

# The Positive Predictive Value of Pediatric Stroke Diagnoses in Administrative Data: A Retrospective Validation Study

Julie Brix Bindslev<sup>1,2</sup>, Soeren Paaske Johnsen<sup>3</sup>, Klaus Hansen<sup>1,4</sup>, Jan Brink Valentin<sup>3</sup>,  
Christina Engel Hoei-Hansen<sup>2,4</sup>, Thomas Truelsen<sup>1,4</sup>

<sup>1</sup>Department of Neurology, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark; <sup>2</sup>Department of Pediatrics, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark; <sup>3</sup>Danish Center for Clinical Health Services Research, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; <sup>4</sup>Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Correspondence: Julie Brix Bindslev, Department of Neurology, University Hospital of Copenhagen, Rigshospitalet, Blegdamsvej 9, Copenhagen, 2100, Denmark, Tel +45 29 92 89 10, Email JBIN0007@regionh.dk

**Background:** This retrospective cohort study aimed to examine the positive predictive value (PPV) of pediatric stroke diagnoses in the Danish National Registry of Patients (DNRP) and the impact of different stroke definitions on the PPV.

**Methods:** We included children registered with a stroke or stroke-related diagnosis in the DNRP between January 2017 through December 2020. Two assessors reviewed medical records and validated cases according to the American Heart and American Stroke Association (AHA/ASA) stroke definition. The level of interrater agreement was examined using kappa statistics. Validation by the AHA/ASA definition was compared with validation according to the definition in the International Classification of Disease 11th version (ICD-11) and the World Health Organization's definition.

**Results:** Stroke was confirmed in 120 of 309 included children, yielding an overall PPV of 0.39 (95% CI: 0.33–0.45). PPV varied across stroke subtypes from 0.83 (95% CI: 0.71–0.92) for ischemic stroke (AIS), 0.57 (95% CI: 0.37–0.76) for unspecified stroke, 0.42 (95% CI: 0.33–0.52) for intracerebral hemorrhage (ICH) to 0.31 (95% CI: 0.55–0.98) and 0.07 (95% CI: 0.01–0.22) for cerebral venous thrombosis and subarachnoid hemorrhage (SAH), respectively. Most non-confirmed ICH and SAH diagnoses were in children with traumatic intracranial hemorrhages (36 and 66% respectively). Among 70 confirmed AIS cases, 25 (36%) were identified in non-AIS code groups. PPV varied significantly across stroke definitions with the highest for the AHA/ASA definition (PPV = 0.39, 95% CI: 0.34–0.45) and the lowest for the WHO definition (PPV = 0.29, 95% CI: 0.24–0.34). Correspondingly, the incidence of pediatric AIS per 100,000 person-years changed from 1.5 for the AHA/ASA definition to 1.2 for ICD-11 and 1.0 for the WHO-definition. The overall interrater agreement was considered excellent ( $\kappa=0.85$ ).

**Conclusion:** After validation, stroke was confirmed in only half of the children registered in the DNRP with a stroke-specific diagnosis. Non-validated administrative data should be used with caution in pediatric stroke research. Pediatric stroke incidence rates may vary markedly depending on which stroke definition is used.

**Keywords:** positive predictive value, stroke diagnosis, stroke definition, children, interrater reliability, Danish National Registry of Patients

## Introduction

Administrative healthcare data are frequently used to monitor the incidence and outcome of pediatric stroke.<sup>1–5</sup> These data have the advantages of cost-efficient acquisition and coverage of large populations for long periods. However, given that administrative data were not intended for research, their accuracy for this purpose has been questioned.<sup>6</sup>

Few studies have examined the accuracy of pediatric stroke diagnoses in administrative data.<sup>6</sup> These have reported positive predictive values (PPVs) ranging between 0.27 and 0.89.<sup>2,7–13</sup> Estimates on the PPV of pediatric stroke diagnoses may vary due to variations in search algorithms and discrepancy in coding practices across different registries. Interrater variation in the diagnosis of stroke may also contribute.<sup>14</sup> In adults, interrater variation in the diagnosis of stroke accounts for 3% of the

differences in estimated stroke incidence.<sup>14</sup> However, to date, no studies have examined the potential impact of interrater variation on the diagnosis of pediatric stroke. Also, the PPV by children's sex and age has rarely been assessed,<sup>11,12</sup> the impact of different stroke definitions has not been explored and previous studies have encompassed a risk of bias since validation was done without blinding to the study population's administrative codes.<sup>6</sup>

Using the Danish National Registry of Patients (DNRP), we sought to address these potential problems. Two researchers who were blinded to the study population's administrative codes assessed the PPV of pediatric stroke diagnoses. We characterized differences in PPV by sex, age and stroke subtype, and we investigated the possible impact of interrater variation on the diagnosis of pediatric stroke. Furthermore, we tested the robustness of our validation results across three different stroke definitions; the definition by American Heart and American Stroke Association (AHA/ASA), the definition in the International Classification of Disease 11th version (ICD-11) and the definition by the World Health Organization (WHO).

## Methods

### Danish National Registry of Patients

The DNRP, a nationwide administrative registry, was established in 1977 and contains data on all discharges from non-psychiatric Danish hospitals. Reporting to the DNRP is mandatory and 99.4% of all discharges are entered in the registry.<sup>15</sup> Data in the DNRP include hospital and department codes, date of admission and discharge and diagnosis codes at discharge defined according to the International Classification of Diseases, 10th edition (ICD-10). Diagnosis codes are registered at discharge by the physician involved in the care of the patient and reported to the DNRP by a medical secretary. Patients can be identified in the DNRP by a unique personal identification number, which is assigned to all Danish citizens at birth or the time of immigration. This personal identification number enables linkage of the DNRP with information from medical records.

### Data Collection

We included children (28 days to 17 years) listed in the DNRP with a first-ever stroke, transient ischemic attack (TIA) or stroke-related diagnosis between January 2017 through December 2020. The following ICD-10 codes were used to identify possible stroke and TIA events; I60 (subarachnoid hemorrhage, SAH), I61-62 (intracerebral hemorrhage, ICH), I63 excl I63.6 (ischemic stroke, AIS), I64 (unspecified stroke), I63.6, I67.6, I676A and DG08 (central venous thrombosis, CVT), DG45 (TIA) and I67-68 excl I67.6 and I67.6A (other cerebrovascular diseases). Information on patients was obtained from the departments where they had been admitted and consisted of medical records, discharge letters, descriptions of Computer Tomography (CT), magnetic resonance imaging (MRI) and angiography reports. Two assessors (TT and JBB) who were blinded to each other's validation and to the ICD-codes registered in the DNRP independently reviewed each event. The assessors were a PhD student with three-year clinical experience and special interest in pediatric stroke research, respectively, a senior researcher and vascular neurologist with profound experience from adult stroke. For children registered with both a stroke-specific diagnosis (ICD-10: I60-64, DG45 and DG08) and a stroke-related diagnosis (ICD-10: I67-68), we validated only the stroke-specific diagnosis. In children with more than one stroke-specific diagnosis, we validated one diagnosis according to the following hierarchy: AIS, CVT, ICH, SAH, unspecified stroke and TIA. Children, where the registered stroke diagnoses referred to a perinatal event (age <28 days), were excluded.

By validation, possible stroke events were categorized as stroke, TIA, traumatic intracranial hemorrhage, other cerebrovascular disease and non-cerebrovascular disease. Stroke was further subcategorized into AIS, ICH, SAH and CVT. Cases in which assessors were in disagreement were evaluated a second time with disagreements resolved by consensus.

### Validation

Possible stroke events were validated in accordance with the AHA/ASA stroke definition.<sup>16</sup> AIS was defined as cerebral cell death caused by ischemia confirmed either by i) imaging or other objective evidence of ischemic injury in a defined vascular territory or ii) clinical evidence of ischemic injury based on symptoms persisting >24 hours or until death, and other etiologies excluded. This AIS definition also includes silent infarction defined as imaging or other neuropathological evidence of CNS infarction, without a history of acute neurological dysfunction attributable to the lesion.

ICH was defined as a collection of blood within the brain parenchyma or ventricular system, not caused by trauma, with or without the presence of neurological symptoms. SAH was defined as bleeding into the subarachnoid space, not caused by trauma, with or without the presence of clinical symptoms. The definitions of ICH and SAH also include silent cerebral hemorrhages defined as a focal collection of chronic blood products within the brain parenchyma, subarachnoid space, or ventricular system on neuroimaging or neuropathological examination that is not caused by trauma and without a history of acute neurological dysfunction attributable to the lesion.

Stroke caused by CVT was defined as infarction or hemorrhage in the brain, spinal cord, or retina caused by thrombosis of a cerebral venous structure. The diagnosis of unspecified stroke was used for an episode of acute neurological dysfunction presumed to be caused by ischemia or hemorrhage, with symptom duration >24 hours or until death, but without sufficient evidence to be classified as one of the above.

## Robustness Test

Validation according to the AHA/ASA stroke definition was compared with validation according to two other common stroke definitions; the definition in ICD-11 and the stroke definition by WHO.

In the ICD-11 definition,<sup>17</sup> AIS was defined as “acute neurological dysfunction caused by focal infarction at single or multiple sites of the brain or retina”. Evidence of acute infarction may come either from: a) Symptom duration lasting more than 24 hours or b) Neuroimaging. ICH was defined as “acute neurological dysfunction caused by hemorrhage within the brain parenchyma or in the ventricular system”. SAH was defined as “acute neurological dysfunction caused by subarachnoid hemorrhage”. Stroke of unknown subtype (ischemic or hemorrhagic) was defined as acute neurological dysfunction lasting more than 24 hours, but subtype of stroke has not been determined by neuroimaging.

According to the WHO stroke definition, stroke was confirmed in the presence of “rapidly developed clinical signs of focal (or global) disturbance of cerebral function, lasting more than 24 hours or leading to death, with no apparent cause other than of vascular origin”.<sup>18</sup> Children with traumatic ICH or SAH were not considered as having stroke.

## Statistics

Positive predictive value (PPV) of stroke diagnoses in the DNPR was calculated by dividing the number of confirmed events with the number of events registered in the DNPR. PPV was calculated for stroke, stroke subtypes, TIA, different stroke definitions and according to age, sex and study-period. 95% confidence intervals (CI) were calculated for PPV assuming normal approximation of the binominal distribution. PPV according to stroke definition was compared using Cochran's Q test. Comparison of PPV was not done for children registered with CVT diagnoses since this diagnosis is not considered a stroke across all of the disease classifications under investigation.

Kappa-statistics was used to test the interrater-agreement between two assessors for confirmation of stroke cases. Kappa-values range from -1 to 1 with -1 corresponding to complete disagreement, 0 to an agreement that is not different from chance and 1 to complete agreement. Kappa values were categorized as  $K < 0$  = poor agreement,  $0-0.20$  = slight agreement,  $0.21-0.40$  = fair agreement,  $0.41-0.60$  = moderate agreement,  $0.61-0.80$  = substantial agreement and  $0.81-1.00$  = excellent agreement. Statistical analyses were done using R version 4.2.2<sup>19</sup> and SPSS version 28.0.0.0.<sup>20</sup>

The results were reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations.<sup>21</sup>

## Ethics

The study was approved by the Danish Data Protection Agency (J. no. P-2021-539) and the Centre for Regional Development, Health Research and Innovation, the Capital Region of Denmark (J. no. R-21048871). Registry-based studies are exempted from ethical approval in Denmark.

## Results

A total of 333 children aged 28 days to 17 years were registered in the DNPR with a first-time stroke ( $n = 232$ ) or stroke-related diagnosis ( $n = 77$ ) between January 2017 and December 2020. Twenty-four were excluded since their registered stroke diagnoses referred to perinatal disease. Medical records were retrieved for the remaining 309 children. The final

sample included 130 girls with a median age of 12 (interquartile range 4–16) and 179 boys with a median age of 10 (interquartile range 4–15). The children were registered in the DNRP with the following diagnoses: I60 subarachnoid hemorrhage (n = 30, 9.7%), I61-62 intracerebral hemorrhage (n = 107, 34.6%), I63 excl I63.6 ischemic stroke (n = 54, 17.5%), I64 unspecified stroke (n = 28, 9.1%), I63.6, I67.6, I676A and DG08 central venous thrombosis (n = 13, 4.2%), DG45 TIA (n = 34, 11.0%), and I67-68 excl I676 and I676A other cerebrovascular diseases (n = 43, 13.9%), [Table 1](#).

Neuroimaging with CT or MRI was performed in 285 (93%) of the children; 97 (31%) were examined with CT, 89 (29%) with MRI and 99 (32%) with both CT and MRI. The proportion of children examined with neuroimaging varied with sex, age and study-period, [Table 2](#).

## Positive Predictive Values

After validation according to the AHA/ASA definition, stroke was confirmed in 120 of 309 children, yielding an overall PPV of 0.39 (95% CI: 0.33–0.45), [Table 1](#). When limiting the analysis to stroke-specific diagnoses (I60-64, I676, I676A, DG08), the overall PPV increased to 0.48 (95% CI: 0.42–0.55). PPV varied across stroke subtypes from 0.83 (95% CI: 0.71–0.92) for AIS, 0.57 (95% CI: 0.37–0.76) for unspecified stroke, 0.42 (95% CI: 0.33–0.52) for ICH to 0.31 (95% CI: 0.55–0.98) and 0.07 (95% CI: 0.01–0.22) for CVT and SAH, respectively, [Table 1](#).

The majority of non-confirmed ICH and SAH diagnoses were in children with traumatic intracranial hemorrhages (36 and 66% respectively). Most non-confirmed CVT diagnoses (n = 7) were in children who had central venous thrombosis but failed to meet the AHA/ASA stroke definition since no infarction or hemorrhage was demonstrated on neuroimaging. Among the 34 children registered with a TIA diagnosis, TIA was confirmed in three while another three were diagnosed with AIS and 28 with non-cerebrovascular diseases.

Among 70 confirmed AIS events, 25 (36%) were identified in non-AIS code groups (ICD-10 code: I60-62, I64, I67-68).

PPV varied significantly across age-groups with the lowest overall PPVs among children aged 28 days to 11 months (PPV = 0.30, 95% CI: 0.16–0.47) and adolescents (PPV = 0.30, 95% CI: 0.22–0.38) and the highest among preschool children aged 1 to 5 years (PPV = 0.61, 95% CI: 0.48–0.74), [Table 2](#). PPV also seemed to vary with sex with the lowest overall PPV in girls (PPV = 0.31, 95% CI: 0.23–0.40) and the highest in boys (PPV = 0.45, 95% CI: 0.37–0.52). However, this difference was not significant.

## Interrater Variation

After medical record review, the assessors disagreed on the stroke diagnosis in 22 out of 309 children, kappa value = 0.85 (95% CI: 0.79–0.91), [Table 3](#). The interrater agreement varied across stroke subtypes from 0.47 (95% CI: –0.13–1.00) for SAH, 0.73 (95% CI: 0.51–0.95) for AIS, 0.83 (95% CI: 0.52–1.00) for CVT to 0.86 (95% CI: 0.66–1.00) and 0.89 (95% CI: 0.80–0.97) for unspecified stroke and ICH, respectively, [Table 3](#).

## Predictive Values According to Stroke Definition

After exclusion of 13 children registered with CVT, a total of 116 children met the stroke definition from AHA/ASA. Among these 116 children, 18 and 31 failed to meet the ICD-11 and WHO-definition, respectively. The 18 children who failed to meet the ICD-11 definitions had silent strokes (n = 17) and retinal infarction (n = 1). The 31 children who failed to meet the WHO definition had silent strokes (n = 17), symptom cessation within 24 hours (n = 13) and retinal infarction (n = 1). The overall PPV varied from 0.29 (95% CI: 0.24–0.34) for the WHO definition to 0.33 (95% CI: 0.28–0.39) for the ICD-11 definition and 0.39 (95% CI: 0.34–0.45) for the AHA/ASA definition, [Table 4](#). Variation in PPV according to stroke definition was demonstrated for all stroke subtypes.

## Discussion

We reviewed medical records for a total of 309 children registered in the DNRP with a stroke or stroke-related diagnosis. After validation according to the AHA/ASA definition, stroke was confirmed in only half of the children registered with a stroke-specific diagnosis. Our findings demonstrate that non-validated administrative data should be used with caution in pediatric stroke research.

**Table 1** The Positive Predictive Value (PPV) of TIA and Stroke Diagnoses in the Danish National Registry of Patients

Diagnosis Registered in the DNRP	Number	Diagnosis According to Medical Record Review									PPV (95% CI)	
		SAH	ICH	AIS	Unspecified Stroke	Stroke Due to CVT	TIA	Traumatic Intracranial Hemorrhage	Other Cerebrovascular Diseases	Non-Cerebrovascular Diseases		
SAH (I60)	30	–	1	1	–	–	–	20	–	8	0.07	0.01–0.22
ICH (I61–I62)	107	1	40	3	1	–	–	39	8	15	0.42	0.33–0.52
AIS (I63 excl. I63.6)	54	–	–	45	–	–	–	1	1	7	0.83	0.71–0.92
Unspecified stroke (I64)	28	–	2	14	–	–	1	–	–	11	0.57	0.37–0.76
CVT (I63.6, I67.6, I676A, DG08)	13	–	–	–	–	4	–	–	7	2	0.31	0.09–0.61
TIA (DG45)	34	–	–	3	–	–	3	–	–	28	0.09	0.02–0.24
Other cerebrovascular diseases (I67–I68, excl I67.6 and I676A)	43	–	1	4	–	–	–	1	17	20	0.12	0.04–0.25
All (I60–I64, I67–I68, DG08, DG45)	309	1	44	70	1	4	4	61	33	91	0.39	0.33–0.45

**Note:** Data presented for consensus validation in accordance with the stroke definition from the American Heart and American Stroke Association.

**Abbreviations:** AIS, arterial ischemic stroke; CVT, central venous thrombosis; DNRP, Danish National Registry of Patients; ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage; TIA, transient ischemic attack.

**Table 2** The Positive Predictive Values (PPV) and Use of Neuroimaging in Subgroups of Children (0–17 Years) in the Danish National Patient Registry (DNPR)

	All, n	Neuroimaging, n(%)	CT, N(%)	MRI, n(%)	PPV, 95% CI
<b>Age-group</b>					
1–11 months	37	32(86.5)	27(73.0)	17(45.9)	0.30 (0.16–0.47)
1–5 years	57	56(98.2)	43(75.4)	37(64.9)	0.61 (0.48–0.74)
6–12 years	81	77(95.1)	48(59.3)	56(69.1)	0.42 (0.31–0.54)
13–17 years	134	120(89.6)	78(58.2)	78(58.2)	0.30 (0.22–0.38)
<b>Sex</b>					
Female	130	118(90.8)	75(57.7)	84(64.6)	0.31 (0.23–0.40)
Male	179	167(93.3)	121(67.6)	104(58.1)	0.45 (0.37–0.52)
<b>Study-period</b>					
2017	82	72(87.8)	42(51.2)	50(61.0)	0.37 (0.26–0.48)
2018	87	78(89.7)	55(63.2)	52(59.8)	0.38 (0.28–0.49)
2019	69	67(97.1)	49(71.0)	42(60.9)	0.44 (0.32–0.56)
2020	71	68(95.8)	50(70.4)	44(62.0)	0.38 (0.27–0.50)
<b>Diagnosis in The DNRP</b>					
SAH (I60)	30	29(96.7)	26(86.7)	6(20.0)	0.07 (0.01–0.22)
ICH (I61-62)	107	104(97.2)	91(85.0)	44(41.1)	0.42 (0.33–0.52)
AIS (I63 excl I63.6)	54	54(100)	36(66.7)	48(88.9)	0.83 (0.71–0.92)
Unspecified stroke (I64)	28	25(89.3)	14(50.0)	23(82.1)	0.57 (0.37–0.76)
CVT (I63.6, I67.6, DG08)	13	13(100)	9(69.2)	12(92.3)	0.31 (0.09–0.61)
TIA (DG45)	34	24(70.6)	5(14.7)	23(67.6)	0.09 (0.02–0.24)
Other cerebrovascular diseases (Di67-68 EXCL. 67.6)	43	36(83.7)	15(34.9)	32(74.4)	0.12 (0.04–0.25)

**Note:** Data presented for consensus validation according to the stroke definition from the American Heart and American Stroke Association.

**Abbreviations:** AIS, arterial ischemic stroke; CVT, central venous thrombosis; DNPR, Danish National Registry of Patients; ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage; TIA, transient ischemic attack.

**Table 3** Interrater Reliability for Stroke Subtypes

Stroke Subtype	N	PPV (95% CI)		Kappa (95% CI)
		Assessor 1	Assessor 2	
SAH (I60)	30	0.03 (0.001–0.17)	0.10 (0.02–0.27)	0.47 (–0.13–1.0)
ICH (I61-62)	107	0.40 (0.31–0.50)	0.44 (0.34–0.54)	0.89 (0.80–0.97)
AIS (I63 excl. I63.6)	54	0.74 (0.60–0.85)	0.83 (0.71–0.92)	0.73 (0.51–0.95)
Unspecified stroke (I64)	28	0.54 (0.34–0.73)	0.54 (0.34–0.73)	0.86 (0.66–1.0)
CVT (I63.6, I67.6, I676A, DG08)	13	0.31 (0.09–0.61)	0.39 (0.14–0.68)	0.83 (0.52–1.0)
TIA (DG45)	34	0.09 (0.02–0.24)	0.06 (0.01–0.20)	0.78 (0.38–1.0)
Other cerebrovascular diseases (I67-I68, excl I67.6 and I676A)	43	0.12 (0.04–0.25)	0.14 (0.05–0.28)	0.48 (0.09–0.87)
All (I60-64, I67-68, DG45, DG08)	309	0.36 (0.31–0.42)	0.40 (0.34–0.46)	0.85 (0.79–0.91)

**Note:** Validation based on the stroke definition from American Heart and American Stroke Association.

**Abbreviations:** AIS, arterial ischemic stroke; CI, confidence interval; CVT, central venous thrombosis; ICH, intracerebral hemorrhage; PPV, positive predictive value; SAH, subarachnoid hemorrhage; TIA, transient ischemic attack.

While the highest PPV was demonstrated for AIS diagnoses, the number of children with AIS was underestimated in the DNRP. A total of 54 children were registered with AIS (ICD 10: I63 excl I63.6); we confirmed 45 of these but identified an additional 25 AIS events in the other code groups (ICD-10 code: I60-62, I64, I67-68). These findings are

**Table 4** The Positive Predictive Value (PPV) of Stroke Discharge Diagnoses in the Danish National Registry of Patients According to Different Stroke Definitions

Diagnosis Registered in the DNRP	N	PPV (95% CI) by Stroke Definition			P value
		AHA/ASA	ICD-11	WHO	
SAH (I60)	30	0.07 (0.01–0.22)	0.03 (0.00–0.17)	0.03 (0.00–0.17)	0.368
ICH (I61-62)	107	0.42 (0.33–0.52)	0.40 (0.31–0.50)	0.36 (0.27–0.45)	0.004
AIS (I63 excl. I63.6)	54	0.83 (0.71–0.92)	0.61 (0.47–0.74)	0.54 (0.40–0.67)	<0.001
Unspecified stroke (I64)	28	0.57 (0.37–0.76)	0.50 (0.31–0.69)	0.39 (0.22–0.59)	0.022
TIA (DG45)	34	0.09 (0.02–0.24)	0.09 (0.02–0.24)	0.06 (0.01–0.20)	0.368
Other cerebrovascular diseases (I67-I68, excl I67.6 and I676A)	43	0.12 (0.04–0.25)	0.12 (0.04–0.25)	0.09 (0.03–0.22)	0.368
All (I60-64, I67-68, DG45, DG08 excl. I63.6, I67.6 and I676A) <sup>†</sup>	296	0.39 (0.34–0.45)	0.33 (0.28–0.39)	0.29 (0.24–0.34)	<0.001

**Notes:** <sup>†</sup>Overall PPV calculated with exclusion of children registered with CVT (n=13) since this diagnosis is not considered a stroke across all of the tested stroke definitions. Comparison Calculated for Children Registered with SAH, ICH, AIS, Unspecified Stroke, TIA and Other Cerebrovascular Diseases (n=296).

**Abbreviations:** AIS, arterial ischemic stroke; ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage.

consistent with those of previous validation studies<sup>8,13</sup> and demonstrate that AIS-specific ICD searches may miss a high number of pediatric AIS cases.

The overall PPV varied significantly with age. This variation may reflect differences in use of neuroimaging. A previous study from Sweden showed that the number of pediatric AIS cases confirmed by neuroimaging increased over time and correspondingly, the number of children incorrectly registered with a stroke diagnosis decreased.<sup>11</sup> Likewise, we found that the subgroups with the highest PPVs were those with the highest number of children examined with neuroimaging.

Studies with adults have demonstrated that interrater variation in the diagnosis of stroke accounts for 3% of the differences in estimated stroke incidence.<sup>14</sup> However, to the best of our knowledge, this is the first study to examine the impact of interrater variation on the diagnosis of pediatric stroke. In our study, two assessors reviewed 309 cases independently from each other. Despite of using the same clinical information and stroke definition, the assessors disagreed in 22 out of 309 cases. Correspondingly, the annual stroke incidence changed with 8.5% from assessor 1 to assessor 2. Our findings emphasize that interrater variation in the diagnosis of pediatric stroke is important to consider when interpreting trend studies.

Few studies, all from high-income countries, have validated pediatric stroke diagnoses in administrative data.<sup>2,7-13</sup> These have reported PPVs of 0.27 to 0.89 for AIS diagnoses, of 0.45 to 0.72 for CVT diagnoses, 0.62 to 0.66 for ICH diagnoses and of 0.52 to 0.83 for SAH diagnoses.<sup>6</sup> In addition to interrater variation, several other factors may contribute to the differences in PPVs across studies, including discrepancies in coding practice in different registries, differences in ICD-versions and variations in applied search algorithms. Differences in stroke definitions may also contribute. In 2013, the AHA/ASA published an imaging-based definition of stroke, which included asymptomatic infarcts and asymptomatic hemorrhages (silent strokes).<sup>16</sup> The AHA/ASA definition also included a symptom-based definition of stroke.<sup>16</sup> However, the inclusion of silent strokes in the AHA/ASA definition is a marked difference from the ICD-11<sup>17</sup> and WHO<sup>18</sup> stroke definitions for which confirmation of stroke requires the presence of neurological symptoms. In our study, we tested the robustness of our validation results across these stroke definitions. We showed that a change from the AHA/ASA definition to the ICD-11 and WHO definition decreased the overall PPV from 0.39 (95% CI: 0.34–0.45) to 0.33 (95% CI: 0.28–0.39), respectively, 0.29 (95% CI: 0.24–0.34). Especially for AIS diagnoses, marked variations in PPVs were demonstrated across stroke definitions (range: 0.54–0.83). Based on our findings, the incidence of pediatric AIS would change from 1.5 per 100,000 person-years for the AHA/ASA definition, to 1.2 per 100,000 person-years for the ICD-11 definition and 1.0 per 100,000 person-years for the WHO-definition. Silent strokes accounted for the majority of events that met the AHA/ASA definition without meeting the ICD-11 and WHO definition. In fact, 13% of events confirmed by the AHA/ASA definition were silent strokes. To allow for valid analyses of temporal and geographic trends in stroke, the AHA/ASA has recommended that symptomatic and silent strokes should be counted separately. Our findings support these recommendations.

Among children registered with ICH and SAH diagnoses, we found a high and even very high number of traumatic intracranial hemorrhages (36 and 66%, respectively). Thus, the PPVs of ICH and SAH diagnoses are below those of other studies.<sup>2,13</sup> All children with traumatic ICH and SAH were correctly diagnosed in the medical records (results not shown) but registered with ICD-codes for spontaneous intracranial hemorrhages (ICD-10: I60-62) instead of codes for traumatic intracranial hemorrhages (eg, ICD-10: DS066, DS068). Therefore, most likely, incorrect coding rather than misdiagnosis was the cause of the low PPVs of hemorrhagic stroke diagnoses in our study.

## Strengths and Limitations

Our study has several strengths, including a consistent validation method. We are the first to estimate variation in PPV due to interrater variation in the diagnosis of pediatric stroke. Moreover, we addressed a lack in the literature by describing variation in PPVs due to differences in stroke definitions. Unlike previous validation studies, we sought to limit the risk of bias by validating with blinding to the study population's administrative codes. Our study was based on a nationwide registry. All residents in Denmark have a unique personal registration number, which enabled a valid linkage of medical records with information in the DNRP. The Danish healthcare system is universal and publicly funded, which ensures that economic capacity is not a bias for referral.

Our study also has a number of limitations. We did not re-evaluate the original neuroimaging recordings. Therefore, interpretation of neuroimaging was limited to descriptions in medical records. Although our search algorithm included all types of diagnoses (referral, admission, discharge, etc.) and diagnoses in any position (primary and secondary diagnoses), these could not be distinguished. Therefore, we were not able to examine differences in PPV according to the type or position of diagnosis. Due to high diagnostic uncertainty early in a disease course, the inclusion of referral diagnoses in search algorithms may be associated with lower PPVs compared to searches based solely on discharge diagnoses. Thus, the inclusion of referral and admission diagnoses in our search algorithm could be an explanation for the low PPVs of hemorrhagic stroke diagnoses. However, in our study, the vast majority of non-confirmed ICH and SAH diagnoses were in children diagnosed with traumatic intracranial hemorrhages after thorough examination during hospital admission (36 and 66% respectively). Therefore, we expect that incorrect registration of traumatic intracranial hemorrhages, rather than our search strategy, was the major cause of the low PPVs of hemorrhagic stroke diagnoses. We were not able to estimate the sensitivity, specificity or negative predictive value of the DNRP since we did not extend our search to include potential false-negative events. Most children with stroke symptoms are likely to be hospitalized; however, stroke events may be missed in the DNRP, eg, in the lack of registration of stroke as a secondary diagnosis in the course of another severe disease. To get further insight into the accuracy of the DNRP, a pediatric stroke registry or another independent data source would be needed which is not established in Denmark.

## Conclusion

The overall PPV of pediatric stroke diagnoses was low in the DNRP. We recommend that efforts should be done to improve the validity of pediatric stroke diagnoses in administrative data before using these data for epidemiological research. In etiological studies, the use of non-validated data may result in attenuation of true associations due to dilution of pediatric stroke cases by non-stroke cases. When comparing trend studies, differences in pediatric stroke incidence may reflect variations in stroke definitions or interrater variation in the diagnosis of stroke.

## Data Sharing Statement

The data that support the findings of this study are available on request from the corresponding author (JBB). The data are not publicly available as their containing information could compromise the privacy of research participants.

## Ethics

The study was approved by the Danish Data Protection Agency (J. no. P-2021-539) and the Centre for Regional Development, Health Research and Innovation, the Capital Region of Denmark (J. no. R-21048871). Registry-based studies are exempted from ethical approval in Denmark.



## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was supported by grants from the Novo Nordisk Foundation (0065517), the Elsass Foundation (21-B01-1477) and the Institute of Clinical Medicine, Copenhagen University. The funders had no role in the design and conduct of the study; in the collection, analysis and interpretation of data; or in the preparation or approval of the manuscript.

## Disclosure

The authors have no conflicts of interest to declare.

## References

1. Chiang KL, Cheng CY. Epidemiology, risk factors and characteristics of pediatric stroke: a nationwide population-based study. *QJM*. 2018;111:445–454. doi:10.1093/qjmed/hcy066
2. Surmava A-M, Maclagan L, Khan F, et al. Incidence and current treatment gaps in pediatric stroke and TIA: an Ontario-wide population-based study. *Neuroepidemiology*. 2019;52(3–4):119–127. doi:10.1159/000493140
3. DeVeber GA, Kirton A, Booth FA, et al. Epidemiology and outcomes of arterial ischemic stroke in children: the Canadian pediatric ischemic stroke registry. *Pediatr Neurol*. 2017;69:58–70. doi:10.1016/j.pediatrneurol.2017.01.016
4. Mallick AA, Ganesan V, Kirkham FJ, et al. Childhood arterial ischaemic stroke incidence, presenting features, and risk factors: a prospective population-based study. *Lancet Neurol*. 2014;13(1):35–43. doi:10.1016/S1474-4422(13)70290-4
5. Tuckuviene R, Christensen A, Helgestad J, Johnsen S, Kristensen S. Paediatric arterial ischaemic stroke and cerebral sinovenous thrombosis in Denmark 1994–2006: a nationwide population-based study. *Acta Paediatr*. 2011;100(4):543–549. doi:10.1111/j.1651-2227.2010.02100.x
6. Brix Bindslev J, Maria Ciochon Klaus Hansen U, Engel Hoei-Hansen C, Truelsen T. Systematic review predictive value of pediatric stroke diagnoses in administrative data: a systematic literature review. *Neuroepidemiology*. 2023;56(6):413–422. doi:10.1159/000527805
7. Golomb MR, Garg BP, Williams LS. Accuracy of ICD-9 codes for identifying children with cerebral sinovenous thrombosis. *J Child Neurol*. 2007;22:45–48. doi:10.1177/0883073807299959
8. Golomb MR, Garg BP, Saha C, Williams LS. Accuracy and yield of ICD-9 codes for identifying children with ischemic stroke. *Neurology*. 2006;67(11):2053–2055. doi:10.1212/01.wnl.0000247281.98094.e2
9. Gerstl L, Badura K, Heinen F, et al. Childhood haemorrhagic stroke: a 7-year single-centre experience. *Arch Dis Child*. 2019;104(12):1198–1202. doi:10.1136/archdischild-2018-316749
10. Srinivasan J, Miller SP, Phan TG, Mackay MT. Delayed recognition of initial stroke in children: need for increased awareness. *Pediatrics*. 2009;124:e227–e234. doi:10.1542/peds.2008-3544
11. Walås A, Svensson K, Gyris M, Bang P, Sundelin HEK. Paediatric ischaemic stroke is a valid diagnosis in the Swedish National Patient Register. *Acta Paediatr*. 2021;110:2179–2186. doi:10.1111/apa.15791
12. Tuckuviene R, Kristensen SR, Helgestad J, Christensen AL, Johnsen SP. Predictive value of pediatric thrombosis diagnoses in the Danish National Patient Registry. *Clin Epidemiol*. 2010;2:107–122. doi:10.2147/CLEP.S10334
13. Agrawal N, Johnston SC, Wu YW, Sidney S, Fullerton HJ. Imaging data reveal a higher pediatric stroke incidence than prior us estimates. *Stroke*. 2009;40:3415–3421. doi:10.1161/STROKEAHA.109.564633
14. Krarup LH, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a national register of patients. *Neuroepidemiology*. 2007;28:150–154. doi:10.1159/000102143
15. Andersen TF, Madsen M, Jørgensen J, Mellekjøer L, Olsen J. The Danish National Hospital Register. A valuable source of data for modern health sciences. *Dan Med Bull*. 1999;46(3):263–268.
16. Sacco RL, Kasner SE, Broderick JP, et al. An updated definition of stroke for the 21st century. *Stroke*. 2013;44(7):2064–2089. doi:10.1161/STR.0b013e318296aeca
17. ICD-11 for mortality and morbidity statistics. Available from: <https://icd.who.int/browse11/l-m/en#/http%3a%2f%2fid.who.int%2fcd%2fentity%2f843843448>. Accessed June 17, 2023.
18. Truelsen T, Begg S, Mathers C. The global burden of cerebrovascular disease. Available from: [https://www.who.int/healthinfo/statistics/bod\\_cerebrovasculardiseasestroke.pdf](https://www.who.int/healthinfo/statistics/bod_cerebrovasculardiseasestroke.pdf). Accessed June 17, 2023.
19. RStudio Team. *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC; 2020. Available from: <http://www.rstudio.com/>. Accessed June 17, 2023.
20. SPSS statistics. IBM. Available from: <https://www.ibm.com/products/spss-statistics>. Accessed June 17, 2023.
21. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806–808. doi:10.1136/bmj.39335.541782.AD

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 & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & 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>