Accepted: 7 November 2017

DOI: 10.1111/ane.12876

ORIGINAL ARTICLE

WILEY Acta Neurologica Scandinavica

Improving identification of idiopathic intracranial hypertension patients in Swedish patient register

A. Sundholm[1,2] | S. Burkill[3] | S. Bahmanyar[3] | A. I. M. Nilsson Remahl[1,2]

[1]Department of Clinical Neuroscience, Karolinska Institutet, Solna, Sweden

[2]Department of Neurology, Karolinska University Hospital, Stockholm, Sweden

[3]Department of Medicine, Centre for Pharmacoepidemiology, Karolinska Institutet, Solna, Sweden

Correspondence
A. Sundholm, Department of Neurology, Karolinska University Hospital, Stockholm, Huddinge, Sweden.
Email: anna.sundholm@sll.se

Funding information
This article received grant from the Swedish Migraine Society.

Objective: Idiopathic intracranial hypertension (IIH) is often misdiagnosed. This can cause problems if conducting register-based studies. The study purpose was to produce algorithms that better identify patients with correct diagnosis of IIH in the Swedish National Patient Register (NPR).

Methods: Patients with ICD-10 code G93.2 for IIH registered in the NPR (2006-2013, Stockholm County) were included and diagnosis validated by medical record reviews. Patients were randomized into two groups: one used to produce the algorithm (n = 105) and one for validation (n = 102). We tested variables possible to extract from registries and used forward stepwise logistic regression which provided a predicted probability of correct diagnosis for each patient.

Results: We included 207 patients of which 135 had confirmed IIH. This gave a positive predictive value of 65.2% (CI: 58.4-71.4). The algorithm produced with variables extracted from registries, that is, age, number of times with diagnosis code G93.2 recorded (>2 times), and acetazolamide treatment, predicted the diagnosis correctly 88.2% (CI: 80.3-93.3) of the time. Excluding treatment data from the algorithm did not change the prediction notably, 86.3% (CI: 78.1-91.7).

Conclusion: We produced two algorithms that with improved accuracy predict whether an IIH diagnosis in the NPR is correct. This can be a useful tool when performing register-based studies.

KEYWORDS

algorithm, epidemiology, idiopathic intracranial hypertension, pseudotumor cerebri syndrome, registry, validity

1 | INTRODUCTION

Idiopathic intracranial hypertension (IIH) is a rare disorder mainly affecting young, obese females. The incidence of IIH is estimated to be between 0.9 and 2.36 per 100 000 individuals.[1-4] In our recent population-based study validating registered diagnosis of IIH by medical record review, we found a slightly lower incidence rate of 0.65 per 100 000 in Stockholm County, Sweden.[5] The disorder is characterized by high intracranial pressure with headache and visual disturbances being the most common symptoms. The pathophysiology behind this disorder is not fully understood.[6] Some potential risk factors for IIH have been proposed; for example, treatment with tetracyclines, vitamin A, and lithium, medical disorders such as hypoparathyroidism, Addison's disease, iron-deficiency anemia, and uremia.[7] Further carefully designed epidemiological studies with enough statistical power are needed to find whether these associations are true or spurious. The use of register data could clarify some of the remaining uncertainty surrounding these associations.

Misdiagnosis of IIH is shown to be common. Research conducted by Fisayo[8] found that 40% of patients with a prior IIH diagnosis did not fulfill the criteria on follow-up. The most common diagnostic error was inaccurate ophthalmoscopic examination in headache patients. Koerner et al[9] also showed a positive predictive value (PPV) of 55% when validating registered diagnosis coding of IIH in inpatient and

© 2017 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd

emergency hospital settings. A recent validation study conducted as part of this project showed a PPV of 65.2% (CI: 58.4-71.4) on registered diagnosis code for IIH (G93.2) in the Swedish National Patient Register (NPR) in Stockholm County registered between 2006 and 2013.[5]

A useful tool for more accurately selecting correct IIH patients if using registry data would be beneficial. The aim of this study was to produce an efficient algorithm for identifying patients with a correct diagnosis of IIH in the NPR using variables possible to extract from Swedish registries.

2 | METHODS

2.1 | Swedish national patient register (NPR), study population, and design

All residents in Sweden have a unique ten digit personal identification number (PIN) which is used in all healthcare registers. All healthcare contacts (with the exception of primary care) have to report to the NPR. The NPR[10] includes data on PIN, age, sex, date of admission and discharge, hospital, clinic, main and up to 21 secondary diagnoses and data on procedures (eg, surgical and radiological procedures). The study population is defined as all patients who received the diagnosis code G93.2 in Stockholm County in the NPR between 2006 and 2013. After approval from each head of department, medical records were obtained, and the modified Dandy Criteria[5,11] used to ascertain whether their diagnosis was correct (see details in previous study[5]). Patients with definite IIH (fulfilling modified Dandy Criteria[5,11]) and probable IIH (some missing data in records meaning the patient did not entirely fulfill the modified Dandy Criteria, (often due to diagnosis in the past where original records were not obtainable) but with an overall clinical description in concordance with IIH) were regarded as IIH patients. Patients with secondary intracranial hypertension (sIH), wrong diagnosis coding, or initial suspicion but later received another diagnosis were grouped as not IIH.

For the purposes of generating the algorithm, the patients were divided randomly into two groups; the "algorithm group" which was used to produce the algorithm, and the "test group" which was used to test how well the algorithm predicts whether a diagnosis is correct or incorrect. If the predicted probability generated by the algorithm was ≥0.5, the algorithm was deemed to have predicted that the patient had true IIH. If the algorithm gave a predicted probability that was <0.5 that meant the algorithm had predicted that the patient did not have IIH.

Ethical approval was granted by the local ethical committee of Stockholm.

2.2 | Statistical analysis

The binary variable for correct or incorrect diagnosis was used as the outcome in a forward stepwise logistic regression model. A forward stepwise model begins with an empty model and tests the fit of the model with the addition of each new variable. The variables which give the most statistically significant improvement to the model are retained. The process is repeated until no additional variables improve the model to a statistically significant degree. The variables available in the national registers (NPR and Prescribed Drug Register (PDR)) which we believed to be useful predictors of a correct IIH diagnosis were included as covariates. This approach meant variables which were not significant were removed from the predictive model. We tested the following variables to produce algorithm 1 (variables that could be drawn from both NPR and the PDR): age, sex, number of times having diagnosis code G93.2 recorded (at least two, three, or five times), having received the diagnosis ever at a neurology department/clinic and if patients had received acetazolamide treatment. Algorithm 2 contained the same variables with the exception of acetazolamide treatment making us independent of the PDR for this algorithm. We obtained predicted probabilities using the outcome of the model for the algorithm group, and applied predicted probabilities to the test group based on patient characteristics for the variables included in the algorithm. The different algorithms produced were evaluated by calculating how well they were predicting both true and incorrect IIH combined (predictive probability value). Positive predictive value (PPV) and negative predictive value (NPV) with 95% confidence intervals (CI) were also evaluated.

3 | RESULTS

3.1 | Study population

A recent validation study conducted as part of this project showed a PPV of 65.2% (CI: 58.4-71.4) on registered diagnosis codes for IIH (G93.2) in the Swedish National Patient Register (NPR) in Stockholm County registered between 2006 and 2013.[5] The total population of Stockholm County for inhabitants ≥18 years old was 1 496 096 in year 2006 and 1 696 445 in 2013 which is some 22% of the Swedish population (2013). We found that 210 patients 18 years or older received the diagnosis code for IIH between 2006 and 2013 in the County of Stockholm.[5] Three were excluded because we were not able to obtain their medical records. The final study population was thus comprised of 207 patients with a recorded IIH diagnosis code. Totally, 135 patients were confirmed as being correctly diagnosed (including 112 patients (83%) with definite IIH and 23 patients (17%) with probable IIH) when validated through their medical records. Of the remaining 72 patients with an incorrect diagnosis, 16 (22%) had secondary intracranial hypertension (sIH), 28 patients (39%) had an incorrect diagnosis code, and 27 patients (38%) had an initial suspicion of IIH but this was later changed to another diagnosis after diagnostic follow-up. A total of 145 patients had a diagnosis code of G93.2 reported at least once from a specialized Neurological Department. In these patients on validation, the diagnosis IIH was correct in 76% of cases. If, however, not receiving a code ever by a specialized neurology department, a correct diagnosis was only correct in 40% of patients.

After randomization, the algorithm group comprised of 105 patients and the test group of 102 patients. See Table 1 for characteristics of the two groups. Characteristics (sex, number of recorded code G93.2 three or more times, and mean age) on the defined groups:

TABLE 1 Characteristics of algorithm group and test group

	Algorithm group: n (%)	Test group: n (%)	Total: n (%)
All	105	102	207
Correct IIH diagnosis			
Yes (%)	67 (64)	68 (67)	135 (65)
No (%)	38 (36)	34 (33)	72 (35)
sIH diagnosis (%)	11 (10)	5 (5)	16 (8)
Sex			
Female (%)	85 (81)	81 (79)	166 (80)
Male (%)	20 (19)	21 (21)	41 (20)
Mean age (SD)	37.3 (1.5)	38.2 (1.6)	

IIH, idiopathic intracranial hypertension; n, number of patients; sIH, secondary intracranial hypertension; SD, standard deviation.

Definite IIH, probable IIH, not IIH, and sIH are added in Appendix 1, all showing significant difference between the IIH group and non-IIH group.

3.2 | Performance results of the algorithms

The variables ultimately selected for the algorithm 1 (using both NPR and PDR data) were as follows: age, number of times with diagnosis code G93.2 being recorded, and acetazolamide treatment. The final algorithm using only NPR data (algorithm 2) included only age and number of times with a diagnosis code G93.2 being recorded. The percentage of correctly identified individuals according to number of diagnosis codes are shown in Table 2. Gender did not contribute to better prediction (even though we saw a significant difference between IIH patients and non-IIH patients), so that variable was removed by the model. We also tried applying ever having the diagnosis set at a specialized neurology department/clinic but that did not give us better prediction and were therefore removed by the statistical model. These results confirmed that algorithm 2(E) (comprising of variables: age and receiving three or more times the diagnosis code G93.2) best served our purpose due to showing the highest PPV and NPV without making us dependent on PDR register data. There was no difference between algorithm 1 (A) (age + receiving two or more times the diagnosis code + acetazolamide treatment) and algorithm 1 (B) (variables: age + receiving three or more times the diagnosis code + acetazolamide treatment). We chose to use algorithm 1(B) for consistency. Algorithm 1 including data from PDR register showed a slighter better prediction (88.2%), compared to algorithm 2 (86.3%). The odds ratio for variables used in algorithm 1(B) was age 0.94 (0.90-0.98), three or more times the code 5.55 (1.73-17.83) and acetazolamide exposure 3.59 (1.12-11.55). The odds ratio for the variables in algorithm 2(E) for age was 0.93 (0.90-0.97) and for three or more times the code 9.62 (3.35-27.66).

3.2.1 | How well did algorithms 1(B) and 2(E)predict on test group?

To analyze which categories were most troublesome for the algorithm to predict correctly, we made some further differentiations. We looked into how well algorithms 1(B) and 2(E) predicted the diagnosis based on the number of times the diagnosis code G93.2 was recorded (results presented in Table 3). Prediction improved steadily

TABLE 3 Algorithm 1(B) and 2(E) predicting correct diagnosis on test group (n=102) according to number of times with a recorded code G93.2

Number times with code G93.2 recorded	Algorithm 1(B) correctly predicted IIH status % (95% CI)	Algorithm 2(E) correctly predicted IIH status % (95% CI)
1	83.8 (68.0-92.6)	81.1 (65.0-90.8)
2	81.8 (48.6-95.5)	72.7 (40.9-91.1)
3-10	87.9 (71.5-95.5)	90.9 (74.9-97.1)
11-30	100	93.8 (65.7-99.2)
31-88	100	100

Algorithm 1(B) comprise of variables: age, exposed to Acetazolamide treatment and receiving ICD-10 code G93.2 three or more times. Algorithm 2(E) comprise of variables age and receiving ICD-10 code G93.2 three or more times.

TABLE 2 Algorithms tested according to frequency of diagnosis codes recorded

	Algorithm given correct prediction % (95% CI)	PPV% (95% CI)	NPV% (95% CI)
Patient and prescription registry* (Algorithm 1)			
A. IIH correctly predicted two or more codes	88.2 (80.3-93.2)	89.7 (79.8-95.0)	85.3 (69.0-93.8)
B. IIH correctly predicted three or more codes	88.2 (80.3-93.2)	89.7 (79.8-95.0)	85.3 (69.0-93.8)
C. IIH correctly predicted five or more codes	80.4 (71.5-87.0)	82.4 (71.3-89.7)	76.5 (59.4-87.8)
Patient registry only** (Algorithm 2)			
D. IIH correctly predicted two or more codes	85.3 (77.0-91.0)	91.2 (81.6-96.0)	73.5 (56.3-85.7)
E. IIH correctly predicted three or more codes	86.3 (78.1-91.7)	91.2 (81.6-96.0)	76.5 (59.4-87.8)
F. IIH correctly predicted five or more codes	82.4 (73.6-88.6)	83.8 (73.0-90.9)	79.4 (62.5-89.9)

PPV, positive predictive value; NPV, negative predictive value.

*Models contain variables: age, ever exposed to Acetazolamide treatment and specified number of times receiving ICD-10 code G93.2.

**Models contain age, specified number of times receiving ICD-10 code G93.2.

TABLE 4 Algorithm 1(B) and 2(E) predicting correct diagnosis on test group according to diagnosis group

Diagnosis group	Algorithm 1(B) correctly predicted IIH status % (95% CI)	Algorithm 2(E) correctly predicted IIH status % (95% CI)
Definite IIH	94.7 (84.4-98.4)	93.0 (82.3-97.4)
Probable IIH	63.6 (28.8-88.3)	81.8 (42.0-96.6)
sIH	60.0 (8.1-96.2)	60.0 (8.1-96.2)
Wrong diagnosis code* or initial suspicion**	90.9 (46.3-99.2)	81.8 (42.0-96.6)

Algorithm 1(B) comprise of variables: age + receiving diagnosis code G93.2 three or more times and exposure to Acetazolamide treatment. Algorithm 2(E) comprise of variables: age + receiving diagnosis code G93.2 three or more times. Definite IIH, idiopathic intracranial hypertension fulfilling modified Dandy Criteria; Probable IIH, overall clinical description in great concordance with IIH but with some missing data in medical records; sIH, secondary intracranial hypertension.

*Patients seeking medical attention for another diagnosis but received code G93.2 (IIH).

**Initial suspicion of IIH that later changed to other diagnosis or not sufficient workup.

with increasing number of times the diagnosis code was documented. Those with two codes were hardest for the algorithm to predict. We also analyzed how well algorithms 1(B) and 2(E) predicted according to subtype of diagnosis (IIH definite, IIH probable, sIH, wrong diagnosis code or initial suspicion, see Table 4). The algorithm had most trouble predicting sIH diagnosis correctly, only 60% correct prediction. Definite IIH was predicted best, with 93%-95% correct prediction.

4 | DISCUSSION

We produced two algorithms that with better accuracy ascertain which patients given the diagnosis code of G93.2 can be considered as true IIH patients in the NPR. There have been other similar attempts to produce algorithms for certain diagnosis groups that have shown a low PPV value in national registries, necessitating the need to produce algorithms which with higher accuracy can find correct patients to include in studies.[12,13] The best predictive algorithm that we found included data from both the NPR and the PDR. With this, we improved prediction of a correct/incorrect IIH diagnosis to 88% accuracy and the PPV improved to almost 90% compared with previous PPV of 65% when using NPR data alone. Using the algorithm which only used variables available in the NPR, results were only slightly lower (86% of diagnoses were correctly classified, and the PPV 91%). Applying an algorithm such as one of the above has the potential to improve register-based research on IIH patients.

In our previous validation study,[5] an incorrect diagnosis coding was seen in 35% of patients, with one-fifth of those patients showing signs of secondary cause behind intracranial hypertension. It was conceivable that the sIH patients would be the hardest for the algorithm to identify. This because they are often treated (acetazolamide treatment) and followed up in the same way as IIH patients, and thereby receiving an IIH code (due to lack of more suitable codes and maybe also lack of new diagnostic re-evaluation on follow-ups) to describe these follow-up visits. This was also seen when observing how well the algorithm predicted sIH patients (only 60% correct prediction). Apart from sIH, the algorithm did to a large extent successfully remove those with a wrong diagnosis from the cohort (82%-91%).

The number of times a patient was given the IIH diagnose was a strong predictor of whether the patient had true IIH or not, with those with more times receiving the diagnose code much more likely to be considered true IIH. The algorithms did less well in predicting probable IIH (64%-82%) compared with definite IIH (93%-95%). We believe this could be due to the fact that patients with probable IIH were generally diagnosed prior to the start of the study. Three quarters of patients received the IIH diagnosis for the first time before 2006 with the majority given their first diagnosis in the 1990s and thereby having less frequent regular follow-up visits. Therefore, the code was registered fewer times for the probable IIH group during the study period (52% of probable IIH had received the diagnosis code only once compared to only 9% in the group of definite IIH). For those with definite IIH, only 43% were not incident patients (first diagnosis before 2006).

4.1 | Limitations and strength of the study

The algorithms have some limitations. The algorithms were not particularly successful at classifying sIH patients as incorrectly diagnosed with IIH. This is likely due to them receiving similar treatment and follow-up visits, thus receiving the code similar number of times as the IIH group. Looking at whether the IIH diagnosis occurred in tandem with a code for a secondary cause of high intracranial pressure (such as viral meningitis) would have been insightful and might if used in the algorithm have filtered out some of the sIH patients to the incorrect diagnosis group. However, this was beyond the scope of the data available for this study; we did not have data on other NPR registered diagnosis codes given to the included patients. It would also have been interesting trying to add data on registered papilledema coding (H47.1) seeing whether that would have given better prediction; however, these data were also not available to us.

Another limitation is that the two randomized groups were not completely similar regarding composition of patients. We tried adding a variable which described which clinic had given the diagnosis code to improve prediction of correct diagnosis. However, if adding neurology department and ophthalmology departments only, we would miss many IIH patients as there are many hospitals in Sweden where neurologists are employed under internal medicine departments. However, when also adding internal medicine departments into the variable list, we got too few patients in each group diagnosed outside these three departments which made the model unstable. As the majority of patients seen in one of these three specialist clinics were seen in a neurology clinic, we included whether the patient had been given the diagnosis code at least once in a neurology clinic to assess whether

this allowed for a more stable model. However, when this variable was included in the stepwise model, it was not retained as one of the useful variables, indicating it was not a significant predictor. This may be because we lacked the statistical power to detect this association.

The strength of this study is that it is based on results from our validation study of diagnosis coding recorded in the Stockholm County, and we had access to details of medical records.[5] We identified all patients diagnosed with IIH (G93.2) in this region during the study period, and we succeeded in collecting the majority of the medical records (99%), which were reviewed by an experienced neurologist (AS). The records which lacked clarity were reviewed by a second neurologist (INR). We then verified whether the diagnosis code given was consistent with IIH according to the modified Dandy Criteria.[11] As a second control of diagnosis coding, 10% of randomly selected medical records were reviewed blindly by a second neurologist (INR) and the consistency in diagnosis coding between the two neurologists were 95%. Using data from the NPR is a strength given the register's extensive coverage of hospitals and clinics in the Stockholm County, meaning we were not limited to patients attending neurology clinics/departments. It is unlikely we missed a large number of individuals given the IIH code due to very high rates of reporting into the registries. The proportion of missing main diagnosis from the NPR has for several years been around 1%.[10] Ludvigsson et al showed good validity of diagnosis registered in the inpatient NPR with a sensitivity of between 85 and 95% depending on the disease.[14] The NPR does not cover primary care settings; however, it is unlikely that patients were diagnosed in such settings as LP pressure measurements and eye examinations are needed, which are handled by doctors in specialized clinics such as neurology departments, internal medicine departments with neurology expertise or ophthalmologists, and sometimes neurosurgery departments (shunt procedures). We therefore expected that the NPR should have good coverage of healthcare contacts which relate to IIH making register studies possible.

5 | CONCLUSION

With the developed algorithms, we improved the accuracy of predicting correct IIH patients in the NPR which is useful if performing epidemiological studies. Our algorithms successfully excluded patients that mistakenly received the diagnosis code; however, the algorithms did not perform well at excluding patients with sIH. Incident cases were more easily correctly identified using these algorithms.

ACKNOWLEDGEMENT

We are grateful for all collaboration and help in collecting medical records from all Head of Departments and their secretaries at various clinics in the Stockholm region.

CONFLICT OF INTEREST

The authors report no conflicts of interest in this work.

ORCID

A. Sundholm http://orcid.org/0000-0002-2321-2423

REFERENCES

1. Durcan FJ, Corbett JJ, Wall M. The incidence of pseudotumor cerebri. Population studies in Iowa and Louisiana. *Arch Neurol*. 1988;45:875-877.
2. Radhakrishnan K, Thacker AK, Bohlaga NH, Maloo JC, Gerryo SE. Epidemiology of idiopathic intracranial hypertension: a prospective and case-control study. *J Neurol Sci*. 1993;116:18-28.
3. McCluskey G, Mulholland DA, McCarron P, McCarron MO. Idiopathic intracranial hypertension in the Northwest of Northern Ireland: epidemiology and clinical management. *Neuroepidemiology*. 2015;45:34-39.
4. Kesler A, Stolovic N, Bluednikov Y, Shohat T. The incidence of idiopathic intracranial hypertension in Israel from 2005 to 2007: results of a nationwide survey. *Eur J Neurol*. 2014;21:1055-1059.
5. Sundholm A, Burkill S, Sveinsson O, Piehl F, Bahmanyar S, Nilsson Remahl AI. Population-based incidence and clinical characteristics of idiopathic intracranial hypertension. *Acta Neurol Scand*. 2017;136:427-433.
6. McGeeney BE, Friedman DI. Pseudotumor cerebri pathophysiology. *Headache*. 2014;54:445-458.
7. Chen J, Wall M. Epidemiology and risk factors for idiopathic intracranial hypertension. *Int Ophthalmol Clin*. 2014;54:1-11.
8. Fisayo A, Bruce BB, Newman NJ, Biousse V. Overdiagnosis of idiopathic intracranial hypertension. *Neurology*. 2016;86:341-350.
9. Koerner JC, Friedman DI. Inpatient and emergency service utilization in patients with idiopathic intracranial hypertension. *J Neuroophthalmol*. 2014;34:229-232.
10. Anonymous. In english - the National Patient Register. http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish: Socialstyrelsen, 2016. Accessed April 21, 2016
11. Smith JL. Whence pseudotumor cerebri? *J Clin Neuroophthalmol*. 1985;5:55-56.
12. Berglund A, Olsen M, Andersen M, et al. Evaluation of ICD-10 algorithms to identify hypopituitary patients in the Danish National Patient Registry. *Clin Epidemiol*. 2017;9:75-82.
13. Egholm G, Madsen M, Thim T, et al. Evaluation of algorithms for registry-based detection of acute myocardial infarction following percutaneous coronary intervention. *Clin Epidemiol*. 2016;8:415-423.
14. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11:450.

How to cite this article: Sundholm A, Burkill S, Bahmanyar S, Nilsson Remahl AIM. Improving identification of idiopathic intracranial hypertension patients in Swedish patient register. *Acta Neurol Scand*. 2018;137:341–346. https://doi.org/10.1111/ane.12876

APPENDIX 1

Specific characteristics distributed on defined patientgroups (definite and probable IIH, not IIH and sIH)

	Definite IIH	Probable IIH	Not IIH	sIH
Sex (P = .04)				
Female (%)	86.6	82.6	69.1	68.8
Male (%)	13.4	17.4	30.9	31.3
Number of codes (P < .000)				
<3 codes (%)	19.6	56.5	98.2	43.8
≥3 codes (%)	80.4	43.5	1.8	56.3
Mean age (SD) (P < .000)	32.5 (1.2)	33.7 (2.5)	49.3 (2.4)	37.7 (2.4)

Definite IIH = fulfilling modified Dandy Criteria. Probable IIH = not entirely fulfilling the modified Dandy Criteria due to some missing data but with overall clinical description in concordance with IIH. Not IIH = wrong diagnosis code chosen, or inital suspiscion of IIH but later changed diagnosis code. sIH= secondary intracranial hypertension.