Construction of the Finnish Army in World War II Database

By Ilari Taskinen

To cite this article: Taskinen, I. (2023). Construction of the Finnish Army in World War II Database. *Historical Life Course Studies, 13*, 44–60. https://doi.org/10.51964/hlcs13565

HISTORICAL LIFE COURSE STUDIES

Content, Design and Structure of Major Databases with Historical Longitudinal Population Data

VOLUME 13, SPECIAL ISSUE 5

GUEST EDITORS
George Alter
Kees Mandemakers
Hélène Vézina

EHPS NETWORK

MISSION STATEMENT

HISTORICAL LIFE COURSE STUDIES

Historical Life Course Studies is the electronic journal of the *European Historical Population Samples Network* (EHPS-Net). The journal is the primary publishing outlet for research involved in the conversion of existing European and non-European large historical demographic databases into a common format, the Intermediate Data Structure, and for studies based on these databases. The journal publishes both methodological and substantive research articles.

Methodological Articles
This section includes methodological articles that describe all forms of data handling involving large historical databases, including extensive descriptions of new or existing databases, syntax, algorithms and extraction programs. Authors are encouraged to share their syntaxes, applications and other forms of software presented in their article, if pertinent, on the EHPS-Net website.

Research articles
This section includes substantive articles reporting the results of comparative longitudinal studies that are demographic and historical in nature, and that are based on micro-data from large historical databases.

Historical Life Course Studies is a no-fee double-blind, peer-reviewed open-access journal supported by the European Science Foundation (ESF, http://www.esf.org), the Scientific Research Network of Historical Demography (FWO Flanders, http://www.historicaldemography.be) and the International Institute of Social History Amsterdam (IISH, http://socialhistory.org/). Manuscripts are reviewed by the editors, members of the editorial and scientific boards, and by external reviewers. All journal content is freely available on the internet at hlcs.nl.

hislives@kuleuven.be

The European Science Foundation (ESF) provides a platform for its Member Organisations to advance science and explore new directions for research at the European level. Established in 1974 as an independent non-governmental organisation, the ESF currently serves 78 Member Organisations across 30 countries. EHPS-Net is an ESF Research Networking Programme.

The European Historical Population Samples Network (EHPS-net) brings together scholars to create a common format for databases containing non-aggregated information on persons, families and households. The aim is to form an integrated and joint interface between many European and non-European databases to stimulate comparative research on the micro-level.
Visit: http://www.ehps-net.eu.

HISTORICAL LIFE COURSE STUDIES
VOLUME 13 (2023), published 23-01-2023

Construction of the Finnish Army in World War II Database

Ilari Taskinen

Tampere University & Visiting researcher Radboud University Nijmegen

ABSTRACT

This article introduces the Finnish Army in World War II Database (FA2W) currently under construction that is being built to study the effects of World War II on Finnish society. The database is a stratified sample of 4,253 representative of the men who served in the Finnish Army in World War II. The data have been gathered from the military service record collection of the Finnish Army, which holds files on practically all draft-age Finnish men of the birth cohort 1903–1926 and around 70% of the birth cohorts 1897–1902. The amount of data is extensive, containing over 60 different variables. The main part of the database consists of men's military careers, comprising longitudinal data on their positions in society and in the army (e.g., civilian/conscript/frontline service), military unit, military branch, task, rank, and service class. Other information includes socio-economic information from the draft and wartime and war experiences, such as wounds, illnesses, medical treatments, death, and honors. In the future the database will be expanded with men's postwar life trajectories to study the long-term effects of the war.

Keywords: Historical database, Demography, Soldiers, Mortality, Social class, World War II, Finland

e-ISSN: 2352-6343
DOI article: https://doi.org/10.51964/hlcs13565

1 INTRODUCTION

Experiences of wartime violence have severe demographic consequences. Most obviously, many people die prematurely, and the effect of these deaths on society has been studied by social historians for decades (Urlanis, 1971; Winter, 1986). More recently, it is better understood that wars severely impact the life course of survivors, as the long-term effects of violence, such as psychological trauma, have emerged in both academic and public discussions (Kivimäki, 2013). The study of these consequences has largely been conducted using contemporary sources in fields like health sciences and sociology (MacLean & Elder, 2007; Modell & Haggerty, 1991). However, as the immensely successful Union Army project on American Civil War soldier data (Fogel et al., 2000) and recent projects in Italy (Fornasin, Breschi, & Manfredini, 2019) and Australia (McCalman, Kippen, McMeeken, Hopper, & Reade, 2019) have shown, demographic historians with their vast historical datasets also have much to offer for the analysis of the human consequences of wars.

Human effects of war were in the focus of the project "Large Databases in Studying the History of War Experiences" (STASKO) led by Ville Kivimäki and conducted in Tampere University, Finland in 2017–2020. The project used digital technologies to create new large-scale datasets for the historical research of World War II Finland. One of its inspirations was the data on the Finnish soldiers who died in that war. The documentation on these approximately 94,000 men is excellent due to practices initiated in wartime. It was exceptional for a nation at war that the Finns did not bury their dead on the battlefield, but brought them back to their hometowns, where they are now at rest in the so-called "hero graves" in prominent locations outside every Finnish church (Kemppainen, 2006). Documentation on these Finnish casualties of war is currently stored in a public database containing socio-economic, military, and cause of death information on each of the fallen (see https://www.sotasampo.fi/en/casualties/). Construction of this highly comprehensive database was begun in the 1980s by private individuals, and it has been managed and updated since the 1990s by the National Archives of Finland. Since 2000, data on fallen soldiers has been available on their internet service, and the raw data is currently available to all interested parties (Ahoranta & Ortamo-Närvä, 2012; Karjalainen, 2014).

The database of the fallen enables statistical analyses of people who died in the war. However, studies of differential mortality (e.g., by social class) are limited by the absence of data on soldiers who survived the war and on men who saw no active service. In the STASKO project we investigated the possibility of taking a representative sample of all Finnish men of the age cohorts that participated in the war. After investigating the military service records of the Finnish Army, we found that this was indeed possible, and this article describes the Finnish Army in World War II Database (FA2W), which is currently under construction. The database consists of a representative sample of 4,253 of Finnish men in the birth cohorts who fought in World War II. The database is based on the military records of the Finnish Army, and contains extensive information on soldiers' social backgrounds, health, military service, and war experiences. In the future the database will be expanded with postwar data. The database is important for understanding the effects of World War II on Finnish society, and it offers excellent opportunities for studying the consequences of war on the life course. Finland may have the most comprehensive and detailed archives on World War II soldiers' wartime service and post war experiences.

2 HISTORICAL BACKGROUND

Before turning to the database, I will briefly review the history of Finland in World War II. This is beneficial for understanding the premises of the database and the opportunities it affords for the demographic study of war.

On the eve of World War II Finland was a small nation of four million people at the top of northern Europe. The nation had been part of the Kingdom of Sweden from the 14th century until 1808 and a grand duchy of the Russian Empire in the period 1809–1917 before gaining independence in 1917. In World War II, the young nation fought three wars. The first of these was the Winter War, which began on November 30, 1939, when the Soviet Union launched an attack to invade her small neighbor which had fallen under her sphere of influence in the secret protocol of Molotov-Ribbentrop Pact signed between the Soviet Union and Nazi Germany. Against all odds, Finland was successful in defending herself and in March 1940 signed a peace treaty after three and a half months of fighting. The nation ceded 10% of its territory in the eastern parts of the country but retained its independence.

Hostilities were resumed when Finland joined Germany to attack the Soviet Union in June 1941 in Operation Barbarossa. This war, known as the Continuation War in Finnish historiography, lasted until September 1944 and saw a successful Finnish offensive in 1941, two and a half years of relatively calm stationary warfare from winter 1942 to summer 1944 and an intense Soviet offensive in 1944, which Finland again warded off. An armistice was signed in September 1944, which led to the third war, the Lapland War, fought from fall 1944 until spring 1945 in northern Finland between the Finns and their former allies, the Germans (Kinnunen & Kivimäki, 2012).

Those familiar with the history of World War II will notice the unique situation of Finland: unlike all other small nations of Eastern Europe, Finland was never occupied. This exceptional and fortunate fate is the basic component of Finnish World War II demographic history, and its consequences are apparent in the nature of the casualties: between 1939 and 1945 Finland lost around 94,000 soldiers but only 2,000 civilians due to the war (Kivimäki, 2019). This ratio is in stark contrast with casualty figures in other European countries, where, without many exceptions, tens of thousands to millions of civilians perished under total warfare and systematic murder (Bessel, 2015). In Finland, however, the war mostly spared civilians. The Soviet Union bombed Finnish cities, but the destruction of these raids was relatively limited and the civilians could continue their lives mostly in safe circumstances, albeit under great stress and material shortages.

This leads us to Finnish men, the subjects of the database described in this article. While the war spared Finnish women and civilians, it was experienced firsthand by a great portion of Finnish men. In order to survive, the small nation was compelled to mobilize the maximum number of men for military tasks (Kurenmaa & Lentilä, 2005). The exact number of mobilized men is one of our research issues, but according to preliminary results, it seems that as many as 750,000–800,000 Finnish men served between 1939 and 1945. This is a considerable number for a nation of four million and means that over 80% of Finnish men of the birth cohorts 1897–1926, the primary age groups from which Finland mobilized men, took part in the war. This extensive participation is a fundamental aspect of the Finnish demographic history of World War II. While in many other countries combat touched only certain social groups, in Finland the war was inherently an experience of the whole nation. The Finnish Army in World War II database reveals how a nation and all its social strata experienced and suffered from the consequences of war.

3 DATA: MILITARY SERVICE RECORDS

Studying the life courses of soldiers and other people in or after a war involves numerous challenges. Often, an obstacle is the lack of nationwide registers, and even in cases where such existed, they may have survived only partially due to inadequate archival practices or destruction caused by outside forces, a danger severely heightened in wartime. In this framework, the opportunities to conduct demographic research in Finland are exceptional thanks to the military service records of the Finnish Army. These records are the basis of the FA2W database, which I will introduce in this section. In the future the database will include data on soldiers' post-war lives, but the current first stage of the database construction focuses on soldiers' pre-war and wartime lives.

Both the survival and richness of the military records is a consequence of Finland's wartime fate. Unlike many other countries, where military archives were lost and destroyed in combat and invasions, in Finland these losses were few because the country was never occupied and not even the Russian bombing of the Finnish home front did much damage. However, Finland needed to make the greatest possible use of her small population to survive, in practice calling to arms every man it could spare. This would not have been possible without meticulous recordkeeping.

The Finnish state adopted universal male conscription after gaining independence in 1917. Between 1918 and 1939, the Finnish Army trained for reserve 509,000 men, which was around 70% of the age cohorts that were called up in drafts (Kronlund, 1988). However, a greater number of men ultimately came into contact with the army because, in desperate need of manpower, the army called up previously exempted men to redrafts during the war. Exemptions from military service were granted only to the very weakest and sickest (Nurminen, 2008).

Information on Finnish men was recorded in various documents during their basic training and wartime service. These records are now organized into several different collections, of which two are used in the construction of the FA2W database. First are the draft records, *kutsuntaluettelot*, compiled for the annual drafts of the army. These drafts were based on information provided by Finnish population register keepers, most notably the Evangelical Lutheran Church of Finland, to which over 95% of the Finnish people belonged, and by other churches and civilian register keepers. Each year before the drafts these authorities provided the regional military authorities with information including the names, places of residence, occupations, education, and marital status of all men of draft age. They compiled this information into draft district specific lists, which were used to call up the men and were updated with additional information such as men's height and weight and the results of draft.[1] These records are currently archived in the National Archives of Finland and, as I will later discuss, they are used as a supplementary source when the corresponding data is missing from the main sources of the FA2W database.

The second collection, and the one that is the base of the FA2W database, is the military service record collection of the Finnish Army. This collection of personnel files contains various documents, although it is mostly known for its main document, the military service record *kantakortti* (see Figure 1). This is a two- to four-page record, depending on its form, which was updated in 1930 and 1945. Every Finnish man who was called up in a draft got such a form which was updated throughout his military service. This record contains extensive information including socio-economic details (e.g., occupation, education, marital status), military training (e.g., military branch, education, evaluation), military service history (units, ranks, tasks, service class) and war experiences (e.g., wounds, illnesses, battles, honors). For officers an additional form contained some further information, such as evaluations of an individual's suitability for different tasks and the occupation of the father.

In addition to the military service record, the personnel files include a wide variety of other documents. A medical inspection record was in principle created for every man in the draft and subsequently updated during his service. This document includes information on medical examinations (e.g., previous illnesses, height, weight), customary procedures (e.g., vaccinations), illnesses, and treatments. Medical information can also be found in individual documents written in military and field hospitals. These hospitals produced reports on the diagnoses and treatments of men which were attached to their personnel files. The files also include a wide variety of other documents gathered by the army, such as disciplinary records and information on their deaths.

The amount of data available in the military service files is vast. On average, there are 14 pages of documentation per man in the sample (see next section), which is an underestimation for men in active service, because those who were exempted may have only two pages. Men with long and eventful military careers may have over 50 pages of documentation. This plethora of data is not merely due to the medical appendix but also to different versions of the military service records. The original pre-war recording principle was to have two copies per man, one of which was kept at the military district headquarters while the other followed a man to his various military units, where it was updated with his service details. When a man completed his military training or returned home during the war, the record was returned to district headquarters, where the more recent information was copied into the main document (Ylönen-Peltonen, 2021). This system seems to have been too difficult to implement in the chaotic circumstances of the war, when men were rapidly transferred between units. When a man joined a new unit during the war, a new service record was often begun. This is fortunate for us, because military service files often include records from men's prewar conscription service as well as from the war, enabling us to gather data on socio-economic status from both periods.

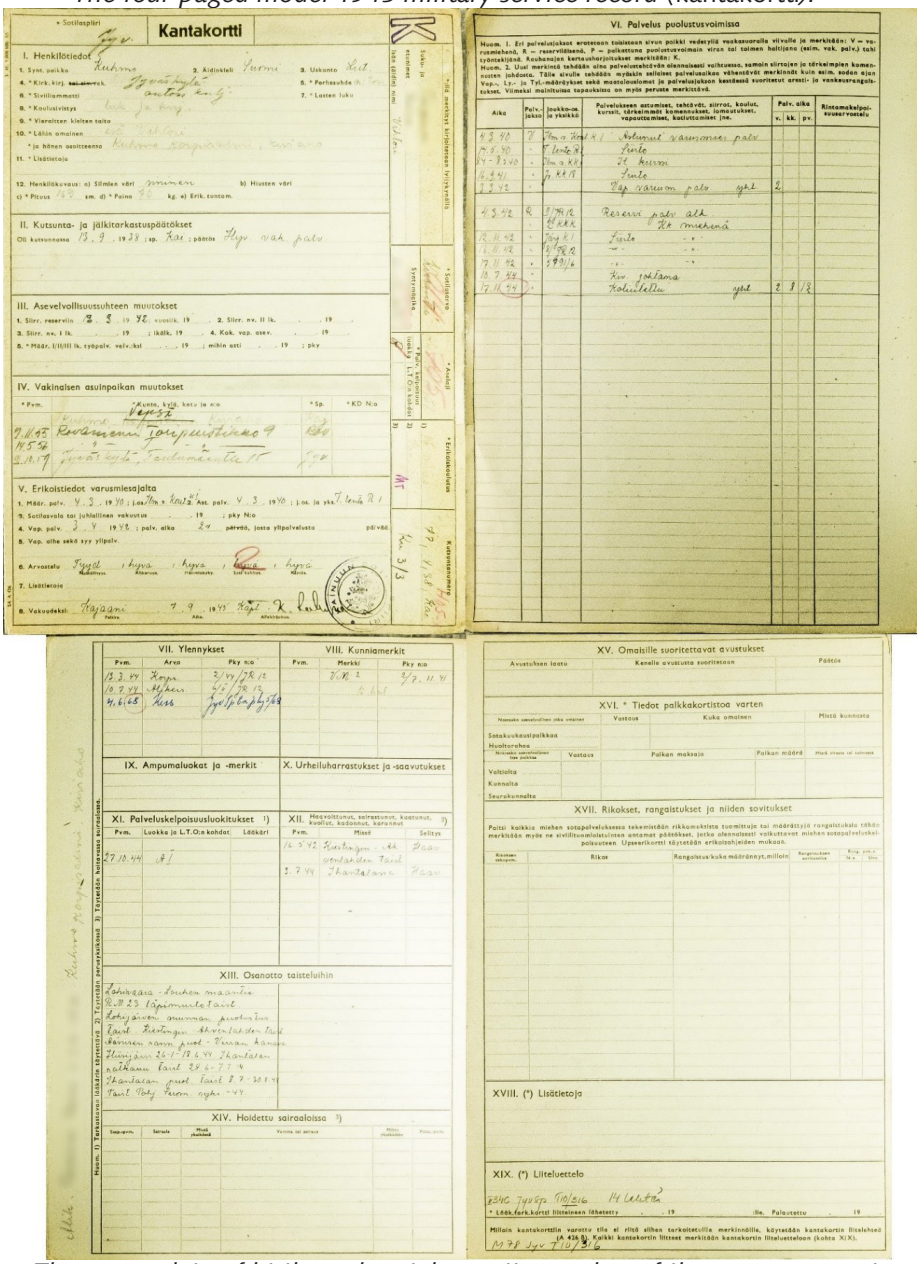
In the Finnish Army, the military districts were responsible for the mobilization of the detachments of the army in the event of war. Military districts kept records of the men on active service and the reservists through the military personnel files. If a man transferred between military districts, his file followed, and this continued for the duration of his status as a reservist. Unless reserve status was discontinued due to poor health, men could stay in the army reserves until age 60, and some information in these files was updated during this time. Since World War II, Finland has not been at war and most men of the war generation did not serve in the army again, except for some refresher training, so postwar documentation is sparse. However, changes in places of residence, information on deaths and social security numbers, introduced into Finland in late 1960s, were frequently updated in the postwar decades.

1 The National Archives of Finland, PLM-33/Ee:3-Ee:5, The draft district regulations and the draft district instruction from the Ministry of Defence, 4 December 1929.

For the youngest men of the war generation, born in the 1920s, the service records may contain entries until the 1980s. After a man was removed from the army's reserves, the service files were first sent to the Central Medical Archive of the Finnish Army, and they were moved in the 1990s to the Military Archive of Finland, which is today part of the National Archives of Finland (Ylönen-Peltonen, 2021).

The meticulousness with which the military service records were stored and updated reveals the importance that Finland, a small nation next to a great superpower, attached to securing her limited population for any national defense effort. The continued importance of this is reflected in the fact that Finland is one of only two nations in Europe to retain universal conscription. At the same time, these archival practices indicate why the military service record collection is not only rich in content but also comprehensive. According to research conducted for the building of FA2W database, the National Archives of Finland stores military service files on nearly every Finnish man born between 1903 and 1926 who survived to draft age. For men born 1897–1902, around 70% were included, omitting 30% who were exempted from army service. Information on this exempted group can be gathered from the draft list to obtain a highly representative sample of all Finnish men of the age cohorts who fought in the war.

Figure 1 *The four-paged model 1945 military service record (*kantakortti*).*



Explanation: The name, date of birth, and social security number of the person are not exposed. The first page contains personal, socio-economic, and conscription service information, the second page career in military service, the third lists promotions, honors, service classes, wounds, illnesses and battles, and the last page contains information on e.g., offences.

4 SAMPLE IN TWO STEPS

The FA2W database is based on a stratified random sample of 4,253 men taken from the military service record collection and the draft lists of the Finnish Army. At the beginning of our project, we lacked a comprehensive overview of the content of the military service record collection, thus, we initially took a less accurate, systematic random sample. After gaining detailed knowledge of the military service record, we could reshape the sample into a more accurate stratified sample.

According to the army guidelines from the early 1930s, a military service record was to be written up for every Finnish man called up in a draft. This meant that a record should have been made for every Finnish man who was alive and of the relevant age because Finland practiced strict universal conscription (Ahlbäck, 2014). When we began to plan a sample from the military service record collections, it was not known how well this principle had been followed and to what extent the records were extant. The archive knew that the collection was very comprehensive and holds altogether 4.5 shelf kilometers of documents. It is the most used collection of the National Archives due to its wide usage in genealogical research (Ylönen-Peltonen, 2021). However, the archive also warned us that some men were missing from the collection. These were presumably most often men of the older age cohorts who had been exempted in drafts or during their conscription service.

At this point there was no way to investigate these limitations in detail. It would have been all but impossible from the kilometer-long rows of file stacks, and it was also difficult from the otherwise very helpful indexes of the collection. The indexes were organized in two parts: a physical one (see Figure 2 for an example) and an electronic one. The electronic file contained only the names of 80% of the men born during 1924–1926. The physical card index contained men born 1897–1923 (and a small number of men from earlier age groups) and remaining 20% of the age cohorts 1924–1926. These cards were kept in 779 boxes each holding around 1,250 cards. The index was in principle organized by birth year and in alphabetical order according to the family name within years (see Table 1). However, the age cohorts 1900–1909, 1910–1918, and all men born before 1898 were organized into separate series, and around one-third of men killed in the war had their own series.

Figure 2 *Military record card index box and index card*

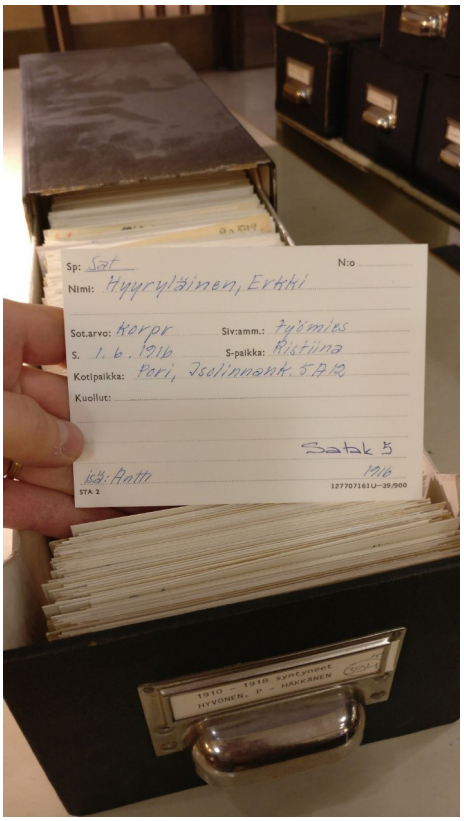


Table 1 *Organization of the indexes of the military service record collection*

Card index		Electronic database	
Series	**Boxes**	**Age cohort**	**Men**
–1897	59		
1898	19		
1899	19		
1900–1909	248		
1910–1918	262		
1919	26		
1920	27		
1921	27		
1922	27		
1923	30		
1924	6	1924	26.996
1925	5	1925	28.612
1926	4	1926	29.319
1st series of fallen	20		
2nd series of fallen	66*		

** Cards of the 1st series of the fallen, holding around one third of the soldiers killed in the war, have been collected from the age cohort series at some point after the original assembly of the card index. After we gathered the initial sample, the archive collected the remaining casualties of the war to a second series of fallen. This series is packed into different, smaller boxes than the rest of the card index.*

Due to these exceptions, it was not possible to make a stratified sample, e.g., to ensure enough cases from each age cohort. We therefore opted for systematic random sampling. This would provide us with a representative sample of the Finnish men who served in the Finnish Army in World War II, under the assumption that soldiers' records would be better preserved than those exempted from the draft. The sample was taken by 1) picking randomly five cards from each of the 779 boxes of the physical card index and 2) picking 336 men from the electronic database, based on the sampling interval calculated by counting the number of cards in the physical card index boxes. After excluding unnecessary reference and duplicate cards picked from index cards, the result was a sample of 4,045 men.

After this initial sampling we started to enter the data, and after completing a quarter, we could analyze the representativeness of the sample. Here we made two important observations: one was that there was a major increase in the number of men in the sample between the age cohorts 1902 and 1903, which suggested that there had been a change in the principle of the creation or preservation of records between these years. The other observation posed a problem for our sample. When comparing the distribution of soldiers killed in the war in our sample with those from the database of the fallen compiled by the National Archives of Finland, we found that there were too many of the fallen in the older age cohorts of the sample and too few in the younger age cohorts.

This observation led us to do further research on the indexes of the collection, which at this point was much easier thanks to the work accomplished by the National Archives of Finland. After our initial sampling, the National Archives of Finland had drawn the index cards of the fallen men still in the main index into their own boxes and added the information on these cards to an electronic database. This new organization of the information revealed, first, that the men included in the database of the fallen were well represented in the military service record collection: around 99% of them had a card in the index and around 97% of their corresponding documents could be found, while the remaining 2% seem to have been lost during archiving. Second, it was now possible to estimate the number of survivors and the total count of men in the index card series. For this, we manually measured the

length of the card rows in 84 card boxes and counted the number of cards in 22 boxes to estimate the number of cards in the index series (see Figure 3).

This new measurement had two main results. First, it revealed the reasons why our initial sample had been distorted. We had erroneously assumed that the card boxes held a consistent number of cards across cohorts, but the measurements revealed discrepancies of 10 to 20% in their card counts. The main reason for the discrepancies was that the series containing one third of the fallen had been collected unevenly, leaving different counts of cards in the age cohort series. Moreover, part of the age cohort series seems to have been condensed into a smaller number of boxes at some point, and these boxes were very tightly packed with relatively large numbers of cards. These findings and the observation that the index cards of the fallen were of a slightly different size than the other cards explained the distortions in our sample.

Second, our measurements of the index cards gave us data about the numbers of men in the military service record collection. As noted, this could be done precisely for the men who died in the war, whose records were confirmed to have been almost entirely preserved. The collection is also very comprehensive for other men. According to our estimates, most age cohorts actually had more cards in the index than the number of Finnish men alive and of draft age in these cohorts (see Table 2). The index included extra cards, such as men who had changed their names. In our initial sample, 4.4% of the cards were of this kind. But even taking these cards into consideration, the collection is very comprehensive.

Figure 3 *Measurement of the length of index card stack*

Measurement of the card index disclosed that the military service record collection was highly representative for the majority of the age cohorts serving in the war. Except for the series of the 1919 age cohort, which curiously contains 12% more index cards than the draft-age population of the cohort, the cohorts 1910–1926 differed only marginally from the expected 100%. Taking account of the imprecision of our measurements, it is safe to say that the military service records were written and archived very meticulously. The small number of missing records should not cause large distortions to the collection.

However, the series 1900–1909 show a notable deficit with 12% of the men missing. This series also marked a significant turning point in the collection as the numbers of men in our sample increased markedly between the age cohorts 1902 and 1903. It seems that military service records were written and preserved on nearly all Finnish men from the age cohort 1903 onwards, although some men who were exempted from drafts might still be missing. Among the older age cohorts, the military service records are missing from around 30% of the men. Our analysis of the initial sample indicates that these were men who had been exempted from the draft or from conscription service, but this was not a strict rule as some of the exempted men did end up in our sample.

Table 2 *Percentages of men in the military service record collection as part of the total number of conscripts per birth year*

–1897	70,115*
1898	73%
1899	72%
1900–1909	88%
1910–1918	103%
1919	112%
1920	100%
1921	101%
1922	96%
1923	99%
1924	103%
1925	102%
1926	101%

Source: Finnish Population Statistics and sample results.

** It was not possible to make the calculation of the –1897 series as it contains men from unknown age cohorts. The estimated number of men in this series is 70,115.*

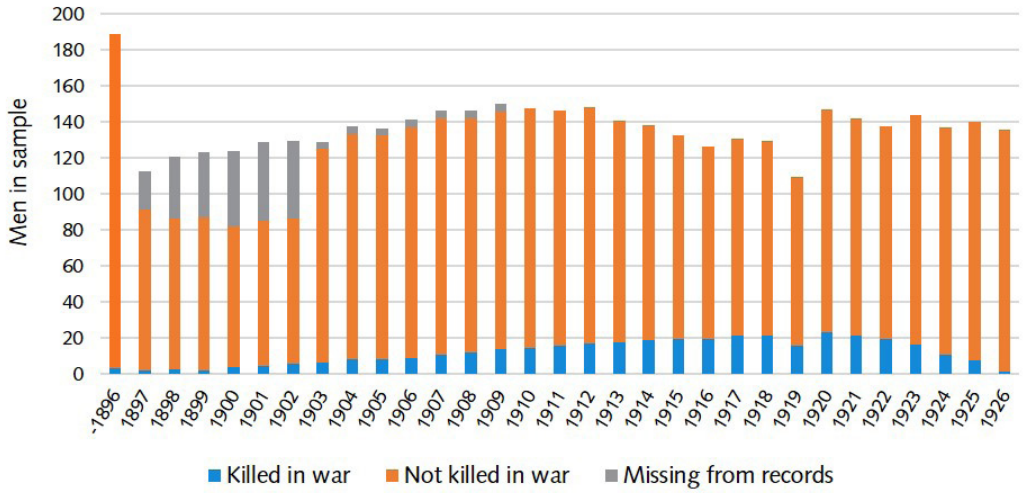
With this research we could now divide the men in the military service record collection into those who died and those who survived the war, and we also knew how many and which types of men were missing from the collection. This new information enabled us to design a better, stratified sample, which would not be based on the record collection like the initial sample, but on population statistics. The new sample was stratified on thirty age cohorts for birth years 1897 to 1926 using the size of each cohort at end of the year preceding the draft.[2] Men in older birth cohorts with a file in the military service record collection were also added to the sample. The size of this sample population is 1,063,141.

Men were further divided into two primary strata within the birth cohorts: those who were killed in the war and those who survived. A third stratum of men who were missing from the military service record collection was added to the birth cohorts 1897–1909. These men were exempt from military service and did not get a military service record. However, they are recorded in draft records (*Kutsuntaluettelot*), introduced in the previous section, from which we will sample these men in the near future. There were 75 strata in all as shown in Figure 4. The sizes of these strata were calculated with information from the database with war deaths compiled by the National Archives and our measurements on the number of men in the series of the military service record collection. The sizes of the strata were calculated using a sampling interval of 1 in 250, which yielded a sample of 4,253 men.

The men selected for our original systematic random sample served as the basis of the stratified sample. When the original sample for a certain stratum did not correspond with the new stratified sample, we removed or added deceased and surviving men to and from the birth cohorts of the sample. Men removed from the sample were randomly selected from the database. Men were added by random selection from the database of fallen soldiers and the card index boxes. Samples for strata with men missing from the service record collection (1897–1909) will be taken from draft records after the new sample is examined to determine why men exempted from the army did not get a military service record.

2 The Finnish Statistical Office reported age cohort specific population statistics that were based on censuses in major cities and towns and civil register-keepers reports in rural Finland during this period once in 10 years (1920, 1930 and 1940). I have calculated cohort sizes between these years by subtracting from these numbers the age cohort specific yearly mortality figures that were reported every year.

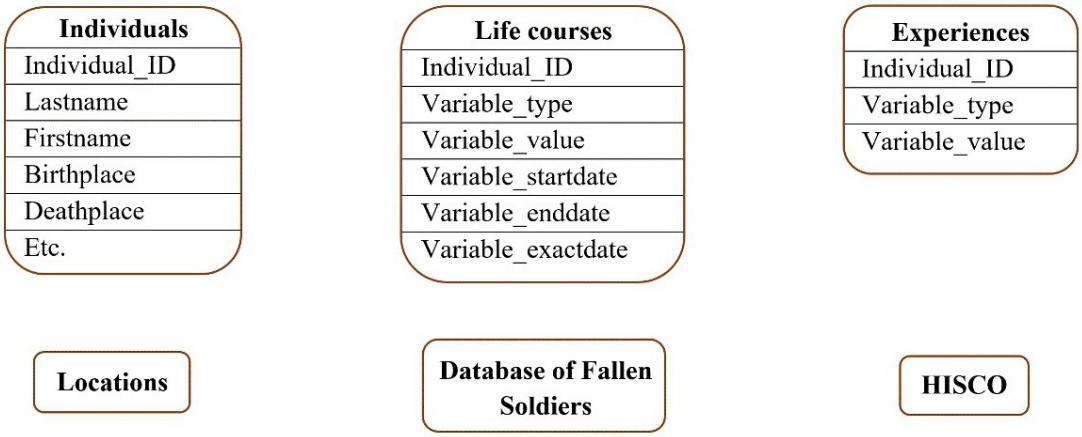
Figure 4 *Strata of FA2W sample*



5 DESIGN OF THE DATABASE AND VARIABLES

During the first stage of the FA2W database construction, we are compiling all core information available in the military service records. The database is broad, consisting of over 60 variables. These include traditional personal and socio-economic information, but its center is in the detailed portrayal of men's military careers and experiences. These variables often consist of standardized values originating from the military classifications of the Finnish Army, but we have also created some ourselves considering our research aims. Ultimately, the data will be stored in a relational SQL database. It is planned that the database would comprise three main tables Individuals, Life courses and Experiences. Furthermore, we will link these tables to three other datasets (see Figure 5), for coding and standardization (Locations for locations, HISCO for occupational titles) and complementary information (Fallen soldiers). The tables Life Courses and Experiences are designed according the principles of the the Entity Attribute Value model (EAV), in which each record entails only one attribute (Stead, Hammond, & Straube, 1982).

Figure 5 *FA2W database design*



In the following three tables we explain the variables that are included in the three main tables. The tables Individuals contains data without an indication of time. The table Life courses includes dynamic data or static data that have a date like birth_date. The table Experiences contains more complex data and variables summarizing values from the first two tables.

Table 3 *Personal and socio-economic data in the FA2W database*

Static data (not dated)	Events with dates	Attributes at the time of draft/conscription	Attributes during the war years 1939–1945
Last Name	Date of Birth	Place of Residence	Place of Residence
First Name	Date of Draft	Civil Status	Civil Status
Place of Birth	Date of Follow-up Examination	Number of Children	Number of Children
Place of Death	Date of Death	Occupation	Occupation
Legitimacy	Illness	Education	Education
Father's Name	Wound	Height	Height
Mother's Name	Offence	Weight	Weight
Mother Tongue	Sentence	Name of Next of Kin	Name of Next of Kin
Religion	Honor	Relationship of Next of Kin	Relationship of Next of Kin
Nationality		Place of Residence of Next of Kin	Place of Residence of Next of Kin
Foreign Language Skills			
Color of Eyes			
Color of Hair			
Social Security Number			

The table "Individuals" includes personal and socio-economic data. This is a long-format table without time stamps and the variables are static (time-constant), see Table 3, first column. The other columns in Table 3 show the variables that are stored in the table "Life Courses" and which are dynamic in time. The ones in the second column are day-specific events. This includes vital events of birth and death and various events of military service starting from draft and ending in wartime experiences, such as wounds, illnesses and honors. In addition to date, we gather values (e.g. type of illness) for these military experiences. In the first phase of the database building, we gather only serious wartime wounds and illnesses requiring hospital treatment. The durations of treatment periods can be obtained from the variable "Position" (see Table 4).

The variables in the third and fourth column are dynamic in a very limited sense, since they are collected from the military records which make only a distinction between the period of conscription service in peace time and serving during war years. So, as this information was not dated in military service records and often updated during service, the starting date and ending date of each type of period is the best accuracy with which the information can be recorded. Subsequently, the time range of men's conscription and wartime service can be calculated from the table "Life Courses". The variable 'Place of Residence' was systematically updated after the war and then the record includes exact dates.

Table 4 *Military career data in the FA2W database*

Periodic Variables	
Variable	**Value**
Position	Civilian/Reservist/Conscript/Army Personnel/Field Army/Home Front Troops/Backup Reservist in Service/Treatment in Military Hospital/Furlough/Secondment/Prisoner/ Prisoner of War/Deserter/Backup Reservist (Classes I, II & III)
Military Unit	Description
Military Branch	Infantry/Field Artillery/Signals/Coastal Artillery/Sapper/Anti-Aircraft/Military Engineer/Home Front Troops/Armored Force/Logistics/Navy/Air Force
Task	Description
Military Rank	Finnish Army Ranks
Service Class	A I, A II, B I, B II, C, D, E

The core part of the table "Life courses" comprises six variables (see Table 4) that form periods in military careers, and for which the start and end dates are exactly dated. "Position" is the main variable of the group consisting of 16 values designed by us to describe a man's position in society and in the army. This variable can be used to distinguish between civilian and military roles. The most important distinction among men serving in the army is between combat roles in the field army and auxiliary roles on the home front. There are also more specific subgroups, such as those in military hospitals and deserters, that offer interesting options for analysis. The "Position" variable builds an uninterrupted life course for men in the sample. Each day of their civilian lives and military careers, they belong to a single category of the "Position" variable.

These military roles can be further specified with the variables "military unit", "military branch", and "task". The first two are crucial because they divide men into units, like infantry, air force, and logistics, that operated in different wartime environments and faced different dangers. "Task," which includes roles like "rifleman", "squad leader", "clerk", and "driver," identifies further differences within military branches. This is particularly important for the infantry, because it distinguishes the men who really did the fighting in the trenches, like riflemen, from the personnel who worked behind the lines, like clerks. "Military rank," which offers a social and hierarchical perspective on work and roles in the army, and "service class," which is based on the army's classification of health and physical capabilities for different forms of service, will be used to analyze changes in the social composition of the army. A military career may include numerous positions if a man served in the army for several years, which was common in the Finnish Army of World War II. On average, a man who fought in the war has 7.6 different positions and 5.7 different units in his career. Half of these men have 10 or more records with changes in the variables position, military unit, or task.

The table "Experiences" contains two types of data (see Table 5). First, it includes undated military career information like the number of battles men fought in, training for military task, and evaluations. Second, it contains the variables we constructed from the two other main tables, summarizing soldiers' military careers and experiences like the duration of wartime service and participation in different stages of the war. These variables cover our prime targets of inquiry and are created at this point to simplify their query during analyses.

The main tables are linked to three other datasets. The most important of these is the database of soldiers killed in the war constructed by the National Archive of Finland. All soldiers in our sample who died in the war can be linked to this database because they were picked from it. This dataset offers additional information regarding soldiers' wartime deaths, such as the location of their mortal wounding and death and its manner (died in action/hospital/went missing/etc.). The table "Locations" provides additional information about the places of birth, residence and death recorded at municipality level in the database. The table includes their wartime population figure, administrative level (municipality, market town, city), province and coordinates. Due to the size of our sample, it is necessary to convert places to provincial level in regional analyses. Lastly, the occupation titles will be linked to the HISCO classification.

Table 5 *Military experience data in FA2W database and constructed variables*

Variable	**Value**
Number of Battles	Number
Military Branch Education	Finnish Army Branches
Military Task Education	Definition
Evaluation (Punctuality, Diligence, Powers of Observation, Military Development, Personal Conduct)	Good/Average/Poor
Membership of Voluntary Militia	Yes/No
Survival	Survived/Killed/Did not Participate
Participation in wars (Winter War, Continuation War, Lapland War)	Yes/No
Military Ranks Classes	Rank and File, NCO, Officer
Duration of Conscription	Number of Days
Duration of Wartime Service	Number of Days
Number of Mobilizations	Number

The FA2W database offers numerous angles from which to study the effects of war, but there are also limitations and challenges inherent in its data. Some variables are affected by recording practices that cannot be elaborated here. However, two clear issues are worth mentioning. Firstly, the data becomes more comprehensive as it becomes more recent. During the 1920s, the first decade of the formation of the Finnish Army, recordkeeping was less systematic and the form of the military service record was less detailed than its later versions. From the 1930s onwards the military records are more comprehensive due to the introduction of a new form and guidelines. For the World War II years, 1939–1945, the data is available in its greatest detail because numerous different documents were used during this active period of service.

Secondly, socio-economic information, such as occupation, marital status, and education, was gathered at two time points: at conscription and during the war years. The military service record form in use from the early 1930s to the end of the war has two sections: peacetime and wartime entries, which were the same for men who joined in wartime. The issues with this information concern the updates of the records and the form. The military service record form was renewed in 1930 and again in 1945, and in both cases information was copied from old forms onto new ones if a man belonged to the reserve. This was not a problem when the army archived old records. However, especially during the upgrade of 1945, some military districts destroyed records from the pre-war and war years after copying information onto new forms until an order was issued for their preservation (Ylönen-Peltonen, 2021). Another challenge is updates during refresher training before or after the war. If a man joined these events, conscription and wartime socio-economic information was sometimes updated to reflect their later status.

These recordkeeping practices mean that information on socio-economic status from both conscription service and the war years is not always accurate, even if it was originally recorded in the military service records. This problem can be overcome in some cases with other documents, such as medical histories, which contain information on place of residence and occupations during the conscription and war periods. However, these recordkeeping practices mean that even if we have occupation information about 93% of the sample and 98% of men who served in the war, we do not have this information in the same detail for both periods. The conscription period data is particularly deficient, and occupation information is missing from one third of the sample. Some of this data can be collected later from the draft lists.

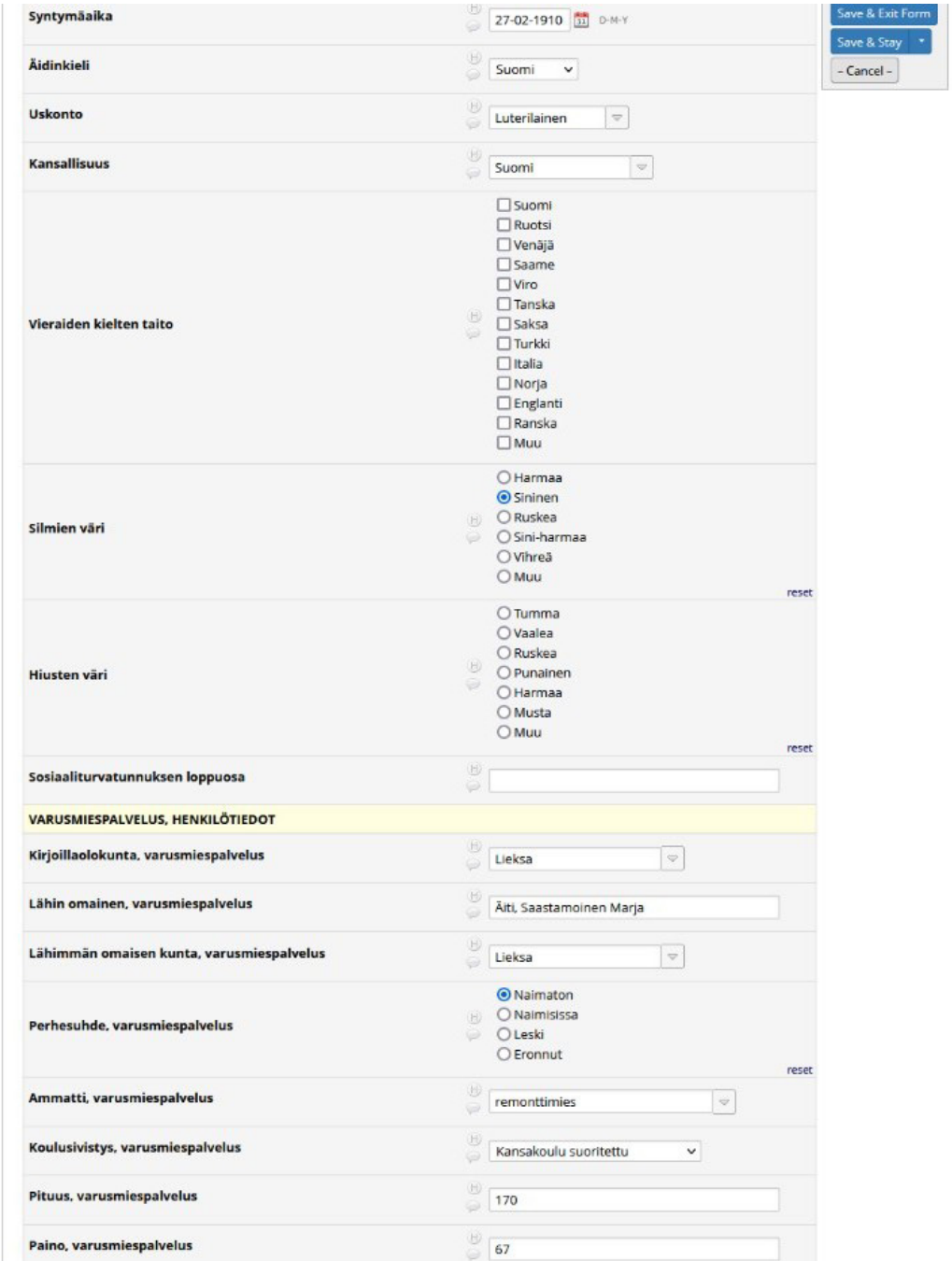
6 DATA ENTRY

The construction of the database began on funding granted by the Finnish Cultural Foundation for the STASKO project in 2017 with the designing of the initial sample. In 2018, over 60,000 pages of documents were photographed. To enter the data, we use the web-based database management platform REDCap (https://projectredcap.org, see also Harris et al., 2009; Harris et al., 2019).

Between 2018 and 2020, salaried research assistants and the author of this article entered one quarter of the data into the database. Since the records include sensitive medical information on people who in rare cases could still be alive, we could not make use of crowdsourcing. Furthermore, crowdsourcing would have been difficult to implement due to the variety and complexity of the data. The military service record files contain numerous different types of documents where recording practices changed over time, meaning that research assistants need extensive training on the material and its historical context to be able to enter the data. The data entry process is relatively time-consuming. With over 60 variables to be collected from an average number of 14 documents, it takes on average 20–30 minutes to enter one man's information into the database. The laboriousness of this work, which is largely due to the richness of the military service record files, was something of a surprise for us and delayed the completion of the database.

After one quarter of the data had been entered, analysis of this data revealed distortions in the sample as described above. These problems were investigated in 2020, and the new stratified sample was designed in 2021. The necessary additional records have since been collected and data entry is slowly progressing. Additional funding permitting, the first phase of the database described in this paper is scheduled for completion in 2024.

Figure 6 *A segment of REDCap data input form*



7 RESEARCH OPTIONS AND FUTURE PLANS

The FA2W database offers many opportunities for both population-level demographic history and life course analysis. A central goal for the database is a better understanding of the impact of World War II on different sections of Finnish society. So far, this question has received relatively little scholarly and public attention. This may be due to the importance of World War II for Finnish national unity, but lack of data has also played a role. The available data has already generated distributions of Finnish war casualties by regions, age cohorts, and social classes. However, few studies have analyzed issues like mortality rates in specific social groups due to the lack of comprehensive data on surviving soldiers. These studies show that these are important issues: there were wide regional differences in mortality rates and the most severely affected seem to have been the rural poor (Kivimäki, 2019; Toivonen, 1998; Waris, 1948).

As a representative sample of all Finnish men of the age cohorts who fought in the war, the FA2W database will make it possible to investigate the direct consequences of war, such as mortality, injuries, and sickness rates among social classes, age cohorts, linguistic groups, and regions. It will enable us to connect these differences with Finnish conscription training and mobilization practices, such as how social groups were trained for different military tasks, who was sent to the trenches, and who could stay home, which crucially determined the distribution of the wartime burden. The FA2W database offers a base to study these inequalities in military sacrifice during the period of total warfare of the World Wars. This issue has been studied previously particularly in the US Army in the post-World War II period (Barnett, Stanley, & Shore, 1992; Kriner & Shen, 2010; Merli, 2000), in which lower classes and minorities have carried the heaviest burden of wars, but it has not gained similar attention in European nations involved in the World Wars that presumably fought total, shared warfare with universally conscripted armies. As nation-states also habitually frame their wars as shared endeavors, the unequal consequences of wars have not been thoroughly publicly acknowledged, and as Kriner and Shen (2016, p. 554) note, it is a topic that "academic scholarship has not seriously explored".

For life course analysis, the main value of the database is that it permits detailed examination of how different forms of wartime service affected soldiers' mortality and survival in war. This can be done with various data such as the duration that the men spent in frontline troops, military branches, military units, and military tasks. This analysis can be further enriched with contextual information about casualty rates in these positions at different phases of the war. We can, for example, calculate how many weeks the sampled men stayed in heavy frontline infantry combat, stationary warfare duty, or rearguard tasks and whether the periods of combat were defensive or offensive in nature. This makes it possible to examine the relationship between mortality and the nature of soldiers' military service and exposure to combat and violence in a very detailed manner. Many military science studies have shown that, for example, infantrymen die in wars more often than artillery personnel (Bellamy, 2000; Buzzell & Preston, 2007), but the FA2W database also enables us to study the forms of service inside the military branches.

The data on wartime service and experiences are also central in the planned future of the database and the research it enables. When the initial work on soldiers' prewar and wartime lives is done, the database will be enriched with postwar data to study the long-term effects of war and exposure to violence. This data will include information about causes of death, which is already partially available in the military records, and information on socio-economic status and health from nationwide records and registers that became available after the war. One plan is to connect the sample to the census records collected every 10 years from 1950 onwards. This will provide immense opportunities to investigate the long-term effects of war and violence. In particular, it will be possible to investigate how much long-term life chances were tied to degree of exposure to violence and combat and to compare the effects of experiencing death face to face in the trenches to service in relatively safe circumstances in auxiliary duties behind the lines. This may elucidate the effects of violence more clearly than studies based only on aggregate data like wartime age cohorts or soldiers of wars in general (Saarela & Finnäs, 2012).

In general, the Finnish case will offer an interesting chance to scrutinize the long-term effects of war. As MacLean and Elder (2007) state in their review of research on the consequences of military service in veterans' post-war lives, it is evident that veterans' fates have varied considerably in different historical circumstances. A comparison between countries like the United States, where soldiers have been drafted mostly from the lower classes, and Finland, where the experience of World War II was shared by all social strata in a spirited and united manner, could offer interesting results. It has been suggested, for example, that PTSD was a lesser problem among Finnish veterans of World War II because this war has been very important in Finnish culture (Hautamäki & Coleman, 2001).

Furthermore, as the database offers a representative picture of the Finnish male generation of the early 20th century, it can also be used to examine questions about the social structure and health of Finnish society and the life courses of the war generation throughout the century. One of the most promising research possibilities is examining connections between health during the conscription period and later life. The military records include rich medical data, which can be added to the database in support of this research.

8 CONCLUSION

In this article, I have introduced the first steps of a database that was born out of curiosity while conducting research on other nearby topics. As this is the first demographic database built by our research group, the process has been a learning experience with constant re-evaluation of our methods and decisions. The foundation of the database is now in place, but many changes are still sure to come when we head into the finalization of its first phase. While constructing a large database has been a laborious task, we believe it to be a worthy undertaking because it promises to provide invaluable empirically grounded insight into human experiences of war, a topic that our research group among many others has examined in recent times, mostly from a cultural historical perspective. When the groundwork is done, we envisage that the database will serve as a foundation to investigate these issues for many years to come.

ACKNOWLEDGEMENTS

The construction of the Finnish Army in World War II Database began thanks to a research grant from the Finnish Cultural Foundation and the Harry Hendunen Fund. The writing of this article was supported by the Society of Swedish Literature in Finland. The construction of the database is led by Ville Kivimäki and the database designed by Ilari Taskinen in Tampere University. Katariina Eskola, Siiri Simppanen, Vikke Niemi, and Miiko Siivonen have worked as research assistants on the project. Raija Ylönen-Peltonen and others in the National Archives of Finland have been truly helpful in assisting us with the military service record collection. Turkka Näppilä deserves thanks for his help with REDCap, Jyrki Ollikainen with his advice on sample design and Jarmo Peltola for sharing his expertise on population statistics and data. I want to further thank Ilkka Jokipii, Joonas Kumpulainen, and Rick Mourits for their advice on different matters of the project and lastly, the Radboud Group for Historical Demography and Family History for hosting my research visit and work with the database.

REFERENCES

Ahlbäck, A. (2014). *Manhood and the making of the military. Conscription, military service and masculinity in Finland, 1917–39*. Farnham: Ashgate.

Ahoranta, T., & Ortamo-Närvä, A.-M. (2012). Arkisto säilyttää sankarivainajien muiston [Archive preserves memory of fallen heroes]. *Sukutieto, 29*(3), 12–13.

Barnett, A., Stanley, T., & Shore, M. (1992). America's Vietnam casualties: Victims of a class war? *Operations Research, 40*(5), 856–866.

Bellamy, R. F. (2000). Why is marine combat mortality less than that of the army? *Military Medicine, 165*(5), 362–367. doi: 10.1093/milmed/165.5.362

Bessel, R. (2015). Death and survival in the Second World War. In M. Geyer & A. Tooze (Eds.), *The Cambridge history of the Second World War. Volume III, Total war: Economy, society and culture* (pp. 252–276). Cambridge: Cambridge University Press. doi: 10.1017/CHO9781139626859.012

Buzzell, E., & Preston, S. H. (2007). Mortality of American troops in the Iraq War. *Population and Development Review, 33*(3), 555–566. doi: 10.1111/j.1728-4457.2007.00185.x

Fogel, R. W., Costa, D. L., Haines, M., Lee, C., Nguyen, L., Pope, C.,… Yetter, N. (2000). *Aging of veterans of the Union Army: Version M-5*. Chicago: Center for Population Economics, University of Chicago Graduate School of Business, Department of Economics, Brigham Young University, and The National Bureau of Economic Research. Retrieved from https://www.nber.org/programs-projects/projects-and-centers/union-army-data/union-army-data-citation-and-use-early-indicators-data

Fornasin, A., Breschi, M., & Manfredini, M. (2019). Deaths and survivors in war: The Italian soldiers in WW1. *Demographic Research, 40*(22), 599–626. doi: 10.4054/DemRes.2019.40.22

Harris, P. A., Taylor, R., Minor, B. L., Elliot, V., Fernandez, M., O'Neal, L.,… REDCap Consortium (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics, 95*. doi: 10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Gonde, J. G. (2009). Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377–381. doi: 10.1016/j.jbi.2008.08.010

Hautamäki, A., & Coleman, P. G. (2001). Explanation for low prevalance of PTSD among older Finnish war veterans: Social solidarity and continued significance given to wartime sufferings. *Aging & Mental Health, 5*(2), 165–174. doi: 10.1080/13607860120038348

Karjalainen, M. (2014). Tietojen vertailua, korjaamista ja asiakaspalvelua: Suomen sodissa 1939–1945 menehtyneiden tietokanta [Comparing information, corrections and customer service: Database of fallen in the Finnish wars 1939–1945]. In P. Happonen (Ed.), *Kleion pauloissa* [In Kleio's spell] (pp. 277–283). Helsinki: SKS.

Kemppainen, I. (2006). *Isänmaan uhrit: Sankarikuolema Suomessa toisen maailmansodan aikana* [The nation's heroes. Military death in Finland during the Second World War]. Helsinki: SKS.

Kinnunen, T., & Kivimäki, V. (Eds.). (2012). *Finland in World War II: History, memory, interpretations*. Leiden & Boston: Brill.

Kivimäki, V. (2013). *Battled nerves: Finnish soldiers' war experiences, trauma, and military psychiatry, 1941–44* (Doctoral dissertation). Åbo Akademi University, Turku. Retrieved from http://www.doria.fi/handle/10024/90586

Kivimäki, V. (2019). Sankariuhri ja kansakunta: Suomalaiset sotakuolemat 1939–1945 [Fallen heroes and nation: Finnish war deaths 1939–1945]. In I. Pajari, J. Jalonen, R. Miettinen & K. Kanerva (Eds.), *Suomalaisen kuoleman historia* [History of Finnish death] (pp. 277–310). Helsinki: Gaudeamus.

Kriner, D. L., & Shen, F. X. (2010). *The casualty gap: The causes and consequences of American wartime inequalities*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195390964.001.0001

Kriner, D. L., & Shen, F. X. (2016). Invisible inequality: The two Americas of military sacrifice. *University of Memphis Law Review, 46*(3), 545–635. Retrieved from https://www.memphis.edu/law/documents/kriner-shen46.pdf

Kronlund, J. (1988). *Puolustusvoimien rauhan ajan historia. Suomen puolustuslaitos 1918–1939* [Peacetime history of defence forces. Finnish defence establishment 1918–1939]. Porvoo: WSOY.

Kurenmaa, P., & Lentilä, R. (2005). Sodan tappiot [Casualties of war]. In J. Leskinen & A. Juutilainen (Eds.), *Jatkosodan pikkujättiläinen* [Small giant of Continuation War] (pp. 1150–1162). Helsinki: WSOY.

MacLean, A., & Elder, G. H. (2007). Military service in the life course. *Annual Review of Sociology, 33*, 175–196. doi: 10.1146/annurev.soc.33.040406.131710

McCalman, J., Kippen, R., McMeeken, J., Hopper, J., & Reade, M. (2019). Early results from the 'Diggers to Veterans' longitudinal study of Australian men who served in the First World War. short- and long-term mortality of early enlisters. *Historical Life Course Studies, 8*, 52–72. doi: 10.51964/hlcs9307

Merli, M. G. (2000). Socioeconomic background and war mortality during Vietnam's wars. *Demography, 37*(1), 1–15. doi: 10.2307/2648092

Modell, J., & Haggerty, T. (1991). The social impact of war. *Annual Review of Sociology, 17*, 205–224. doi: 10.1146/annurev.so.17.080191.001225

Nurminen, T. (2008). Muuttuva armeija [Changing army]. In J. Kulomaa & J. Nieminen (Eds.), *Teloitettu totuus — Kesä 1944* [Executed truth — Summer 1944] (pp. 47–75). Helsinki: Ajatus Kirjat.

Saarela, J., & Finnäs, F. (2012). Long-term mortality of war cohorts: The case of Finland. *European Journal of Population, 28*(1), 1–15. doi: 10.1007/s10680-011-9246-x

Stead, W. W., Hammond, W. E., & Straube, M. J. (1982). A chartless record – Is it adequate? *Proceedings of the Annual Symposium on Computer Application in Medical Care, Nov 2*, 89–94. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2580254/

Toivonen, T. (1998). War and equality: The social background of the victims of the Finnish Winter War. *Journal of Peace Research, 35*(4), 471–482. doi: 10.1177/0022343398035004

Urlanis, B. Ts. (1971). *Wars and population* (L. Lempert, Trans.). Moscow: Progress Publishers.

Waris, H. (1948). *Suomalaisen yhteiskunnan rakenne* [Structure of Finnish society]. Helsinki: Otava.

Winter, J. M. (1986). *The Great War and the British people*. London: Macmillan.

Ylönen-Peltonen, R. (2021). Sotapolun jäljillä arkistossa: Talvi- ja jatkosodan henkilöhistoriallisista arkistolähteistä Kansallisarkistossa [Tracing soldier's war path: Personal history records from Winter War and Continuation War years in national archive]. *Sotahistoriallinen aikakauskirja, 41*, 201–219.