Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

METHODOLOGY

Algorithm linking patients and general practices in Denmark using the Danish National Health Service Register

This article was published in the following Dove Press journal:
Clinical Epidemiology
9 August 2016
Number of times this article has been viewed

Clinical Epidemiology downloaded from https://www.dovepress.com/
For personal use only.

Maiken Ina Siegismund Kjaersgaard[1]
Peter Vedsted[2]
Erik Thorlund Parner[1]
Bodil Hammer Bech[3]
Mogens Vestergaard[2,4]
Kaare Rud Flarup[2]
Morten Fenger-Grøn[2]

[1]Section for Biostatistics, [2]Research Unit for General Practice, [3]Section for Epidemiology, [4]Section for General Medical Practice, Department of Public Health, Aarhus University, Aarhus, Denmark

Background: The patient list system in Denmark assigns virtually all residents to a general practice. Nevertheless, historical information on this link between patient and general practice is not readily available for research purposes.

Objectives: To develop, implement, and evaluate the performance of an algorithm linking individual patients to their general practice by using information from the Danish National Health Service Register and the Danish Civil Registration System.

Materials and methods: The National Health Service Register contains information on all services provided by general practitioners from 1990 and onward. On the basis of these data and information on migration history and death obtained from the Civil Registration System, we developed an algorithm that allocated patients to a general practice on a monthly basis. We evaluated the performance of the algorithm between 2002 and 2007. During this time period, we had access to information on the link between patients and general practices. Agreement was assessed by the proportion of months for which the algorithm allocated patients to the correct general practice. We also assessed the proportion of all patients in the patient list system for which the algorithm was able to suggest an allocation.

Results: The overall agreement between algorithm and patient lists was 98.6%. We found slightly higher agreement for women (98.8%) than for men (98.4%) and lower agreement in the age group 18–34 years (97.1%) compared to all other age groups (≥98.6%). The algorithm had assigned 83% of all patients in the patient list system after 1 year of follow-up, 91% after 2 years of follow-up, and peaked at 94% during the fourth year.

Conclusion: We developed an algorithm that enables valid and nearly complete linkage between patients and general practices. The algorithm performs better in subgroups of patients with high health care needs. The algorithm constitutes a valuable tool for primary health care research.

Keywords: general practice, medical record linkage, patient list, primary care, registers

Introduction

The Danish health registers constitute a unique resource for health research as data can be linked at an individual level using the patient's personal identification number (a Civil Personal Registration [CPR] number).[1,2] Denmark has tax-funded public health care with free and equal access for all citizens, and the general practitioners (GPs) play a vital role as the first line in the health care system by serving as gatekeepers to specialized care. The GPs perform the main part of the initial diagnosis and treatment, and they are responsible for the majority of all drug prescriptions.[3] Virtually all Danes (98%) have chosen to join a patient list. The patient lists make up the patient

Correspondence: Maiken Ina Siegismund Kjaersgaard
Section for Biostatistics, Department of Public Health, Aarhus University, Bartholins Allé 2, Aarhus C 8000, Denmark
Tel +45 8716 7998
Email maiken@ph.au.dk

© 2016 Kjaersgaard et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

list system, and we refer to these as Group-1-insured patients. This implies that they are assigned to a specific general practice, which carries the responsibility for all aspects of everyday health care without costs to the patient.[4] Thus, the majority of Danes are exposed only to a particular general practice, which might have implications for several aspects of the care provided. However, few studies have focused on the impact of primary health care on patient outcomes and the influence of patient characteristics on general practice because information on links between patient and general practice are not readily available for research.

All persons residing in Denmark are registered centrally with a CPR number, which forms the basis for public registrations relating to income taxation, postal address, etc. The number is also used in the patient list system for per-capita remuneration of GPs for health care services provided. The patient lists are continuously updated with information, for example, death and practice change. A national IT supplier is commissioned to continuously update the records held in an administrative database for the Danish municipalities and regions. The general practice patient lists are stored with a history of a maximum of 5 years. The Patient List Database, which was created by the Research Unit for General Practice at Aarhus University, contains information based on the data from the national IT supplier on all Group-1-insured persons who have resided in Denmark since 2002. The database is, however, not publicly accessible for research purposes. Furthermore, many studies based on health registers require inclusion of data from before 2002, which necessitates linkage between patients and general practices longer back in time.

The Danish National Health Service Register, henceforth denoted as service register, contains information on weekly invoices for all services provided by GPs since January 1, 1990, including the CPR number of consulting patients.[5] Each practice is identified by a specific provider number. Therefore, this register could potentially be used to assess linkages over a longer time span as this information is available through the research services that provide data access and infrastructure for register-based research in Denmark.[6,7]

In this paper, we developed an algorithm linking Group-1-insured patients to the general practice with which they were listed using information from the service register and information on migration history and death obtained from the Danish Civil Registration System. We evaluated the performance of the algorithm by comparing with the Patient List Database over the years 2002–2007 as we had access to overlapping data from both the service register and the Patient List Database.

Materials and methods

Data sources for the algorithm

The service register contains invoices for services provided by health professionals to patients. We will here solely focus on health services delivered by GPs, who are remunerated by a mix of capitation and fee for service. Remuneration is based on centrally negotiated agreements between the Danish trade union for GPs and Danish Regions, or the corresponding former organizational entities known as counties, until the end of 2006. The services recorded in the service register are categorized as basic service, preventive service, additional service, or laboratory test. The service register has been described in detail elsewhere.[5] No data breaches have been observed during the life span of the register, and the completeness of the register is considered high due to the economic incentive for the GPs to register services for remuneration purposes. The service register contains variables identifying provider specialty, including general practice, type of service provided, the patient's health insurance group, and type of patient. Patient type could be one of four categories: "own patient", "passerby" (eg, travelers needing immediate medical care), "other practice's patient" (similar to the previous category, but this term is used when collaborating practices cover for each other, eg, during holiday periods), and "Group-2-insured" (patients who are not registered in the patient list system and thus may choose any general practice against payment of part of the fee). For each delivered service, information on week and year of service delivery is available, and month and year of invoice are also stated. Furthermore, the service register contains information on whether the provided services were delivered during daytime or out of hours. Standard opening hours in Denmark are Monday through Friday from 8 am to 4 pm, and patients are referred to out-of-hours doctors outside this time interval.[8]

Two aspects regarding the service register need special attention. First, the services provided to children younger than 16 years of age were recorded under an adult's CPR number, typically the mother's, until 1996.[9] This means that the proposed algorithm is not applicable to children younger than 16 years before 1996. Second, the general practice providing the service can be identified through a provider number, but the individual doctor delivering the service cannot be identified in partnership practices. Currently, there are approximately 3,600 GPs in Denmark organized in 2,200 practices.[4]

Information used in the algorithm is summarized on the left-hand side of the Table S1. The patient's CPR number enables us to follow a particular individual's contacts with

primary care over time and link these data with individual information from all national registers, such as the Danish Civil Registration System. The Civil Registration System was established in 1968 and contains individual information on, for example, the resident's sex, date of birth, and continuously updated information on vital status.[10] The study was approved by the Danish Data Protection Agency (2009-41-3471). No ethics approval is needed for register-based studies in Denmark and informed consent is not required.

Study population of the algorithm

The study population consisted of all Group-1-insured individuals in the Patient List Database during the years 2002–2007. For this period, we also had access to data from the service register, thus enabling us to validate the algorithm within that time period. Individuals were considered for the algorithm from their first visit with a GP within that time period and followed until death, emigration, or December 31, 2007, whichever came first. We allowed for emigration and subsequent immigration by excluding time periods between date of emigration and date of immigration. An emigrated person had to immigrate and have a contact with a GP on or following the date of immigration in order to be allocated to a general practice again.

Algorithm

We considered data for 10% of the Danish population from the service register for the development of the algorithm, whereas evaluations of the algorithm performance are based on the complete data. The exact steps of the algorithm are described in Table 1. The algorithm has been implemented in the statistical software Stata (StataCorp LP, College Station, TX, USA); the code is available on a website.[11] Below we will elaborate on details regarding some of the steps.

Selection of services

Data from the service register were restricted to general practice services (C_SPECIALE=80) for Group-1-insured patients (C_SIKGRP=1) during daytime (C_TIDSKODE=1). Services in children reported with the CPR number of an adult should be excluded (C_SIKKON=B). For this particular study, we did not have access to the variable C_SIKKON. Thus, services in children registered

Table 1 Outline of the steps in the algorithm linking patients and general practices in Denmark using data from the National Health Service Register and the Civil Registration System

Step number	Title and description
Step 1	**Restrict services and recode time of services.** Restrict data from the National Health Service Register and recode time of service as described in the sections on "Selection of services" and "Recoding the time of services."
Step 2	**Exclude services coded during emigration and after death.** Exclude invoices on or after patients' date of death or emigration until reimmigration, if applicable.
Step 3	**Exclude service weeks with multiple practices.** If invoices pertaining to more than one general practice are registered for the same patient within the same week, exclude all the invoices for the patient during that week.
Step 4	**Code preliminary practice time intervals.** Code preliminary practice time intervals for each patient–general practice relationship, starting from the date of first contact with a practice until the day before contact with another practice or end of follow-up.
Step 5	**Recode practice time intervals taking into account emigration and death.** If a patient emigrates (including moving to Greenland) or dies within a preliminary practice time interval, recode the end date of the interval to the date of emigration or death, whichever came first.
Step 6	**Recode practice time intervals taking into account practice closure.** For each general practice, define the closure date as the last date with registered services, unless the date falls within the last month before end of follow-up. Recode the end date to closure date for all patient intervals concerning this general practice and including this date. If applicable, recode the starting date for the same patient's subsequent practice time interval to the next day.
Step 7	**Recode practice time intervals taking into account patients moving.** For patients moving between municipalities in Denmark (codes 101-861; 275 municipalities) within an interval, recode the end date to the moving date and the subsequent starting date to the following day. For patients moving more than once, use the latest moving date.
Step 8	**Drop small (<31 days) practice time intervals.** Drop patient–general practice intervals shorter than 31 days, unless the interval starts on the first of a month and ends on the last day of a month, and include the time in the preceding interval.
Step 9	**Recode practice time intervals to monthly intervals.** Recode the starting date to the first date of the month. Recode the end date to the last date of the preceding month, unless the interval ends on the date of death or emigration or on the last day of the month.

with the CPR number of an adult could not be excluded. We considered this problem negligible as only 0.04%–0.1% of all services during the time period 2002–2007 were provided for children reported under the CPR number of an adult,[12] and this adult person was also likely to be a parent listed at the same general practice as the child. Next, we wanted to restrict to services performed in the GPs' own patients. This information is contained in a variable for type of patient. According to documentation of the type of patient variable, categories 0 and 00 are used for "own patient", whereas categories 1 and 01 are used for "own patient (Group-1)", but documentation of the distinction between the two does not exist or has been lost.[13] After restricting to services in Group-1-insured patients during daytime, only few services with categories 0 and 00 remained, and we excluded these (n=6). Thus, services were further restricted to services in Group-1-insured own patients (C_PATGRP=1 or C_PATGRP=01). Basic fees are paid for basic services and some preventive services. Any additional services, including laboratory tests, provided along with basic fee services will not contribute with more information, and these were thus excluded. The list of included basic fee services is given in Table S2. Further discussion on generation of the list is given in the Supplementary materials. The proportions of correctly allocated services, referred to as the positive predictive values (PPVs), were generally high for the patient type "own patient (Group-1)" (Table S3).

Recoding the time of services

The dates of service provision were generally extracted from week number and year of provision. However, to identify the correct week of service, the following rule should be applied: if year of service is identical to year of invoice, and week of service is 51–53, whereas month of invoice is January or February, then the year of service is recoded to year of service minus 1 year.[14] Further discussion on identifying week of service is given in the Supplementary materials. A few services during the life span of the service register (<1,000) were coded with 53 as week of service in a year that did not have 53 weeks nor succeeded a year with 53 weeks, and these should be excluded. For this particular study, we did not have information on invoice month and could, therefore, not apply the aforementioned rule. However, previous investigations have shown that only week 52 in 2003 constituted a problem during the years 2002–2007.[14] Consequently, we excluded services during week 52 provided and invoiced in 2003 (1% of all services in 2003).

The weekday on which a service was performed in a given week is not registered. We coded date of treatment as Wednesday of the week the service was provided. Other choice of weekday did not change the performance of the algorithm. Some additional imprecision relating to the dates must be acknowledged as, in some periods of the existence of the service register, the provided services were not invoiced weekly but rather compiled and invoiced on a monthly basis.[15] This is part of the reason why some services appear after patient emigration or death (n=117,052; 0.07%), and these were excluded in this work.

Service weeks with multiple practices

Some weeks had several invoices pertaining to the same patient in more than one general practice (n=199,077 person-weeks; 0.1% of all person-weeks registered in the service register). We excluded invoices during weeks pertaining to more than one general practice. However, we also investigated how to best place each patient with a nonunique week at a specific general practice so as to retain the information in the dataset. An alternative strategy is described in the Supplementary materials, but the achievable improvement in the allocation of patients to general practices was negligible compared to the exclusion strategy.

Monthly practice intervals

The algorithm allocated patients to a general practice on a monthly basis because some practices compile and invoice on a monthly rather than weekly basis for services provided[15] and because changes from one practice to another occur on the first of a month, except when patients move to a new address. To enable allocation on a monthly basis, we chose to disregard intervals shorter than 31 days, except if the interval started on the first of a month and ended on the last day of a month. Steps 8 and 9 can potentially be left out of the algorithm if one is not concerned with allocating patients to general practices on a monthly basis. In Step 9, an interval from February 4, 2002 to August 20, 2006 is recoded to an interval from February 1, 2002 to July 31, 2006, and an interval from February 4, 2002 to August 31, 2006 is recoded to an interval from February 1, 2002 to August 31, 2006.

Performance assessment

We evaluated the performance of the algorithm by estimating the agreement, the proportion allocated, and the proportion correctly allocated according to the Patient List Database. The agreement with the Patient List Database was defined as

the percentage of months for which the algorithm allocated patients to the correct general practice. The agreement may depend on certain patient characteristics as some patient groups (eg, elderly, sick, and women) more often visit with a GP. Thus, the algorithm was also evaluated when stratified by age, sex, and calendar time to identify possible deviances in algorithm coverage. As individuals are allocated continuously over time from their first visit with a GP, we evaluated the completeness of the algorithm by computing the proportion allocated, which was defined as the proportion of the total population in the Patient List Database captured within 1 and 2 years from the onset of the algorithm. The proportion correctly allocated was defined as the proportion of the total population in the Patient List Database allocated to the correct general practice at a given time. In a sensitivity analysis, we estimated how well the algorithm linked patients and general practices in different situations, that is, at the time of death of patients, at the time of somatic hospital admissions, and at the time of first sessions with a private practicing psychologist (provider specialty 63; services 0101, 0110, and 1101) or psychiatrist (provider specialty 24; service 0110). Somatic hospitalizations were obtained from the Danish National Patient Register,[16] and consultations with private practicing psychologists and psychiatrists were obtained from the service register.

Results

For the data available from 2002, the algorithm allocated 83.5% of the total Group-1-insured practice population to a general practice after 1 year and 91.2% after 2 years with almost perfect allocation; 82.7% and 89.9% of all Group-1-insured were correctly linked after 1 and 2 years, respectively (Figure 1). We managed to allocate up to 93.9% of the patient list (Figure 1). The overall agreement of the algorithm with the patient lists over the years 2002–2007 was 98.6%. We assessed 95% confidence intervals using cluster robust variance estimation to account for the correlation between observations on the same patient in different months.[17] However, as most intervals in the present context are extremely narrow (<0.01%), we report only the estimated proportions. When stratifying by age, sex, and calendar time, we obtained the proportions summarized in Tables 2 and 3. Agreement for all age categories was 98.6% or above, except for 18–34 year olds, with 97.8% in women and 96.4% in men. In addition, agreement was somewhat lower in 35–49 year-old men (98.6%) compared to women (99.1%). Overall, the agreement for women (98.8%) was higher than the agreement for men (98.4%). When we stratified by time in years, the agreement declined from 99.2% in 2002 to a minimum of 98.3% over the years 2002–2007. The algorithm seems to stabilize at ~98.6% (Table 3). We estimated how well the algorithm

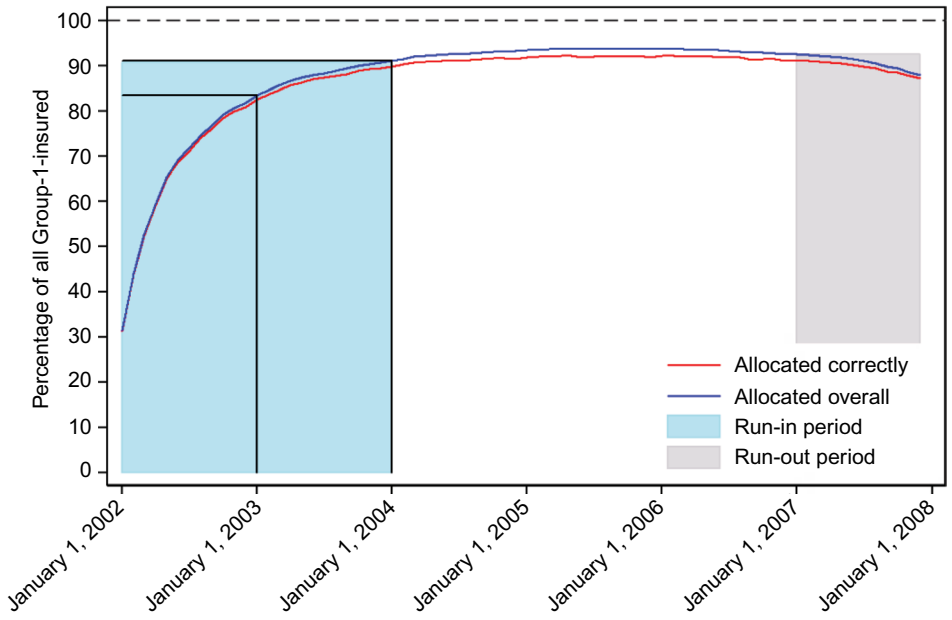


Figure 1 Coverage by time since January 1, 2002.
Notes: The blue curve depicts the percentage of the total Group-1-insured practice population allocated to a general practice for each month. The red curve depicts the percentage of patients allocated correctly in accordance with the Patient List Database. The straight, solid lines represent coverage within 1 year and 2 years from onset.

Table 2 Percentage of agreement stratified by age and sex

Age (years)	Men (%)	Women (%)	Total (%)
0–17	98.6	98.6	98.6
18–34	96.4	97.8	97.1
35–49	98.6	99.1	98.9
50–64	99.2	99.3	99.2
65–79	99.4	99.5	99.4
≥80	99.5	99.5	99.5
Total	98.4	98.8	98.6

Table 3 Percentage of agreement stratified by calendar time

Year	Total (%)
2002	99.2
2003	98.8
2004	98.4
2005	98.3
2006	98.4
2007	98.7
Total	98.6

linked patients and general practices in different situations (not shown in tables). First, we observed 326,634 deaths in Group-1-insured patients during the years 2002–2007; the algorithm linked up to 97.0% of the deceased patients to a general practice at the time of death with 99.1% agreement. Second, we observed 19,871,780 somatic hospitalizations in 4,319,451 Group-1-insured patients. The algorithm linked up to 96.5% of the hospitalized patients to a general practice at the time of hospital admission with 98.8% agreement. Third, we observed 199,525 first sessions with a psychologist in 148,408 Group-1-insured patients and 145,937 first sessions with a psychiatrist in 119,300 Group-1-insured patients. The corresponding algorithm linkage was up to 99.0% and 98.8%, and the agreement was 98.8% and 98.6%, respectively. Within 6 months, the algorithm allocated 91.9% (91.4% correctly) of all deaths, 96.9% (95.9% correctly) of all first sessions with a psychologist, and 97.4% (96.3% correctly) of all first sessions with a psychiatrist. Within 1 year, the algorithm allocated 93.9% (93.1% correctly) of all somatic hospitalizations.

Discussion

We developed, implemented, and evaluated the performance of an algorithm linking Group-1-insured Danish residents to the general practice with which they were listed and had to consult for medical advice. This linkage was based on the services that the individuals had received from the general practice. The 98.6% agreement with the Patient List Database demonstrates that the validity of the algorithm is high.

As the algorithm is based on information on services provided, it cannot identify an allocation until the first visit of a patient with a GP during the follow-up. Furthermore, the algorithm cannot identify the correct allocation after a change of practice until the patient has contacted the new practice for the first time. Thus, the quality of the linkage between patients and general practices depends on the frequency of the patient's GP visits. This was reflected in better allocation for women than for men and better allocation with increasing age. An additional reason for the slightly lower agreement among the younger patients (18–34 years) might be higher rates of moving in this group, which implies frequent change of practice. In Denmark, people moving more than 15 km (10 km in the capital city) away from their general practice must generally change to a local general practice. These findings correspond well with existing literature on consultation behavior. Wang et al found that the overall consultation rate was lower for men than for women and that the gap in consultation rates between men and women varied across the life course, with no differences between the sexes observed in early and later life.[18] The algorithm performed better in subgroups of patients with mental or physical illness, which mirrors higher consultation rates in people with mental or physical illness, for example, Kontopantelis et al[19] or Hunt et al.[20]

The run-in period of the algorithm when applied for the total population of all Group-1-insured individuals was about 2 years when we chose to attain 90% coverage. However, the run-in period was no more than half a year for people who died or had a consultation with a psychologist or psychiatrist and 1 year for people with somatic hospitalization. Thus, as expected, the run-in period was shorter for patient groups with high medical needs.

Because of the recoding of date in Steps 6 and 7, the algorithm allocation on a given date will, for a small proportion of persons, depend on information registered at a later time. This is reflected by a slight decrease in the performance of the algorithm over the last year of the study period. In Step 8, we used future information when excluding contacts with a practice followed by contact with another practice within 31 days. We do not consider exclusion of such short time intervals to be a serious source of bias, but we refrained from letting change of practice depend on two or more consecutive visits, which would have increased the time conditioned on considerably. To some extent, we may circumvent these problems by using an appropriate run-out period; 1 year of additional data seems to be adequate.

The agreement between the allocation of the algorithm and the Patient List Database was remarkably high throughout the entire study period; highest in 2002 (99.2%) and lowest in 2005 (98.3%). The slight decline in agreement was expected because the PPV of the first service at a practice was high, but patients migrate and some practices close over time. Still, as the PPVs of the services provided for the patient type categorized as "own patient" did not amount to 100%, we would not expect to be able to reach perfect agreement with the Patient List Database. The 16 Danish counties were transformed to five regions in 2007. As the system with providers and remuneration was not changed, we have no reason to believe that the presented algorithm will perform different from 2007 and onward.

A limitation of this study is that the algorithm was only validated for the subset of the services performed during the years 2002–2007. However, we expect the PPVs of all remaining services used outside the years 2002–2007 to be high as they resemble the services validated for the algorithm. We considered using all services performed during the years 2002–2007, but we found large variations in the PPVs (range: 6%–100%) for the different services. This was our main argument for selecting specific services for inclusion in the algorithm.

We used 10% of the data for the development of the algorithm. However, evaluations of the algorithm performance are based on the complete data instead of the remaining 90%, but overfitting on the complete data is considered minor.

In Step 1 of the algorithm, we explicitly kept only the services performed during daytime and thereby excluded the time of service during week category "not relevant". This category was almost exclusively used by the counties Frederiksborg, Ribe, Ringkøbing, and Viborg during 1992 and 1994,[21] and this is a limitation of our data source. Therefore, it could be considered to include this category only for these particular counties for these specific years.

As almost all Danes are exposed to one particular general practice at any given time, it makes sense to study, for example, the GP's propensity toward prescribing medication, referring patients, and ordering diagnostic tests. Linking patients and general practices thus enables studies of variations in general practice and patient outcomes using the Danish nationwide health registers. A limitation of the data source, and thus to variation studies using the algorithm, would be that the individual GPs in partnership practices cannot be identified.

Conclusion

We developed an algorithm that enables valid and nearly complete linkage between patients and general practices in Denmark. The algorithm performs better in subgroups of patients with high health care needs. The algorithm constitutes a valuable research tool because it facilitates studies that evaluate the effort of general practice on the outcome of their patients and the influence of patient characteristics on general practice.

Acknowledgments

This study was supported by the Danish Cancer Society, the Novo Nordisk Foundation, the Program for Clinical Research Infrastructure (PROCRIN) established by the Lundbeck Foundation and the Novo Nordisk Foundation and administered by Danish Regions, and the Danish Research Foundation for General Practice. MISK, MV, and MFG are supported by unrestricted grants from the Lundbeck Foundation. We thank Lone Niedziella at the Research Unit for General Practice, Aarhus University, for language editing.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. *Scand J Public Health.* 2015;43(4):333–339.
2. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol.* 2014;29(8):541–549.
3. Moth G, Olesen F, Vedsted P. Reasons for encounter and disease patterns in Danish primary care: changes over 16 years. *Scand J Prim Health Care.* 2012;30(2):70–75.
4. Pedersen KM, Andersen JS, Søndergaard J. General practice and primary health care in Denmark. *J Am Board Fam Med*. 2012;25(Suppl 1):S34–S38.
5. Andersen JS, Olivarius Nde F, Krasnik A. The Danish National Health Service Register. *Scand J Public Health.* 2011;39(7 Suppl):34–37.
6. The Danish Health Data Authority. *Registers and services*. Available from: http://sundhedsdatastyrelsen.dk/da/registre-og-services. Accessed March 4, 2016.
7. Statistics Denmark. *Data for Research*. Available from: www.dst.dk/en/TilSalg/Forskningsservice.aspx. Accessed March 4, 2016.
8. Christensen MB, Olesen F. Out of hours service in Denmark: evaluation 5 years after reform. *BMJ*. 1998;316(7143):1502–1505.
9. Pedersen PA, Hollnagel H, de Fine Olivarius N, Reusch S, Sørensen M, Thorsen H. Eget sygesikringsbevis til børn. Nye muligheder for epidemiologisk forskning i den primære sundhedssektor [Individual registration of children in The Danish National Health Service Register. New possibilities for epidemiologic research in primary health care]. *Ugeskr Laeger.* 1999;161(46):6351–6354. Danish.
10. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health.* 2011;39(7 Suppl):22–25.
11. Kjaersgaard et al. Algorithm implementation. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Algorithm.pdf. Accessed June 21, 2016.
12. Program for Clinical Research Infrastructure. Børnemarkering [Indication for service in child]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Boernemarkering_v20160115.pdf. Accessed March 4, 2016. Danish.

13. Program for Clinical Research Infrastructure. Patientgruppe [Type of patient]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Patientgruppe_v20160115.pdf. Accessed March 4, 2016. Danish.
14. Program for Clinical Research Infrastructure. Honorering og afregning [Provision and invoicing]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Honorering_og_afregning_v20160115.pdf. Accessed March 4, 2016. Danish.
15. Program for Clinical Research Infrastructure. Honoraruge [Week of provision]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Honoraruge_v20160209.pdf. Accessed March 4, 2016. Danish.
16. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449–490.
17. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics*. 2000;56(2):645–646.
18. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. *BMJ Open*. 2013;3:e003320.
19. Kontopantelis E, Olier I, Planner C, et al. Primary care consultation rates among people with and without severe mental illness: a UK cohort study using the Clinical Practice Research Datalink. *BMJ Open*. 2015;5:e008650.
20. Hunt K, Ford G, Harkins L, Wyke S. Are women more ready to consult than men? Gender differences in family practitioner consultation for common chronic conditions. *J Health Serv Res Policy*. 1999;4(2):96–100.
21. Program for Clinical Research Infrastructure. Tidskode [Time of service during week]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Tidskode_v20160209.pdf. Accessed March 4, 2016. Danish.

Supplementary materials

List of basic fee services

A list of provided services remunerated with a basic fee during standard opening hours in general practice, for the time period 1990-2015, was generated based on documentation from the Program for Clinical Research Infrastructure.[1,2] The out-of-hours services 0471, 0501, and 0602 were disregarded from the list of services identified by Program for Clinical Research Infrastructure. For the time period 1990-2015, the basic fee services shown in Table S2 was used; the services provided in 2002-2007 are indicated by numbers in bold font. For each of the services provided in 2002-2007, it was evaluated whether the providing practice was the practice where the patient was listed at the given time according to the Patient List Database. The proportion of correctly allocated services, referred to as the positive predictive value, is shown in Table S3. The positive predictive values were over 99% for the services provided for the type of patient "own patient (Group-1)", with the exception of a single child care service (98.3%) and home visits (range: 90.9%–98.7%). However, home visits were kept in the list in order to capture various populations, especially the elderly. No differences were observed between the individual years.

Rule for recoding the time of services

In the service register, weeks are numbered according to the ISO 8601 standard, which implies that week one includes the first Thursday of a given year, whereas some statistical software packages may handle week numbering differently.[3] Services are invoiced according to a fixed invoice schedule; month of invoice covers up to and including the second-to-last week of the month with a Wednesday. Each service week is thus paired with a unique invoice month, except for week 52, as this week pertains to both the beginning and the end of the year for some years, with corresponding invoice months January and December, respectively. However, as apparent from the data, it is unclear whether a given service in weeks 51-53 was provided during the year that the week belongs to or the year before. The month and year of invoice, together with week and year of service provision, can be used to ensure unambiguity. For example, if a service is noted to have an identical year of provision and year of invoice, say 2003, and week of provision is 52 and the month of invoice is December, then we believe the service was performed in 2003, and no recoding of year of service provision is needed. On the other hand, if the week of provision had been 52 and month of invoice had been January, then year of service provision should be recoded from 2003 to 2002 to correctly place the time of service. These considerations have led to the rule described in the section "Recoding the time of services".

Optimal strategy for choosing between multiple general practices within same service week

In Step 3 of the algorithm, we excluded services during weeks with invoices pertaining to more than one general practice for the same patient. This section describes the best strategy for placing each patient with a nonunique week at a specific general practice if we choose to recode instead of excluding nonunique weeks. In such case, one can recode multiple practices within the same service week by implementing and substituting the following for Step 3 in the algorithm.

If one or more succeeding, but not necessarily consecutive, weeks involve invoices pertaining to more than one general practice, then identifying both the unique practice preceding the weeks and the unique practice succeeding the weeks is required. Invoices from different general practices during the same week are handled according to the following hierarchical coding scheme:

1. If information on the preceding or the succeeding practice is missing, exclude the nonunique week.
2. If a nonunique week includes a service at the practice identical to the identified unique practice preceding the week(s), and the nonunique week does not include a service by the identified unique practice succeeding the week(s), then replace the nonunique week by the identified unique practice preceding the week(s).
3. Otherwise, replace the nonunique week by the identified unique practice succeeding the week(s).

By restricting the population to individuals with one or more nonunique weeks and running the algorithm with the original Step 3 and the alternative Step 3 on this restricted cohort, the number of months for which the two approaches allocated patients differently was compared. The percentage of agreement was 44.1% for the original Step 3 and 55.3% for the alternative approach. Thus, the difference in allocation is small, and no difference can be seen on the overall percentage of agreement when substituting the alternative for the original Step 3 and running the algorithm for the entire population for the years 2002-2007.

Table S1 Overview of data sources

Source	Information	Remarks	Variable name
National Health Service Register	Patient's identification number		PNR (Encrypted V_CPR)
	Provider number		YNR (Encrypted C_YNR)
	Provider speciality	Use only general practice	C_SPECIALE
	Services provided in general practice	Use only basic fee services	C_YDELSESNR
	Week and year service was provided		V_HONUGE and V_HONAAR
	Month and year service was invoiced		V_AFRMDR and V_AFRAAR
	Time of service during week	Use only Monday–Friday 8 am–4 pm	C_TIDSKODE
	Type of patient	Use only own patients	C_PATGRP
	Health insurance group	Use only Group-1-insured	C_SIKGRP
	Indication for service in child	Exclude children	C_SIKKON
Civil Registration System	Patient's identification number		PNR (Encrypted V_PNR)
	Patient's date of death		C_STATUS and D_STATUS_HEN_START
	Patient's migration history	Current and historical information including Greenland	D_UDREJSE_DATO and D_INDREJSE_DATO
	Patient's movement history	Current and historical information	D_TILFLYT_DATO, D_FRAFLYT_DATO, C_ANNKOR and C_KOM

Table S2 List of basic fee services for the years 1990-2015 to be used in the algorithm

Service code
0101, **0102**, **0103**, **0104**, **0105**, **0106**, **0107**, **0108**, **0109**, 0110, 0120, 0121, 0122, **0201**, 0202, 0203, **0411**, **0421**, **0431**, **0441**, **0451**, **0461**, **0491**, **4003**, **4021**, **4022**, 4023, 4024, 4025, 4026, 4027, **4050**, **4063**, **4106**, 4247, 4248, 4249, **6101**, **8110**, **8120**, **8130**, **8140**, 8142, 8143, 8144, 8145, 8146, 8147, 8148, 8150, 8160, 8201, 8202, 8203, 8204, 8205, 8206, 8207, 8208, **8210**, **8211**, **8212**, **8213**, **8214**, **8215**, **8216**, **8217**, 8310, **8317**, **8318**, **8319**, **8320**, **8326**, **8327**, 8328, 8329, 8330, 8334, 8335, 8336, **8701**, **8702**, 8703, 8704, **8705**, **8706**, **8707**, 8708, **8901**, **8920**, **8921**, **8922**, **8923**, **8924**, **8925**, 8935, 8936, 8937, 8938

Note: The services provided in 2002-2007 are indicated by numbers in bold font.

Table S3 List of basic fee services with PPV for patient type "own patient (Group-1)"

Type of service	Service code	No of services	PPV (%)
Ordinary consultations	0101	102,023,368	99.6
	0102	3,179	99.3
	0103	204,301	99.8
	0104	200,364	99.9
	0105	1,499,362	99.6
	0106	1,260,314	99.9
	0107	1,939	99.8
	0108	164	100
	0109	68	100
Telephone	0201	81,627,440	99.4
Home visits	0411	1,932,053	97.4
	0421	427,655	97.2
	0431	221,209	96.8
	0441	59,181	96.0
	0451	16,511	93.8
	0461	7,927	90.9
	0491	305,516	98.7
Talk therapy	4003	153,172	99.7
	4021	38,396	99.3
	4022	32	100
	4050	31,232	99.4
	4063	13,452	99.7
	4106	50,494	99.7
	6101	1,164,460	99.6
Antenatal care	8110	401,744	99.5
	8120	332,346	99.5
	8130	316,516	99.5
	8140	293,705	99.7
Child care	8210	58,780	98.3
	8211	342,186	99.6
	8212	360,921	99.7
	8213	364,182	99.7
	8214	298,853	99.8
	8215	277,045	99.8
	8216	260,234	99.8
	8217	339,701	99.8
Hepatitis B vaccinations	8317	3,114	99.6
	8318	305	100
	8319	670	99.0
	8320	143	100
	8326	133	100
	8327	630	99.8
Children's vaccinations	8701	537,154	99.8
	8702	24,495	99.6
	8705	19,636	99.6
	8706	273,386	99.6
	8707	31,973	99.8
Other vaccinations	8901	1,530	99.6
	8920	1,185,584	99.9
	8921	84,822	99.8
	8922	20,553	99.8
	8923	1,436	99.7
	8924	40,430	99.8
	8925	1,215	100

Abbreviaton: PPV, positive predictive value.

References

1. Program for Clinical Research Infrastructure. *Grundydelser* [Basic services]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Grundydelser_v20160209.pdf. Accessed March 4, 2016. Danish.
2. Program for Clinical Research Infrastructure. *Samtaleterapi* [Talk therapy]. Available from: http://cap.au.dk/fileadmin/cap.au.dk/Documents/Samtaleterapi_v20160209.pdf. Accessed March 4, 2016. Danish.
3. Cox NJ. Stata tip 68: week assumptions. *The Stata Journal*. 2010; 10(4):682–685.

Clinical Epidemiology

Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: https://www.dovepress.com/clinical-epidemiology-journal