



University of Southern Denmark

## Intracerebral hemorrhage

### positive predictive value of diagnosis codes in two nationwide Danish registries

Hald, Stine Munk; Kring Sloth, Christine; Hey, Sabine Morris; Madsen, Charlotte; Nguyen, Nina; García Rodríguez, Luis Alberto; Al-Shahi Salman, Rustam; Möller, Sören; Poulsen, Frantz Rom; Pottegård, Anton; Gaist, David

*Published in:*  
Clinical Epidemiology

*DOI:*  
10.2147/CLEP.S167576

*Publication date:*  
2018

*Document version:*  
Final published version

*Document license:*  
CC BY-NC

*Citation for pulished version (APA):*

Hald, S. M., Kring Sloth, C., Hey, S. M., Madsen, C., Nguyen, N., García Rodríguez, L. A., Al-Shahi Salman, R., Möller, S., Poulsen, F. R., Pottegård, A., & Gaist, D. (2018). Intracerebral hemorrhage: positive predictive value of diagnosis codes in two nationwide Danish registries. *Clinical Epidemiology*, 10, 941-948. <https://doi.org/10.2147/CLEP.S167576>

Go to publication entry in University of Southern Denmark's Research Portal

#### Terms of use

This work is brought to you by the University of Southern Denmark.  
Unless otherwise specified it has been shared according to the terms for self-archiving.  
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.  
Please direct all enquiries to [puresupport@bib.sdu.dk](mailto:puresupport@bib.sdu.dk)

# Intracerebral hemorrhage: positive predictive value of diagnosis codes in two nationwide Danish registries

Stine Munk Hald<sup>1,2</sup>  
 Christine Kring Sloth<sup>1,2</sup>  
 Sabine Morris Hey<sup>1,2</sup>  
 Charlotte Madsen<sup>1</sup>  
 Nina Nguyen<sup>3</sup>  
 Luis Alberto García Rodríguez<sup>4</sup>  
 Rustam Al-Shahi Salman<sup>5</sup>  
 Sören Möller<sup>2,6</sup>  
 Frantz Rom Poulsen<sup>7</sup>  
 Anton Pottegård<sup>8</sup>  
 David Gaist<sup>1,2,6</sup>

<sup>1</sup>Department of Neurology, Odense University Hospital, Odense, Denmark; <sup>2</sup>Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; <sup>3</sup>Department of Radiology, Odense University Hospital, Odense, Denmark; <sup>4</sup>Spanish Centre for Pharmacoepidemiologic Research (CEIFE), Madrid, Spain; <sup>5</sup>Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK; <sup>6</sup>Odense Patient Data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark; <sup>7</sup>Department of Neurosurgery, Odense University Hospital, Odense, Denmark; <sup>8</sup>Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark

Correspondence: David Gaist  
 Department of Neurology, Odense University Hospital, J.B. Winsløvsvej 4, 5000 Odense, Denmark  
 Tel +45 6541 2485  
 Fax +45 6541 3389  
 Email dgaist@health.sdu.dk

**Purpose:** The purpose of this study is to establish the validity of intracerebral hemorrhage (ICH) diagnoses in the Danish Stroke Registry (DSR) and the Danish National Patient Registry (DNPR).

**Patients and methods:** We estimated the positive predictive value (PPV) of ICH diagnoses for a sample of 500 patients from the DSR (patients recorded under ICH diagnosis) and DNPR (International Classification of Diseases, version 10, code I61) during 2010–2015, using discharge summaries and brain imaging reports (minimal data). We estimated PPVs for any ICH (a-ICH) and spontaneous ICH (s-ICH) alone. Furthermore, we assessed PPVs according to whether patients were recorded in both or only one of the registries. Finally, in a subsample with ICH diagnoses with access to full medical records and original imaging studies (extensive data, n=100), we compared s-ICH diagnosis and hemorrhage location after use of extensive vs minimal data.

**Results:** In the DSR, the PPVs were 94% (95% CI, 91%–96%) for a-ICH and 85% (95% CI, 81%–88%) for s-ICH. In the DNPR, the PPVs were 88% (95% CI, 84%–91%) for a-ICH and 75% (95% CI, 70%–79%) for s-ICH. PPVs for s-ICH for patients recorded in both registries, DSR only, and DNPR only were 86% (95% CI, 82–99), 80% (95%CI, 71–87), and 49% (95%CI, 39–59), respectively. Evaluation of extensive vs minimal data verified s-ICH diagnosis in 98% and hemorrhage location in 94%.

**Conclusion:** The validity of a-ICH diagnoses in DSR and DNPR is sufficiently high to support their use in epidemiologic studies. For s-ICH, validity was high in DSR. In DNPR, s-ICH validity was lower, markedly so for the small subgroup of patients only recorded in this registry. Minimal data including discharge summaries and brain imaging reports were feasible and valid for identifying ICH location.

**Keywords:** stroke, epidemiology, register-based research, intracranial hemorrhage

## Introduction

Danish nationwide registries represent a useful resource for epidemiological studies. While affording the obvious advantages of large numbers and a high degree of coverage, many of these registries primarily serve administrative purposes. This includes the Danish National Patient Registry (DNPR) where the validity of the diagnosis codes varies substantially.<sup>1</sup> The Danish Stroke Registry (DSR), a clinical database established in Denmark in 2003 to monitor the quality of care provided to stroke patients, offers an attractive alternative data source for epidemiological studies. It is mandatory to report standardized detailed information on all acute admissions for stroke at hospitals in Denmark to the DSR.<sup>2,3</sup> Although the DNPR and the DSR have been used extensively for research, few studies have focused on the validity of stroke

diagnosis codes in the registries.<sup>4–10</sup> Only three of these studies specifically addressed the validity of intracerebral hemorrhage (ICH) codes in the DNPR and low-to-moderate positive predictive values (PPVs) of ICH codes were reported of 66%–76%.<sup>4,5,10</sup> All three studies concern data prior to 2010 and do therefore not necessarily reflect the validity of ICH codes in more recent years, where increased focus on stroke may have resulted in greater attention to correct coding of the disorder. No previous studies have specifically assessed the validity of ICH diagnoses in DSR.

The site of the hemorrhage in the brain can be used to distinguish between the two most frequent types of nontraumatic ICH (ie, deep [nonlobar] vs lobar ICH). The etiology of these two types of spontaneous ICH differs, with hypertensive angiopathy largely correlating with deep ICH and sporadic cerebral amyloid angiopathy (CAA) with lobar ICH.<sup>11</sup> Distinction between the two ICH types has important ramifications for prognosis and use of preventive treatment, since CAA has a high recurrence rate, particularly in patients on anticoagulants,<sup>12</sup> and possibly antiplatelets.<sup>13</sup> Also, lobar bleeds confer twice the risk of developing subsequent dementia compared with nonlobar ICH.<sup>14</sup> In spite of its clinical importance, information on ICH type is not routinely collected in Danish medical registries. However, we believe that cost-effective methods of correctly classifying ICH can be developed in a Danish setting that will enable the conduct of large-scale epidemiologic studies. As a first step in this direction, we wished to assess whether a combination of selected medical record data (discharge summary and the brain imaging study report) can be used to correctly classify patients by ICH type. Depending on the outcome of this simple approach, we envisioned rapid and cost-efficient classification of large ICH patient cohorts.

In summary, we undertook the current study to establish the validity of overall ICH diagnosis codes in the DSR and the DNPR and to furthermore assess the feasibility of using sparse medical record data to correctly classify ICH by type, ie, lobar vs nonlobar.

## Patients and methods

We defined ICH as a symptomatic event (new headache, altered level of consciousness, or neurological symptoms), with or without new neurological signs, referable to a focal collection of blood within the brain parenchyma seen on brain imaging with signal characteristics consistent with the time of symptom onset; we defined spontaneous ICH (s-ICH), as ICH not attributable to prior trauma or hemorrhagic transformation of an ischemic stroke or an alternative

explanation (eg, tumor or vascular malformation).<sup>15</sup> Although similar to the World Health Organization stroke definition,<sup>16</sup> the abovementioned definition also allows the inclusion of patients based exclusively on symptoms (eg, severe sudden onset headache), where imaging supports new onset ICH.

## Setting and data sources

We based this study on data from hospital contacts of residents of the Region of Southern Denmark (RSD; 1.2 million inhabitants), a geographically defined region in Denmark. Patients suspected of a stroke are principally admitted or transferred to one of the four dedicated stroke units at neurology departments in the region, which also hosts a single neurosurgery department. All hospitals in RSD report data in a standardized format to the DNPR.<sup>1</sup> The regional authorities have copies of all hospital electronic medical records (EMRs) in RSD, which can be used for research purposes, provided consent is obtained from the heads of departments involved in patient care.

For the period January 1, 2007, to December 31, 2015, we retrieved all admissions to hospitals in RSD with primary discharge codes (in Danish, “*aktionsdiagnose*”) for ICH (code I61, International Classification of Diseases, version 10 [ICD-10]) on patients who resided in RSD at the time of their admission. In DSR, stroke diagnoses are recorded as “hemorrhagic stroke” (ie, ICH), “ischemic stroke”, or “unspecified”. We retrieved data on all hospital contacts recorded under “hemorrhagic stroke” in DSR for the period January 1, 2003, to December 31, 2015, using the same residency criteria as described previously. We wished to compare coding for the same events for patients recorded in both registries. Therefore, within each registry sample, we limited records to each patient’s first hospital contact in the period 2010–2015. We focused on validating diagnoses in this time period as we believed it would more accurately reflect current coding practice of ICH diagnoses. Furthermore, medical records and brain imaging studies were more likely to be available for the chosen period.

## Selection of cases for validation

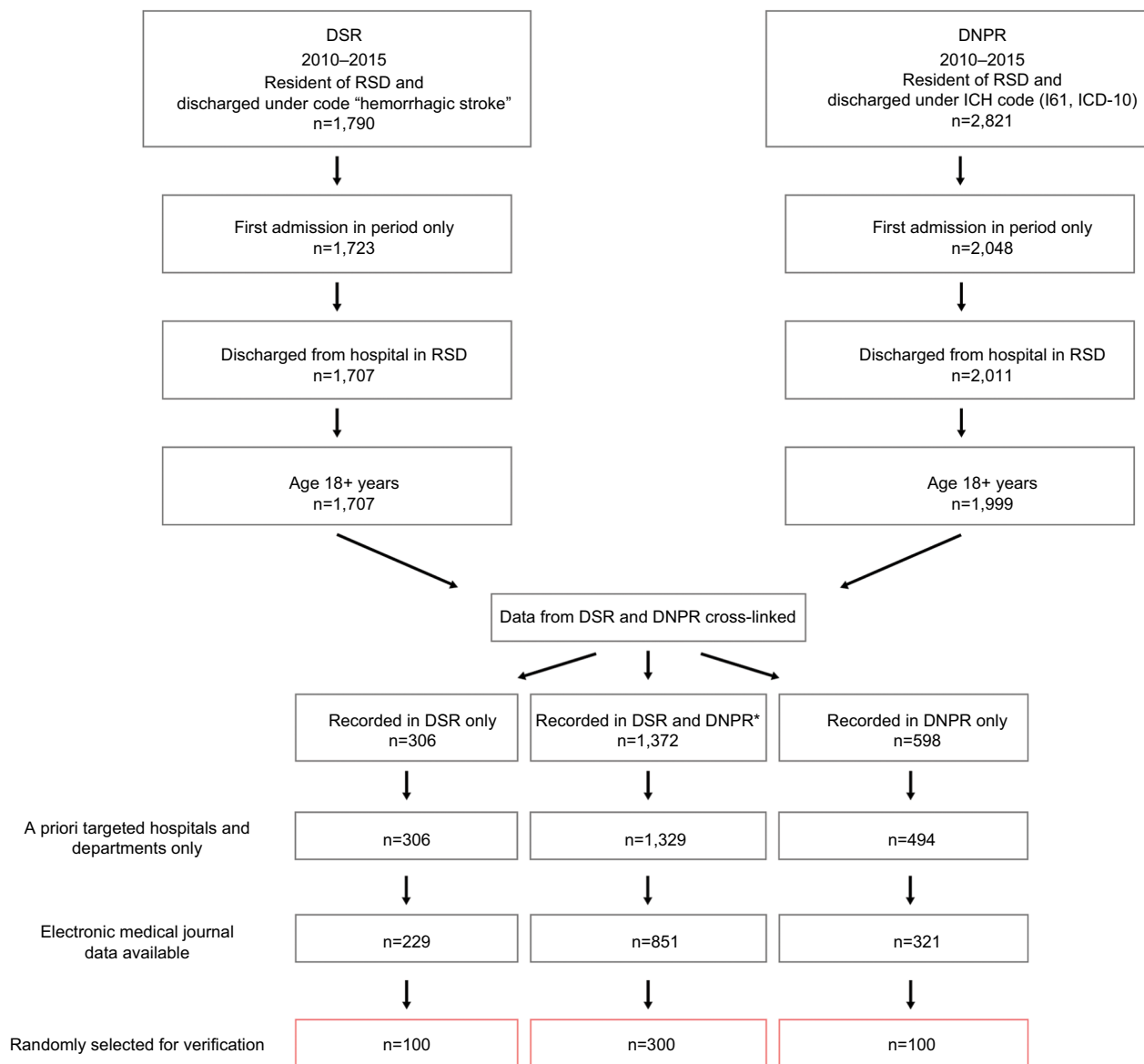
### Primary aim

The two registries are frequently used as independent data sources for epidemiological research, ie, as a single data source for identification of patients with ICH. Combining information from the registries might enhance the ability to detect all true ICH cases, which could prove useful, eg, in studies of ICH incidence rates, and could enable the development of reliable algorithms for identifying ICH. In

accordance with these considerations, we designed the present validation study to calculate the PPV of ICH diagnoses in the registries when used as single data sources or in combination. The researchers performing the assessments described subsequently were blinded with regard to the register source.

We identified the sample for validation as follows (Figure 1). First, we retrieved data on all patients recorded under ICH diagnoses in the DNPR or the DSR in 2010–2015 as described previously. For each sample, we identified the first admission during the period (to minimize capture of readmissions for same ICH event) and only included patients admitted to hospitals in RSD who were older than 18 years

(DSR does not record data on patients 0–17 years of age). Second, we merged the resulting data and classified patients into three mutually exclusive groups: 1) recorded in both registries; 2) recorded in DNPR only; and 3) recorded in DSR only. For patients recorded in both registries, to enhance the likelihood of studying the same event, we further requested that admission dates in the two registries be separated by no more than 15 days (98% of cases). Third, we limited the sample to patients who were admitted to any of the 42 a priori selected hospital units (hospital departments, emergency wards, and radiology units) all of which granted us permission to access medical records; the selected hospital



**Figure 1** Identification of cases for validation from the DSR and the DNPR.

**Notes:** \*Data of 29 patients were not included due to gap between recorded dates of admission in the two registries of 15 days or more.

**Abbreviations:** DNPR, Danish National Patient Registry; DSR, Danish Stroke Registry; ICH, intracerebral hemorrhage; RSD, Region of Southern Denmark.

Clinical Epidemiology downloaded from <https://www.dovepress.com/> by 130.226.87.174 on 23-Oct-2018  
For personal use only.

departments included 9 of 10 regional stroke units that had reported 99% of ICH admissions recorded in the DSR. Fourth, as hospitals joined the regional EMR system in a staggered fashion, we limited data contingent on the date the individual hospital joined the EMR. Finally, we randomly selected patients from each of these groups (group 1: 300, group 2: 100, and group 3: 100) for validation.

Data on these 500 patients recorded under ICH diagnoses were primarily validated based on information from discharge records and brain imaging study reports provided to us by the regional authorities. To ensure coverage of transfers between units (eg, from emergency ward to stroke unit) and of imaging work-up for secondary causes (eg, follow-up imaging after incident ICH), we requested discharge records and brain imaging study reports for a period spanning 1 week before to 5 months after the admission data of the index event. Four study physicians supervised by a neurologist with a special interest in stroke assessed this information and abstracted data to a structured form. Information collected included ICH diagnosis verified (“yes”, “no”, or “unclassifiable”), whether it was s-ICH, and location of s-ICH (ie, “lobar”, “non-lobar”, “infratentorial”, “unclassifiable” [due to large ICH], or “information insufficient to classify”). In cases of doubt regarding the diagnosis, full medical records were retrieved.

### Secondary aim

A secondary aim of our study was to assess to what extent ICH can be correctly classified by location of the hemorrhage (ie, lobar ICH vs deep ICH) based on selected medical record information (discharge summary and brain imaging study reports; referred to as *minimal data*). To achieve this purpose, among the 500 patients retrieved as described previously, we identified a sample of 100 patients, where original digital imaging studies and full medical records were available. Further, due to data restrictions on access to imaging studies, we only included patients admitted to one of two hospitals in the region, Odense University Hospital (OUH; 1,400 bed hospital with neurology and neurosurgery departments) and Svendborg Hospital (large nonuniversity hospital with dedicated stroke unit). We also required that patients resided in the primary catchment area of the two hospitals at the time of their stroke to avoid overrepresentation of secondary ICH (eg, due to arteriovenous malformations, where OUH is regional treatment center).

For this subset of 100 patients, we carried out the validation in a predefined three-step procedure. Step 1 corresponded to the procedure used in the main study (ie, evaluation based

on minimal data); in step 2, a study neurologist evaluated full medical records for relevant admission(s) and revised ICH diagnosis and location accordingly. Finally, in step 3, classification was further evaluated based on the addition of the results of assessment of original brain imaging studies by the study neurologist in conference with an experienced radiologist with full access to the clinical information. This stepwise approach enabled us to assess to what degree the validation of the larger sample based on minimal data might have profited by access to full medical record information (step 1 vs step 2) or visual assessment of the brain imaging studies (step 1 vs step 3).

### Validity of other stroke diagnoses in DSR

To assess to what extent patients with ICH were misclassified under other stroke diagnoses in DSR, we used the same inclusion criteria as described previously to identify a random sample of patients recorded in DSR under diagnoses of “ischemic stroke” (n=100) or “stroke unspecified” (n=100). We verified diagnoses based on the evaluation of discharge summaries and brain imaging reports.

### Statistical analyses

We calculated PPV for the ICH diagnosis for each of the registries (DNPR vs DSR) and for each of the three groups (ie, DNPR and DSR, DNPR only, or DSR only) calculating 95% CIs by the maximum likelihood method. We intentionally sampled 300 random patients from DNPR and DSR and 100 patients only from each of the other groups (ie, DNPR only and DSR only). To take this into account, we applied inverse probability weighting to adjust for different proportions sampled from each group. All analyses were performed using STATA 15.0.

The study was approved by the Danish Data Protection Agency (Approval ID 15/53398) and the Danish Health and Medicines Authority.

### Results

A total of 1,790 and 2,821 discharges were recorded under ICH diagnosis codes in DSR and DNPR, respectively. After application of three of the inclusion criteria (first admission in period, older than 18 years, and admissions to hospital in RSD), we cross-linked the data from the two registries. This resulted in a total of 2,276 patients, after exclusion of 29 patients recorded in both registries with admission dates more than 15 days apart. In all, 60% were recorded in both registries, 13% in DSR only, and 26% in DNPR only (Figure 1). Of the two remaining inclusion criteria, demanding availability

of EMR data produced the most pronounced reduction in eligible cases (Figure 1).

As mentioned previously, diagnoses were primarily verified based on discharge summaries and brain imaging reports delivered electronically by central authorities. However, in some cases, discharge summaries (n=16) or brain imaging reports (n=73) could not be identified in the central data repository and were instead retrieved manually from medical records of the departments involved. Furthermore, in 15 cases, minimal data were deemed insufficient to reach a diagnosis and full medical records were requested, mainly to clarify whether the ICH was spontaneous or secondary to other causes (n=12; all classified as non s-ICH).

In DSR, the PPV for any ICH (a-ICH) diagnosis was 94% (95% CI, 91–96) and for s-ICH 85% (95% CI, 81–88). The corresponding values for DNPR were 88% (95% CI, 84–91) and 75% (95% CI, 70–79), respectively. Underlying diagnoses for non-ICH cases frequently represented other types of intracranial hemorrhages, eg, subdural hematoma, or ischemic stroke with hemorrhagic transformation (Table 1). Year of admission, age, and sex had little impact on PPV of a-ICH in either registry (Table 2). Cases recorded in both the DSR and the DNPR had the highest PPVs (a-ICH 95%, 95% CI, 92–97; s-ICH 86%, 95% CI, 82–89), whereas the lowest PPVs were observed in cases only recorded in DNPR (a-ICH 72%, 95% CI, 62–80; s-ICH 49%, 95% CI, 35–59) (Table 3).

**Table 1** PPV of register codes for ICH and underlying diagnoses in nonverified cases in the DSR and the DNPR

|  | DSR (n=400) | DNPR (n=400) |
|--|-------------|--------------|
| <b>Diagnosis ICH verified</b>              |             |              |
| <i>Any ICH</i>                             |             |              |
| Number                                     | 374         | 358          |
| PPV (95% CI)                               | 94% (91–96) | 88% (84–91)  |
| <i>Spontaneous ICH</i>                     |             |              |
| Number                                     | 338         | 307          |
| PPV (95% CI)                               | 85% (81–88) | 75% (70–79)  |
| <b>Diagnosis ICH not verified, no. (%)</b> |             |              |
| Subdural hematoma                          | 4 (15.4)    | 14 (33.3)    |
| Subarachnoid hemorrhage                    | 6 (23.1)    | 9 (21.4)     |
| Hemorrhagic infarct                        | 8 (30.8)    | 9 (21.4)     |
| Ischemic stroke                            | 4 (15.4)    | 3 (7.1)      |
| Code error                                 | 3 (11.5)    | 6 (14.3)     |
| Other <sup>a</sup>                         | 1 (3.9)     | 1 (2.4)      |

**Note:** <sup>a</sup>“Other”: one case with insufficient information and one case with intraventricular hemorrhage in connection with ventricular catheter insertion.

**Abbreviations:** CI, confidence interval; DNPR, Danish National Patient Registry; DSR, Danish Stroke Registry; ICH, intracerebral hemorrhage; PPV, positive predictive value.

A total of 387 cases of s-ICH were identified in either of the registries with the hemorrhage located as follows: lobar (35%), deep (34%), infratentorial (19%), large unclassifiable ICH (9%), and insufficient information (3%).

### Assessment of diagnosis and ICH location based on minimal vs extensive data

In a subsample of 100 cases evaluated with minimal data, we formally assessed the effect of access to additional information to verify ICH diagnosis and location in a stepwise fashion, as described previously. Using minimal data, a diagnosis of a-ICH was reached in 92 cases (PPV 92%, 95% CI, 85–96) and a diagnosis of s-ICH in 82 cases (PPV 82%, 95% CI, 73–89) (step 1). Diagnoses verified through use of full medical records (step 2) were in agreement with those established through minimal data in 90 of 92 cases for a-ICH and 80 of 82 cases for s-ICH. Compared with minimal data, visual inspection of brain scans (step 3) resulted in agreement in 91 of 92 cases of a-ICH and in 80 of 82 s-ICH cases. A single case classified as non-s-ICH, based on minimal data, was reclassified as s-ICH based on visual inspection of the brain imaging studies. Finally, regarding location of hemorrhage, brain imaging study inspection results were in agreement with original evaluation in 78 of 82 cases classified as s-ICH according to minimal data (Table 4).

### Validity of “ischemic stroke” and “stroke unspecified” in DSR

The PPVs for ischemic stroke were 96% (95% CI, 90–99) for “ischemic stroke” and 81% (95% CI, 72–88) for “stroke unspecified”. Verified ischemic strokes included two cases of hemorrhagic transformation. Nonverified cases under “ischemic stroke”/“stroke unspecified” were classified as transient ischemic attack (2/6), carotid artery dissection (0/1), occlusion of retinal artery (0/1), stroke suspicion not confirmed (0/1), subarachnoid hemorrhage (0/1), various noncerebrovascular disorders (1/7), or insufficient information (1/2). No cases of ICH were recorded under these diagnoses.

## Discussion

In this study, based on data from 2010 to 2015, we found a high validity of a-ICH diagnosis in DSR and DNPR using each of these registries as a single data source. This was also the case for s-ICH in DSR, whereas in DNPR, s-ICH validity was lower. PPVs of patients recorded under ICH diagnoses in both registries, or in DSR only were also high, whereas for patients recorded in the DNPR only, validity was low. Use of



**Table 2** PPV of register codes for any ICH stratified by year, age, and sex in the DSR and the DNPR

|                          | DSR (n=400)      |              | DNPR (n=400)     |              |
|--------------------------|------------------|--------------|------------------|--------------|
|                          | Confirmed/sample | PPV (95% CI) | Confirmed/sample | PPV (95% CI) |
| <b>Year of admission</b> |                  |              |                  |              |
| 2010–2012                | 133/141          | 94% (89–97)  | 117/131          | 88% (80–93)  |
| 2013–2015                | 241/259          | 94% (90–96)  | 241/269          | 88% (83–92)  |
| <b>Age (years)</b>       |                  |              |                  |              |
| 20–54                    | 39/42            | 93% (80–98)  | 45/54            | 83% (70–91)  |
| 55–64                    | 48/50            | 96% (86–99)  | 46/48            | 92% (73–98)  |
| 65–74                    | 99/106           | 94% (89–97)  | 86/89            | 96% (88–99)  |
| 75–84                    | 122/131          | 93% (88–97)  | 119/138          | 86% (80–91)  |
| 85+                      | 66/71            | 93% (84–97)  | 62/71            | 86% (75–92)  |
| <b>Sex</b>               |                  |              |                  |              |
| Men                      | 188/202          | 94% (90–96)  | 173/191          | 89% (83–93)  |
| Women                    | 186/198          | 94% (90–97)  | 185/209          | 88% (82–92)  |

**Abbreviations:** CI, confidence interval; DNPR, Danish National Patient Registry; DSR, Danish Stroke Registry; ICH, intracerebral hemorrhage; PPV, positive predictive value.

**Table 3** PPV of register codes for any ICH according to whether diagnosis was recorded in one or both registries

|                         | DSR and DNPR (n=300) | DSR only (n=100) | DNPR only (n=100) |
|-------------------------|----------------------|------------------|-------------------|
| <b>Any ICH verified</b> |                      |                  |                   |
| Number                  | 286                  | 88               | 72                |
| PPV (95% CI)            | 95% (92–97)          | 88% (80–93)      | 72% (62–80)       |
| <b>Spontaneous ICH</b>  |                      |                  |                   |
| Number                  | 258                  | 80               | 49                |
| PPV (95% CI)            | 86% (82–89)          | 80% (71–87)      | 49% (39–59)       |

**Abbreviations:** CI, confidence interval; DNPR, Danish National Patient Registry; DSR, Danish Stroke Registry; ICH, intracerebral hemorrhage; PPV, positive predictive value.

discharge summaries and brain imaging reports was sufficient to validly evaluate ICH diagnosis and hemorrhage location.

Two studies of the validity of stroke diagnoses (primary or secondary) in DNPR based on data collected prior to year 2000 reported PPVs for ICH of 66% and 74%, respectively.<sup>4,5</sup> A more recent study reported a PPV for primary ICH code in DNPR of 76%.<sup>10</sup> The reported PPVs presumably concern

s-ICH, as stroke was defined according to World Health Organization in these studies. Only one of these studies explicitly stated that traumatic ICH was excluded.<sup>10</sup> In spite of certain methodological differences, we conclude that the results of the present study are in line with previous reports on the PPV of s-ICH in DNPR. We further provide novel information regarding DSR validity of ICH coding and conclude that the PPV for this diagnosis is of similar magnitude in the two registries.

Our study has several strengths. We used nationwide registries where cross-linkage was simple and accurate due to the unique and permanent personal identifiers allotted to all residents of Denmark. We were able to retrieve minimal data on all subjects, and ascertainment of diagnosis was performed by physicians that could adjudicate with a neurologist with considerable stroke experience. Finally, we validated our main approach of using minimal data in a subsample, where we used all available information, including inspection of original brain-imaging studies.

**Table 4** Agreement with regard to spontaneous type and hemorrhage location in 100 patients assessed through discharge records and brain imaging reports (minimal data) vs full medical record and visual inspection of original brain imaging studies (extensive data)

| Minimal data – location of spontaneous ICH | Extensive data – location of spontaneous ICH |      |                             |                          |                          |                                   |
|--|--|------|-----------------------------|--------------------------|--------------------------|-----------------------------------|
|  | Lobar  | Deep | Infratentorial <sup>a</sup> | Unclassifiable large ICH | Insufficient information | Not classified as spontaneous ICH |
| Lobar                                      | 24   | 0    | 0                           | 1                        | 0                        | 1                                 |
| Deep                                       | 0  | 34   | 0                           | 1                        | 0                        | 0                                 |
| Infratentorial <sup>a</sup>                | 0  | 0    | 10                          | 0                        | 0                        | 0                                 |
| Unclassifiable large ICH                   | 0  | 1    | 0                           | 9                        | 0                        | 0                                 |
| Insufficient information                   | 0  | 0    | 0                           | 0                        | 1                        | 0                                 |
| Not classified as spontaneous ICH          | 1  | 0    | 0                           | 0                        | 0                        | 17                                |

**Notes:** Shaded cells represent agreement between both methods. <sup>a</sup>Includes intraventricular hemorrhages.

**Abbreviations:** CI, confidence interval; DNPR, Danish National Patient Registry; DSR, Danish Stroke Registry; ICH, intracerebral hemorrhage; PPV, positive predictive value.

Our study also has some limitations. While our inclusion criterion of availability of data through EMR provided several logistic advantages, it also led to reduction of cases eligible for evaluation and resulted in almost twice as many cases being evaluated in the last 3 years of the study period, compared with the first 3 years of the study. As the overall study period was short, we find it unlikely that this had any major impact on our findings. Another limitation of our approach is the underlying assumption that all patients with stroke are recorded in either of the registries. We have no data on patients with ICH who die prior to hospitalization. Furthermore, it is likely that some patients with stroke are either not recorded in registry (eg, surgical patients suffering a stroke while hospitalized) or incorrectly coded under different diagnoses, or simply not recognized as ICH by the treating physicians.<sup>8</sup> As we lacked data on these possible missed or misclassified cases, we were not able to assess completeness (sensitivity) of the registries,<sup>17</sup> ie, the number of patients with a verified diagnosis registered in either the DSR or the DNPR divided by the total number of patients with a verified ICH diagnosis. Our findings for “ischemic stroke” and “stroke unspecified” in DSR indicate that misclassification of patients with ICH under other stroke codes is rare in this registry. The completeness of overall stroke diagnosis in DSR has recently been reported to be high.<sup>8</sup> Combined with our findings, this indicates that completeness of ICH diagnosis is also high in DSR. For DNPR, however, we did not ascertain non-ICH stroke codes for the presence of misclassified ICH cases and did not include secondary ICH codes. Although not assessed in this study, we find it likely that completeness of ICH diagnosis in DNPR is lower than that of DSR, as was recently reported for stroke of any type.<sup>8</sup>

The distribution of location of hemorrhage in patients with s-ICH in the main sample was as expected from the literature.<sup>15,18,19</sup> In the subsample of 100 cases, use of extensive data largely replicated the assessment of both diagnosis and hemorrhage location based on minimal data. However, the neurologist evaluating the diagnoses based on extensive data was not blinded to the diagnoses based on minimal data only. Furthermore, the cases selected for extensive evaluation were discharged from two units with access to radiologists with extensive experience in neuroradiology, which may have influenced the quality and detail of brain imaging reports, as well as other parts of the material available for validation. In spite of these caveats, we consider the results to be reassuring and that they support the use of the simple method (minimal data) to collect information on ICH type and location in patients identified through DSR and DNPR. Along

with the possibility of linkage of DSR and DNPR to other Danish data sources, including registries with information on prescription drug use, this approach may open new avenues for large-scale epidemiologic studies of subtypes of s-ICH, eg, lobar ICH. Although lobar ICH is associated with CAA, it is not always caused by this condition.<sup>20</sup> Efforts are currently being directed toward identifying clinical, imaging, and genetic characteristics that may allow correct identification of CAA-related s-ICH.<sup>21–23</sup> Further work is needed to evaluate to what degree these, or similar approaches, can be used to allow valid identification of patients with CAA-related s-ICH within the framework of routinely collected databases such as DSR and DNPR.

## Conclusion

PPV of a-ICH diagnosis is sufficiently high in both registries to support their use in epidemiologic analytical studies (eg, risk or prognosis studies). This is also true of s-ICH in DSR. PPV for s-ICH in DNPR overall was moderate, compared with DSR, and this difference was mainly driven by the low validity of ICH coding of the small subgroup of patients recorded in DNPR only. Enriching these data sources with information on s-ICH from discharge summaries and brain imaging reports is feasible and valid. DSR and DNPR may therefore also be a useful resource for large-scale studies of subtypes of ICH, eg, lobar ICH.

## Acknowledgments

The authors wish to thank Kristian Kvist, Department of Business Intelligence, RSD, and Inge Lise Udbye Christiansen, Department of Business Intelligence, RSD, for contributing to the collection of data.

## Disclosure

The activities of Dr David Gaist are supported by a grant from Odense University Hospital. The other authors report no conflicts of interest in this work.

## References

- Schmidt M, Schmidt SAJ, 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.
- Mainz J, Krog BR, Bjørnshave B, Bartels P. Nationwide continuous quality improvement using clinical indicators: the Danish National Indicator Project. *Int J Qual Health Care*. 2004;16(Suppl 1):i45–i50.
- Olsen TS, Dehlendorff C, Andersen KK. Sex-related time-dependent variations in post-stroke survival – evidence of a female stroke survival advantage. *Neuroepidemiology*. 2007;29(3–4):218–225.
- Johnsen SP, Overvad K, Sørensen HT, Tjønneland A, Husted SE. Predictive value of stroke and transient ischemic attack discharge diagnoses in the Danish National Registry of Patients. *J Clin Epidemiol*. 2002;55(6):602–607.



5. Krarup L-H, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a national register of patients. *Neuroepidemiology*. 2007;28(3):150–154.
6. Thygesen SK, Christiansen CF, Christensen S, Lash TL, Sørensen HT. The predictive value of ICD-10 diagnostic coding used to assess Charlson comorbidity index conditions in the population-based Danish National Registry of Patients. *BMC Med Res Methodol*. 2011;11:83.
7. Frost L, Andersen LV, Vestergaard P, Husted S, Mortensen LS. Trend in mortality after stroke with atrial fibrillation. *Am J Med*. 2007;120(1):47–53.
8. Wildenschild C, Mehnert F, Thomsen RW, et al. Registration of acute stroke: validity in the Danish Stroke Registry and the Danish National Registry of Patients. *Clin Epidemiol*. 2014;6:27–36.
9. Gaist D, Vaeth M, Tsiropoulos I, et al. Risk of subarachnoid haemorrhage in first degree relatives of patients with subarachnoid haemorrhage: follow up study based on national registries in Denmark. *BMJ*. 2000;320(7228):141–145.
10. Lühndorf P, Overvad K, Schmidt EB, Johnsen SP, Bach FW. Predictive value of stroke discharge diagnoses in the Danish National Patient Register. *Scand J Public Health*. 2017;45(6):630–636.
11. Meretoja A, Strbian D, Putaala J, et al. SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. *Stroke*. 2012;43(10):2592–2597.
12. Rosand J, Hylek EM, O'Donnell HC, Greenberg SM. Warfarin-associated hemorrhage and cerebral amyloid angiopathy: a genetic and pathologic study. *Neurology*. 2000;55(7):947–951.
13. Biffi A, Halpin A, Towfighi A, et al. Aspirin and recurrent intracerebral hemorrhage in cerebral amyloid angiopathy. *Neurology*. 2010;75(8):693–698.
14. Moulin S, Labreuche J, Bombois S, et al. Dementia risk after spontaneous intracerebral haemorrhage: a prospective cohort study. *Lancet Neurol*. 2016;15(8):820–829.
15. Samarasekera N, Fonville A, Lerpiniere C, et al; Lothian Audit of the Treatment of Cerebral Haemorrhage Collaborators. Influence of intracerebral hemorrhage location on incidence, characteristics, and outcome: population-based study. *Stroke*. 2015;46(2):361–368.
16. The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): A Major International Collaboration. WHO MONICA project principal investigators. *J Clin Epidemiol*. 1988;41(2):105–114.
17. Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol*. 1996;25(2):435–442.
18. Lavados PM, Sacks C, Prina L, et al. Incidence of lobar and non-lobar spontaneous intracerebral haemorrhage in a predominantly Hispanic-Mestizo population – the PISCIS stroke project: a community-based prospective study in Iquique, Chile. *Neuroepidemiology*. 2010;34(4):214–221.
19. Cordonnier C, Rutgers MP, Dumont F, et al. Intra-cerebral haemorrhages: are there any differences in baseline characteristics and intra-hospital mortality between hospital and population-based registries? *J Neurol*. 2009;256(2):198–202.
20. Banerjee G, Carare R, Cordonnier C, et al. The increasing impact of cerebral amyloid angiopathy: essential new insights for clinical practice. *J Neurol Neurosurg Psychiatry*. 2017;88(11):982–994.
21. Charidimou A, Fox Z, Werring DJ, Song M. Cerebral amyloid angiopathy research: on the verge of an explosion? *Int J Stroke*. 2015;10(5):E47–E48.
22. Charidimou A, Boulouis G, Roongpiboonsopit D, et al. Cortical superficial siderosis multifocality in cerebral amyloid angiopathy: a prospective study. *Neurology*. 2017;89(21):2128–2135.
23. Rodrigues MA, Samarasekera N, Lerpiniere C, et al. The Edinburgh CT and genetic diagnostic criteria for lobar intracerebral haemorrhage associated with cerebral amyloid angiopathy: model development and diagnostic test accuracy study. *Lancet Neurol*. 2018;17(3):232–240.

## Clinical Epidemiology

### 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 preventive initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

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

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.

Dovepress