

Fra folketellinger og kirkebøker til norsk befolkningsregister

Gunnar Thorvaldsen

Registreringsentral for historiske data, Universitetet i Tromsø

Et befolkningsregister tar sikte på å holde en løpende oversikt over befolkningen i et gitt geografisk område. Det tar gjerne utgangspunkt i en folketelling, og blir så oppdatert med opplysninger om demografiske hendelser (fødsler, vielser, begravelser) i tillegg til flytting. Artikkelen omhandler et tema som er aktuelt av to grunner. For det ene i et samtidsperspektiv fordi Norges sentrale folkeregister nå har avløst det systemet med folketellinger og kirkebøker som holdt oversikt over vår befolkning gjennom nær tre hundreår. For det andre i et historisk perspektiv fordi det er mulig å knytte sammen informasjon fra dataregistrerte folketellinger og kirkebøker slik at vi kan skape et historisk befolkningsregister for Norge. Både det løpende, sentrale folkeregistret og eksisterende lokale, historiske befolkningsregistre inneholder viktige grunnlagsdata for lokal- og regionalhistorisk forskning.

Sammen med de andre nordiske land har Norge et av de eldste og mest fullstendige system i verden for å holde rede på sin befolkning. Kirkebøker med innførsler om de døpte, viede og døde har vært ført fra 1623, og føringen ble påbudt over hele landet på 1680-tallet. Siden 1700-tallet har vi bevart kirkebøker fra et stadig økende antall prestegjeld – pga kirkebranner og hull i avleveringsplikten mangler vi riktignok kirkebøker for et fåtall prestegjeld også på 1800-tallet. Dessuten må vi regne med at et varierende antall døpte, viede og begravde mangler i de eldste listene, en manko på mellom ti og tjue prosent kan være et rimelig anslag for 1700-tallet. Enda mer mangelfulle er de flyttelistene som ble påbudt ført fra 1820, men allikevel holder våre kirkebøker høy standard ved internasjonal sammenligning.¹

Det samme gjelder våre folketellinger, især de som ble ført med navn fra 1801 og 1865 – det er mer problematisk å utnytte manntallene fra 1660-årene og 1701 og de statistiske tellingene fra 1769 og 1815 til 1855. Fra folketellingene har Statistisk sentralbyrå publisert statistikk om lag hvert tiende år, mens oppgaver fra kirkebøkene og annen innrapportering av vitale hendelser er publisert i *Folkemengdens bevegelse* og av Michael Drake. Denne inndelingen i tverrsnittsdata og vitalstatistikk har vært vanlig i vest-europeiske land. Mens det aggregerte tallmaterialet gjerne kombineres i analyser, ble individdataene ført punktvis og separat i de ulike kildeseriene, det være seg i folketellinger, kirkebøker osv.

Denne artikkelen behandler imidlertid ikke slike kilder med enkeltstående innførsler, men befolkningsregistre hvor personer kan følges mer systematisk over tid innenfor et gitt administrativt område. Moderne registre er gjerne nasjonale, mens de historiske som regel dekker et lite geografisk område, gjerne et prestegjeld. Min definisjon inkluderer befolkningsregistre som i prinsippet følger hele befolkningen i hele eller deler av et administrativt område, men ikke slike som bare følger spesielt utvalgte grupper av personer. Dermed kommer studier som anvender familierekonstitusjonsmetoden med, mens kohortanalyser er utelatt. Eksempler på de sistnevnte er studier som følger konfirmantene i Ullensaker eller de som ble født i Sør-Audnedal i bestemte år. De teller ikke som befolkningsregistre fordi de kun omfatter mer eller mindre tilfeldige utvalg av personer fra et geografisk område. Heller ikke databaser som følger befolkninga gjennom et kort tidsrom, faller inn under definisjonen av befolkningsregister. Det

gjelder for eksempel det prosjektet ved Historisk institutt, Universitetet i Bergen, der to hovedoppgaver tok utgangspunkt i lenking av folketellinga 1801 og kirkebøkene i et par omkringliggende år.² For enkelte prestegjeld fins sjeleregistre som ble oppdatert ved husbesøk på samme måte som husförhörslängdene i Sverige.³ Disse burde telle som befolkningsregistre dersom de er noenlunde fullstendige. Jeg har imidlertid ikke forsøkt å skille mellom slike gårds- og slektshistoriske verk som dekker hele befolkninga og slike som utelater mange av de eiendomsløse – det kunne være tema for en separat artikkel.

Mot et løpende befolkningsregister

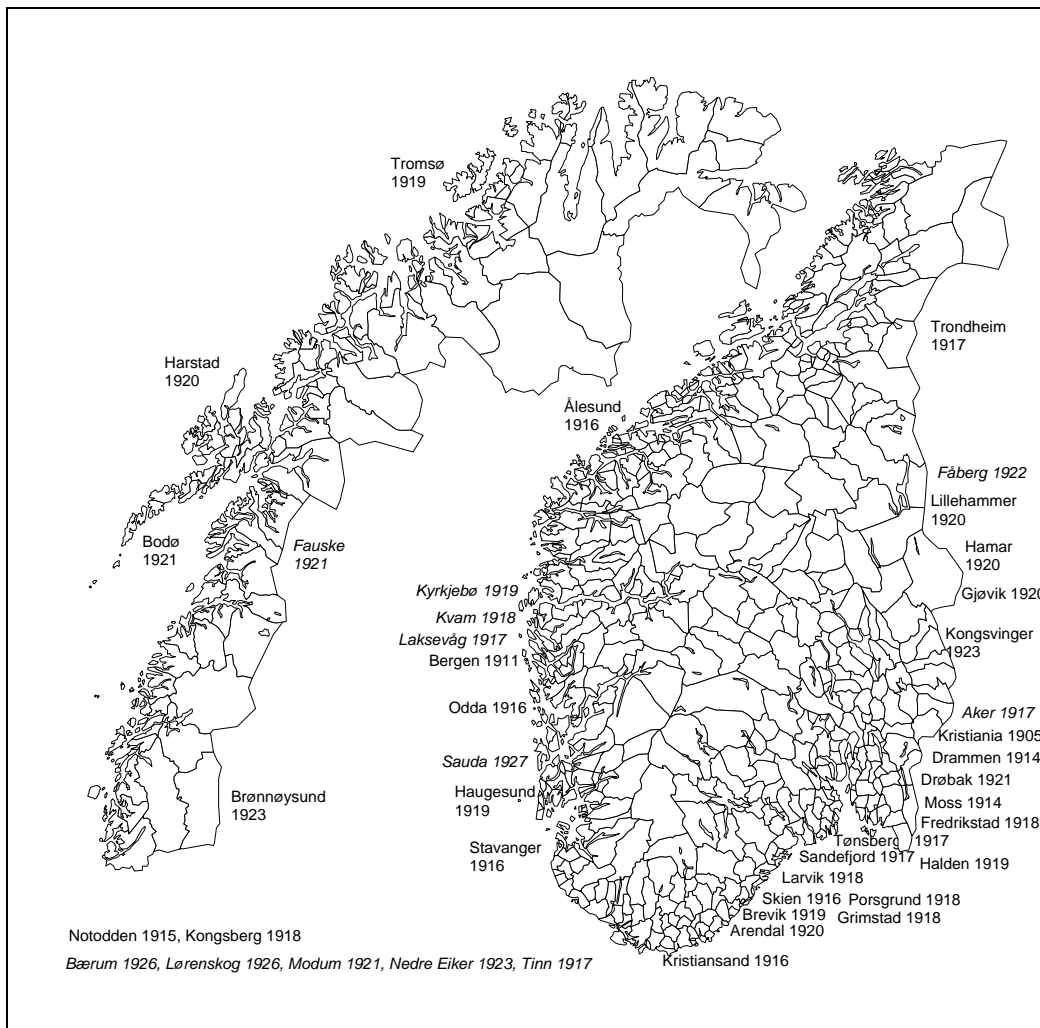
Statistisk sentralbyrå arrangerte fullstendige folketellinger med – stort sett – ti års mellomrom fram til 1980, i 1990 nøyde SSB seg med å telle et representativt utvalg av norske hushold, men så var folketellingenes æra over her til lands. I 2001 begrenset Byrådet seg til å avholde bolig-telling. Disse hadde fra 1875 blitt et stadig viktigere supplement til folketellingene, men var altså ved årtusenskiftet blitt seg selv nok. Bakgrunnen for reformen var at SSB mente å finne tilstrekkelig informasjon om rikets innbyggere i det sentrale folkeregistret og andre registre, slik at utsendte skjemaer strengt tatt bare trengtes for å knytte individene til sine boliger. Tiltaket var ikke spesielt radikalt: Danmark avholdt sin siste skjemabaserte folketelling i 1970, Finland i 1980.⁴ Og folkeregister er intet nytt tiltak i Norge; satt på spissen kan vi si at metoden feiret hundreårsjubileum ved helt å ta over som grunnlagsmateriale for befolkningsstatistikken ved årtusenskiftet.

For selv om det sentrale, elektroniske personregistret med personnummer som identifikator ble opprettet på 1960-tallet, har det norske røtter i form av manuelle registre tilbake til tidlig 1900-tall. Internasjonalt er ideen om et løpende oppdatert befolkningsregister enda eldre. Den ble lansert under den franske revolusjon i 1791, men først realisert i Nederland, Belgia og Italia midt på 1800-tallet og fra 1876 i Stockholm.⁵ Kristiania folkeregister ble opprettet på nyåret



Gustaf Amneus. Foto: Portrettsamlingen, By-museet i Bergen.

1906.⁶ Ildsjelen bak tiltaket var Gustaf Amneus, sjef for kommunens statistiske kontor, og bakgrunnen var noe av den samme som i Stockholm, en av de byene han hadde besøkt på en studietur i 1902. Tiårige folketellinger og hendelsesbaserte kirkebøker strakk ikke til for å holde oversikt over den store og flyktige befolkninga i storbyen – folketallet i Kristiania var nær tredoblet siden 1875. 1800-tallets liberale befolkningsregime var forlatt, og skulle den nye løsgjengerloven, fattigloven og fremmedloven fra hundreårsskiftet håndheves, måtte folkebokføringen styrkes. Den lov om folkeregister som stortinget fulgte opp med i 1905, var imidlertid basert på frivillighet, hver kommune kunne selv avgjøre ut fra egne behov. I tillegg til å drive igjennom loven – Amneus var jurist liksom ledelsen i Statistisk sentralbyrå – innførte Amneus nye metoder i Kristianas folkeregister. Protokollsystemet fra Stockholm ble forlatt og erstattet med fire registre ført på kort: Hovedregister, adresseregister samt inn- og utflyttingsregistre. Helt nytt var ikke dette, Amneus kan meget vel ha hentet inspirasjon fra et tilsvarende system som ble diskutert innført i Nederlands folkere-



Innføring av folkeregistre i norske kommuner 1905 til 1927. Landkommuner med kursiv. Kilde: Thor Andersson: "Folkeregister i Norge." *Nordisk statistisk tidsskrift* 1927, 515f.

gister fra 1890-tallet. Men Kristiania var først ute med kortsystemet, og Amneus har blitt kalt det moderne folkeregistrets far.⁷

Neste by ut var Bergen, hvor folketellinga i 1910 viste at myndighetene hadde mistet oversikten over innflyttingen og dermed antall innbyggere i byen. Amneus, som ble spurt til råd, mente trygt å kunne anbefale sitt system. Ikke bare hadde det rasjonalisert føring av skattemanntall, valgmannantall, oversikten over skolepliktige barn og andre manntall; utgiftene til folkebokføringen hadde samlet sett blitt redusert. Han mente og at innkrevingen av skatt had-

de blitt så effektiv at han tok mye av æren for reduksjonen i Kristianias skatteøre! I 1912 gjennomførte Bergen en kommunal folketelling som i tillegg til innflytternes fødested ba dem besvare hvor lenge de hadde vært i byen. Dermed fikk også Bergen et forbedret grunnlag for å føre folkeregister, med særlig fokus på migrasjonen inn og ut av byen. Kanskje var det nærheten til hovedstaden som gjorde at Moss ble den tredje og siste byen med kommunalt folkeregister før første verdenskrig.

Både under første og annen verdenskrig ble det tatt viktige initiativ for å forbedre folkebokfø-

ringen. Grunnet for å føre nasjonal statistikk ble reformert ved innføring av et sivilt fødselsregister fra 1916, altså fra da av et alternativ eller supplement til kirkebøkernes dåpslister når vi skal velge kildemateriale for et historisk befolkningsregister. Under første verdenskrig og i et par år etter fikk en rekke byer folkeregister. Det gjelder naturlig nok de fire største byene etter Kristiania og Bergen: Trondheim, Stavanger, Drammen og Kristiansand. Dernest ble flere forstadskommuner inspirert av nabobyene: Aker, Bærum, Lørenskog og Laksevåg. Videre hadde noen raskt voksende, ensidige industristeder sett behovet for bedre oversikt: Fauske, Odde, Sauda og Tinn. Dermed hadde folkeregistret tatt et viktig steg i riksdekkende retning – det var også innført i landkommuner, om enn slike med viktige tettsteder og noe bypreg. Et motiv i tillegg til det generelle behovet for oversikt var nok å holde oppsikt med fremmede i urolige tider. Odelstinget behandlet i 1915 en revidert lov om folkeregistre sammen med et tillegg til fremmedlovgivingen.⁸

På 1920- og 1930-tallet ble folkeregister innført i ytterligere 64 kommuner slik at de løpende, lokale registrene omfattet 49 byer og 42 herreder med halvparten av landets befolkning da annen verdenskrig brøt ut. Det var i hovedsak mindre bykommuner som innførte registrene, og naturlig nok sto økonomiske argumenter i fokus i de økonomiske trengselstidene. Det gjaldt for det ene å få oversikt over den hjemmehørende befolkning slik at fremmede kunne nektes støtte fra stedets fattigkasse. For det andre var effektiv skatteinndrivning viktigere enn noensinne. Virkelig effektive kunne folkeregistrene vanskelig bli på frivillig basis, og fra 1920 arbeidet Byforbundet og andre for å gjøre registrene landsdekkende. Hvilken inspirasjon talsmennene hentet fra Danmark er ukjent, der ble obligatoriske folkeregistre innført i alle kommuner fra 1924.⁹ Til tross for at administrative hensyn ble støttet med argumenter fra befolkningsstatistikere som syntes de fikk mistenkelig få oppgaver fra dissentermenighetene, lot det seg ikke gjøre å få gjennomslag for tanken om landsdekkende registre før krigen. At reformen ble støttet av elementer på ytterste høyre

fløy med rasistiske tilbøyeligheter og ønsker om informasjon og omfattende kontroll med «fremmede», kan kanskje noteres som en bjørnetjeneste – deres motiver fant liten forståelse hos de bevilgende myndigheter.¹⁰

Okkupasjonsmakten tok i februar 1942 initiativ til opprettelsen av et sentralt folkeregister.¹¹ Tyskernes primære siktemål var å kontrollere persontrafikken også i landets mange småkommuner, men registret kunne i tillegg brukes til å føre statistikk over fødsler, giftermål og dødsfall, noe Statistisk sentralbyrå understreket i sine kommentarer til reformen. Reichskommissariat lovfestet obligatoriske folkeregistre i alle kommuner fra 1 mars 1943, og overførte samtidig registreringer av sivilstand dit fra kirken. Registrerens nazistiske bakgrunn var et problem også da ordningen skulle forlenges etter krigen. Lov om folkeregistre ble vedtatt av Stortinget i 1946, noe som var nødvendig for at folketellinga det året skulle kunne brukes som grunnlagsmateriale. Etter krigen var folkeregistrene av vekslende kvalitet rundt om i landets kommuner, og det var behov for et nytt kontrollpunkt i form av en folketelling for å oppdatere registrene.

For å redusere inntrykket av å bare strø sand på vedtak fattet av nazimyndighetene, ble ansvaret for sivilstandsregistreringen ført tilbake til de geistlige myndigheter. Det ble heller ikke tillatt for kommunene å bygge videre på folkeregistre opprettet under krigen. Når 29 representanter stemte imot å gjøre folkeregistrene obligatoriske, var vel helst motivet at små, fattige kommuner skulle slippe ytterligere administrative forpliktelser – rotet i mange av de kommunale registrene fristet vel heller ikke til utvidelse av ordningen. Men Stortinget vedtok altså at det skulle forbli obligatorisk å føre registre og at folketellinger kunne brukes som grunnlagsmateriale. Men også etter kontroll mot 1946-tellinga viste det seg at folkeregistrene var mangelfulle.

Det var noe av grunnen til at det ble avholdt ny, regulær folketelling i 1950 og at denne tellinga ble spesielt utformet med tanke på oppdatering av de kommunale folkeregistrene. For å lette kontrollarbeidet inneholder denne tellinga såkalt retrospektive variabler; det har ellers vært uvanlig å spørre etter fortidige forhold i norske

folketellinger. Personer som var flyttet inn til kommunen etter 3 desember 1946, skulle oppgi når de flyttet inn, hvor de sist var bosatt og hvor de var bosatt ved forrige tellingstidspunkt. Etterpå ble det gjort et stort arbeid i alle landets kommuner for å oppdatere folkeregistrene med informasjon fra 1950-tellinga. Det forenklet arbeidet med oppdatering av de kommunale registrene at de nå var obligatoriske slik at inn- og utflyttere kunne spores andre steder.

Innføringen av det elektroniske folkeregistret i 1964 hadde nær sammenheng med at Statistisk sentralbyrå utviklet systemet med ellevesifrede personnummer. Disse to reformene forutsatte hverandre, både for administrative og såkalt arkivstatistiske formål, dvs som grunnlag for kombinasjon av informasjon fra ulike registre i statistikkproduksjonen.¹² Behovet for et sentralt befolkningsregister kan også sees i sammenheng med innføringen av forskuddsskatt i 1957, et inntrykk som forsterkes av at ansvaret for det sentrale folkeregister i 1991 ble overført til Skattedirektoratet. Dette ble gjort til tross for at Statistisk sentralbyrå stadig var en viktig bruker av registret fordi de nå i prinsippet kunne slutte å avholde skjembaserte folketellinger. Data fra kompletterende databaser kan lenkes til det sentrale befolkningsregistret på individnivå med det unike personnummeret som koblingsnøkkel. I tillegg til at registermetoden sparer kostnader, er dens største fortrinn at man med små ekstra kostnader kan la datamaskinen trekke ut statistiske oversikter så ofte man måtte ønske både over befolkningens sammensetning og de fortløpende endringene. En svakhet ved registrene er at noen typer opplysninger er vanskelig å oppdatere uten å sende ut spørreskjemaer, for eksempel data om utdanning tatt i utlandet eller detaljerte opplysninger om arbeidsfunksjoner internt i bedriftene. Vi skal ikke her gå nærmere inn på den omfattende diskusjonen som oppsto om personvernet da det viste seg at personnummeret ble anvendt i langt flere sammenhenger enn opprinnelig forutsatt.¹³

Historiske befolkningsregistre

Mens det løpende folkeregisteret er landsdekkende og har blitt oppdatert kontinuerlig siden 1964, er det bare Island som hittil har laget et landsdekkende historisk befolkningsregister med hele befolkninga tilbake til middelalderen der kildemateriale er tilgjengelig.¹⁴ Norske forsøk på å knytte sammen historiske individdata til kollektive biografier er mer begrensede både i tid og rom, de dekker for det meste deler av 1700- og 1800-tallet, og lenker sammen data bare om personer med tilknytning til én kommune. Som nevnt ble det moderne folkeregistret laget på grunnlag av informasjon fra folketellingene. De historiske befolkningsregistrene er derimot i utgangspunktet laget på basis av kirkebøkene, gjerne ved hjelp av familierekonstitusjon eller tilsvarende metoder.

Familierekonstitusjonsmetoden

Nominativt kildemateriale fra mange norske kommuner har blitt lenket sammen for å beskrive livsløpene til innbyggerne der, især på 1700- og 1800-tallet. Den mest standardiserte metoden for slik sammenlenking av individdata er familierekonstitusjon basert på dåps-, vielses- og begravelleslistene i kirkebøkene. Kjernen i familierekonstitusjonsmetoden, er å knytte sammen data fra ulike kilder med individ og familie som koblingspunkt. Metoden ble utviklet av Louis Henry i Frankrike på 1950-tallet for å beregne ulike demografiske rater med utgangspunkt i kirkebøkens lister over døde, viede og begravede.¹⁵ Et hyppig forekommende problem i demografiske analyser er at populasjonens størrelse er ukjent, for eksempel at vi kun kjenner antall døde personer, mens folketall og aldersgruppering er ukjent fordi vi mangler folketellinger. Da blir det umulig å beregne rater for å sammenligne dødeligheten eksempelvis i to ulike kirkesogn eller i ulike sosiale lag.

Det lar seg imidlertid gjøre å rekonstruere familiene og dermed befolkninga ved å gruppere all informasjon om fødte, viede og døde i kirkebøkene – nærmest som et stort puslespill. Ved å kombinere data fra vielseslistene med ekteparenes barnefødsler slik de er registrert i dåpsliste-

ne, blir familiene rekonstituert, dvs satt sammen til en beskrivelse av deres vitale hendelser. I tillegg hentes informasjon fra begravelleslistene for å finne ut når familiemedlemmene forlot befolkninga. I teorien skal vi dermed få en fortløpende oversikt over den befolkninga som var til stede i prestegjeldet gjennom den perioden kirkebøkene dekker. Siden er metoden tilpasset den noe ulike kildesituasjonen i andre land; imponerende mye er gjort med de problematiske kirkebøkene i Storbritannia. Våre svenske kollegaer liker å påpeke at de var først ute i og med Edins studier fra 1915 og Hyrenius sine fra 1942 på grunnlag av det utmerkede svenske kirkeboksmaterialet. Da må vi huske at Henrys metode både innebærer å koble sammen individ og familier på bestemte måter og å beregne en rekke demografiske rater. Især på det siste punktet presenterte verken Edin eller Hyrenius noen ensartet standard, slik Henry gjorde. De tyske Ortsfamilienbücher eller Dorfsippenbücher kan også danne grunnlag for familierekonstitusjon; de har mye til felles med norske bygdebøker.

I Norge ble familierekonstitusjonsmetoden introdusert på 1960- og 1970-tallet med Ståle Dyrvik og Sølvi Sogner som de ledende utøvere. Statsarkivar Lajos Juhász' metodeartikkel i *Heimen* kom samtidig med at de første lokale norske undersøkelser som anvendte metoden ble presentert fra 1970 av, de første og de fleste ved Universitetet i Bergen. Mange hovedoppgaver har anvendt metoden, men bare to norske doktoravhandlinger, nemlig Sølvi Sogners om Rendalen og Rasmus Sundes om Arnafjord-delen av Vik i Sogn. Sistnevnte hadde hovedfokuset på å rekonstituere familiene til emigrantene fra Vik til Midt-Vesten i USA, slik også Olav Tysdal har fulgt emigranter fra Skudenes. Siden det er så vidt mye vanskeligere å følge individer og familier i byene enn på landsbygda pga større migrasjon, er det bare fire hovedoppgaver som hentet materiale fra by: Moss, Trondheim, Stavanger og Bergen, sistnevnte som ledd i en utvalgsundersøkelse fra Søndre Bergenhus amt. Prestegjeld på Vestlandet har fått stor oppmerksomhet idet halvparten av undersøkelsene hentet materiale fra Rogaland, Hordaland, Sogn og Fjordane samt Møre og Romsdal fylker, de fleste

veiledet av Ståle Dyrvik ved Historisk institutt, Universitetet i Bergen. Fra fylkene Oppland, Vestfold, Troms og Finnmark har ingen anvendt kirkebøkene til familierekonstitusjon, og metoden er kun anvendt for én kommune i Nord-Norge. Det demografisk svært interessante fler-etniske området i nord står dermed udekket mht den type undersøkelser som krever longitudinelle befolkningregistre. På svensk side av grensa bygges derimot slike registre systematisk nettopp i kommuner med samisk befolkning.¹⁶ Se for øvrig oversikten i vedlegget.

Selv om de skandinaviske kirkebøkene er bedre enn tilsvarende materiale fra andre land, er det prinsipielt interessant å gruppere de kritiske innvendingene mot familierekonstitusjonsmetoden i tre hovedgrupper: 1) Sensurering; dvs at det ikke har lyktes å rekonstruere hele befolkninga i undersøkelsesområdet. Dette kan skyldes flytting, mangler ved kildene eller at forskeren ikke har greidd å knytte sammen opplysninger som egentlig fins der. 2) De vitale hendelsene i undersøkelsesområdet er mangelfullt registrert. Den mest vanlige kritikken av norske kirkebøker er at det mangler innførsler av døde spedbarn i tida før 1820. 3) Små tall kan gi tilfeldige resultater. Familierekonstitusjonsmetoden er arbeidskrevende selv ved anvendelse av moderne informasjonsteknologi. Dette begrenser størrelsen på det materialet som i praksis lar seg anvende. Når vi da deler inn materialet i analysegrupper, etter kjønn, sosial status, alder osv, blir gruppene i hver tabellcelle i minste laget. Vi kan kompensere ved å slå sammen materiale fra en lang periode, men dette kan selvsagt skjule den historiske utvikling som det er vår oppgave å avdekke. Mye av løsningen på problemene ligger i å komplettere med data fra annet kildemateriale enn kirkebøkene.

Rendalsdatabasen

Dette historiske befolkningsregistret er en utvidelse av Sølvi Sogners familierekonstitusjon ved at data fra en rekke komplementære kilder ble lagt til. Rendalsdatabasen dekker Øvre og Ytre Rendalen gjennom tidsrommet fra 1733 til 1900 og bygger i hovedsak på kirkebøker og fol-

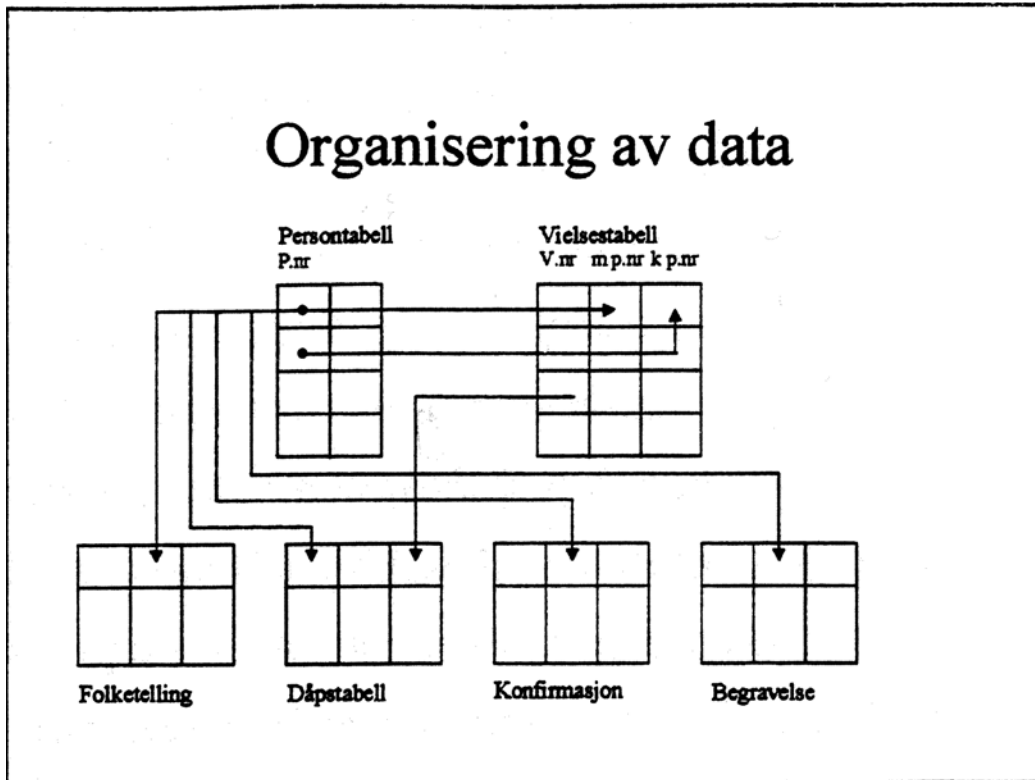


Fig. 3. Organisering av data i Rendalsdatabasen. Kilde: Gjelseth 2000:68.

ketellinger, men med viktige innslag fra annet materiale som emigrasjonslistene, skattemantallene, pantebøkene og matriklene. Disse kildene er delvis transkribert hos Registreringsentral for historiske data og Digitalarkivet, dels lagt inn og kvalitetssikret i Rendalsdatabasen direkte. Opplysningene i primærkildene er i noen grad supplert med informasjon fra bygdebøkene, men enda viktigere er at prestegjeldet fra før var inngående studert ved anvendelse av familierekonstitusjonsmetoden fram til 1828, slik at data om størstedelen av befolkninga allerede var lenket sammen. Sølvi Sogner valgte Rendalen som undersøkelsesområde fordi det ligger avsides og var mindre påvirket av migrasjon enn kommuner flest. Dermed ble dette registret mer komplett enn hva som er realistisk for norske kommuner flest selv om det bare omfatter et lokalt avgrenset område.

I og med at utgangspunktet var de rekonstituerte familiene ble det naturlig å lage databasen

ved å lenke sammen vielsespostene med dåpspostene i kirkebøkene, basert på informasjon om par og foreldre slik de opptrer i de to kildetype. For å håndtere de uekte fødslene ble det konstruert falske eller virtuelle vielser i databasen for å få med informasjon om begge foreldrene. Det innebærer at en kvinne eller mann som fikk barn utenfor ekteskap med ulike partnere, er registrert i flere kunstig konstruerte vielser. Samme kvinne eller mann kan ellers pga gjengifte forekomme i flere dataposter basert direkte på vielseslistene.

Det ble opprettet personnummer, i prinsippet for alle individer som har bodd i Rendalen i tidsrommet. Det er kurant for dem som ble døpt der, men mer problematisk for innflytterne, for det er da risiko for å gi samme individ to ulike nummer pga inkonsekvent føring av navn eller alder, og dessuten fare for å gi samme nummer til to forskjellige innflyttere som vi ikke kjenner fødselsdato for. I alt ble det opprettet identiteter for

14687 personer og av disse ble 4506 ikke funnet i dåpslistene, men første gang i vielsesliste, inn- og utflytterliste, konfirmasjonsliste, dødfødtliste eller begravelsesliste. Databasen er organisert med en tabell for hver av de disse kildene og med en tilleggstabell for personnummerne. Mellom individpostene i disse tabellene er det opprettet pekere slik at livshistorier kan lages ved å kombinere informasjon fra de ulike tabellene. Det ble ikke anvendt programvare for automatisk lenking i prosjektet, lenkingen måtte dermed foretas interaktivt. Valget av et forholdsvis enkelt relasjonsdatabasesystem (Microsoft Access) synes vellykket, men det synes like klart at mye arbeid kunne vært spart dersom alle navn hadde vært standardisert før arbeidet med lenking startet.¹⁷

Ovenstående data fra kirkebøkene ble supplert med data fra ekstraskattmannntallet 1762 og folketellingene 1801, 1865, 1875 og 1900. Her ble det oppnådd svært gode resultater ved sammenlenking av kildene på individnivå. I folketellingene fra 1801, 1865 og 1900 ble fra 93,4 til 97,1 prosent av individene fra kirkebokstabellene identifisert, mens resultatet for 1875-tellinga, som ikke er transkribert for hele landet, var at 90,1 % ble identifisert. Siden alder mangler i manntallet, var lenkingen til dette mer problematisk, slik at kun to tredjedeler av de registrerte mennene fra kirkebøkene ble gjenfunnet i første omgang. Ved å supplere med informasjon fra pantebøkene kunne imidlertid hele 97,5 % lenkes til manntallet.¹⁸ For å kunne beregne ulike demografiske rater er det av stor betydning at livsløp kan avsluttes på en entydig måte. I så måte er det betryggende at hele 3427 av de 3716 personene i folketellinga 1900 kunne gis personnummer ut fra kirkebøkene.

Også mellom de ulike kirkebøkene ble det oppnådd gode lenkingsresultater. For å teste dette ble en kohort valgt ut: Alle som var observert i Rendalen fra 1815 til 1824. Siden dette var midt i den perioden databasen dekker, skulle sjansene være gode for å finne kohortens medlemmer bde i dåpsliste og begravelsesliste. Det viser seg at nesten halve kohorten er identifisert i disse to kildene så vel som i vielseslistene. I tillegg er nesten en femtedel identifisert i dåps- og

begravelsesbøkene, men ikke som viede. En stor del av disse var unge, sannsynligvis døde de fleste av dem som ugifte. Om lag femten prosent er identifisert i dåps- og vielsesliste, mens under fem prosent er identifisert som viede og begravede i Rendalen. Dette kan settes i sammenheng med at det var mindre innflytting til Rendalen enn utflytting fra stedet. At kirkebøkernes flyttelister er mangelfulle, kunne i noen grad kompenseres ved å lenke mot emigrantlistene og de landsdekkende folketellingene 1865 og 1900.¹⁹ Nesten ingen ble kun identifisert i begravelseslistene.

Rendalsdatabasen har blant annet vært anvendt til studiet av overgangen fra arrangerte ekteskap til de unges egne valg og til studiet av nedgangen i spedbarndødeligheten på landet.²⁰ I og med at folketellinga 1910 inneholder fødselsdato og nå er transkribert, er det sannsynlig at Rendalsdatabasen kan føres lenger fram i tid med tilsvarende sikker gjenfinning av personer forutsatt at også kirkebøker fra 1900-tallet registreres. Å bygge ut denne databasen med informasjon fra folketellinga 1910 og kirkebøkene i tidlig 1900-tall synes dermed å være en kurant oppgave – mye av lenkingen kan med fordel automatiseres siden fødselsdato fins i de fleste kildene fra denne perioden og navnene er standardisert. Studiet av mortalitetsnedgangen, fallende fødselstall og inngåelse av ekteskap kan dermed føres inn i en ny fase av den demografiske overgangen. Migrasjonsstudier derimot vil i stor grad være avhengig av en database som dekker større områder.

Asker og Bærum – Demolink

Den digitale demografiske databasen for Asker og Bærum er ulik Rendalsdatabasen blant annet fordi den ble bygget opp uten forutgående manuell lenking og fordi folketellingene var det sentrale kildematerialet. Eli Fure valgte dette lokale området for sine demografiske studier først og fremst fordi prestegjeldet har bevart nominative utgaver av folketellingene 1815, 1825 og 1835. I instruksene til folketellingene 1815 til 1855 var det ingen regler om innsending av navnelister til Tabellkontoret i Indredepartementet,

og derfor er slike kladder som regel gått tapt.²¹ Folketellingene 1801, 1865 og 1875 var også utgangspunkt for lenkingen, det samme gjelder dåps-, vielses- og begravelleslistene i kirkebøkene 1814 til 1878 og gårdsmatriklene fra 1826, 1838, 1866 og 1886. Den befolkninga som vokste fra 4600 innbyggere i 1801 til nærmere 10 000 mot slutten av undersøkelsesperioden, er representert i kildene med om lag 100 000 personinnførsler, og enda flere i databasen fordi personer med doble fornavn eller etternavn ble representert med to dataposter. Slik gjøres lenkingen sikrere ved at Kari Anne Hansdatter kan finnes både under Anne og under Kari i de sorterte listene på dataskjermen. For å lette samsorteringen er navnene standardisert ved å fjerne stavforskjeller som skiller mellom ulike variantstavinger av hva vi vanligvis oppfatter som samme navn. Sorteringen vil for eksempel ikke skille mellom navn som 'Ane' og 'Anne'. Det er med andre ord enighet om at navnene bør standardiseres før lenkingsarbeidet begynner.

Datasystemet Demolink ble utviklet av informatikeren Lars Nygaard og anvendt av Eli Fure til å lenke individdata fra Asker og Bærum gjennom perioden 1801 til 1878. Systemet lenker ikke personpostene automatisk, men sorterer og ordner dem slik at forskeren får god oversikt og kan bestemme hvilken informasjon som vedrører samme individ via kommandoer på datamaskinens tastatur. Slik avansert sortering gjør arbeidet med interaktiv lenking mer effektivt fordi systemet presenterer mulige lenker for historikeren, som så kan avgjøre hvilken informasjon som bør lenkes innenfor den aktuelle delen av databasen. Dette sparer så mye tid at det blir mulig å lenke sammen flere kilder parallelt og behandle materialet for store kommuner eller prestegjeld – Asker og Bærum hadde hele 9480 innbyggere i 1875. Den lenkede befolkninga er altså mer enn dobbelt så stor som Rendalens, til gjengjeld dekker databasen for Rendalen et nesten tre ganger så langt tidsrom. Dette avspeiler et av historikerens evige dilemma: Et lenket befolkningsregister utvider det repertoaret av problemstillinger som lar seg studere, men er en så arbeidskrevende metode at det studerte område og tidsrom må avgrenses. En annen interessant

forskjell mellom de to databasene fra henholdsvis Hedmark og Akershus er at den førstnevnte anvender relasjonsdatabase slik man gjør for de svenske longitudinelle databasene ved Demografiska databasen i Umeå, mens den sistnevnte blir analysert med samme programvare som anvendes for tilsvarende datasett ved Stockholms historiska databas.²²

Det fins lenkesystemer som tar utgangspunkt i ektepar og andre par som f eks mor og barn under lenkingen, men Demolink er individorientert. Fure valgte å først lenke alle menn, deretter alle kvinner. De kandidatene som ble vist fram på skjermen for lenking, hadde samme fornavn og patronymikon, noe som gjør lenkingssystemet sårbart for feil i navnene. Kandidatene for lenking kan deles inn i enda mindre grupper for de store gruppene som bar de vanligste navnene. Vielsespostene var et godt utgangspunkt for å lenke informasjonen om ett og samme individ fra flere kilder, for her fins informasjon både om brudeparet selv og deres fedre slik at det var mulig å lenke til informasjon fra andre kilder om både opprinnelsesfamilien og den nye familien som ble skapt i og med vielsen. Informasjonsrike kildeinnførsler vil ofte inneholde data som gjør det mulig å avgjøre hvorvidt andre, knappere innførsler hører sammen. For eksempel kan opplysninger om Hans Hansens far i vielseslista gjøre det mulig å identifisere ham i en folketelling selv om flere personer med dette navnet hadde samme fødselsår. Dette er en av grunnene til at det jevnt over er vanskeligere å lenke sammen informasjon om enslige individer enn dem som lever sammen med sin familie.

Når mange kildeinnførsler ses i sammenheng vil man uvegerlig finne en del feil i kildene. Dette kan føre til at to sett med innførsler må slås sammen til ett livsløp fordi det viser seg at alle postene omhandler samme person, men det kan og hende at enkelte kildeinnførsler må fjernes fra livsløp fordi de ikke passer sammen med korrigert tilleggsinformasjon. For en enkelt kommune vil det som regel være overkommelig å nøste opp de endringer som er nødvendige i forbindelse med hver revisjon av lenkingsarbeidet. En fordel med en slik iterativ eller kumulativ lenkingsprosedyre er at lenkingen kan gjøres på

Fra folketellinger og kirkebøker til norsk befolkningsregister

R Knr Fornavn Patronym Bosted Kirkesogn F.Å. R År Beslektet person

The screenshot displays the Demolink software interface. At the top, a header row defines the columns: R (role), Knr (number), Fornavn (first name), Patronym (patronym), Bosted (residence), Kirkesogn (parish), F.Å. (year of birth), R (role), År (year), and Beslektet person (related person). Below this is a large table with columns for ID, name, role, address, parish, birth year, role, year, and related person. The table contains numerous entries, some with asterisks indicating specific relationships. Below the table are navigation buttons like 'Cancel selection', 'Select IERno.', and 'Go to IERno.'. On the right side, there are buttons for 'Previous ind. event', 'Next ind. event', 'Correct data', 'Finish messages', 'Previous message', and 'Next message'. Below these are windows for 'Source-entry window' and 'Main Menu'. The 'Source-entry window' shows details for a specific individual, including their name, birth date, and parents. The 'Main Menu' contains buttons for 'Test life list', 'Find life list', 'Ancest life list', 'Compose ID rec', 'Show life list', 'Merge life lists', 'Show household', 'Correct source entry', 'Split life list', 'Show source entry', and 'Quit'. At the bottom right, an 'Information window' shows 'Program: Demolink' and 'Selection mode'.

Fig. 4. Skjerm bilde i Demolink. R = rolle i begivenheten eller type begivenhet. Knr = Nummer i kilde. F.Å. = Fødselsår, merket B hvis beregnet. Beslektet person: E = ektefelle, * = barn, F = far, M = mor. Vinduer ovenfra og fra venstre: Individvindu, meldingsvindu, kildevindu, hovedvindu og informasjonsvindu.

nytt for de kildeinnførselene som ble til overs i første omgang. Da vil man ofte se innførsler som hører sammen til tross for at de har avvikende navn eller alder pga feil i kildene. I tradisjonell kortbasert familierekonstitusjon var det imidlertid så arbeidskrevende å gjenta lenkingen ut fra stadig nye kriterier at dette i praksis forbød seg selv. Og hvis lenkingen skjer for et stort område som et fylke eller et helt land vil det fort skje at man mister oversikten, fordi en endring vil føre til en lang kjede av endringer som det vil ta svært lang tid å nøste opp. Lenkene kan med andre ord påvirke hverandre analogt med en rekke dominobrikker.

Lenkingen ble som nevnt foretatt ut fra fornavn, etternavn, alder og bosted, det siste defi-

nert som tilhørighet til gårder med samme matrikelnummer. Yrke ble ikke brukt, det ville gjøre det lettere å lenke innehavere av sjeldne yrker og dermed føre til at den lenkede delen av befolkninga fikk skjev sosial sammensetning. Bosted er tilsvarende problematisk som kriterium for lenking fordi det favoriserer lenking av dem som ikke flyttet. Når det allikevel er forsvarelig å lenke ved hjelp av bosted, er det fordi mange personer med vanlige kombinasjoner av for- og etternavn ellers ikke lar seg identifisere. Vi vil gjerne tro at Hans Olsen på gården Dal født ca 1850 var den samme både i 1865- og i 1875-tellinga, men sikkert er det ikke når det fins hundrevis av kildeinnførsler med en slik navnekombinasjon. Den som lenker historiske

kilder blir stilt overfor en rekke tvilstilfeller og blir nødt til å velge å enten prioritere sikkerhet eller å lenke så mange som mulig. Problemet ble mindre fordi den gode oversikten fra mange kildeinnførsler i Demolink ga ekstra informasjon som reduserte antall tvilstilfeller, men Fure vektla i siste instans sikre lenker.

Databasen for Asker og Bærum inneholder vel 100 000 personbegivenheter for perioden 1801 til 1878. Da er for eksempel en dåp regnet som tre personbegivenheter, en for mor, en for far og en for barnet. Optimistisk uttrykt er resultatet at 84 prosent av disse kunne lenkes til minst én annen personbegivenhet. Best var resultatet for lenking mot folketellingene 1825 og 1835 med 95 prosent lenkede, disse ligger da også midt i perioden. Svakest resultat ble oppnådd for fedre i vielseslistene og for de døde etter 1866 da det var få kilder å lenke sammen. Resultatet av lenkingen er at 31230 ulike individer angivelig skal ha blitt registrert i kirkebøker, folketellinger og/eller matrikler minst en gang i løpet av perioden. Mer pessimistisk må vi konstatere at hele 45 prosent av disse kun er registrert én gang. Hovedgrunnen er sikkert migrasjon, det aktuelle området ligger nær Kristiania, som vokste raskt gjennom store deler av perioden. Dette gjør at lenkingsprosenten blir lavest i perioder med få kilder, som for eksempel fra 1835 til 1865, da det ikke fins mellomliggende, nominative folketellinger. Mortaliteten påvirker også lenkingsresultatet, mange av personene som er registrert i folketellinga 1801, var døde da de kirkebøkene som er dataført fra og med 1814 opprinnelig ble nedtegnet. Manglende lenker kan imidlertid også skyldes at det ikke var tilstrekkelig informasjon i en kilde til å finne en sikker makkerpost i supplerende kildemateriale. Det gjelder især begravelleslistene hvor den døde ofte er ført uten henvisning til noe familiemedlem og hvor alderen kunne bli ført noe omtrentlig.

En analyse av de lenkede postene viser som forventet at lenkingsresultatet var noe svakere for personer med vanlige navn. Sosial status for lenkede og ulenkede individer kunne sammenlignes i folketellingene og avslører en viss favorisering av bøndene mht antall lenkede personbegivenheter. Det skyldes nok både at disse var

mindre geografisk mobile og at de faktisk hadde flere demografiske begivenheter registrert om seg i kirkebøkene. Den interaktive lenkingen i Demolink-systemet viser nok sin største styrke i en mellomstor kommune som Asker og Bærum. I en liten kommune vil et vanlig, hyllevare-databaseprogram kunne brukes med godt resultat fordi det spares mindre tid ved å skreddersy programvaren. Skal man lenke en stor by, et fylke eller et land er det nødvendig å automatisere betydelige deler av lenkingen. Så får tilgangen til arbeidskraft avgjøre i hvilken grad den automatiske lenkingen kan suppleres med manuelle kontroller, justeringer og tillegg. Den lenkede databasen for Asker og Bærum har blitt brukt til å studere både spedbarndødelighet og navneskikker.²³

Busetnadssoge

Spesielle programpakker er utviklet for produksjon av gårds- og slektshistoriedelen i bygdebøker. «Busetnadssoge» bygger idemessig på det eldre systemet «Ættesoge», men er nyutviklet teknologisk sett slik at det både er et kraftigere verktøy og mer brukervennlig. Systemet tas med her fordi det peker framover mot lokalhistoriske prosjekter hvor det samme datagrunnlaget kan anvendes både til tradisjonell bygdebok og til demografiske studier. I utgangspunktet var formålet med systemet å rasjonalisere produksjonsprosessen fra kildegrunnlag til ferdig trykt bygdebok, og systemet har vist seg anvendelig i arbeidet med gårds- og slektshistoriebind for både Sula, Volda, Aurskog-Høland og Herøy kommuner. «Busetnadssoge» kan importere individdata fra andre transkriberte kildeutgaver, for eksempel dem vi finner i Digitalarkivet eller hos Registreringssentral for historiske data. Det gjelder blant data lagt til rette for Lesja, se oversikten i vedlegget.

Egentlig er lite ren automatikk innebygget i «Busetnadssoge», men programmet kan standardisere navn, samsortere individopplysninger fra ulike kilder og så la brukeren bestemme hvilke kildeposter som omhandler samme gård, familie og individ. I tillegg til kirkebøker og folketellinger kan transkribert informasjon fra pante-

registre, matrikler, skifteprotokoller og tingbøker og andre kilder integreres, og dermed få med bosetting og deler av befolkningen også for perioden før kirkebøkene startet. Etter ca 1800 kan det være mest realistisk å gjøre et utvalg blant kildene. Ved utnyttelse av muntlige kilder og lokalt tilgjengelig kildemateriale er det mulig å følge befolkningen helt fram til vår tid. Resultatet er en søkbar database hvor opplysningene er lenket slik som i Demolink og i Rendalsprosjektet, men i tillegg inneholder «Busetnadssøge» skreddersydde verktøy for å lage layout til den fotosatsen som brukes som trykkegrunnlag for bygdebøkene.

Navnet «Busetnadssøge» er et bevisst valg idet programvaren gjør det mulig å få med data om ulike former for bosetting i kommunen, altså ikke bare oppsitterne på gårdene og deres slekt slik det var vanlig tidligere. Bruk av informasjonsteknologi er i denne sammenhengen demokratiserende fordi det blir realistisk å få med husholdene på husmannsplassene i tidligere tider så vel som innbyggerne i den moderne bygdebyen. Et viktig grep for å få dette til kan være å gjøre grenda til det sentrale nivået i et hierarkisk topografisk system, hvor kommunen er overordnet nivå og gårder, plasser og annen type bosetting er underordnede nivåer. Innenfor denne geografiske strukturen har historikeren Arnfinn Kjelland og informatikeren Ole Martin Sørungård lagt til rette for at i prinsippet alle historiske og nålevende personer skal kunne finne sin plass – et stort puslespill som krever mange arbeidsøkter framfor datamaskinen. Her kommer det godt med at programmet inneholder rutiner for navnestandardisering og effektive søk etter personer. Sluttproduktets renommé må ivaretas gjennom å prioritere sikre lenker framfor usikre lenker, selv om de sistnevnte kunne øke representativiteten. Kryssreferanser opprettes mellom personbegivenhetene slik at individene kan kobles sammen i familiegrupper og knyttes til sine respektive bosteder. Det er enkelt å legge til informative merknader om personer, gårder og grender, for eksempel den type opplysninger som fins i mer tradisjonelle gårds- og slektshistorier.

Kommuner som skal i gang med bygdebokprosjekter, bør vurdere Busetnadssøge som alternativ til tradisjonell bokproduksjon, både fordi det sparer ressurser på veien fram mot bokmanus, og fordi det skapes en verdifull database som kan stilles til disposisjon for innbyggerne og som kan holdes ved like for framtida. Busetnadssøge har allerede vist sin slagkraft som alternativ til de tradisjonelle framstillingene av gårds- og slektshistorie for norsk landsbygd. I byene har vi ingen tilsvarende tradisjon for bosetningshistorie, men det arbeidet som er gjort med Odenses historie i Danmark, tyder på at det også i urbane strøk er rasjonelt å knytte innbyggerne til sine bosteder for å lage et digitalt kildegrunnlag for lokalhistoriske framstillinger.²⁴ Kjelland arbeider for å legge datainnholdet i Busetnadssøge til rette for befolkningshistoriske analyser som del av den allmenne lokalhistoria eller i andre studier av den demografiske utviklinga. Nyskapende vil det og være å utnytte den gode stedfestinga av demografiske begivenheter til å kartlegge flytting på mikroplanet, dvs innenfor den enkelte kommune.²⁵

Migrasjon i Troms fylke på 1800-tallet

Et annet grep er valgt av Gunnar Thorvaldsen (1995) i hans studie av migrasjon i Troms gjennom annen halvdel av 1800-tallet, den perioden da Nord-Norge utviklet seg fra å være et område med netto innflytting til å ha netto utflytting.²⁶ Her studeres for første gang befolkningen i et helt fylke ut fra folketellingsdata på individnivå. For å få informasjon om intern flytting i kommunene og sammenhengen mellom sosial og geografisk mobilitet var det nødvendig å følge grupper av enkeltindivider over tid. For dette formålet ble det utviklet programvare for lenking av to og to kilder slik at de kan utnyttes statistisk i såkalt bitudinal analyse. Forsøk med å følge personer over tre folketellinger ga så vidt magre resultater at dette ikke ble implementert i lenkingsprogrammene. Utflytterne til resten av fylket er med i prosessen. Lenkingsprosessen ble automatisert både for å kunne dokumentere den bedre, for å gjøre lenking overkommelig for

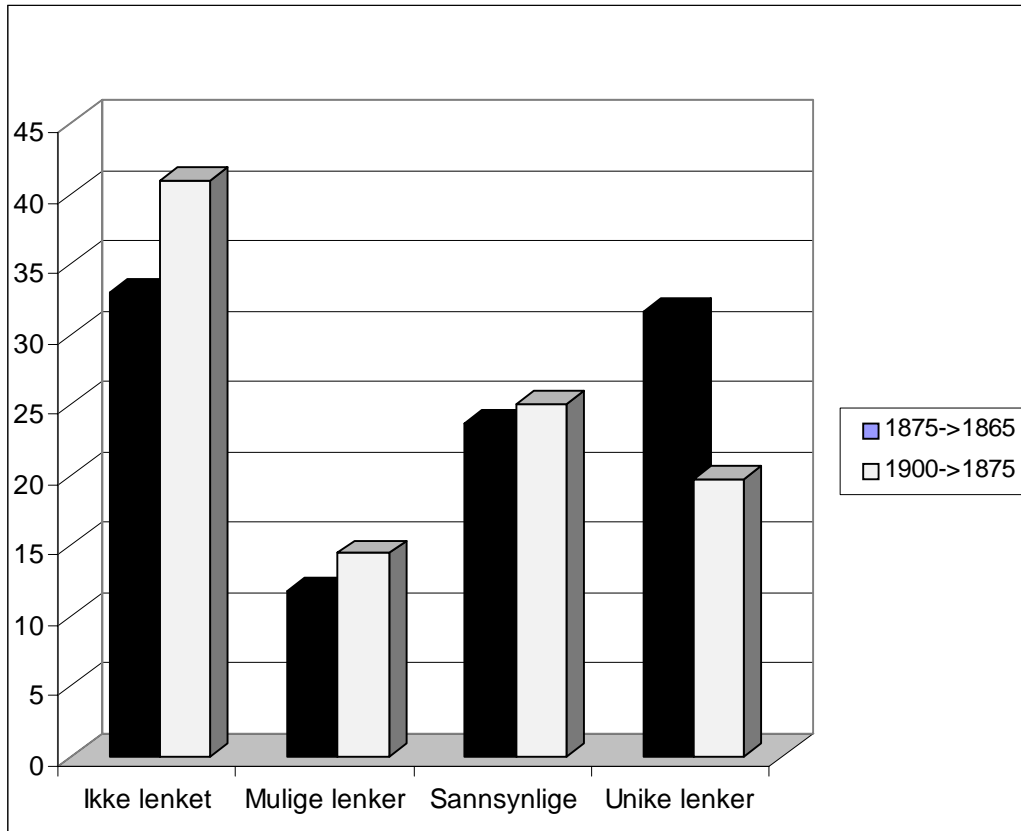


Fig. 4. Andel individer lenket fra 1900 til 1875 og fra 1875 til 1865 i folketellingene for Målselv, Balsfjord og Karlsøy. Barn født etter første telling utelatt.

flere kommuner og for å teste ut denne type teknikker på et norsk kildemateriale for første gang.

Lenkingen tok utgangspunkt i en kombinasjon av fulltekst og kodet versjon av den enkelte folketelling,²⁷ i form av filer med oppslag ut fra individenes rekkefølge i originalkilda og en søkenøkkel sammensatt av personenes kjønn og navn. Navna er standardisert ifølge en rekke regler og en unntaksliste. En radikal normalisering skal sørge for at personer blir kandidater for lenking selv om skrivemåten av navna i utgangspunktet er relativt ulike, mens en mer moderat form skal gi mulighet for å skille mellom personer.

Lenkingsprogrammet leste inn grupper av personer med samme kjønn og samme forbokstav i navnet fra to kilder. Det forsøkte først å finne fram til par av personposter fra de to kildene som

både er unike og like, dvs har stor grad av likhet mellom kildene og stor grad av ulikhet i forhold til andre poster i samme kilde. Disse ble merket som «sikre» lenker. Deretter ble resten av postene gjenstand for såkalt «matchscoring» ved at det gis positive poengvekter for likhet og negative for ulikhet mellom alle par av poster fra de to kildene, og postene i kilde én lenkes med de poster i kilde to som gir flest poeng. Ved poenglikhet avgjør grad av likhet i original form av navn, eller det ble foretatt et tilfeldig valg. Det krevdes et minimum av poeng for at par av personposter kunne lenkes. Programmet ble først kjørt med utgangspunkt i fornavn og deretter med utgangspunkt i etternavn for å kunne lenke personer hvor et av navna avvek mye fra en kilde til en annen. Til sist ble det brukt spesielle regler som ser på ektepar hvor bare den ene har blitt

lenket og slik at den andre ektefellen fikk en ny mulighet for å komme med i det lenkede utvalget.

Etter lenkinga ble kodede data fra de sammenlenkede folketellingene overført til en statistikkpakke slik at det blir mulig å lage statistiske oversikter med utgangspunkt i den lenkede del av befolkninga, eller befolkninga i hver telling. Dette gjorde det mulig å undersøke i hvilken grad den delen av befolkninga som ble lenket, er representativ for hele befolkninga på tellings-tidspunktet mht variabler som kjønn, alder, yrker, fødesteder mv. Ser vi framover i tid, blir folk fra den første tellinga borte både pga dødsfall og utflytting. Dødeligheten ble i noen grad kompensert for ved å utnytte informasjon fra begravelleslistene for Målselv og Balsfjord. Utflyttere som var født i kommunen og oppholdt seg i Troms, er også med i grunnlagsmaterialet, mens andre utflyttere ikke kan lenkes. Ser vi bakover i tid, vil innflyttere som har kommet etter at den første folketellinga ble avholdt, ikke komme med i prosessen. Dessuten må vi se bort fra dem som var født i tidsrommet mellom tellingene. I tillegg til dette «naturlige» bortfallet av personer fra det lenkede materialet, kommer et antall individer som programvaren ikke klarer å lenke fordi datagrunnlaget ikke er tilstrekkelig presist eller programmene ikke klarer å utnytte den informasjon som finnes på en adekvat måte. Dessuten vil noen personer feilaktig komme med som lenkede fordi algoritmene skaper falske lenker pga forveksling.

Ser vi bakover i tid, har det lyktes å lenke om lag en tredjedel av personene i folketellinga 1875 til personposter i forrige folketelling. Sett framover i tid blir lenkingsprosenten en smule lavere, noe som skyldes manglende informasjon fra begravelleslistene kombinert med større ut- enn innflytting fra området. Mellom 1875- og 1900-tellingene er den lenkede andelen ca 8% lavere både pga mer flytting, flere personer med samme navn og større avvik i skrivemåte av navn, aldersoppgaver osv. En klar majoritet av de lenkede personopplysningene kommer i kategoriene «sikre» eller «sannsynlige», mens 11% til 14% av lenkene kalles «mulige» og kan betraktes som tvilstilfeller. Andelen lenkede var

høyest i Målselv og lavest i Karlsøy, noe som både skyldes at der var mindre flytting blant befolkninga på innlandet og at folk her hadde større spredning mht navn og dermed var enklere å lenke. En sammenligning med tall for fødte og døde fra *Folkemengdens bevegelse* viser nemlig at bortfallet av personlenker må skyldes lenkingsproblemer i tillegg til flytting. Siden vi kun har sikre tall for netto ut- og innflytting, lar ikke bortfallet seg fordele nøyaktig på de to kategoriene. Stikkprøver tyder allikevel på at hovedårsaken er flytting.

I motsetning til familierekonstitusjon basert på kirkebøker, har vi kjennskap til hele populasjonen når lenkinga baserer seg på folketellinger. Det viser seg at utvalget av lenkede personer ikke er noe representativt utvalg av hele populasjonen i folketellingene. Der er en klar overrepresentasjon av etablerte og bofaste mennesker med jordeiendom og tilknytning til egen kjernefamilie. Unge, ugifte mennesker med tjener- eller arbeiderstatus havner typisk blant de ulenkede. Dette har nær sammenheng med disse gruppens større geografiske mobilitet, men det kan også skyldes at de har egenskaper som gjør dem vanskelige å lenke, som for eksempel vanlige navn. Bruk av teknikker for lenking basert på relasjoner til andre personer er vanligvis til liten nytte ved lenking av disse gruppene. Der var liten forskjell på hvilken andel av hhv same- og de norske som lot seg lenke, mens bare noe over halvparten av kvenene ble lenket. Med unntak av Målselv var kvinnene noe vanskeligere å lenke enn menn, kanskje et resultat av at kvinnene flyttet mer og hadde mindre navnespredning utenfor bøndenes rekke. I Karlsøy var det større bortfall enn vanlig blant folk over 60, noe som kan ha sammenheng med at det var vanskeligere å greie seg som fisker i så høy alder. Tilknytninga til jordeiendom gjorde søringene nesten like lette å lenke som de innfødte, med unntak av i Karlsøy. Folk født i Nordland og Finnmark var vanskeligst å identifisere i to tellinger, antakelig fordi de flyttet inn eller ut av Troms mellom tellingene.

Juridiske begrensninger

Det vil være begrenset hvor langt fram i tid informasjon fra et historisk befolkningsregister kan gjøres allment tilgjengelig. Personregisterloven hjemler egentlig bare personvern for avdøde personer, mens Statistikkloven er strengere fordi den klausulerer bruk av statlig innsamlet kildemateriale i ett hundre år. Slik har det seg at folketellinga 1910 for Norge ikke kan gjøres allment tilgjengelig før i 2010, mens folketellinga 1912 for Bergen by er i allmenn bruk. Personregisterloven omtaler noen typer informasjon som særlig følsom med et spesielt vern, det gjelder blant annet data om etnisk opphav. Selv om loven altså strengt tatt ikke gjelder for avdøde, er det rimelig å vente med å gjøre følsomme data allment tilgjengelig inntil en rimelig lang periode har gått, for eksempel en generasjon eller 30 år etter dødsfallet. Særlig følsomme data, for eksempel om sinnssykdom, bør neppe legges ut før etter hundre år, dvs samme klausuleringsperiode som i Statistikkloven.

Noe annerledes stiller dette seg for forskning. For statistiske formål kan man ofte klare seg med anonymiserte data. De som lenker sammen befolkningsregistret, må selvsagt ha tilgang til alle data – fulle navn, adresser, fødselsdato osv – og være underlagt streng taushetsplikt, men dette kan fjernes før forskerne får tilgang til det ferdig lenkede datasettet. For sikkerhets skyld må de i tillegg erklære at ingen identifiserbar informasjon skal bringes videre til tredjepart. Bruken av historiske personregistre må langt på vei følge mer generelle juridiske regler for anvendelsen av personhistorisk kildemateriale. Disse er behandlet i et spesialnummer av *Heimen* i 2005, se især Tor Breiviks artikkel.

Et nasjonalt historisk befolkningsregister?

Både det løpende, nasjonale folkeregistret og de lokale, historiske registrene har vært og er grunnlag for en rekke forskningsprosjekter. På bakgrunn av de mulighetene et sammenhengende register over befolkningen gir til å legge komplette livsløp til grunn for analysene, har Norsk demografisk forening tatt initiativ til å ut-

rede hvordan et landsdekkende befolkningsregister kan utvikles bakover i tid. En gruppe med representanter for Riksarkivet, universitetene, Norsk regnesentral og Statistisk sentralbyrå arbeider med planene, sammen med en referansegruppe med videre representasjon. Det er enighet om at folketellinga 1910 vil stå sentralt ved opprettelsen av et slikt register fordi den er transkribert til maskinleselig format og vil bli stilt til allmenn disposisjon i 2010. Denne tellinga inneholder fødselsdato, og all erfaring viser at dette gjør det lettere å knytte til data fra andre kilder. Mange dåps-, vielses- og konfirmasjonslister fra 1800-tallet inneholder fødselsdato, og i den grad disse er transkribert kan de lenkes til 1910-tellinga. Mye av dette arbeidet er det mulig å automatisere fordi vi også har tilgang til lister med standardiserte navn, men det er også aktuelt å la slektsgranskere og andre interesserte bidra til lenkingen via Internett.

Mens folketellingene altså er sperret for annet enn statistisk bruk i ett hundre år, kan kirkebøkene anvendes etter seksti år (dåpslistene til 1930). Dette gjør det aktuelt å transkribere de kirkebøkene fra perioden fram mot 1930 som nå er gjort tilgjengelig via Digitalarkivets hjemmesider slik at de kan inngå i det historiske befolkningsregistret.²⁸ Slik kan vi både bygge et historisk befolkningsregister som dekker hele landet og som åpner opp for historisk-demografisk forskning etter år 1900. Tidlig 1900-tall er en understudert periode i vår befolkningshistorie til tross for at avgjørende viktige demografiske endringer fant sted, det er nok å nevne nedgangen i fødselstallene og den siste delen av fallet i spedbarndødeligheten. Selvsagt er det ønskelig å inkludere informasjon fra folketellingene etter 1910 også, men om det er god forskningsøkonomi avhenger av om materialet kan stilles til disposisjon for statistisk forskning slik dette er hjemlet i Statistikkloven.

Gunnar Thorvaldsen

Professor

Adr.: Registreringsentral for historiske data

Universitetet i Tromsø

9037 Tromsø

E-post: gunnar.thorvaldsen@sv.uit.no

Noter

- 1 Thorvaldsen 1996; Wrigley et al 1997.
- 2 Haavet 1982; Engelsen 1983.
- 3 Helgheim 1977.
- 4 Thorvaldsen 2007.
- 5 Janssens 1993; Thorvaldsen 1995b.
- 6 Soltvedt 2004.
- 7 Soltvedt 2004; Kok 2006.
- 8 Soltvedt 2004: 165.
- 9 Johansen 2002: 178.
- 10 Soltvedt 2004: 167f.
- 11 Sjøbye 1998.
- 12 Bjerve og Nordbotten 1969. Se også <http://www.ssb.no/vis/emner/02/02/om.html>
- 13 Soltvedt 2004: 183f.
- 14 Det såkalte Decode-prosjektet. Se http://en.wikipedia.org/wiki/DeCODE_Genetics
- 15 Dyrvik 1983: 93ff.
- 16 Se <http://www.ddb.umu.se/>. Registrering av församlingar, Norra regionen. Lest 23–01–08.
- 17 Gjølseth 2000: 67f og 83.
- 18 Bull 2006: 28.
- 19 Bull 2006.
- 20 Bull 2006; Sogner et al 2002.
- 21 Tranberg 1994.
- 22 For Rendalens del anvendes relasjon databasen Microsoft Access, mens Asker og Bærum-dataene ble behandlet med statistikk systemet SAS.
- 23 Fure 1990; 1999; 2000.
- 24 Johansen 1999.
- 25 <http://tilsett.hivolda.no/ak/BSS/Busetnadssoge.html> inneholder supplerende informasjon. Se også Andersen 2006 og <http://www.rhd.uit.no/ahc/abstracts.html#fr20>. Jeg takker for kommentarer fra Arnfinn Kielland.
- 26 http://www.arkivverket.no/URN:kb_read
- 27 Thorvaldsen 1995a.
- 28 Erikstad og Thorvaldsen 2006.

Litteratur

- Andersen, K. A. (2006). «Busetnadssøge.» *Heimen* 43 (3): 244–247.
- Anderson, M. (1977). «Some problems in the use of census type material for the study of family and kinship systems.» *Time, Space and Man. Essays on Microdemography*: 69–80.
- Bean, L. L. et. al. (1980). «The Genealogical Society of Utah as a Data Resource for Historical Demography.» *Population Index* 46 (1): 6–19.

- Bouchard, G. (1992). «Current Issues and New Prospects for Computerized Record Linkage in the Province of Québec.» *Historical Methods*, 67–73.
- Breivik, T. (2005). «Arkivaretikk og brukeretikk. Bruksregler for offentlige arkiver i et etisk perspektiv.» *Heimen* 42 (2): 99–109.
- Bull, H. H. (2006). *Marriage decisions in a peasant society: the role of the family of origin with regard to adult children's choice of marriage partner and the timing of their marriage in Rendalen, Norway, 1750–1900*. Faculty of Humanities, University of Oslo. Nr. 268.
- Davenport, D. P. (1984). «Tracing rural New York's out-migrants, 1855–1860.» *Historical Methods*.
- Dijk, H. v. (1977). «Longitudinal cohort analysis of the population of Eindhoven.» *Time, Space and Man. Essays on Microdemography*. S. Söderlund. Umeå: 245–250.
- Drake, M. (1969). *Population and society in Norway 1735–1865*. Cambridge, Cambridge University Press.
- Dyrvik, S. (1983). *Historisk demografi. Ei innføring i metodane*. Bergen.
- Engelsen, R. (1983). «Mortalitetsdebatten og sosiale skilnader i mortalitet.» *Historisk tidsskrift* (2): 161–202.
- Erikstad, M. og G. Thorvaldsen (2006). «Statistikk basert på individdata fra folketellingene, nye muligheter.» *Heimen* 43 (1): 41–54.
- Fure, E. (1990). Oppkalling og familiementalitet. *Historisk tidsskrift*, 69, 146–162.
- Fure, E. (1999). «Nedgangen i dødeligheten i Asker og Bærum 1733–1878.» *Heimen* 36 (3): 177–186.
- Fure, E. (2000). «Spedbarndødelighet og sosiale forskjeller i Asker og Bærum 1814–1878. En metode for studier på individnivå.» *Heimen* 37 (4): 293–304.
- Fure, E. (2000). «Interactive Record Linkage. The Cumulative Construction of Life Courses.» *Demographic Research* 3.

- Genealogy. (2005). «The Genealogical Society of Utah as a Data Resource for Historical Demography.» Retrieved 2 Nov, 2007, from <http://www.genuki.org.uk/big/Linking/index.html>.
- Gjelseth, M. (2000). *Relasjonsdatabaser som verktøy i en historisk-demografisk studie*. Historisk institutt. Oslo, Universitetet i Oslo.
- Gutmann, M. P., Ed. (1992). *Historical Methods: Theme issue on record linkage*.
- Helgheim, J. «Sjeleregistre som kjeldemateriale». *Heimen 1977-1*: 269–274.
- Hogan, D. P. K., David I. (1985). «Longitudinal Approaches to Migration in Social History». *Historical Methods*. 20–29.
- Haavet, E. (1982). *Avvik eller uhell? Ugifte foreldre omkring 1800 – en sosial analyse*. Historisk institutt. Bergen, Universitetet i Bergen.
- Janssens, A. (1993). *Family and social change. The household as a process in an industrializing community*. Cambridge.
- Johansen, H. C. (1975). *Befolkningsudvikling og familiestruktur i det 18. århundre*. Odense.
- Johansen, H. C. (1984). *Record-linkage units: Persons, families, dwellings?* Seminarinnlegg ved Umeå Universitet.
- Johansen, H. C. (1999). «Urban social and demographic reconstitution. The case of eighteenth century Odense.» *History and Computing* 11 (1–2): 115–128.
- Johansen, H. C. (2002). *Danish population history*. Odense.
- Katz, M. og J. Tiller (1972). «Record-Linkage for Everyman: A Semi-Automated Process.» *Historical Methods Newsletter* 5: 144–150.
- Kitt, A., D. Daulton & E. Reis. *The Reconstitution of Viana do Castelo*. Southampton, Department of Social Statistics, University of Southampton, 1990.
- Kok, J. (2006). «Sources for the historical demography of The Netherlands in the 19th and early 20th centuries». *Fertility developments in Taiwan and the Netherlands, 1850–1950*. Amsterdam.
- Kruken, K. (1982). Frå reformasjonen til den nordiske namnerenessansen. *Norsk personnamnleksikon*, Oslo.
- Langholm, S. (1976). «On the scope of micro-history.» *Scandinavian journal of history*: 1–24.
- Lien, E. (uten år). *Navnematerialet i folketellinga fra 1801*. Trondheim.
- Matchware, T. *Autostan. Generalized Standardization System*. Burtonsville, Maryland, 1996.
- McCaa, R. (1989). «Populate. A Microcomputer Projection Package for aggregative Data Applied to Norway, 1736–1970.» *Annales de démographie historique*: 288–298.
- Miller, R. og G. Thorvaldsen (1997). «Beyond Record linkage: Longitudinal Analysis of Turn-of-the-century Inter-urban Swedish Migrants.» *History and Computing* 9: 106–121.
- Nygaard, L. (1985). *Fra historiske kilder til persondatabase*. Universitetet i Oslo.
- Nygaard, L. (1992). «Name standardization in record linkage: an improved algorithmic strategy.» *History and Computing* (2).
- Oldervoll, J. (1982). *Maskinell kopling av eit historisk materiale*.
- Pitkänen, K. (2003). *Användning av kommunionsböcker att utarbete en folkräkning för år 1890*. Helsinki.
- Pouyez, C. R., Raymond og F. Martin (1981). «The Linkage of Census Name Data: Problems and Procedures,» *Journal of Interdisciplinary History*, XIV:I (Summer): 129–152.
- Roberts, E. (2003). «The North Atlantic Population Project: An Overview.» *Historical Methods* 36 (2): 80–88.
- Ruggles, S. (2003). «Linking Historical Censuses: A New Approach.» *IMAG workshop*, <http://www.nappdata.org/imagpapers/ruggles.pdf>.
- Ruggles, S. et. al. (1996). «Distributing Large Historical Census Samples on the Internet.» *History and Computing* 9: 145–159.

- Schofield, R. (1992). «Automatic Family Reconstitution. The Cambridge Experience.» *Historical Methods*: 75–79.
- Schurer, K. et al (2003). *The Victorian Panel Survey – A scoping study for the ESRC*. Colchester, UK Data Service, University of Essex.
- Singapore, Department of Statistics. (2001). *Combining Survey and Administrative Data for Singapore's Census of Population 2000*. 53rd session of the International Statistical Institute, Seoul.
- Sogner, S. et. al. (2002). «The rural reward. Infant mortality in Norway during the demographic transition. A case study.» *Historical Studies in Mortality Decline*. W. Hubbard et al (eds). Oslo, Det Norske Videnskaps-Akademi/Novus forlag. 3: 79–95.
- Soltvedt, K. (2004). «Folkeregistre og personnummersystemer i Norge fra 1905 til 2001.» *Folketellinger gjennom 200 år*. K. Soltvedt. Oslo, Statistisk sentralbyrå: 159–189.
- Steckel, R. H. (1988). «Census matching and migration. A research strategy.» *Historical methods*: 52–60.
- Szreter, S. (2007). «The right of registration: Development, identity registration, and social security – A historical perspective.» *World Development* 35 (1): 67–86.
- Søbye, E. (1998). «Jødeforfølgelsene under den annen verdenskrig: Et mørkt kapittel i statistikkens historie?» *Samfunnsspeilet* 12 (4): 2–17.
- Thorvaldsen, G. (1981). «Normalisering av personnavn», i Eirik Lien (red). *De nordiske datalingvistikkdagene*. Trondheim.
- Thorvaldsen, G. (1995). «The Encoding of Highly Structured Historical Sources.» *Computers and the Humanities* 28: 301–305.
- Thorvaldsen, G. (1995a). *Migrasjon i Troms i annen halvdel av 1800-tallet. En kvantitativ analyse av folketellingene 1865, 1875 og 1900*. Registreringsentral for historiske data. Universitetet i Tromsø.
- Thorvaldsen, G. (1995b). «Longitudinal sources and longitudinal methods – Studying migration at the Stockholm Historical Database.» *Swedish urban demography during industrialization*. A. T. Brändström, L-G. Umeå, Umeå universitet.
- Thorvaldsen, G. *Håndbok i registrering og bruk av historiske persondata*. Oslo, Tano Aschehoug 1996.
- Thorvaldsen, G. (1998). «Historical Databases in Scandinavia.» *The History of the Family. An International Quarterly* 3 (3): 371–383.
- Thorvaldsen, G. *Databehandling for historikere*. Oslo, Tano Aschehoug 1999.
- Thorvaldsen, G. (2006). «Away on Census Day. Enumerating the Temporarily Present or Absent.» *Historical Methods* 39 (2): 82–96.
- Thorvaldsen, G. (2007). «An International Perspective on Scandinavia's Historical Censuses.» *Scandinavian Journal of History* 32 (3): 237–257.
- Tranberg, A. (1994). *Folk og fant. – 3. utg.* Oslo, Norsk lokalhistorisk institutt.
- Tysdal, O. (1990). *Befolkningsutviklingen i Sandnes og deler av Høyland 1801 til 1875*. Bergen.
- Tysdal, O. (2003). «Skudenesdatabasen: en dokumentasjon av emigrasjon til Nord-Amerika 1865–1965». Stavanger, *Migrasjonsprosjektet i Rogaland*: 49–60.
- Vestre, M. (1985). *FONEQ, et navnestandardiseringsystem for personnavn i historiske kildetekster*. Universitetet i Oslo.
- Vinje, F.-E. *Norske familienamn, Oslo* NRK, 1973.
- Wrigley, E. A. *English population history from family reconstitution 1580–1837*. Cambridge University Press 1997.
- Wrigley, E. A. and Roger Schofield (1973). «Nominal record linkage by computer and the logic of family reconstitution.» *Identifying People in the Past*. E. A. Wrigley (ed). London.
- Wåhlin, S. (1977). *Variantstavningar av släktnamn*. Göteborg.

Fylke	Prestegjeld / kommune	Periode	Periode og referanse
Østfold	Moss	1776–1825	Jannecke Lahn, 1986
Akershus	Asker og Bærum	1801–1878	Eli Fure, 2004
Akershus	Asker	1801–1835	Frøydis Dahl Bryhn, 1986 Oslo
Akershus	Nesodden	1710–1825	Bente Nordbø, 1980
Akershus	Ullensaker	1733–1845	Björg Halvorsen og Kari Indseth, 1975 Steinar Kneppen, 1976
Akershus	Gjerdrum		Lajos Juhasz, Ståle Dyrvik
Hedmark	Rendalen	1732–1900	Sølvi Sogner, 1976/Hans H Bull 2006
Oppland	Lesja	1820–1900	Arnfinn Kjelland, 1987–96
Buskerud	Sandsvær	1750–1801	Dag Kristoffersen, 1983
Telemark	Bø	1727–1815	Hege Brit Randsborg, 1979
Agder	Bygland	1801–1900	Reidar Vollen, 1988
Rogaland	Avaldsnes	1758–1801	Sven Gabrielsen Øygarden, 1979
Rogaland	Sandnes og Høyland	1801–1875	Olav Tysdal, 1990
Rogaland	Skudenes	1865–1965	Olav Tysdal, 2003 Stavanger
Rogaland	Del av Petri sogn, Stavanger	1835–1875	Marit Karin og Jan Alsvik, 1982
Hordaland	Etne	1665–1801	Ståle Dyrvik, 1971
Hordaland	Os	1660–1801	Jan Oldervoll, 1970
Hordaland	Ølen	1700-tallet	Anders Haugland, 1976
Hordaland	Lindås	1660–1801	Erik Thomassen, 1973
Hordaland	Lindås	1801–1875	Lars Bastesen, 1976
Hordaland	Stord	1666–1801	Svein Gunnar Koppang, 1974
Hordaland	Hamre	1701–1801	Anne Østerbø Kåstad, 1981
Hordaland	Halsnøy	1875–1920	Merete Hofslett, 1980
Sogn og fjordane	Luster	1732–1800	Per Assev, 1991 Oslo
Sogn og fjordane	Amafjord, Vik	1801–1900	Rasmus Sunde, 2001
Møre og Romsdal	Sula	1690–2007	Ole Martin Sørungård, 2004–06
Møre og Romsdal	Volda	1742–2007	Olav Myklebust
Møre og Romsdal	Norrdal	1737–1865	Terje Linge, 1977
Romsdal	Nesset	1665–1801	Bjame Skorgen, 1973
Sør-Trøndelag	Baklandet	1727–1801	Nina Hveem Carlsen, 1986 Oslo
Sør-Trøndelag	Vår Frue sogn, Trondheim	1743–1825	Margunn Skjei Knudtsen, 1997, Trh
Nord-Trøndelag	Overhalla	1664–1801	Torkel Christiansen, 1971
Nordland	Vega	1765–1865	Kirsti Holand, 1986 Oslo

Vedlegg: Norske prestegjeld hvor det er laget historiske befolkningsregister ved hjelp av familie-rekonsitusjon eller tilsvarende metoder. Uten utgiversted: Bergen.