

Forsvarets arkiver som kunnskapsressurs - digitalisering og transkribering av militære helsedata

*Kayla Tungodden, forskningskoordinator, Arkivverket,
kayhol@arkivverket.no*

Elin Anita Fadum, seniorrådgiver Forsvarets helseregister og seniorforsker ved Legeforskningsinstituttet, efadum@mil.no

*Anna Malmø-Lund, seniorrådgiver ved Arkivverket,
annmal@arkivverket.no*

*Jeanette Engeland, forsker og fagkonsulent ved Nasjonalt senter for aldring og helse, Sykehuset i Vestfold,
jeanette.engeland@aldringoghelse.no*

Eirik Jansen, tidligere seniorrådgiver ved Forsvarets veteran­tjeneste.

*Vegard Skirbekk, professor ved Folkehelseinstituttet; Nasjonalt Senter for Aldring og Helse; UiO og Columbia University,
vegard.skirbekk@fhi.no*

*Leif Åge Strand, seniorrådgiver ved Forsvarets helseregister,
leifstrand@mil.no*

*Kristine Vejrup, forskningskoordinator, Forsvarets sanitet,
kvejrup@mil.no*

Sammendrag

Arkivverket har store arkiver avlevert av Forsvaret som dekker mange ulike militære avdelinger og typer informasjon. Forsvarets arkiver inneholder nasjonalt dekkende data som gir et representativt bilde på viktige helse­mål ved sesjon for den mannlige populasjon tilbake til fødselskohorter fra 1941.

For å tilgjengeliggjøre informasjonen som ligger lagret i arkiver er det behov for å digitalisere dokumenter. Arkivverkets Digitalarkiv gir en enestående mulighet for å digitalisere og tilgjengeliggjøre dokumenter for ulike formål. For at data skal kunne benyttes videre i forskningsprosjekter må de hentes ut, enten ved manuell transkribering, eller ved bruk av kunstig intelligens-modeller. Forsvaret og Arkivverket har gjennom et samarbeidsprosjekt sett på ulike løsninger for å lokalisere og tilgjengeliggjøre helseinformasjon fra Forsvarets arkiver.

Forsvarets helseregister har hjemmel for å lagre informasjon om helse og å benytte den videre med muligheten for tilkobling av informasjon fra andre administrative dataregister på individnivå. Dette representerer en enestående muligheter for forskning i områder knyttet til en rekke fagområder, som mental og somatisk helse, kognitive og nevrologiske endringer gjennom livsløpet, og sosiale, demografiske og økonomiske forhold. Forskningen kan gi vitenskapelige bidrag på en rekke områder innen epidemiologi, studier av kognisjon gjennom livsløpet og populasjonsgenetikk.

Nøkkelord: digitalisering, forsvarets arkiver, historiske helseopplysninger, forskning, tilgjengeliggjøring, kunstig intelligens

Abstract

The significance of the Armed Forces' archives as a knowledge resource - digitization and transcription of military health data

The National Archives has many of the Norwegian Armed Forces' archives, covering a range of military departments. The Norwegian Armed Forces' archives include key health data collected during conscription, which are representative of the male population nationwide as far back as birth cohorts from 1941.

Digitizing the archives is an important step in making this information available for further use. The National Archives' platform, the Digital Archives, offers a secure means of making digitized documents accessible for further use. For such data to be available and utilized in research, it must be extracted by manual transcription or machine learning models. The Norwegian Armed Forces and the National Archives have collaborated to explore methods for identifying and extracting health data from the Norwegian Armed Forces' archives.

The Norwegian Armed Forces Health Register has the authority to store and use this information, including connecting this information with other administrative registers. This presents a unique opportunity for research on a wide range of topics such as mental and somatic health, cognitive and neurological changes, and the impact of social, demographic, and economic conditions. Such research may support scientific contributions in a variety of areas within epidemiology, cognitive lifetime studies, and population genetics.

Key words: digitization, Norwegian Armed Forces archives, historical health information, research, access, artificial intelligence

Innledning

Dokumentene som er lagret i de historiske arkivene er viktig for å belyse aktuelle problemstillinger. De kan gi oss informasjon om hendelser og aktiviteter som har skjedd i tidligere tider og fortelle oss historier om enkeltpersoner og institusjoner. Dette kan øke vår forståelse av hvordan samfunnet vårt har utviklet seg, og gi informasjon om utviklingen i helsen hos befolkningen. Informasjonen som har blitt samlet inn, og nå er lagret i arkiver og registre, ble vanligvis samlet inn for et annet formål enn forskning. Grundig kvalitetskontroll og tilrettelegging av data må utføres før data fra historiske arkiver kan introduseres i forskning.

Dokumentene i Forsvarets arkiver ved Riksarkivet inneholder informasjon om mannlig fødselskull fra 1941 til 1957. Disse arkivene inneholder personell og

helseopplysninger som ikke er registrert i Forsvarets helseregister, og kan derfor i noen tilfeller være den eneste kilden til informasjon om helsen til tidligere forsvarspersonell. Arkivene kan gi innsikt i historikk om Forsvarets virke og operative oppgaver, og gi en unik mulighet til å utvide informasjonen i Forsvarets helseregisteret og samle informasjon om tidlige internasjonale operasjoner som Norge har deltatt i.

Digitaliserte og transkriberte datasett vil gjøre det mulig å beskrive utviklingen av helsen hos norske menn tilbake til 1941 – noe som må sies å være unikt. Datasett på eldre fødselskohorter vil gi grunnlag for befolkningsstudier og longitudinelle kohortstudier som undersøker prevalens og endringer over tid av for eksempel overvekt, allergi, epilepsi eller andre helseplager. Data kan kobles med andre register som Norsk pasientregister og hjerte-kar registeret for å undersøke hvordan blodtrykk, puls, fysisk form og kroppsvekt henger sammen med risiko for hjerte- og karsykdom senere i livet. Utvikling av infrastruktur for å digitalisere og transkribere informasjon fra arkivmateriale vil muliggjøre tilgjengeliggjøring av opplysninger fra Forsvarets arkiver som ikke bare brukes i helseforskning, men også gi grunnlag for historiske og samfunnsrettede prosjekter. I denne artikkelen vil vi beskrive hvordan arkivene har vært en kilde til nye data som kan benyttes til forskning.

Om Forsvarets helseregister

Forsvarets helseregister ble opprettet ved konsesjon i 2001 for å kunne samle og ta vare på relevante data om helsen til Forsvarets personell. I 2005 ble registeret lovfestet som ett av de nasjonale helseregistrene i Norge. Registerets formål er å føre tilsyn med helsen til Forsvarets personell, samt å gi grunnlag for forskning. Opplysningene i registeret er hentet fra Forsvarets medisinske journalsystem, Forsvarets personelldatabase og Forsvarets spørreundersøkelser som omhandler helse. I registret ligger også kobling til opplysninger fra Folkeregisteret og Dødsårsaksregisteret. Forsvarets helseregister aggregerer fire kategorier data: persondata, tjeneste- og administrativ data, helsedata og miljødata. Det er sammenstillingen av disse ulike kategorier av data som gjør det mulig å overvåke helseutvikling og dermed identifisere og avdekke risikofaktorer hos forsvarspersonell.

Om Digitalarkivet

Digitalarkivet er en nasjonal fellesløsning for mottak, bevaring og tilgjengeliggjøring av historiske data. Digitalarkivet brukes til å lagre, sikre, dele og gi tilgang til åpne og klausulerte arkiver. Forskere kan via Digitalarkivet få tilgang til statlige og private arkiver som er avlevert til Arkivverket. De kan også få tilgang til arkiver fra andre arkivinstitusjoner, bibliotek og museer. Over 50 institusjoner bruker Digitalarkivet som publiseringskanal for sine arkiver. Over 200 kommuner bruker Digitalarkivet som bevarings- eller publiseringsløsning.

Forsvarets arkiver i Arkivverket

Forsvaret er pålagt å avlevere sine arkiver til Arkivverket, som har store arkiver fra mange militære avdelinger. Arkivverket har militære arkiver fra kommandoer, staber og avdelinger fra infanteri, kavaleri, artilleri, sjøforsvar og flyvåpen. I tillegg omfatter arkivene militært justisvesen, utskrivingsvesen, forsyningsvesen, pensjonskasser, festninger og ingeniørvåpen. De eldste arkivene går tilbake til 1640. Som hovedregel oppbevares arkiv fra militære avdelinger i perioden 1818 til 1940 ved statsarkivene. Materiale fra departement, overkommando og sentrale staber, samt fra tiden før 1818 og etter 1945, oppbevares i Riksarkivet. Forsvarets avleverte arkiver er beskrevet i arkivportalen.no og i interne kataloger ved Arkivverket.

Forsvarets personellmapper er lagret i ulike arkiver, og det er stor variasjon i hvordan arkivene er systematisert, samt hvilken informasjon som er lagret i personellmapper. Flere av arkivene inneholder personellmapper med rulleblad og helseopplysninger for vernepliktige, militært og sivilt personell. Et militært rulleblad er et standardisert skjema med opplysninger om mannskapets navn, grad, tjenestetid, utmerkelser, informasjon om helse og fysisk og psykisk kapasitet. Personellmapper avleveres fra mannskapenes siste rulleførende enhet, som for de fleste menige og korporaler er Vernepliktsverket. Personellmapper til befal ligger lagret i arkiver tilknyttet Divisjoner. Arkivene er ordnet per årsklasse, som gjør at de fleste som er født et bestemt år er å finne i samme arkiv. Årsklassene 1941-1957 er avlevert fra alle de seks geografiske

avdelingskontorene; Oslo, Hamar, Trondheim, Bergen, Krigskommissariatet i Nord-Norge og Kristiansand. Personellmapper for eldre mannskaper finnes bare sporadisk, og mangler ofte helsekort. Personellmapper for enkelte mannskaper født 1941 til 1957 mangler imidlertid av ulike årsaker. Dette kan for eksempel være fordi personen ble kjent udyktig ved sesjon eller på rekruttskole, fordi personen døde, utvandret, eller ble overført til Heimevernet eller Politireserven. Andre årsaker kan inkludere kassering av mapper eller ukomplette avleveringer for de aktuelle årgangene. En del eldre rulleblad og forsvarsarkiver har nå blitt digitalisert.



Tysklandsbrigaden, øvelse Skandia III. Kildereferanse: Fotograf ukjent. (1952). NTBs krigsarkiv. Arkivverket/Riksarkivet, RA/PA-0797/U/Ua/L0004/0116.

Tilgang til digitale data fra Forsvarets historiske arkiver

Alle dokumenter som digitaliseres fra Forsvarets historiske arkiver lagres i Arkivverkets Digitalarkiv. Dokumentene som kan gjøres allment tilgjengelig i tråd med personvernlovgivingen er publisert åpent på Digitalarkivet.no. For nyere arkiver må enkeltpersoner søke Arkivverket om innsyn for å få tilgang til digitale dokumenter i Digitalarkivet. For tilgang til digitaliserte data fra Forsvarets historiske arkiver til forskningsøyemed, vil dette være avhengig av at data er transkribert, kvalitetssikret og overført til Forsvarets helseregister. Det må søkes om tilgang til data i henhold til rutiner for datatilgang i Forsvarets helseregister, og vil utleveres i henhold til Forskriften til Forsvarets helseregister, og etisk godkjenning fra REK.

Etiske og juridiske vurderinger for bruk av Forsvarets arkiver

Forsvarets helseregister, som et nasjonalt helseregister, har en lovpålagt forpliktelse til å sørge for at data er komplette og korrekte. Digitaliseringen av eldre helsedata fra Forsvarets arkiver legger til rette for at Forsvarets helseregister kan oppfylle sine forpliktelser ved å gjøre kvalitetsgjennomganger og tilføre opplysninger som har gått tapt.

Arkivverket er med hjemmel i arkivloven forpliktet til å bevare og sikre statlige arkiver som har betydelig kulturelt eller forskningsmessig verdi, eller som inneholder rettslig eller viktig forvaltningsmessig dokumentasjon, og gjøre dem tilgjengelige for ettertiden. Dette gjelder også Forsvarets arkiver. Arkivverket har fattet bevaringsvedtak for personellmapper i Forsvaret som er omfattet av dette prosjektet. Bevaring og avlevering av Forsvarets arkiver er hjemlet i arkivloven §§ 9 og 10. Forsvaret har søkt og fått tilgang til avleverte arkiver med hjemmel i arkivforskriften § 23.

Sjøforsvarskohortene

Opprettelse av sjøfartskohortene er et godt eksempel på helsefaglig forskning basert på informasjon hentet ut av Forsvarets historiske arkiver. Rundt siste tusenårsskifte var det en rekke oppslag i norske medier om kreft og andre helseproblemer blant Sjøforsvarets personell. Disse oppslagene dreide seg i hovedsak om medfødte misdannelser blant barn av offiserer som hadde tjenestegjort på en missiltorpedobåt (KNM Kvikk), dødsfall blant tidligere ubåtoffiserer, asbesteksponering på fartøyene og ved marineverftene, og om krefttilfeller blant sivilt ansatte på kystfortene på Rødbergodden og Meløyvær. For å undersøke påstandene nærmere ble «Prosjekt HMS Sjø» startet, og det ble etablert kohorter av sivilt og militært personell som hadde tjenestegjort i Sjøforsvaret tilbake til 1950. Hensikten var å koble disse mot Kreftregisteret og Dødsårsaksregisteret. Da Forsvarets personelldatasystem (P3) kun hadde opplysninger om tjenestetid tilbake til rundt 1990, var det nødvendig å innhente opplysninger bakover i tid fra papirbaserte kilder. Registreringsarbeidet ble utført av et team innleid fra Manpower og pågikk i tre år. Informasjon fra 39 000 personellmapper innhentet fra mobiliserings-

rullene (oversikter over personell som er mobiliseringspliktige), fra de enkelte sjøforsvarsdistrikter og marinebaser, samt fra Arkivverket, ble registrert inn i P3.

Kohortetableringen er beskrevet i egen artikkel (Strand et al., 2017). Blant de viktigste forskningsresultatene var at menn i sjøforsvaret hadde lavere dødelighet enn norske menn generelt, og at denne forskjellen syntes å vare livet ut. På den annen side hadde mannskaper på fartøyene langt større risiko for å bli diagnostisert med



Ubåt, K.N.M. KyA. Kildereferanse: Strand, R. (1955). Arkiv etter billedbladet NÅ. Arkivverket/Riksarkivet, RA/PA-0797/U/Ua/L0047/0739.

kreft, og for å dø av kreft og andre sykdommer enn annet militært personell i Sjøforsvaret (Strand et al., 2020, Strand et al., 2023). Blant sivile kvinner bidro økt forekomst av kreft i ovarier og i lunge til en generell forhøyet kreftrisiko. I en undergruppe av renholdere ble det observert økt risiko for blære- og lungekreft, noe som kan settes i sammenheng med kontakt med rengjøringsmidler, men også med røyking (Strand et al., 2018). Det er så langt publisert fem vitenskapelige artikler på kreft og dødelighet på disse kohortene, og flere er under arbeid. Resultatene er også presentert på nasjonale og internasjonale epidemiologikonferanser.

Pilotprosjekt for å vurdere muligheter for bruk av informasjon fra Forsvarets arkiver

I 2019 ble det gjennomført et pilotprosjekt som skannet og transkriberte helsedata fra Forsvarets arkiver for vernepliktige menn født i 1950. Prosjektet var et samarbeid mellom Forsvarets helseregister og Folkehelseinstituttet der hensikten var å kartlegge militære arkiver i Arkivverket, og undersøke om opplysningene var så godt bevart at de kunne overføres til Forsvarets helseregister som elektronisk søkbare data.

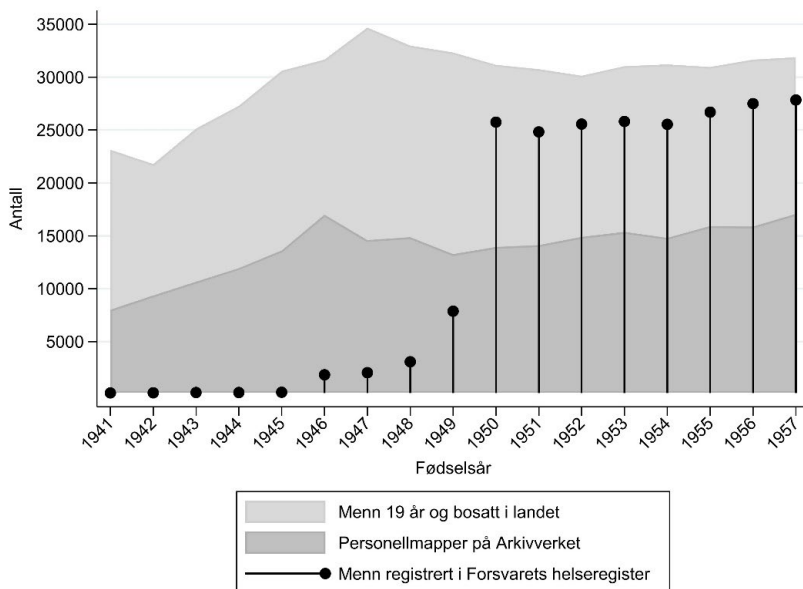
Prosjektet fokuserte på tre helsedokumenter som inneholdt målte data om blant annet høyde, vekt, blodtrykk, puls og evnenivå i 18-19 års alder, samt selvrapporterte opplysninger om egen sykdom og sykdom i familien. Disse tre helsedokumentene ble skannet, og det ble gjort en manuell transkribering av utvalgte data. Arbeidet er nærmere beskrevet i egen rapport (Fadum et al., 2020). Et utvalg av de transkriberte data er allerede brukt i forskningsprosjektet Akershus hjerteundersøkelse 1950 (Akershus hjerteundersøkelser 1950 (2023) <https://ace1950.no/>).

Data ble nøye gjennomgått før overføring til Forsvarets helseregister for å sikre at opplysningene var riktig avlest, transkribert og knyttet til rett person. Kvalitetssikringen er beskrevet i egen artikkel (Vejrup 2023 - artikkel sendt til tidsskrift). Konklusjonen etter kvalitetssikring av data transkribert fra Forsvarets arkiver er at den manuelle transkriberingsprosessen gir godt samsvar mellom kilde materialet og det transkriberte materialet. Dataene har også godt samsvar med det

som er registrert tidligere i Forsvarets helseregister og vil tilføre både informasjon på eksisterende individer i registeret, samt tilføre nye individer til registeret.

Pilotprosjektet fant at Arkivverket oppbevarer personellmapper for rundt halvparten av alle menn som var født i 1941-1957 (Figur 1). Pilotprosjektet viste at papirarkivene inneholder helsedata av god kvalitet som ikke er lagret i Forsvarets helseregister og la et godt dokumentert grunnlag for å søke økonomisk støtte fra Forsvarsdepartementet til videreføring av arbeidet med å digitalisere innholdet i personellmapper i militærarkivene i Arkivverket (Tabell 1).

Tabell 1. Oversikt over antall personellmapper for vernepliktig personell avlevert fra Forsvarets personell og vernepliktsenter til Arkivverket.



Figur 1: Estimat gjort i pilotprosjekt av antall militære personellmapper som kan lokaliseres ved Arkivverket.

Fødselsår	Antall bokser	Anslag over antall personellmapper	Innhold
1941	395 + (2?)	9131	Skjema for militær legeundersøkelse - sesjon
1942	425 + (2?)	9821	Ikke undersøkt
1943	484	11 132	Ikke undersøkt
1944	543	12 489	Ikke undersøkt
1945	619	14 237	Ikke undersøkt
1946	773	17 779	1 + 3
1947	663	15 249	Ikke undersøkt
1948	676	15 584	Ikke undersøkt
1949	603	13 869	1 + 2 + 3
1950	634	14 582	1 + 2 + 3
1951	642	14 766	Ikke undersøkt
1952	677	15 571	1 + 2 + 4: a.60-80%, b. <5%
1953	699	16 077	1 + 2 + 4: a. 80%, b. <5%
1954	673	15 479	1 + 2 + 4: a. 80-90%, b. 70%
1955	724	16 652	Ikke undersøkt
1956	721	16 583	1 + 2 + 4: a. 85%, b. 75%
1957	776	17 848	Ikke undersøkt

Inneholder:

1. Helsekort med psykologiske tester, skjema for militær legeundersøkelse – Sesjon (Blankett 4212), audiogram, militært tannkort, sykekort, skademelding
2. Egenerklæring helse
3. “Rulleblad” med høyde, fødsels- og personnummer, psykologiske tester, teknisk innsikt, regneferdigheter
4. “Rulleblad” med høyde og vekt, fødsels- og personnummer (mangler ofte i 1941), psykologiske tester, teknisk innsikt
5. Isometrisk muskelstyrke
6. Maksimalt oksygenopptak

Prosjekt for utvikling av infrastruktur for digitalisering og identifisering av veteraner

I 2021 fikk Forsvarets Sanitet (FSAN) bevilget midler fra Forsvarsdepartementet til videre digitalisering av Forsvarets papirarkiver. Oppdraget bestod av å videreutvikle en infrastruktur basert på den første digitaliseringen, men også å identifisere personellmapper til veteraner som deltok i internasjonale operasjoner fra 1947 til 1967. Ville det være mulig å etablere veterankohorter i Forsvarets helseregister basert på opplysninger i disse arkivene? Dette satte i gang et samarbeid mellom Forsvaret, Arkivverket, Sykehuset i Vestfold og Folkehelseinstituttet for å utforske og beskrive muligheter for å identifisere og hente ut informasjon om tidligere militært personell fra arkivene. Prosjektet hadde varighet fra 2021 til 2023. Andre aktører har vist interesse og bidratt inn i prosjektet, deriblant Handelshøgskolen i Bergen og Forsvarets veteran-tjeneste. Prosjektet har utviklet en infrastruktur for å skanne og laste opp alle dokumenter i personellmapper i Forsvarets eksisterende papirarkiv til Arkivverkets Digitalarkivet. Arkivene med Fødselskohorter 1941-1946 avlevert fra Vernepliktsverket var skannet og lastet opp til Digitalarkivet ved prosjektslutt. Dette utgjør rundt 75 000 personellmapper som er tilgjengelig gjennom Digitalarkivet.

Digitaliseringen er blitt utført av ansatte i Forsvarets helseregister. Skanning er gjennomført i Arkivverkets lokaler på Sognsvann i tråd med Arkivverkets digitaliseringsløype og lagret sikkert i Digitalarkivet. Arbeidet med å gjøre skannede filer søkbare ved hjelp av kunstig intelligensbasert maskinell transkripsjon er gjennomført i tråd med personvernlovgivingen. Data er beskyttet og håndtert i henhold til arkivloven, personvernloven, helseregisterloven, forskrift for Forsvarets helseregister og sikkerhåndbok for Forsvarets helseregister. Gjeldende retningslinjer for etikk i biomedisinsk forskning er blitt fulgt.

Veteraner som deltok i internasjonale operasjoner 1947-1967

Muligheter for å rekonstruere informasjon om deltagere i tidlige internasjonale operasjoner ved hjelp av Forsvarets arkiver avlevert til Arkivverket ble undersøkt ved hjelp av Arkivverkets avleveringslister og manuelle søk i arkivene. Det kan være av både helsemessig og historisk interesse å se på materialet fra samlede eldre internasjonale militære operasjoner.

Veteraner som det er lite informasjon om i Forsvarets helseregister er blant annet militært personell som deltok i Kongo ONUC 1960-1964. Forsvarets veteranstjeneste har samlet personellister fra oppdraget som viser at det var 1173 personer som tjenestegjorde i operasjonen. Det finnes ikke et eget arkiv for ONUC operasjonen, og materialet er spredt over flere ulike arkiver.

I lokaliseringsarbeidet ble det tatt utgangspunkt i mannskaplisten fra ONUC som ble koblet sammen med Arkivverkets avleveringslister som inneholder navn på over 300 000 personer som har tjenestegjort i ulike regiment. Samtidig ble det gjort manuelle søk i depot innenfor aktuelle årstall og regiment etter treff på veteraner. Ved å gjøre strukturerte søk på denne måten sikrer vi også at vi ikke utelukker personer som har hatt tjeneste i ONUC. Rett person blir bekreftet i rulleblad eller ved andre dokumenter i mappe, og søket er avgrenset til å bekrefte person, lokasjon og fødselsnummer.

Mange av veteranene som tjenestegjorde i ONUC var piloter og flyplasspersonell. Søkemetoden har tatt utgangspunkt i å starte med de ulike luftvernsarkivene der vi har ansett det som sannsynlig å lokalisere flest mulig veteraner. Mengden av dokumentasjon som foreligger i mappene varierer fra person til person, fra rikholdige mapper til mapper som knapt har innhold. Veteraner fra Luftvernet har ofte en lang karriere i Forsvaret og mye innhold i personellmappen.

Arbeidet med å lokalisere og samle veteraner er møysommelig og tidkrevende. Det er antatt at vi finner rundt 50% av veteraner som tjenestegjorde i ONUC i arkivene. I de tilfellene man finner personell som har tjenestegjort i ONUC skannes hele arkivboken som personellmappen ligger i og lastes opp til Digitalarkivet. Her vil personellet merkes slik at man kan gjøre søk og samle disse i en kohort. Videre vil utvalgt helseinformasjon og informasjon om tjenesten i ONUC transkriberes, og denne informasjonen vil samles i en datafil som overføres til Forsvarets helseregister og vil danne grunnlaget for forskning. Denne søkemetoden kan også tas i bruk for å identifisere veterankohorter fra andre tidlige internasjonale operasjoner, som for eksempel FN operasjonen i Gaza.

Bruk av kunstig intelligens for å hente ut data fra digitale dokumenter

Maskinlæring og digitalisering muliggjør kvantitativ forskning i større omfang. Prosjektet utforsker hvordan maskinlæring kan brukes for å trekke ut utvalgte helseopplysninger i Forsvarets arkiver. På denne måten kan data i Forsvarets arkiver bli relevant som kilder for nye forskningsmiljøer. Samarbeidet mellom Forsvaret, Arkivverket, Folkehelseinstituttet og aktører i helsesektoren bidrar til utvikling på området. Forsvaret samler fremdeles inn helsedata på papir eller som skannede vedlegg til journal, og dette samarbeidet kan bidra med kunnskap om teknologi som Forsvarets helseregister kan benytte for å overføre digitale opplysninger til søkbare data i registeret. Dette vil gjøre data mer tilgjengelige for forskning.



*Tysklandsbrigaden, øvelse Skandia III. Kildereferanse: Fotograf ukjent. (1952).
NTBs krigsarkiv. Riksarkivet, RA/PA-0797/U/Ua/L0004/0116b.*

I prosjektet utforsker Arkiverket muligheter for maskinell transkribering av det digitale materialet fra Forsvarets arkiver ved hjelp av kunstig intelligens. Dette kan åpne for at helseopplysninger kan gjenkjennes og transkriberes maskinelt, slik at materialet blir både elektronisk søkbart og mulig å strukturere som veterankohorter fra tidlige internasjonale operasjonene. Det arbeides i første omgang med modeller for å gjenkjenne og ta ut aktuelle helsedata som kan overføres til Forsvarets helseregister. Videre vil det utforskes om kjerneinformasjonen (personidentifikasjon) i dokumentene kan digitaliseres. Dette vil kunne gi en full digital arkivindeks på personnivå som kraftig effektiviserer imøtekomelse av individuelle forespørsler om innsyn. Kjerneinformasjonen i papirarkivet vil kunne integreres i det eksisterende digitale Forsvarets helseregister. Informasjonen vil også være mulig å koble til en mengde andre nasjonale registre. For årganger der fødselsnummer ikke er registrert vil dette kunne tilordnes via lenking og identifisering innenfor rammene av Historisk befolkningsregister (Historisk befolkningsregister (2023) <https://histreg.no/>).

Hvorfor er informasjon i Forsvarets arkiver verdifull?

Forsvarets arkiver har en verdi innenfor flere ulike områder. Gode arkiver er essensielle for å kunne gjennomføre historiske og samfunnsvitenskapelige forskningsprosjekter. Forsvarsarkiver har en naturlig plass innenfor slektsforskning. Arkiver etter andre verdenskrig har lenge vært viktige på dette området, men det er en merkbar økning i interesse for hva en far eller bestefar bedrev i Tysklandsbrigaden i 1950. Det er naturlig å tenke seg at dette vil gjelde tjeneste på 1960-tallets Gaza eller i 1980-tallets Libanon, ettersom nye generasjoner vokser opp. Men Forsvarets arkiver er også høyaktuelle til helseforskning. Arkivene er med andre ord sentrale innenfor et relativt vidt spenn av ulike felt med betydning for enkeltpersoner, beslutningstakere og folkehelse. Vernepliktskohortene i Sverige og Danmark har begge store fødselskull, inkludert menn født hovedsakelig fra 1950 og fremover til i dag. Dette er informasjonskilder som gir gode muligheter for Skandinavisk samarbeide. I de nordiske landene har alle innbyggere et unikt personlig identifikasjonsnummer som skaper muligheter for registerbasert forskning. De nordiske nasjonale helse- og sosioøkonomiske registre er ekstremt rike på data på individnivå om den skandinaviske befolkning, og noen av registrene har informasjon helt tilbake til 1950-tallet.

Verdi for Forsvarets virksomhet

Forsvarets arkiver tjener operative og praktiske formål for Forsvarets virksomhet. Ikke minst er det betydningsfullt at arkivene kan brukes av Forsvaret i forbindelse med erfaringsoverføring og læring innenfor operasjonsplanlegging og annen type operativ virksomhet (Haaland, 2011). Arkivene er også viktige i forbindelse med dokumentasjon knyttet til erstatningssaker og pensjonssaker, og arkiver fra og med slutten av 1970-tallet brukes som grunnlag for beslutning i saker angående stridsdekorasjoner. I en søknad om erstatning på grunn av fysisk eller psykisk skade på grunn av tjeneste i en internasjonal operasjon er dokumentasjon fra arkivene en viktig del av beslutningsgrunnlaget. Det å kunne dokumentere at en hendelse skjedde i eksempelvis Sør-Libanon i 1984 har betydning for soldaten som sliter med senskader i dag. Et annet eksempel er når Forsvarets dekorasjonsråd skal vurdere om en soldat skal dekoreres eller

ikke for innsats i en operasjon. Da vil det ofte være nødvendig å undersøke hva som finnes av rapporter om en hendelse. God dokumentasjon vil gjøre jobben til rådsmedlemmene lettere.

Bruk av Forsvarets arkiver i moderne kunnskapsutvikling

De fleste epidemiologiske studier, spesielt de som fokuserer på helsehendelser som oppstår senere i livet, har ufullstendige eller lite data på tidlige livsforhold. Digitaliserte forsvarsdata, fra for eksempel sesjon, vil gjøre det mulig for forskere å utnytte mer representative og større datasett som reduserer risikoen for utvalgsskjevhet. Tilgang til eldre fødselsårganger og mer detaljert informasjon for eksisterende årganger vil gi mulighet for å generere ytterligere forskning. Brukeren av slike data er i all hovedsak kvantitativt orientert og er helt avhengige av at informasjonen foreligger digitalt. Så vidt vi vet finnes det ingen tilsvarende datakilde for slike studier i Norge. Fordi opplysningene er tilgjengelige fra arkivene for minst halvparten av alle menn født mellom 1941 og 1957 (Fadum et al., 2020), kan digitaliseringen gi grunnlag for befolkningsstudier som undersøker prevalens og endringer over tid av for eksempel overvekt, allergi, epilepsi eller andre helseplager. Vi kan også konstruere longitudinelle kohortstudier for å undersøke hvordan blodtrykk, puls, fysisk form og kroppsvekt henger sammen med risiko for hjerte- og karsykdom senere i livet. Slike studier er gjennomført med data fra svenske vernepliktige og bør absolutt gjøres i Norge ved å koble helseopplysninger fra sesjon mot opplysninger fra Norsk pasientregister og hjerte- og karregister (Henriksson et al., 2020). Andre norske helseregistre inneholder data først fra 60-tallet, og det er derfor mangel på eldre helsedata. Digitalisering og transkribering av Forsvarets tidligere arkiver vil kunne supplere andre helseregistre og skape mulighet for multigenerasjonsforskning.

Det finnes eksempelvis ingen tilsvarende data i forskningen for å forstå livsløpsfaktorer som predikerer Alzheimer's sykdom eller variasjon etter alder i normal kognitiv funksjon eller levealder. Slike studier avhenger av representative, harmoniserte og store datasett som ofte ikke er tilgjengelige (Čukić I et al., 2017). Det vil og kunne gi innsikt i hvilken grad senere fødte kohorter har

bedre funksjon, helse, og kognisjon – noe som kan gi viktig informasjon om nåværende og framtidige eldre (Bonsang et al., 2017; Skirbekk et al., 2012). Digitaliseringen vil også gi mer detaljerte data om fysisk form, som vil være svært verdifulle i Forsvarets forskning (Aandstad et al., 2019).

For at opplysninger fra arkivene skal kunne brukes i forskning må dokumenter skannes og lastes opp i Arkivverkets digitale arkiv, for så å transkriberes. Dette har vært et sentralt fokus i et samarbeid mellom Arkivverket, Forsvarets sanitet og Sykehuset i Vestfold i perioden 2021-2023. Når arkivmaterialet blir digitalisert kan transkriberte data overføres til Forsvarets helseregister, forvaltes etter gjeldene lovverk og danne grunnlag for fremtidig forskning og helseovervåking gjennom for eksempel kobling til andre datakilder og registre.

Oppsummering - tilrettelegger for unik forskning

Forsvarets arkiver inneholder nasjonalt dekkende data som gir et representativt bilde på viktige helse mål ved sesjon, med muligheten for tilkobling av informasjon fra andre administrative dataregister på individnivå. Dette vil gjøre det mulig å gjennomføre en rekke banebrytende studier, hvor målinger i tidlige livsfaser gir unike muligheter for kausalstudier av hendelser senere i livsløpet. Dette vil kunne gi vitenskapelige bidrag på en rekke ulike områder som epidemiologi, studier av kognisjon gjennom livsløpet og populasjonsgenetikk.

Omfanget og detaljene i data tilsier at materialet har et stort potensial i forskning og er sjeldent i internasjonal sammenheng. Tilgang til data og norske helseregistre er blitt omtalt som «Norges gull», og helseregistrenes verdi i vitenskapelig sammenheng er godt dokumentert gjennom godt publiserte og anerkjente arbeider på en rekke felt de senere år.

Nye opplysninger fra Forsvarets arkiver har potensiale for å øke ytterligere informasjon om personell og data ved å tilføre helseregisteret opplysninger som

er manglende for deler av årskullene født fra 1941 til 1957 samt gi informasjon om veteraner som deltok i tidlige internasjonale fredsbevarende operasjoner. Digitalisering og overføring av data til Forsvarets helseregister vil kunne tilgjengeliggjøre et rikt omfang av nye, relevante data til flere forskningsmiljøer.

Referanser

- Aandstad, A. & Hagberg R. (2019). Reliability and Validity of a maximal treadmill test for predicting aerobic fitness in Norwegian prospective soldiers, *Military Medicine*, 184 (3-4), e245–e252, <https://doi.org/10.1093/milmed/usy195>
- Akershus hjerteundersøkelse 1950. (2023). <https://ace1950.no/>
- Bonsang, E., Skirbekk, V., & Staudinger, U. M. (2017). As you sow, so shall you reap: Gender-role attitudes and late-life cognition. *Psychological Science*, 28(9), 1201-1213. <https://doi.org/10.1177/0956797617708634>
- Čukić I., Brett C.E., Calvin C.M., Batty G.D. & Deary I.J. (2017). Childhood IQ and survival to 79: Follow-up of 94% of the Scottish Mental Survey 1947. *Intelligence*, 63, 45-50. <https://doi.org/10.1016/j.intell.2017.05.002>
- Fadum E.A., Häberg I., Rønning, K., Ariansen I. & Strand L.Å. (2020). *Digitalization of health data from military conscription – a pilot project*. Norwegian Institute of Public Health. <https://www.fhi.no/en/publ/2020/digitalization-of-health-data-from-military-conscription--a-pilot-project/>
- Fadum E.A., Strand L.Å., Rudvin I., Hæreid M.L. & Borud E.K (2022). The Norwegian Armed Forces Health Registry conscription board health examinations 1968–2018. *Scandinavian Journal of Public Health*, 50(2), 153-160. <https://doi.org/10.1177/14034948209204>
- Haaland, Torunn L. (2011). Erfaringsbasert læring i det norske forsvaret i Afghanistan. *Luftkrigsskolens tidsskriftserie*, 23, 77–81.
- Henriksson H., Henriksson P., Tynelius P., Ekstedt M., Berglind D., Labayen I., Ruiz J.R., Lavie C.J. & Ortega F.B. (2020). Cardiorespiratory fitness, muscular strength, and obesity in adolescence and later chronic disability due to cardiovascular disease: a cohort study of 1 million men. *European Heart Journal*, 41(15), 1503–1510. <https://doi.org/10.1093/eurheartj/ehz774>
- Historisk befolkningsregister. (2023). <https://histreg.no/>

- Skirbekk V., Loichinger E. & Weber D. (2012). Variation in cognitive functioning as a refined approach to comparing aging across countries. *Proceedings of the National Academy of Sciences*, 109(3), 770-774. <https://doi.org/10.1073/pnas.1112173109>
- Strand L.A., Koefoed V.F., Oraug T.M. & Grimsrud T.K. (2008). Establishment of the Royal Norwegian Navy personnel cohorts for cancer incidence and mortality studies. *Military Medicine*, 173(8), 785–791. <https://doi.org/10.7205/milmed.173.8.785>
- Strand L.A., Martinsen J.I., Fadum E.A. & Borud E.K. (2020). Temporal trends in the healthy soldier effect in a cohort of Royal Norwegian Navy servicemen followed for 67 years. *Occupational and Environmental Medicine*, 77, 775-781. <https://doi.org/10.1136/oemed-2020-106475>
- Strand L.A., Martinsen J.I., Rudvin I., Fadum E.A. & Borud E.K. (2018). Cancer incidence and all-cause mortality among civilian men and women employed by the Royal Norwegian Navy between 1950 and 2005. *Cancer Epidemiol*, 57, 1-6. <https://doi.org/10.1016/j.canep.2018.08.011>
- Strand L.A., Rudvin I., Martinsen J.I., Fadum E.A. & Grimsrud T.K. (2023). Lifelong differences in cancer incidence and mortality between subgroups of similar socioeconomic status in the Royal Norwegian Navy, *Occupational and Environmental Medicine*, 80, 514-521. <https://doi.org/10.1136/oemed-2023-108877>