

Cohort profile: CROSS-TRACKS – a population-based open cohort across healthcare sectors in Denmark

Authors: Anders Hammerich Riis^{1,2}, Pia Kjær Kristensen¹, Matilde Grøndahl Petersen¹, Ninna Hinchely Ebdrup^{1,3}, Simon Meyer Lauritsen^{2,3}, Marianne Johansson Jørgensen¹

Affiliations:

¹Horsens Regional Hospital, Horsens, Denmark

²Enversion A/S, Aarhus, Denmark

³Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

Corresponding author:

Anders Hammerich Riis (ORCID: 0000-0002-6684-4068)

Horsens Regional Hospital, Sundvej 30, 8700 Horsens, Denmark

Email: anderiis@rm.dk

Abstract

Purpose

This paper describes the open cohort CROSS-TRACKS, which was set up in Denmark in 2015. The cohort comprises population-based data from primary care, secondary care, and national registries. In addition, the paper outlines the new research opportunities provided by the cohort; the possibility to study patient pathways and transitions across sectors while adjusting for sociodemographic characteristics. The new data source is especially interesting for studies of potentially preventable admissions and readmissions.

Participants

A total of 221,283 individuals resided in the four Danish municipalities that constituted the catchment area of Horsens Regional Hospital in 2012-2018. A total of 96% of the population used primary care, 35% received at least one transfer payment, and 66% was in contact with a hospital at least once in the period. Additional clinical information is available for hospital contacts (eg alcohol intake, smoking status, body mass index, and blood pressure). A total of 27% (n=9,986) of individuals aged ≥ 65 years had at least one potentially preventable hospital admission, and 69% (n=6,881) of these individuals had more than one.

Findings to date

The cohort is currently used for research projects in epidemiology and artificial intelligence. These projects comprise a prediction model for potentially preventable hospital admissions, a clinical decision support system based on artificial intelligence, prevention of medication errors in the transition between sectors, health behaviour and sociodemographic characteristics of men and women prior to fertility treatment, and a recently published study applying machine learning methods for early detection of sepsis.

Future plans

The CROSS-TRACKS cohort will be expanded to comprise the entire Central Denmark Region consisting of 1.3 million residents. The cohort can provide new knowledge on how to best organise interventions across healthcare sectors and prevent potentially preventable hospital admissions. Such knowledge would benefit both the individual citizen and society as a whole.

Strengths and limitations of this study

- The CROSS-TRACKS cohort is a population-based open cohort containing routinely collected data from both primary and secondary care combined with sociodemographic register data.
- The cohort is readily available for research projects: the data sources have already been linked, patient pathways have been identified, and data have been remodelled and prepared for cutting-edge use of artificial intelligence and to enable easier and simpler querying across source systems.
- The cohort is currently limited to the catchment area of Horsens Regional Hospital, but it covers both urban and rural areas and it has a total population of 221,283 individuals by 2018.
- The cohort will be expanded to include additional catchment areas in the Central Denmark Region, comprising a total population of 1.3 million individuals.
- The cohort offers new opportunities for health research by tracking the patients' pathway across healthcare sectors to provide new knowledge on the best way to organise future interventions.

Introduction

An increasingly ageing population and limited resources place new demands on society and health care[1]. The global population of individuals aged 65+ years is projected to increase from 8.2% in 2015 to 15.9% in 2050, corresponding to >1.5 billion people[2]. The elderly population has higher needs for health care compared to younger citizens[3,4]. For example, the number of hospital admissions have increased by 45% from 2006 to 2018 among Danes aged 65+ years[5]. Therefore, Denmark and many other countries have reorganised the healthcare systems; they have reduced the number of hospital beds and shortened the length of hospital stays. The total number of days in a hospital bed has decreased by 16% from 2006 to 2018[5]. Therefore, an increasing part of the healthcare, treatment, and rehabilitation services are provided by general practitioners (GP) and the municipalities in the primary care setting.

One focus has been directed towards reducing the number of potentially avoidable readmissions[6]. In the US, almost 20% of discharged patients are readmitted to hospital within 30 days[7]. This is costly to the individual citizens and to society. Special attention has been turned towards potentially preventable admissions and readmissions. These are often seen for ambulatory care sensitive conditions (ACSCs), which comprise conditions that can usually be treated and prevented in the primary care setting[8]. More than 20% of hospital admissions among elderly US residents in nursing homes are estimated to concern ACSCs[9]. In Denmark, 15% of all hospitalisations of the elderly population are expected to be preventable[10]. The rates of potentially preventable hospital admissions are high in Denmark compared to other European countries[11]. Furthermore, a recent study found higher rates of potentially preventable hospitalisations and readmissions for chronic medical conditions in Denmark compared to the US[12]. Consequently, there seems to be room for improvement in our approach to preventable admissions and readmissions.

Several interventions have aimed to reduce the risk of readmissions[13] and to improve the care for discharged patients in the primary care setting[14]. Most of these focused on the pre-discharge period (eg patient education and discharge planning) or the post-discharge period (eg follow-up telephone calls and home visits), whereas only few interventions bridged over both sectors (eg patient-centred discharge instructions and transition coaching). The interventions bridging across sectors were most effective[13]. This calls for studies of patient transitions across sectors in a population-based setting.

CROSS-TRACKS[15] is an open cohort comprising routinely collected administrative data from primary care, secondary care, and national registries. The combined data sources allow us to explore key factors such as

sociodemographic characteristics, labour market affiliation, social transfers, home-based nurse care, GP contacts, and potential links between these factors and the use of services in secondary care.

The cohort is a valuable tool for studying patient pathways across sectors, including the patterns that characterise potentially preventable hospital admissions. This paper describes the new opportunities in healthcare research that are emerging from the comprehensive CROSS-TRACKS cohort.

Cohort description

Setting

The tax-funded healthcare system in Denmark provides free access for all citizens to healthcare services. The five Danish regions are the backbone of the healthcare system, and they are responsible for organising and operating the somatic and psychiatric hospitals (Figure 1). Furthermore, the regions reimburse GPs and specialists in private practice (eg physiotherapists, psychologists, or dentists) for dispensed prescriptions and services provided[16]. The GPs can refer citizens to a wide range of welfare services (eg home care and rehabilitation) that are provided in the primary care setting by the 98 municipalities under the regions[16]. The catchment area usually comprises the municipalities in the geographical area around the hospital.

The GP is the primary entry point into the healthcare system for Danish citizens, and the GP acts as a gatekeeper to secondary care[17], but the citizens can contact the prehospital directly for emergencies. Citizens have unrestricted access to dental care, which is free of charge for citizens below the age of 18. After hospital discharge, patients may be referred to the GP or a triage manager in the municipality for further treatment, care, or rehabilitation.

All Danish citizens receive a unique identifier at birth or immigration. This number is also referred to as the civil registration (CPR) number[18]. The CPR number enables electronic linkage for all data in the cohort across registries and databases. The cohort data is securely stored at the data warehouse of the Central Denmark Region in accordance with data protection regulations. The cohort is readily available for research projects as the data sources have already been linked and patient pathways have been identified. Further, the data have been remodelled according to the dimensional modelling principles laid out by Kimball and Ross[19], to provide a simple and consistent query interface across source systems.

Study population

The CROSS-TRACKS cohort consists of all citizens aged 18+ years residing in the catchment area of Horsens Regional Hospital, which comprises Hedensted, Horsens, Odder and Skanderborg municipalities (Figure 1). The inclusion period started on 1 September 2012 (cohort entry date) and will last through 2022 (Figure 2). Individuals are included on their 18th birthday or when moving to one of the four included municipalities. Individuals are followed from the date of inclusion until five years after the end of the inclusion period (31 December 2027) or date of death. Individuals moving away from these municipalities are followed for five years after the date of moving.

Patient characteristics are time-dependent because disease develops over time, and this affects the number and frequency of contacts with the healthcare system. The use of nationwide data sources ensures optimal adjustment for potential differences in patient characteristics, even for individuals living outside the included municipalities during part of the follow-up period. Moreover, we use look-back periods to assess covariates prior to cohort entry date[20]. The length of look-back periods can be found in Figure 2.

Data collection

Individuals are included in the cohort according to the inclusion criteria shown in the flow chart in Figure 3.

Administrative and health registries

The CPR number originates from the Danish Civil Registration System (CRS)[18], which also provides date of birth, gender, vital status, civil status, migration, and place of residence. The CRS is updated on a daily basis with information on vital status and place of residence; this ensures complete follow-up for all individuals. The CROSS-TRACKS cohort includes data from the CRS from 1 September 2002, which enables a 10-year look-back period prior to each individual's date of inclusion.

Weekly recordings of social benefits and other transfer incomes are retrieved from the Danish Register for Evaluation of Marginalization (DREAM)[21]. DREAM can be used to categorise individuals into unemployed, on sick leave, receiving part-time benefit, receiving disability pension, or receiving retirement pension[22].

Primary care

The municipality-based electronic care provision record (ECPR) holds data on the health services granted to the citizens. These services comprise municipality-based practical help, rehabilitation assistance, personal

care, and home visits by a community nurse. The data includes the period (start and end dates) and total amount of time assigned to the services.

The Danish National Health Service Register (DNHSR) holds information on contacts with healthcare professionals (eg GP, dentist), type of consultation (eg ordinary consultations, home visits, telephone consultations, electronic consultations), services provided, and laboratory tests performed[23].

The Danish National Database of Reimbursed Prescriptions (DNDRP) contains detailed information on prescribed drugs that are partially reimbursed by the regions. This includes information on date of the redemption, number of pills, dose, number of defined daily doses, name and type of drug according to the Anatomical Therapeutic Chemical classification system[24].

The prehospital electronic record holds information on treatment and transportation of patients by ambulance, medical cars, or helicopters. Registrations include urgency level, date and time of departure and arrival, treatments, medications, and a free-text field for additional information that may help ensure optimal transfer of patients to the hospital[25].

Secondary care

The Danish National Patient Registry (DNPR) encompasses all hospital contacts, including admission date, discharge date, and outpatient visits. Diagnoses, surgical procedures, and examinations are registered according to the International Classification of Diseases[26]. The main reason for each hospital contact is registered as the primary diagnosis, whereas secondary diagnoses are stated as other diseases in relation to the contact. Diagnoses registered in the DNPR can be used to identify individuals admitted to hospital with a primary ACSC-related diagnosis. In addition, the Charlson Comorbidity Index (CCI) score is calculated based on 19 conditions[27].

The electronic health record (EHR) in the Central Denmark Region consists of healthcare-related registrations linked with date and time. The DNPR is based on registrations in the EHR, but this record contains additional data that are not transferred to the DNPR. This includes registrations on triage, patient's pulse, blood pressure, respiration frequency, and microbiology test results. Some of these are used for early detection of critical illness. The EHR can also be used to identify patient-related lifestyle factors, such as diet, smoking and alcohol habits, and physical activity. Furthermore, hospital-based drug treatment is registered in the EHR by the doctors prescribing the drug and the nurses administering the drug by dispensing the medicine to the patients.

The clinical laboratory information system (LABKA) consists of test results from blood and urine samples drawn by health professionals (in both primary care and hospitals) and analysed in clinical laboratories[28]. Registrations include date of sample, health professional or hospital department requesting the analysis, requested assay or biomarker, and corresponding Nomenclature, Properties and Units (NPU) code.

An overview of data sources and availability can be seen in Figure 2 and Table 2. The codes used to extract the characteristics of citizens can be found in Table 1 in the supplementary material.

Although the EHR data in the data warehouse of the Central Denmark Region is updated on a daily basis, the combined cohort consisting of all data sources is only updated annually. The cohort is dynamic in the sense that the data quality is continuously improved as new research projects build on new sections of the cohort. See Table 2 in the supplementary data for a description on how registrations on body mass index (BMI) was cleaned before use.

Population characteristics

On the cohort entry date, 166,970 individuals aged 18+ years resided in the four municipalities. From 1 September 2012 through 2018, 24,235 individuals turned 18 years of age, and 30,078 moved to one of the municipalities. This yields a total population of 221,283 individuals in the open cohort (Figure 3). The main characteristics of citizens in the cohort are shown in Table 1. Individuals aged below 18 years appear in the table, but they are not included in the cohort before their 18th birthday.

The catchment area of Horsens Regional hospital contained one health centre (located in Skanderborg), 69 GPs, 20 specialists in private practice, 11 chiropractors, 52 dentists, 45 physiotherapists, 25 podiatrists, and 25 psychologists.

The largest municipality was Horsens with 78,758 citizens, corresponding to 36% of the population on 1 September 2012. Skanderborg had 49,160 (22%) citizens, Hedensted had 39,981 (18%) citizens, and Odder had 19,493 (9%) citizens. In the beginning of the inclusion period, Horsens was the only urban municipality, but Skanderborg went from rural to urban during the period.

The median age on cohort entry date was 41 years (IQR: 25-58 years) with an equal distribution of gender. The proportion of citizens who were married or living in a registered partnership was 44%, and 44% of the citizens aged 18 years and above were living alone on cohort entry date.

During the inclusion period from 1 September 2012 through 2018, 35% of the total population received some kind of transfer payment. Sick leave was the predominant reason with 10%, which was followed by unemployment among 9% of the population.

A total of 21% had a CCI score of 1 or above at the end of the inclusion period. For the part of the population (66%) with a hospital contact in the period from 1 September 2012 through 2018, we used registrations in the EHR to obtain the clinical characteristics. Early detection of critical illness was registered in the EHR among 43% of the population with at least one hospital contact. Registrations of blood pressure were used to identify patients with at least one registration of high blood pressure (49%). Smoking status was recorded in 48% of individuals with at least one hospital contact, and 53% of these reported any smoking. Alcohol intake was registered in 40% with at least one hospital contact, and 30% of these had at least one incidence of intake above the recommendations of the health authorities, which is up to 7 drinks per week for females and 14 drinks per week for males[29]. Of the 90,385 patients with information on BMI, 50,222 (56%) were overweight (BMI ≥ 25 kg/m²).

Healthcare utilisation

Almost everyone in the population received at least one service provided from primary care during the inclusion period. Daytime GP services were utilised by 95% of the population, dental services by 77%, and psychologist services by 8% during the inclusion period. Municipality-based care was provided to 13% of the citizens receiving rehabilitation assistance (8%), and nursing care (8%) was the most frequent service.

Antibiotics was the most frequently used type of prescription medication; 57% of the population redeemed at least one antibiotics prescription during the inclusion period. NSAID prescriptions was redeemed by 37%, and blood pressure lowering medication was used by 24% of the population.

A total of 42,827 patient transportations were made by the prehospital in the period from 2017 to 2018; 26% were acute, 23% were urgent, and 49% had unknown urgency level. The transports concerned 18,671 unique individuals, and 7,786 individuals had more than one transport. Overall 66% of the population had at least one hospital contact during the inclusion period, 33% had an inpatient contact, 61% had an outpatient visit, and 27% had an emergency room visit.

In the population of individuals aged 65+ years, 9,986 (27%) had at least one potentially preventable hospital admission during the inclusion period; 6,881 (69%) of these had more than one potentially preventable hospitalization.

Almost three million medications were ordered, and about 12 million medications were dispensed in the hospitals. Overall, 4.5 million biological samples were drawn, and 52 million laboratory tests were performed on these samples.

Findings to date

The CROSS-TRACKS cohort was established to gain greater research-based knowledge on the causes of potentially preventable (re)admissions by tracking patient pathways across healthcare sectors. The new knowledge can be used to strengthen the collaboration across sectors and to target joint efforts.

The cohort is currently being used for several ongoing health research projects. These include a prediction model for potentially preventable hospital admissions, a clinical decision support system based on artificial intelligence, prevention of medication errors in the transition between sectors, and health behaviour and sociodemographic characteristics of men and women in fertility treatment. Another study has applied machine learning methods for early detection of sepsis and has identified potential for initiating interventions earlier[30]. In continuation of this study, the authors used explainable artificial intelligence to develop an early-warning score system for detection of acute critical illness[31]. This study adds transparency in artificial intelligence by explaining which relevant electronic health records data predicts critical illness.

Strengths and limitations

CROSS-TRACKS is the first cohort to provide readily available data from primary and secondary care for research projects: the data sources have already been linked, and patient pathways have been identified. A major strength of the cohort is the completeness and the high quality of the available data, which is obtained through the prospective collection of population-based data. An additional strength is that the unique Danish CPR number allows linkage across registers and complete follow-up for all individuals; this offers the possibility to monitor disease progression over time. The clinical information about the patients can be followed through the EHR when they live in the Central Denmark Region. If they move away, the DNPR can be used to follow the individual's disease development. Additionally, the 10-year look-back period and the follow-up through 2027 hold the possibility of following citizens for up to 25 years.

Given the detailed information on medication in both sectors, this cohort presents the complete history of registry-based drug use for each individual; this fosters rich opportunities for adjusting for confounders. Furthermore, the combination of psychiatric inpatient contacts, antidepressants and antipsychotics prescriptions, and talk therapy by GPs and psychologists offers potential for identifying psychiatric disease.

In addition, several validation studies have reported high validity of the Danish registry data[21,26,32]. The cohort itself forms a basis for easy validation as the free-text registrations in the EHR are usually regarded as the gold standard to which registry information is compared.

Nevertheless, we note some limitations of the cohort. Not all data sources have yet been validated, and there are missing data on lifestyle characteristics, as they may not have been reported for all citizens in contact with a hospital. However, investigator bias is limited since much of the data is recorded by clinicians during diagnosis and treatment, and not for research purposes. The information on medication use in primary care is based on reimbursable redeemed prescriptions, whereas no information on over-the-counter medication is available. However, citizens with regular medication use have economic incentives to acquire the prescribed drugs because of the reimbursement option. Furthermore, a study on the use of aspirin and NSAIDs found that the potential for identifying medication use through prescription registries in Denmark is high[33].

Another limitation is the missing data on treatments by private psychologists, physiotherapists, and chiropractors; these services are not covered by the healthcare system. However, the private healthcare sector is small in Denmark and is mostly based on services eligible for reimbursement by the regions, and these services are thereby included in the registries. Additionally, private hospitals report to the DNPR[16], which also contributes to maintaining the high completeness of registry data.

The municipalities use different suppliers for reporting to the ECPR. This means that it requires extensive data management to streamline the information, and this could be a potential limitation. The inclusion of ECPR data is conducted in close collaboration with the municipalities and their data managers to ensure correct interpretation and high validity of the data.

Only the catchment area for Horsens Regional hospital is included in the cohort, and this may compromise the generalisability. However, the catchment area includes both urban and rural municipalities. In addition, citizen contacts with Aarhus University Hospital, which is located in the Central Denmark Region, are also included in the cohort through the recordings in the EHR. Because of the structure of the Danish national registries and databases, the cohort can easily be extended to comprise additional municipalities and regions and ultimately all of Denmark. Starting in 2020, we will expand the cohort to include all 1.3 million residents in the Central Denmark Region. This region consists of one university hospital, one psychiatric hospital, four hospitals with 24-hour emergency department, four hospitals without 24-hour emergency department, and ten health centres and acute clinics[16].

The CROSS-TRACKS cohort provides valuable new opportunities for health research across primary and secondary care in the future; this is expected to benefit both the individual citizens and society as a whole. The cohort can contribute with new knowledge on how to organise interventions across sectors to reduce the number of potentially preventable (re)admissions.

Collaboration

Access to CROSS-TRACKS data can be granted for research and analysis projects if certain criteria are met, and an application has been submitted to the steering committee. The committee strongly encourages national and international collaboration. For more information on access criteria and application form, see the website www.tvaerspor.dk (currently in Danish).

Further details

Contributors

AHR, PKK, and MJJ wrote the first draft of the manuscript. MGP was responsible for data access. AHR performed the statistical analyses. All authors provided critical review of the manuscript and approved the final version.

Funding

The project was supported by Horsens Regional Hospital, Fund for the advancement of health research in Central Denmark Region, the development and resource funds under Danish Regions (*Udviklings- og forskningspuljen ved Danske Regioner*), the Danish Health Confederation (*Sundhedskartellet*), Public Health in the Central Denmark Region – a collaboration between municipalities and the region (*Folkesundhed i Midten*), and Circular Co-Creation under the MedTec Innovation Consortium.

Competing interests

None declared.

Patient and public involvement

Patients and the public were not involved in the development of this research cohort.

Patient consent

Not required.

Ethical approval

Register-based studies require no ethical approval in Denmark. The study was approved by the Danish Data Protection Agency (file numbers 2007-58-0010 and 2012-58-006) and the Danish Patient Safety Authority (file number 3-3013-1380/1, 3-3013-1380/2, and 3-3013-2741/1).

Provenance and peer review

Not commissioned; externally peer reviewed.

Data sharing statement

See Collaboration section.

Acknowledgements

We gratefully thank the CROSS-TRACKS steering committee for initiating the project and providing access to the cohort. We also thank the Enversion A/S team for data acquisition, modelling, validation, and support.

Supplementary material

Table 1. Codes for defining population characteristics and healthcare utilisation

Table 2. Cleaning of body mass index registrations from the electronic health record

Table 3. Cleaning of blood pressure registrations from the electronic health record

References

- 1 Beard JR, Officer AM, Cassels AK. The World Report on Ageing and Health. *Gerontologist* 2016;**56**:S163–6. doi:10.1093/geront/gnw037
- 2 United Nations. Department of Economic and Social Affairs. Population dynamics. World population prospects 2019. <https://population.un.org/wpp/Download/Probabilistic/Population/> (accessed 1 May 2020).
- 3 Christensen K, Doblhammer G, Rau R, *et al.* Ageing populations: the challenges ahead. *Lancet (London, England)* 2009;**374**:1196–208. doi:10.1016/S0140-6736(09)61460-4
- 4 Sander M, Oxlund B, Jespersen A, *et al.* The challenges of human population ageing. *Age Ageing* Published Online First: 2015. doi:10.1093/ageing/afu189
- 5 Statistics Denmark. Statistics Bank. <https://www.statistikbanken.dk/statbank5a/default.asp?w=3440> (accessed 1 May 2020).
- 6 van Walraven C, Jennings A, Forster AJ. A meta-analysis of hospital 30-day avoidable readmission rates. *J Eval Clin Pract* 2012;**18**:1211–8. doi:10.1111/j.1365-2753.2011.01773.x
- 7 Jencks SF, Williams M V., Coleman EA. Rehospitalizations among patients in the medicare fee-for-service program. *N Engl J Med* Published Online First: 2009. doi:10.1056/NEJMsa0803563
- 8 Purdy S, Griffin T, Salisbury C, *et al.* Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public Health* Published Online First: 2009. doi:10.1016/j.puhe.2008.11.001
- 9 McAndrew RM, Grabowski DC, Dangi A, *et al.* Prevalence and patterns of potentially avoidable hospitalizations in the US long-term care setting. *Int J Qual Heal Care* 2015;**28**:104–9. doi:10.1093/intqhc/mzv110
- 10 Kommunernes Landsforening [Local Government Denmark]; Danske Regioner [Danish Regions]; Finansministeriet [Danish Ministry of Finance]; Ministeriet for Sundhed og Forebyggelse [Danish Ministry of Health and Prevention]. Forebyggelse af indlæggelser. Synlige resultater. [Prevention of hospital admissions. Visible results]. 2014.<https://sundhedsdatastyrelsen.dk/da/tal-og-analyser/analyser-og-rapporter/sundhedsvaesenet/synlige-resultater> (accessed 1 May 2020). Danish
- 11 Thygesen LC, Christiansen T, Garcia-Armesto S, *et al.* Potentially avoidable hospitalizations in five

- European countries in 2009 and time trends from 2002 to 2009 based on administrative data. *Eur. J. Public Health*. 2015. doi:10.1093/eurpub/cku227
- 12 Schiøtz M, Price M, Frølich A, *et al*. Something is amiss in Denmark: A comparison of preventable hospitalisations and readmissions for chronic medical conditions in the Danish Healthcare system and Kaiser Permanente. *BMC Health Serv Res* 2011;**11**:347. doi:10.1186/1472-6963-11-347
- 13 Hansen LO, Young RS, Hinami K, *et al*. Interventions to Reduce 30-Day Rehospitalization: A Systematic Review. *Ann Intern Med* 2011;**155**:520–8. doi:10.7326/0003-4819-155-8-201110180-00008
- 14 Hesselink G, Schoonhoven L, Barach P, *et al*. Improving Patient Handovers From Hospital to Primary Care: A Systematic Review. *Ann Intern Med* 2012;**157**:417–28. doi:10.7326/0003-4819-157-6-201209180-00006
- 15 TVÆRSPOR [CROSS-TRACKS]. <http://www.tvaerspor.dk/> (accessed 1 May 2020). Danish
- 16 Schmidt M, Schmidt SAJ, Adelborg K, *et al*. The Danish health care system and epidemiological research: from health care contacts to database records. *Clin Epidemiol* 2019;563–91. doi:10.2147/CLEP.S179083
- 17 Pedersen KM, Andersen JS, Sondergaard J. General Practice and Primary Health Care in Denmark. *J Am Board Fam Med* Published Online First: 2012. doi:10.3122/jabfm.2012.02.110216
- 18 Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur. J. Epidemiol*. 2014. doi:10.1007/s10654-014-9930-3
- 19 Kimball R, Ross M. *The Data Warehouse Toolkit, The Definitive Guide to Dimensional Modeling*. Wiley 2013.
- 20 Riis AH, Johansen MB, Jacobsen JB, *et al*. Short look-back periods in pharmacoepidemiologic studies of new users of antibiotics and asthma medications introduce severe misclassification. *Pharmacoepidemiol Drug Saf* 2015;**24**:478–85. doi:10.1002/pds.3738
- 21 Hjollund NH, Larsen FB, Andersen JH. Register-based follow-up of social benefits and other transfer payments: Accuracy and degree of completeness in a Danish interdepartmental administrative database compared with a population-based survey. *Scand J Public Health* Published Online First: 2007. doi:10.1080/14034940701271882
- 22 Fonager K, Lundbye-Christensen S, Andreasen JJ, *et al*. Work Status and Return to the Workforce

- after Coronary Artery Bypass Grafting and/or Heart Valve Surgery: A One-Year-Follow Up Study. *Rehabil Res Pract* 2014;631842. doi:10.1155/2014/631842
- 23 Sahl Andersen J, De Fine Olivarius N, Krasnik A. The Danish national health service register. *Scand J Public Health* Published Online First: 2011. doi:10.1177/1403494810394718
- 24 Johannesdottir SA, Horváth-Puhó E, Ehrenstein V, *et al.* Existing data sources for clinical epidemiology: The Danish National database of reimbursed prescriptions. *Clin Epidemiol* 2012.
- 25 Lindskou TA, Mikkelsen S, Christensen EF, *et al.* The Danish prehospital emergency healthcare system and research possibilities. *Scand. J. Trauma. Resusc. Emerg. Med.* 2019 Nov 4;27(1):100. doi:10.1186/s13049-019-0676-5
- 26 Schmidt M, Schmidt SAJ, Sandegaard JL, *et al.* The Danish National patient registry: A review of content, data quality, and research potential. *Clin. Epidemiol.* 2015 Nov 17;7:449-90. doi:10.2147/CLEP.S91125
- 27 Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- 28 Grann AF, Erichsen R, Nielsen AG, *et al.* Existing data sources for clinical epidemiology: The clinical laboratory information system (LABKA) research database at Aarhus University, Denmark. *Clin Epidemiol* Published Online First: 2011. Apr 1;3:133-8. doi:10.2147/CLEP.S1790
- 29 Sundhedsstyrelsens udmeldinger om alkohol [The Danish Health Authority’s statements about alcohol intake]. <https://www.sst.dk/da/viden/alkohol/alkoholforebyggelse/sundhedsstyrelsens-udmeldinger-om-alkohol> (accessed 1 May 2020). Danish
- 30 Lauritsen SM, Kalør ME, Kongsgaard EL, *et al.* Early detection of sepsis utilizing deep learning on electronic health record event sequences. *Artif Intell Med* Published Online First: 2020. doi:10.1016/j.artmed.2020.101820
- 31 Lauritsen SM, Kristensen M, Olsen MV, *et al.* Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun* (Under review 2020). Preprint available: <https://arxiv.org/pdf/1912.01266v1.pdf> (accessed 1 May 2020)
- 32 Riis AH, Erichsen R, Ostfeldt EB, *et al.* Validating registry data on statins prescriptions by blood measurements. *Pharmacoepidemiol Drug Saf* 2019;28. doi:10.1002/pds.4700
- 33 Schmidt M, Hallas J, Friis S. Potential of prescription registries to capture individual-level use of

aspirin and other nonsteroidal anti-inflammatory drugs in Denmark: Trends in utilization 1999-2012.

Clin Epidemiol 2014. May 12;6:155-68. doi:10.2147/CLEP.S59156

- 34 Pedersen HS, Fenger-Grøn M, Bech BH, *et al.* Frequency of health care utilization in the year prior to completed suicide: A Danish nationwide matched comparative study. *PLoS One* 2019. Mar 27;14(3). doi:10.1371/journal.pone.0214605
- 35 Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health* 2011;39:30–3. doi:10.1177/1403494811401482

Tables and figures

Table 1. Sociodemographic characteristics, clinical characteristics, and healthcare utilization among citizens in the four included Danish municipalities (Hedensted, Horsens, Odder and Skanderborg) in 2012-2018

Sociodemographic characteristics	N (%)
Age at cohort entry date (1 September 2012)	
Median age at cohort entry date, years (IQR)	41 (25-58)
<18 ^A	24,235 (11)
18-39	80,434 (36)
40-64	80,288 (36)
65-79	28,196 (13)
80+	8,130 (3.7)
Gender	
Female	109,267 (49)
Male	112,016 (51)
Municipality (on 1 September 2012)	
Hedensted	39,981 (18)
Horsens	78,758 (36)
Odder	19,493 (8.8)
Skanderborg	49,160 (22)
Other (individuals moving to the four municipalities after 1 September 2012)	33,891 (15)
Individuals moving away from the municipalities	38,967 (18)
Marital status	
Married or registered partnership	97,216 (44)
Divorced or dissolved partnership	17,859 (8.1)
Widowed or longest living in a registered partnership	11,077 (5.0)
Not married	77,815 (35)
Unknown	17,316 (7.8)
Living alone on cohort entry date ^B	72,967 [44]
Transfer payments	
Any kind of transfer payment	77,410 (35)
Unemployment	19,849 (9.0)
Sick leave	21,873 (9.9)
State education grant	13,021 (5.9)
Maternity/paternity leave	8,356 (3.8)

Retirement pension	8,266 (3.7)
Disability pension	1,047 (0.5)
Social security benefits	6,996 (3.2)
Deaths during follow-up	11,470 (5.2)
Clinical characteristics	
Charlson Comorbidity Index score at the end of the inclusion period	
0 (no comorbidity)	174,159 (79)
1 (low)	21,226 (9.6)
2 (medium)	13,382 (6.0)
3+ (high)	12,516 (5.7)
Citizens with at least one hospital contact (recorded in the electronic health records)	147,087 (66)
Early detection of critical illness, total score ^C	63,258 [43]
Diastolic blood pressure ^C	102,048 [69]
Systolic blood pressure ^C	102,085 [69]
High blood pressure (at least one registration of diastolic > 90 mm Hg or systolic > 140 mm Hg) ^C	71,029 [48]
Pulse ^C	100,589 [68]
Respiratory rate ^C	71,904 [49]
Oxygen saturation ^C	81,096 [55]
Temperature ^C	71,000 [48]
Smoking status ^C	70,557 [48]
No smoking ^D	32,978 [47]
Any smoking ^D	37,579 [53]
Current smoker (at last registration) ^D	17,484 [25]
Occasional smoker (at last registration) ^D	1,216 [1.7]
Previous smoker (at last registration) ^D	18,879 [27]
Alcohol intake ^C	58,833 [40]
Within recommendations; ≤7 drinks for females, ≤14 drinks for males ^E	41,320 [70]
Above recommendations; >7 drinks for females, >14 drinks for males (any time) ^E	17,513 [30]
Body mass index (at least one registration) ^C	90,384 [61]
<18.5 kg/m ^{2F}	3,202 [3.5]
18.5 – 24.9 kg/m ^{2F}	36,961 [41]
25 – 29.9 kg/m ^{2F}	30,568 [34]
30 – 34.9 kg/m ^{2F}	13,068 [14]
35 – 39.9 kg/m ^{2F}	4,451 [4.9]

$\geq 40 \text{ kg/m}^2$ ^F	2,134 [2.4]
Triage of patients (at least one registration) ^C	43,431 [30]
Immediate resuscitation (red) ^G	1,474 [3.4]
Very urgent (orange) ^G	9,568 [22]
Urgent (yellow) ^G	17,516 [40]
Standard (green) ^G	14,549 [33]
Non-urgent (blue) ^G	324 [0.75]
Healthcare utilization	
Any health services from the municipalities	29,158 (13)
Rehabilitation (eg physiotherapy)	17,118 (7.7)
Practical help (eg home cleaning)	11,692 (5.3)
Personal care (eg weekly bath)	12,419 (5.6)
Home visits by a community nurse (eg injection of medicine)	16,709 (7.6)
Any primary healthcare contact	212,952 (96)
Daytime face-to-face contact with general practitioner	210,718 (95)
Out-of-hours face-to-face contact with general practitioner	82,684 (37)
Blood tests at general practitioner	157,689 (71)
Haemoglobin test at general practitioner	106,915 (48)
Spirometer test at general practitioner	36,164 (16)
Electrocardiograph at general practitioner	81,034 (37)
Talk therapy at general practitioner	26,869 (12)
Urine test at general practitioner	92,254 (42)
Rapid strep test at general practitioner	46,179 (21)
C-reactive protein test at general practitioner	132,728 (60)
Other service at general practitioner	208,597 (94)
Chiropractor	46,345 (21)
Dentist	171,333 (77)
Physiotherapist	66,788 (30)
Podiatrist	7,170 (3.2)
Psychologist	17,527 (7.9)
Prescription medication	
At least one prescription of antibiotics	126,280 (57)
At least one prescription of antidiabetics	12,577 (5.7)
At least one prescription of aspirins	21,613 (9.8)
At least one prescription of blood pressure lowering medication	53,127 (24)

At least one prescription of statins	33,837 (15)
At least one prescription of proton pump inhibitors	47,845 (22)
At least one prescription of systemic glucocorticoids	25,654 (12)
At least one prescription of inhaled corticosteroid therapy	30,999 (14)
At least one prescription of NSAIDs	80,944 (37)
At least one prescription of opioids	34,893 (16)
At least one prescription of antidepressants	30,491 (14)
At least one prescription of antipsychotics	9,767 (4.4)
Urgency level among transports in 2017-2018, total number	42,827
Acute (A) ^H	11,181 [26]
Urgent (B) ^H	9,637 [23]
Scheduled (C) ^H	865 [2.0]
Supine transport (D) ^H	23 [0.1]
Other service (E) ^H	72 [0.2]
Urgency level unknown ^H	21,049 [49]
Hospital contacts	
At least one hospital contact	147,087 (66)
At least one inpatient hospital contact	73,791 (33)
At least one inpatient hospital contact with acute admission	62,729 (28)
At least one outpatient hospital contact	134,330 (61)
At least one emergency room visit	59,382 (27)
Emergency call to the prehospital resulting in assistance by medical car	5,624 (2.5)
Emergency call to the prehospital resulting in assistance by helicopter	110 (0.05)
Emergency call to the prehospital resulting in no transportation	339 (0.15)
Potentially preventable admissions among individuals aged 65+ years	9,986 [27]
In-hospital medication ordering	2,778,432
In-hospital medication dispensing	12,240,606
Biological samples	4,562,264
Laboratory tests in clinical laboratories	52,607,667

Data are given as number (percentage of total population) or [percentage of subpopulation] during the inclusion period (1 September 2012 to 31 December 2018), unless otherwise specified.

^AIncluded from their 18th birthday

^BAmong 166,970 individuals living in the municipalities aged 18 years and above.

^CAmong 147,087 individuals with at least one hospital contact.

^DAmong 70,557 individuals with at least one registration of smoking status.

^EAmong 58,833 individuals with at least one registration of alcohol intake.

^FLast registration among 90,385 individuals with at least one registration of body mass index.

^GLast registration among 43,431 individuals with at least one registration of triage.

^HAmong 42,827 transports in 2017-2018.

Table 2. Administrative and health registries included in the CROSS-TRACKS cohort

Name and type of data source and year of initiation	Description of data source	Main variables	Period of data inclusion in the CROSS-TRACKS cohort
Danish Civil Registration System[18], administrative registry, 1968	Information on all individuals residing in Denmark	<ul style="list-style-type: none"> • CPR number (allowing linkage across all data sources) • Date of birth • Gender • Vital status (updated daily) • Civil status (updated daily) • Migration (updated daily) • Place of residence (updated daily) 	1 September 2002 – 31 December 2027
Danish Register for Evaluation of Marginalization[21,22], administrative registry, 1991	Weekly recordings of transfer payments	<ul style="list-style-type: none"> • Reasons for receiving transfer payment (Unemployed, sick leave, part-time benefit, disability pension, or retirement pension) • Dates for starting and ending transfer payment 	1 September 2007 – 31 December 2027
Municipality-based electronic care provision record, health database, 2012	Health services provided by the municipalities to citizens	<ul style="list-style-type: none"> • Type of service (assistance for home cleaning, personal care, home visits by a community nurse, or rehabilitation) • Subtype of service (eg community nurses can inject medicine, treat wounds, or measure vital parameters) • Location of service provided (eg citizen's own home or nursing home) • Dates for starting and ending the service • Total amount of time assigned to the service 	1 September 2012 – 31 December 2027
Danish National Health Service Register[23,34], administrative registry, 1990	Contacts with healthcare professionals	<ul style="list-style-type: none"> • Service provider (eg GP, physiotherapist, dentist) • Type of consultation (ordinary consultations, home visits, telephone consultations, or electronic consultations via email or a secure website) • Service provided • Laboratory tests performed • Amount of reimbursement • Date of the service • Dates for starting and ending treatment 	1 September 2007 – 31 December 2027
Danish National Database of Reimbursed Prescriptions[24,32],	Filled reimbursable prescriptions from community pharmacies	<ul style="list-style-type: none"> • Name of prescribed drug • Type of drug according to the Anatomical Therapeutic Chemical classification system • Number of pills 	1 September 2007 – 31 December 2027

administrative database, 2004		<ul style="list-style-type: none"> • Dose • Number of defined daily doses • Date of redemption 	
Prehospital electronic record[25], health database, 2012	Patient records from emergency medical services	<ul style="list-style-type: none"> • Type of transport (ambulance, medical car or helicopter) • Urgency level • Distances • Addresses • Detailed patient journal included from 2017: Information on treatments, medications, and journal notes as free text • Date and time stamps for take offs and arrivals 	1 September 2012 – 31 December 2027
Danish National Patient Registry[26,35], administrative registry, 1977	Hospital contacts	<ul style="list-style-type: none"> • Type of contact (inpatient or outpatient) • Acute vs. scheduled • Hospital and department code • Diagnoses, surgical procedures, and examinations according to the International Classification of Diseases • Primary diagnosis and secondary diagnoses in relation to the contact • Course of admission (since 2019) • Costs associated with the hospital contacts, treatments, and procedures • Dates of admission and discharge for inpatients • Date of contact for outpatients 	1 September 2002 – 31 December 2027
Electronic health record, health database, 2012	Secondary data source for the Danish National Patient Registry	<ul style="list-style-type: none"> • Diagnoses, treatments, and examinations • Structured registrations of health care content • Triage • Microbiology • Standardized registrations on patient's pulse, blood pressure, respiration frequency etc. used for early detection of critical illness • Registrations of patients' lifestyle factors regarding diet, smoking, alcohol and physical activity • Recordings of patient's height and weight used for calculating body mass index. See supplementary data table 2 for details on data cleaning. • In-hospital medication: Doctor's orderings and nurses' dispensing's 	1 September 2012 – 31 December 2027

		<ul style="list-style-type: none"> • Meta data on updates of patient’s electronic medication records shared by all Danish health professionals • Journal notes as free text • Date and time stamps of registrations 	
Clinical laboratory information system[28], health database, 2000	Laboratory test results from clinical laboratories	<ul style="list-style-type: none"> • Type of biological material • Identification of health professional or hospital department requesting the analysis • NPU codes for types of tests requested (a blood specimen can be analysed for several different contents) • Test results and units • Date and time of sample and test results 	1 September 2012 – 31 December 2027

Figure legends

Figure 1. The five Danish regions and the 98 municipalities. The black circle indicates the four municipalities (Hedensted, Horsens, Odder and Skanderborg) included in the CROSS-TRACKS cohort, and the black dot indicates the location of Horsens Regional Hospital.

Figure 2. Data availability for the CROSS-TRACKS cohort in Denmark.

Figure 3. Flow chart of individuals included in the CROSS-TRACKS cohort in Denmark 2012-2018.





