+	1900	FT1900	0718-002-0075-00-001
+	1910	FT1910	0706-005-0038-05-001

Figur 2: : Tidslinje med referanser til fire folketellinger for Thorvald Mikkelsen. Den kan utvides for hver folketelling på <https://rhd.uit.no/folketellinger/tidslinje.aspx?idi=3177832>

Tidslinjefunksjonen gjør det mindre tidkrevende å følge grupper av mennesker over tid. Kohorter av personer som deler de samme egenskapene kan defineres i det avanserte brukergrensesnittet, slik at søkeresultatene blir mer tilpasset statistiske formål. Det er imidlertid ingen innebygde statistiske prosedyrer, og brukeren må selv lage kategoriene og ivareta representativiteten. Et eksempel på en mer kompleks tidslinje, som også inneholder opplysninger fra kirkebøkene om fisker Haldor Hansen (1850-1922) finnes på <https://rhd.uit.no/folketellinger/tidslinje.aspx?idi=8231280>.

Internasjonal kontekst

De transkriberte kildene i databasene ved UiT er del av en betydelig internasjonal samling med individdata med over to milliarder historiske personinnførsler i mange av landene i den industrialiserte del av verden. Både metodeutviklingen og den empiriske forskningen skjer i nettverk omkring konferanser og tidsskrifter som *Historical Methods*, *Social Science History*, *Historical Life Course Studies* etc. Et konkret resultat er integreringen av norske data fra folketellingene 1865 til 1910 i databasen *North Atlantic Population Project* ved Minnesota Population Center, se ipums.org. *European Historical Population Samples Network* har betydd mye for vår metodeutvikling og publisering, se figur 3 og ehps.net.eu.

Forskningseksempel 3: Emigrasjon

Ved hjelp av RHDs kodete folketellinger fra Norge og ditto materiale fra USA lenket på individnivå, publiserte en gruppe økonomer forskning om motivene bak emigrasjonen fra Europa til USA og hvorvidt emigrasjonen lønte seg. I masseemigrasjonens tid opprettholdt USA en nesten åpen grense, noe som gjør det mulig å studere en migrasjonsprosess uhindret av adgangsbegrensninger. Ved hjelp av lenkede data om 50 000 norske menn studerte de effekten

av rikdom for intern eller internasjonal migrasjon i perioden 1850–1913. Her utnyttet de variasjoner i foreldrenes formue og i forventet arv etter fødselsrekkefølge, kjønns-sammensetning på søsken og bostedsregion. De konkluderte med at relativ velstand gjorde beslutningen om å emigrere mindre sannsynlig i denne epoken. Forskerne antyder at de fattige kunne være mer tilbøyelige å flytte hvis migrasjonsrestriksjoner ble opphevet i dag, og diskuterer implikasjonene av slike historiske funn for utviklingsland. Forskerne estimerte også i hvilken grad emigrasjonen lønte seg ved å sammenligne utvandrere fra Norge til USA med brødre som bodde i Norge mot slutten av 1800-tallet. Sammenligningen av bofaste og emigrerte søsken tyder på at lønnsomheten ved emigrasjon var relativt lav, og de ser dette i sammenheng med at relativt flere fattige fra urbane strøk enn fra rurale områder emigrerte (Abramitzky, Boustan, & Eriksson, 2012, 2013). Resultater fra Ullensaker-prosjektet indikerer at denne forskjellen ikke gjelder norske distrikter, selv om dette muligens kan forklares med spesielle by-land- og periodeeffekter (Koren, 1979). Vi trenger derfor ytterligere forskning med longitudinelle data for å teste økonomenes konklusjoner om at relativ velstand motvirket emigrasjon og at emigrasjon bare i mindre grad ga positive økonomiske resultater.

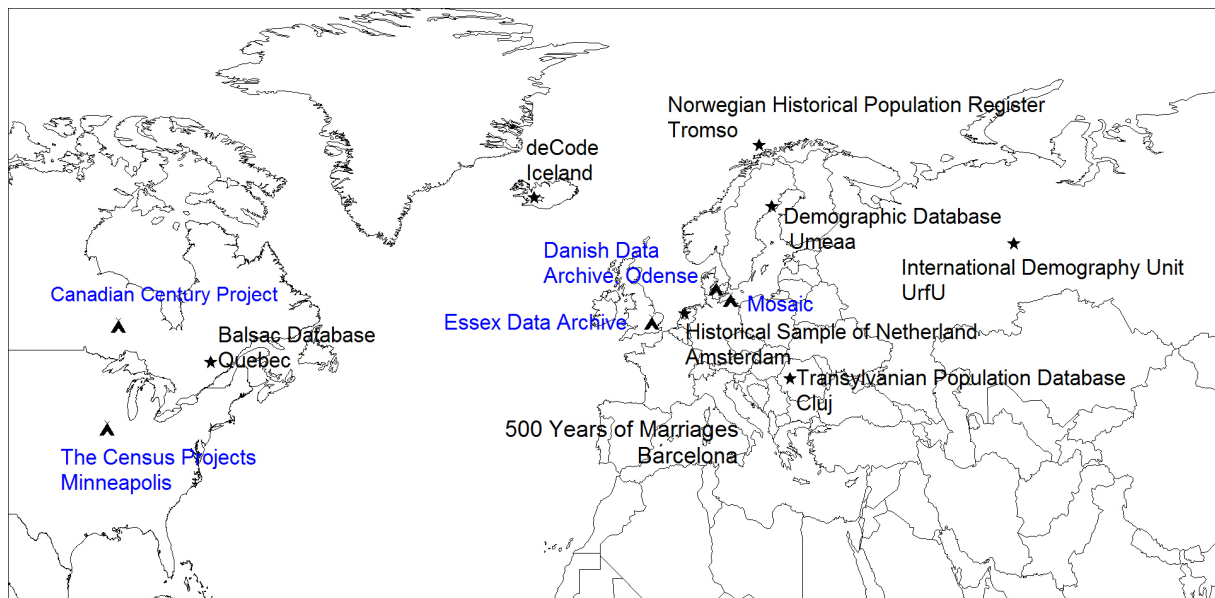
Forskningseksempel 4: Intergenerasjonell spedbarndødelighet

I en komparativ studie av likhet og ulikhet mellom generasjonene mht spedbarndødelighet, deltok forskerteam fra Nederland, Belgia, Sverige og Norge med sine respektive datasett. I første del av prosjektet ble det gjennomført regionale analyser, men med felles teoretisk og metodisk design. Resultatene fra bruk av lenkede kirkebøker- og folketellinger i Historisk folkeregister for Troms fylke i Norge viser at det var påvirkning mellom generasjonene med hensyn til risiko for å oppleve spedbarndød. En kvinnes barn hadde større risiko for å dø før ettårsdagen hvis bestemora deres hadde flere spedbarndødsfall. Risikoen for spedbarns død blant barn av døtre fra slike høyrisikofamilier var minst 30 prosent høyere enn blant spedbarn født av døtre til mødre som hadde opplevd null døde spedbarn. I motsatt ende finner vi flertallet, der 60-70 prosent av familiene på tvers av generasjoner, aldri opplevde å miste noe spedbarn (Sommerseth, 2018). Lignende funn ble også gjort i analysene til de andre landene, og samlet sett peker disse funnene i retning av en differensiell forståelse av hvordan dødeligheten påvirket de yngste. Spedbarnsdødelighet i historiske populasjoner påvirket ikke alle hus og familier. Neste steg i prosjektet har vært å sette sammen alle landenes datasett til en felles database og gjennomføre en felles statistisk analyse. Resultatene bekreftet funn fra første runde og representerer et pionerarbeid for videre internasjonalisering (Quaranta & Sommerseth, 2018).

Det internasjonale Intermediate Data Structure (IDS)-formatet ble brukt for å gjøre datasettene fra ulike land sammenlignbare.

Punktvis kan vi oppsummere slik:

- Myte 1: At historieforskning med folketellinger og kirkebøker bare er kvantitativ.
- Myte 2: At arbeidet stort sett er lokallhistorisk – historiske protokolldata som folketellinger og kirkebøker bearbeides i et stort internasjonalt nettverk.
- Perioden fra 1800 er dekket med transkriberte, søkbare kirkebøker og folketellinger, men deler må kvalitetssikres.
- Historisk befolkningsregister har fått sin annen bevilgning fra Norges forskningsråd, og lenker sammen data om Norges befolkning på individ- og gruppenivå fra år 1800.
- Forskning vha kildene gjøres innen sosialhistorie, medisinhistorie, mortalitet, fertilitet, migrasjon, befolkningsutvikling, etnografisk feltarbeid, tjeneres antall og status osv.
- Datasett som kombineres på individnivå gir sikrere forskningsresultater enn sammenstilling av aggregater.



Figur 3: Forskningssentere som bearbeider historiske individdata og samarbeider i European Historical Population Samples Network.

Referanser

- Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, 102(5), 1832–1856. DOI: <https://doi.org/10.1257/aer.102.5.1832>
- Abramitzky, R., Boustan, L. P., & Eriksson, K. (2013). Have the poor always been less likely to migrate? Evidence from inheritance practices during the age of mass migration. *Journal of Development Economics*, 102, 2-14. DOI: <https://doi.org/10.1016/j.jdeveco.2012.08.004>
- Alhaug, G. (2005). Støtte frå Noregs forskingsråd til standardisering av namn i folketeljningane [Standardizing census names with Research Council support]. *Nytt om namn, Nr 41* (2005), 47-49.
- Fure, E. (1990). Personnavn og tidsånd [Personal names and the spirit of the times]. *Namn og Nemne*, 8, 35-55.
- Fure, E. (2004). – en besynderlig Regelmæssighed: dødeligheten i Asker og Bærum på 1800-tallet med særlig vekt på spedbarnsdødeligheten (Doktoravhandling). [- a curious Regularity: mortality in Asker and Bærum in the 19th century with particular emphasis on infant mortality (Doctoral dissertation).]. Oslo: University of Oslo.
- Glavatskaya, E., & Thorvaldsen, G. (2013). Sergej Sergel's Field Research in Northern Norway and Finland: Contextualizing Early 20th-Century Sami. *Arctic Anthropology*, 50(1), 105-119. DOI: <https://doi.org/10.3368/aa.50.1.105>
- Holden, L., Boudko, S., & Thorvaldsen, G. (2020). Lenking og kobling i Historisk befolkningsregister [Record linkage and family pointers in the Historical Population Register]. *Heimen*, 57(3), 216-229. DOI: <https://doi.org/10.18261/issn1894-3195-2020-03-04>
- Hovland, E. (1977). *Folket, bygda og historia [The population, the community and the history]*. Bergen: Universitetsforlaget.
- Kjelland, A. (2018). Mapping and Analysing Remigration Based upon Norwegian Farm- and Genealogical History Projects. . *Journal of Migration History*, 4(2), 314–329. DOI: <https://doi.org/10.1163/23519924-00402005>
- Koren, E. (1979). *Utvandringen fra Ullensaker 1867-99 : en sosialhistorisk undersøkelse*. (Cand.philol Hovedoppgave), University of Oslo, Oslo.
- Langholm, S. (1976). On the scope of micro-history. *Scandinavian journal of history*, 1-24. DOI: <https://doi.org/10.1080/03468757608578894>

- Nygaard, L. e. a. (1995). *Histform - Norsk standard for registrering og utveksling av nominative folketellingsdata for årene 1865-1910 [Histform - Norwegian standard for registration and exchange of nominative census data for the years 1865-1910]*. Tromsø: Registreringssentral for historiske data, Universitetet i Tromsø.
- Quaranta, L., & Sommerseth, H. (2018). Introduction: Intergenerational Transmissions of Infant Mortality using the Intermediate Data Structure (IDS). *Historical Life Course Studies*, 7(2). DOI: <https://doi.org/10.51964/hlcs9288>
- Sommerseth, H. (2018). The Intergenerational Transfer of Infant Mortality in Northern Norway during the 19th and Early 20th Centuries. *Historical Life Course Studies*, 7(2). DOI: <https://doi.org/10.51964/hlcs9284>
- Thorvaldsen, G. (1984). Befolkningsutvikling i Tromsø 1866 til 1900. *Heimen*, 95-106.
- Thorvaldsen, G. (1994). The Encoding of Highly Structured Historical Sources. *Computers and the Humanities*, 28(4/5), 301-305. DOI: <https://doi.org/10.1007/BF01830278>
- Thorvaldsen, G. (2008). Hushjelper og jordbrukstjenere – når kom nedgangen i tjenertallene? [When did the number of domestic servants decline in Norway?]. *Historisk tidsskrift*, 87(3), 451-464. DOI: <https://doi.org/10.18261/ISSN1504-2944-2008-03-04>
- Thorvaldsen, G. (2018). *Censuses and Census Takers: A Global History*. London: Routledge. DOI: <https://doi.org/10.4324/9781315148502>
- Thorvaldsen, G., Pujadas-Mora, J., Andersen, T., Eikvil, L., Lladós, J., Fornés, A., & Cabré, A. (2015). A Tale of Two Transcriptions. Machine-Assisted Transcription of Historical Sources. *Historical Life Course Studies*, 2, 1-19. DOI: <https://doi.org/10.51964/hlcs9355>
- Thorvaldsen, G., Sommerseth, H. L., & Holden, L. (2020). Anvendelser av Norges historiske befolkningsregister [Interfaces to Norway's Historical Population Register]. *Heimen*, 57(3), 230-243. DOI: <https://doi.org/10.18261/issn.1894-3195-2020-03-05>
- Thorvaldsen, G., Kilder og metoder for å identifisere byfolk på tidlig 1900-tall. *Heimen* 60(1-2) 80-97. DOI: <https://doi.org/10.18261/heimen.60.1-2.6>
- Vick, R., & Huynh, L. (2011). The Effects of Standardizing Names for Record Linkage: Evidence from the United States and Norway. *Historical Methods*, 44(1), 15-24. DOI: <https://doi.org/10.1080/01615440.2010.514849>
- Winkler, W. E. (1990). *String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage* 354–359. Paper presented at the Proceedings of the Section on Survey Research Methods. American Statistical Association.