

Silent Brain Infarction: A Missed Opportunity for Prevention

Tsadok Avgil M, PhD¹, Streifler YJ, MD², Benbenishty G, MBA³, Man N, MA³, Balancer R, MD, PhD^{1,4}, Makori A, MD, MHA⁵, Feldman SB, PhD¹ and Gilutz H, MD^{1,4*}

¹Clalit Research Institute, Clalit Health Services Tel-Aviv, Israel

²Department of Neurology, Sackler Faculty of Medicine Tel Aviv University, Tel Aviv, Israel

³Business Intelligence, Division of Computing and Information, Clalit Health Services, Tel Aviv, Israel

⁴Faculty of Health Sciences at Ben Gurion University of the Negev, Beer Sheva, Israel

⁵Director of Imaging Informatics, Clalit Health Services, Tel Aviv, Israel

*Corresponding author: Harel Gilutz, MD, ESCF, Professor of Cardiology, Clalit Research Institute, Clalit Health Services, Shoham 2, Ramat Gan, Israel, Tel: +972-3-6925800, Fax: +972-3-6925821, Email: gilutz@bgu.ac.il

Citation: Tsadok AM, Streifler YJ, Benbenishty G, Man N, Balancer R, et al. (2019) Silent Brain Infarction: A Missed Opportunity for Prevention. *Academia J Stroke* 1:005.

Volume 1	Issue 1
Pages	25-30
Received	📅 November 11, 2019
Accepted	📅 December 16, 2019
Published	📅 December 20, 2019

Abstract

Introduction: Clinical silent brain infarction (SBI) is prevalent in 10%-30% of the elderly, predicting future stroke and cryptogenic dementia. SBI could be "administrative" due to communication errors between radiologists and caregivers.

Objective: To study the documentation gap between a brain infarction in a computerized brain tomography (brain-CT) report and diagnosis in the Electronic medical record (EHR).

Methods: This is a multicenter retrospective observational study, using Clalit health maintenance organization integrated data. We studied unstructured data of 1,339 brain-CT reports of patients aged 50-80 years. Patients were categorized by the presence or absence of brain infarction in the brain CT report and then compared to the presence of stroke diagnoses in the corresponding patient EHR. We divided the patient cohort into four groups: with/without brain infarction in the brain-CT report and with/without brain infarction diagnoses in the EHR.

Results: Among cases with brain infarction in the brain-CT reports, the rate of missed documentation in the EHR was 38.1%. Missing documentation was more likely to be associated with the identification of lacunar infarctions (p-value = 0.01), inadequate documentation of other cardiovascular comorbidities and risk factors as well as lower rates of secondary prevention medication (p-value = 0.01). Factors showing significant association with missing documentation in the EHR were a non-neurological indication (p-value = 0.001) if a brain infarction was mentioned only in the report text (and not in summary) (p-value = 0.001), and lacunar infarctions (p-value = 0.009).

Conclusions: A brain infarction can be silent both clinically and administratively. In particular lacunar infarctions are overlooked as their potential long term harmful effect is not well recognized. The main informational barrier is most likely due to interdisciplinary miscommunication. Unstructured computerized report processing can help close this gap and potentially improve care.

Keywords: Silent brain infarction, Unstructured-data

Introduction

Silent brain infarctions (SBI) are ten times more frequent than symptomatic brain infarcts [1] and are common among the older population with a worldwide prevalence of 10%-30% [2,3]. SBI is an independent predictor for future stroke [4] and is associated with an increased risk of dementia, motor deficits, gait impairment, psychiatric disorders and impairments in daily activities [5,6].

Missed documentation of an SBI in an electronic health record (EHR) refers to a situation in which a lesion compatible with brain infarction is observed in a brain computerized tomography report, yet is not documented in a patient's EHR.

Neuroradiology reports are essential sources of information that are usually recorded in EHRs as unstructured data (free text) that are not transformed into structured codes. Therefore, the information embedded in the unstructured brain computerized tomography

(brain-CT) report may be missed, unlike structured, coded data that is easily documented. Besides, poor communication between radiologists and patients' caregivers have been described as potential radiological missed information [7,8]. Moreover, unclear report writing may also lead to a missed documentation of the right diagnosis in a patient's HER [9]. An example of diagnosis mismatch has been found when comparing stroke registries and hospital discharge documents [10-12].

In the current study, we used Clalit Health Services' (CHS) integrated EHR data to evaluate the gap between the documentation of a patient's brain infarction in the brain-CT report and the diagnosis in the EHR and to identify factors associated with this gap.

Methods

Setting

CHS is the largest integrated payer-provider health care organization in Israel and provides care to more than 4.4 million members through an integrated network of 14 hospitals and over 1,500 community clinics. CHS's EHR has been fully computerized since 2002. The CHS comprehensive and integrated database includes EHR data on individual socio-demographic data, administrative records, clinical diagnoses and procedures obtained during hospitalizations and community visits, medication prescriptions and purchasing, laboratory test results, and imaging data. Membership turnover within CHS is less than 2% annually [13].

Study design and data source

The study was a retrospective multicenter study, using data from the CHS's Radiology Information System between January 1st to December 31st, 2017. We randomly selected 5,000 brain-CT reports from five different CHS hospitals (1,000 reports from each hospital). The date in which the brain-CT image was taken was considered as the index date. The brain-CT data was compared with coded diagnoses, demographic and clinical profiles. Following, potential reasons for missed diagnosis were studied.

Parameters

Demographic information included age, sex, and ethnicity. Clinical data were comorbidities, the presence of cardiovascular diseases (CVD) (ischemic heart disease, peripheral vascular disease, stroke), CVD risk factors (hypertension, diabetes, body mass index), dementia and atrial fibrillation. Smoking status was not included as it is an unreliable parameter in the administrative data.

Diagnoses were identified in the EHR using the International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes and were taken from the hospital and the community settings. For ischemic stroke, the following codes were used: 433.x1, 434.x. For hemorrhagic stroke, the following codes were used: 430.x, 431.x. The selection of these stroke-specific codes was shown to have the best positive predictive value for stroke diagnosis [14]. For transient ischemic attack (TIA), the following codes were used: 435.x, 362.34 and for healed or old brain infarctions we used V12.54.

We documented the presence of ischemic heart disease (410-413), peripheral vascular disease (443.9), hypertension (401-404), diabetes (250), dementia (290, 294.1, 294.2, 331.1, 331.2, and 331.82) and atrial fibrillation (427.31).

Primary care EHR was evaluated for the presence of stroke/TIA/brain infarction written as noncoded free text diagnosis.

If a patient has purchased one dose of statins, antiplatelet, or oral anticoagulation at any time before the index brain-CT he was considered as treated with these medications.

Brain-CT reports were processed as follows: We have defined the indication for the brain-CT as suspected stroke, suspected TIA, other neurological complaints, trauma, neurosurgery, and unknown or other indications. We have evaluated whether patients were referred to brain-CT from the emergency department and whether the patient had ever prior brain-CT. Mentioning brain infarction in the report section was classified as only in summary, only in text or both in text and summary.

Study population

We included all patients with brain-CT reports restricted to individuals aged 50-80 years to avoid bias induced by the atypical finding at a younger age and on the other hand the very high prevalence (35%) in older patients [2]. We have also excluded patients who were not members of CHS for at least two consecutive years before the index date. If a patient had more than one brain-CT (duplicates), the one with the finding of brain infarction was chosen for the study.

We excluded non-CHS members, patients with head CT that were not brain-CT (e.g. of the internal auditory canal) and patients that have left the CHS within three months from the index date.

Analysis

For each person who had a brain-CT, we retrieved the radiology free-text report that includes imaging data, clinical indication for the test, unstructured data describing normal and abnormal findings and a summary of the results. A clinician read the radiology reports. A second reviewer conducted cross-validation of 200 cases, and there was an agreement of 97%.

The neuroradiology findings that were manually read were categorized into ischemic non-lacunar infarctions, lacunar infarctions, both ischemic and lacunar brain infarction, and spontaneous intracerebral hemorrhage. Brain infarction was not considered to be present when the report was indecisive. Finally, we documented whether the diagnosis of any infarction was reported in the body of the text section of the report and at the report bottom line, the summary.

We then evaluated the presence of a stroke diagnosis recorded in the EHR at any time before the index date of the brain-CT scan through to three months following the scan (to allow time for the EHR to be updated after the CT scan).

The following four study groups were then created based on the concordance between the brain-CT report and EHR diagnosis documentation:

1. Brain infarction identified within the brain-CT report (R+ for report) and in the EHR system as either stroke or old brain infarction (D+ for diagnosis); (R+D+).
2. Brain infarction identified within the brain-CT report but not in the EHR system (R+D-), which is the "administrative missed diagnosis" group.
3. Brain infarction not identified within the brain-CT report but identified (i.e., brain infarction or stroke) in the EHR system (R-D+).
4. Brain infarction not detected within the brain-CT report and not in the EHR system (R-D-).

Baseline characteristics were assessed across these four study groups and were extracted in proximity to the index date. The variables included socio-demographic data, comorbidities, and medication purchases.

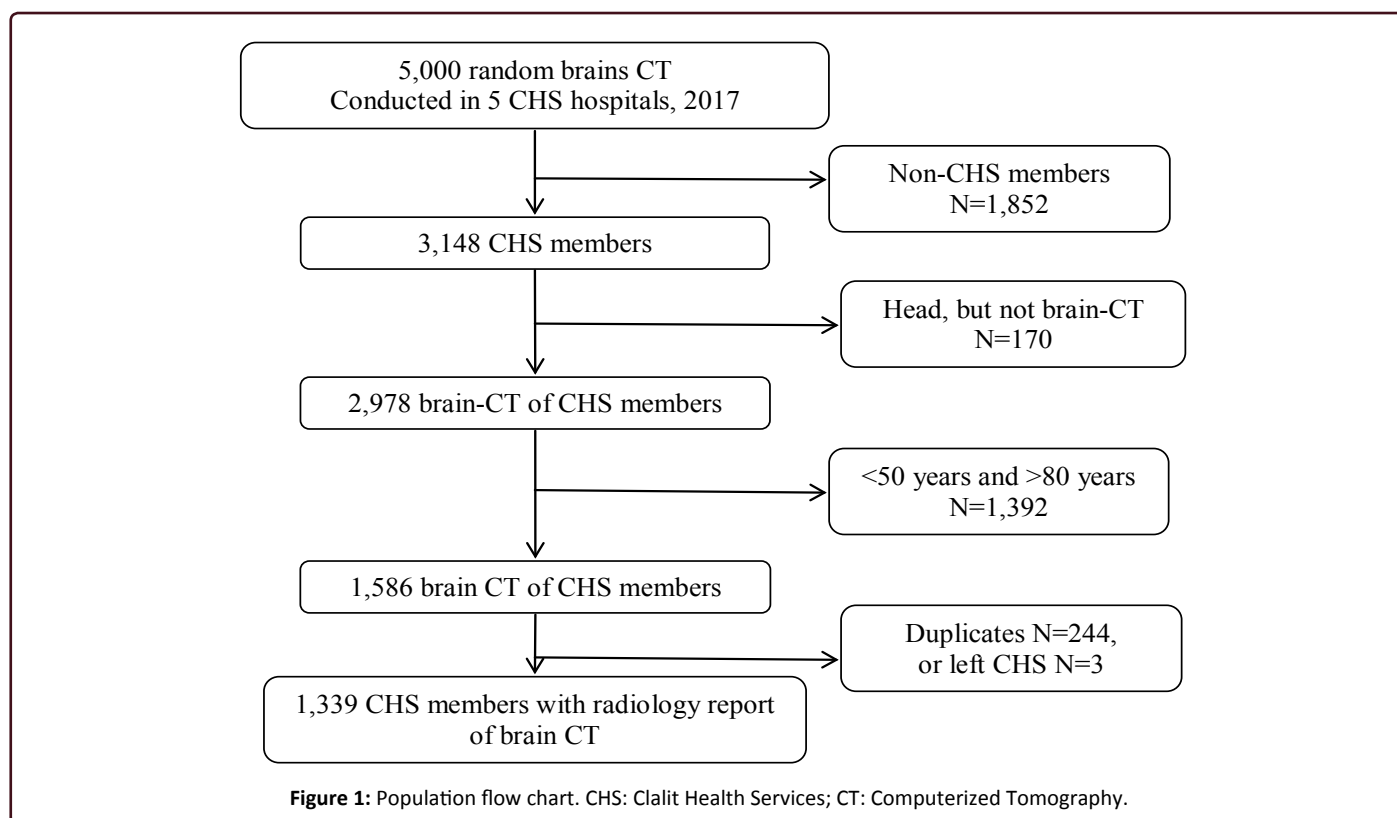
Continuous variables were presented as means and standard deviations (sd), and dichotomous variables were presented as percentages. All statistical analyses were performed using R version 1.0.143.

The organization's institutional review board approved this study and the use of these data.

Unpublished anonymized data will be shared by request from any qualified investigator.

Results

We initially evaluated 5,000 brain-CTs reports, of which 1,339 individuals with a brain-CT were included in the study (Figure 1). This group of 1,339 was divided into four groups according to the concordance of a brain infarction diagnosis (either as brain infarction or stroke) in the EHR and with or without a finding of brain infarction on the CT report. There were 263 (19.6%) patients who had both a brain infarction finding in the brain-CT report and a written diagnosis in the EHR (R+D+) and 162 (12.1%) patients with a reported finding of brain infarction in the brain-CT report but without a documented diagnosis in the EHR (R+D-).

**Table 1:** Socio-demographic and clinical characteristics among the four study groups.

Study group, n (%)	All	R+D+ (a)	R+D- (b)	R-D+ (c)	R-D- (d)	p value
	1,339 (100)	263 (19.6)	162 (12.1)	197 (14.7)	717 (53.5)	(a) vs. (b)*
Demographics, n (%)						
Age, mean (SD)	68.0 (8.7)	70.2 (7.9)	71.8 (7.7)	68.4 (8.0)	66.3 (9.0)	
Women	697 (52.1)	113 (43.0)	75 (46.3)	118 (59.9)	391 (54.5)	
Arabs	154 (11.5)	39 (14.8)	15 (9.3)	25 (12.7)	75 (10.5)	
Comorbidities, n (%)						
IHD	409 (30.5)	119 (45.2)	54 (33.3)	58 (29.4)	178 (24.8)	0.0152
PVD	75 (5.6)	34 (12.9)	6 (3.7)	10 (5.1)	25 (3.5)	0.0016
Hypertension	934 (69.8)	237 (90.1)	116 (71.6)	153 (77.7)	428 (59.7)	0.0348
Diabetes	621 (46.4)	176 (66.9)	77 (47.5)	92 (46.7)	276 (38.5)	< 0.001
Dementia	89 (6.6)	26 (9.9)	13 (8.0)	20 (10.2)	30 (4.2)	
BMI, mean (SD)	28.2 (5.5)	28.6 (6.0)	27.7 (5.2)	28.4 (5.8)	28.2 (5.3)	
AF	169 (12.6)	54 (20.5)	32 (19.8)	25 (12.7)	58 (8.1)	
Medications, n (%)						
OAC [†]	183 (13.7)	56 (21.3)	30 (18.5)	32 (16.2)	65 (9.1)	
OAC with AF	117 (69.2)	43 (79.6)	21 (65.6)	18 (72.0)	35 (60.3)	
Any anti-platelets [‡]	1,005 (75.1)	251 (95.4)	127 (78.4)	175 (88.8)	452 (63.0)	< 0.001
Any anti-thrombotic [§]	1,016 (75.9)	254 (96.6)	130 (80.2)	175 (88.8)	457 (63.7)	
PVD/IHD with ASA	476 (35.6)	148 (56.3)	63 (38.9)	71 (36.0)	194 (27.1)	< 0.001
Statins	1,068 (79.8)	243 (92.4)	134 (82.7)	176 (89.3)	515 (71.8)	0.0022

†: Warfarin or NOACs; ‡: ASA or Clopidogrel; §: ASA or Clopidogrel or any OAC; *: Statistically significant differences between the R+D+ (a) vs. R+D- (b) groups ($p < 0.05$); AF: Atrial fibrillation; ASA: Aspirin; BMI: Body Mass Index; CHS: Clalit Health Services; IHD: Ischemic Heart Disease; PVD: Peripheral Vascular Disease; CVD: Cardiovascular Disease: PVD and/or IHD; OAC: Oral Anticoagulation; R+/-: Brain infarction documentation in the report; D+/-: Brain infarction documentation in the EHR.

There were 197 (14.7%) patients without brain infarction in the brain-CT report but with a diagnosis in the EHR (R-D+), and 717 (53.5%) patients with neither a reported brain infarction in the brain-CT imaging nor a diagnosis in the EHR (R-D-).

The demographic and clinical backgrounds of the patients are presented in **Table 1**. Patients with recorded brain infarction in the brain-CT reports with or without a record in the EHR (R+D+ and R+D- groups) were older, less likely to be women and had a higher rate of comorbidities compared to patients with neither brain infarction reported on the brain-CT nor in the EHR (R-D-).

Patients who had a brain infarction finding on the brain-CT report and a written diagnosis in the EHR (R+D+) compared to those without a written diagnosis in the EHR (R+D-) had comparable mean ages (70.2 vs. 71.8 years, respectively) (**Table 1**). However, patients with a finding of brain infarction in the CT and no diagnosis recorded in the EHR (R+D-) had significantly lower rates of documented ischemic heart disease, (33.3% vs. 45.2%), peripheral vascular disease (3.7% vs. 12.9%), hypertension (71.6% vs. 90.1%) and diabetes (47.5% vs. 66.9%). The R+D- group, was less often treated with aspirin or clopidogrel (78.4% vs. 95.45), even in the presence of CVD other than stroke (56.3% vs. 38.9%), and less often purchased anticoagulation in the presence of documented AF (80.2% vs. 96.6%).

Indications for the brain-CT imaging and their findings are presented in **Table 2**. Patients who had a finding of brain infarction in their CT without a recorded diagnosis (R+D-), as compared to those who did have a diagnosis (R+D+), were less often suspected to have a stroke (12.3% vs. 43.7%, $p < 0.001$). Among the R+D- group, an indication for brain-CT due to trauma or other neurological presentation (such as dizziness, syncope, or headache) was more frequent compared to the R+D+ group ($p = 0.05$ and $p = 0.001$, respectively). Patients in the R+D- group less often reported to have ischemic non-lacunar infarctions than the R+D+ group (34.0% vs. 48.7%, $p < 0.001$), while they had a higher rate of lacunar infarctions (72.2% vs. 60.5%, $p < 0.01$). Concordant brain infarction diagnosis (R+D+) was associated with reporting

the brain infarction finding in the summary as compared to reporting only in the text in the R+D- group ($p < 0.01$).

Discussion

This study, conducted with individuals aged 50-80 years, found that 12.1% of the population that was referred for a b-CT for any reason, had an administrative missed diagnosis of SBI, meaning a BI finding in the CT report was not transformed to a written diagnosis in the patient's EHR. That is presenting administrative missed diagnosis or "administrative SBI". In a recent meta-analysis, SBI was shown to be an independent predictor for stroke, with a hazard ratio of 2.08,4 increasing the risk for cryptogenic stroke [15], motor deficits and impairments in activities of daily living [5,6]. Importantly, the risk of recurrent stroke after silent lacunar infarct is similar to that of most other types of stroke [3,16,17].

Lacunar infarcts show a favourable prognosis in the short term, but an increased risk of death, stroke recurrence, and dementia in the mid- and long term; thus requiring preventive treatment [18].

A "Scientific Statement for Healthcare Professionals" from the American Heart Association/American Stroke Association for primary stroke prevention suggested that it may be reasonable to follow the guidelines for primary stroke prevention in individuals with SBI [19]. Indeed; it was shown recently that aspirin significantly reduced stroke recurrence in patients with lacunar stroke [20].

Patients whose EHRs are not updated following a brain infarction finding in brain-CT are potentially at risk for poor clinical outcomes, as they may not receive appropriate measures for stroke prevention. Thus, detecting and documenting the presence of SBI is a significant step toward effective prevention.

Other examples of missed diagnoses were found in a study

Table 2: Indication and findings of the computerized brain tomography scan among the four groups.

Study group, n (%)	All	R+D+ (a)	R+D- (b)	R-D+ (c)	R-D- (d)	p value
	1339 (100)	263 (19.6)	162 (12.1)	197(14.7)	717 (53.5)	(a vs. b)*
Indication for brain-CT, n (%)						
Suspected stroke	255 (19.0)	115 (43.7)	20 (12.3)	58 (29.4)	62 (8.6)	0.001
Suspected TIA	29 (2.2)	8 (3.0)	3 (1.9)	9 (4.6)	9 (1.3)	
Other neurological presentations	382 (28.5)	56 (21.3)	65 (40.1)	44 (22.3)	217 (30.3)	< 0.001
Trauma	299 (22.3)	44 (16.7)	40 (24.7)	31 (15.7)	184 (25.7)	0.05
Neurosurgery	113 (8.4)	17 (6.5)	10 (6.2)	25 (12.7)	61 (8.5)	
Unknown	237 (17.7)	22 (8.4)	19 (11.7)	28 (14.2)	168 (23.4)	
Other	24 (1.8)	1 (0.4)	5 (3.1)	2 (1.0)	16 (2.2)	
Referred from ED, n (%)	197 (14.7)	49 (18.6)	24 (14.8)	34 (17.3)	90 (12.6)	0.06
Had prior brain-CT, n (%)	430 (32.1)	143 (54.4)	59 (36.4)	84 (42.6)	144 (20.1)	
Location of brain infarction in the report, n (%)						
Only in summary	247 (18.4)	141 (53.6)	106 (65.4)	-	-	
Only in text	25 (1.9)	13 (4.9)	12 (7.4)	-	-	< 0.001
Both in text and summary	153 (11.4)	109 (41.4)	44 (27.2)	-	-	< 0.001
Radiology findings, n (%)						
Ischemic non-lacunar infarctions	183 (13.7)	128 (48.7)	55 (34.0)			0.001
Lacunar infarctions	276 (20.6)	159 (60.5)	117 (72.2)			0.009
Intracerebral hemorrhage	20 (1.5)	17 (6.5)	3 (1.9)			0.03

*P values are between the two groups with brain infarction finding: R+D+ vs. R+D-; CHS: Clalit Health Services; Brain-CT: Computerized Brain Tomography; CT: Computerized Tomography; ED: Emergency Department; TIA: Transient Ischemic Attack; R+/-: Brain infarction documentation in the report; D+/-: Brain infarction documentation in the EHR.

verifying the presence of stroke diagnosis in the Scottish Stroke Care Audit, which compared administrative hospital discharge databases. This study found a 25% discrepancy between the two databases, i.e., missed the diagnosis [11]. Similarly, the Dijon Stroke study identified a 22.9% missed diagnosis between registry data and discharge diagnosis [9,12]. We suggest that the difference in the magnitude of the administrative missed diagnoses in our study was due to better integration of the CHS system, with many unified computerized systems between the hospitals and the community settings.

We identified several factors of appropriate brain infarction documentation in the EHR. Among patients with a brain infarction finding in the brain-CT report that was not documented in the EHR, we found lower rates of cardiovascular disease, hypertension, and diabetes. Patients whose diagnoses were not recorded in the EHR were also less often treated with antiplatelet medication in the presence of documented CVD. Even among patients with AF, those in the administrative missed diagnosis group were less likely to receive anticoagulation treatment. This unexpected finding generates the hypothesis that neglecting documentation may represent a poor handling and deficient EHR diagnoses, consequently leading to less than optimal treatment.

EHR information is captured by entering data in structured or unstructured formats. Structured data is created through constrained choices in pre-determined models that make information easily searchable, aggregated, analyzed, and reported, yet often refer to the primary clinical problem, neglecting other significant clinical information. Unstructured clinical data, on the other hand, exist in the form of free-text narratives. Unstructured reports have the potential to improve patient care when transformed into a structured form [21,22]. Exploring the Radiology Information System's free-text, semi-structured reports had the potential to close the gap between text narratives and formal coded diagnoses. We manually analyzed brain-CT reports, token key linguistic elements, and transformed the text into a binary decision in a rule-based algorithm. The algorithm took the brain-CT report text and eventually classified whether there is or there is no identified brain infarction in the text with excellent accuracy, thus able to routinely detect missed diagnoses in brain-CT reports.

Limitations

This study had a few limitations. First, we relied on administrative and clinical data from an EHR system to detect stroke diagnosis. Routine real-world data collection has a risk of stroke misclassification. Second, we could not evaluate the reproducibility of the brain-CT reports introducing potential errors. However, typically, these images were routinely reviewed by the department's radiologist, and some were corrected. Third, we addressed reports from the hospital setting, many from the emergency department. Thus, there may have been referral bias compared to brain-CT in the community. However, the strength of our study was in using a population-based design that received routine clinical practice, thus presenting real-world results.

Forth, we have found a group of patients with a diagnosis of stroke in the EHR, but brain infarction was not detected in their brain-CT reports. There are few possible reasons for this; some of these patients had a brain-CT performed outside of the CHS healthcare system, and therefore were not available for analysis. Others had either TIA or underwent a brain-CT in the very early stage of the stroke, too early for brain infarction to appear, yet the diagnosis was given on clinical grounds. Additionally, clinical strokes are not always detected in brain-CTs. Finally, some patients might have had a false positive diagnosis.

Conclusions

The study presented a significant incidence of administrative

missed diagnosis of existing brain infarction on brain-CT. "Administrative SBI" was found to be associated with probable inadequate documentation of other comorbidities and with a lower rate of secondary prevention measures. A further study verifying that and the impact on the patient's prognosis is warranted.

Our result also showed that lacunar infarction on a brain-CT was less likely to be transformed into diagnosis in the patient EHR, which could mean that the significance of the clinical implication of lacunar infarction was underestimated.

The main barrier to the documentation of brain infarction in the study was most likely due to interdisciplinary communication. Processing unstructured data, preferably by computerized automated reading by using natural language processing or artificial intelligence, may decrease administrative missed diagnosis and potentially improve care.

Study Funding

None.

Disclosure

The authors report no disclosures.

References

1. Leary MC, Saver JL. Annual incidence of first silent stroke in the United States: a preliminary estimate. *Cerebrovasc Dis* 2003;16:280-5.
2. Vermeer SE, Koudstaal PJ, Oudkerk MM, Hofman A, Breteler MM. Prevalence and risk factors of silent brain infarcts in the population to based Rotterdam Scan Study. *Stroke* 2002;33:21-5.
3. Chen DW, Wang YX, Shi J, Zhang WQ, Yang F, Yin YW, Ma LN. Multiple Silent Brain Infarcts Associated with Severer Stroke in Patients with First to Ever Ischemic Stroke without Advanced Leukoaraiosis. *J Stroke Cerebrovasc Dis* 2017;26(9):1988-1995.
4. Gupta A, Giambone AE, Gialdini G, Finn C, Delgado D, Gutierrez J, Wright C et al. Silent Brain Infarction and Risk of Future Stroke: A Systematic Review and Meta to Analysis. *Stroke* 2016;47:719-25.
5. Prins ND, van Dijk EJ, den Heijer T, Vermeer SE, Jolles J, Koudstaal PJ et al. Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain*. 2005;128(Pt 9):2034-41.
6. Au R, Massaro JM, Wolf PA, Young ME, Beiser A, Seshadri S et al. Association of white matter hyperintensity volume with decreased cognitive functioning: the Framingham Heart Study. *Arch of Neurol* 2006;63:246-50.
7. McCreadie G, Oliver TB. Eight CT lessons that we learned the hard way: an analysis of current patterns of radiological error and discrepancy with particular emphasis on CT. *Clin Radiol* 2009;64:491-9.
8. Kim YW, Mansfield LT. Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors. *Am J Roentgenol* 2014;202:465-70.
9. Bruno MA, Walker EA, Abujudeh HH. Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction. *Radiographics* 2015;35:166 -76
10. Vermeer SE, den Heijer T, Koudstaal PJ, Oudkerk M, Hofman A, Breteler MM. Incidence and Risk Factors of Silent Brain Infarcts in the Population to Based Rotterdam Scan Study. *Stroke* 2003;34:392-396.
11. Turner M, Barber M, Dodds H, Dennis M, Langhorne P, Macleod MJ. Agreement between routine electronic hospital discharge and Scottish Stroke Care Audit (SSCA) data in identifying stroke in the Scottish population. *BMC Health Serv Res* 2015;15:583.
12. Aboa-Eboulé C, Mengue D, Benzenine E, Hommel M, Giroud M, Béjot Y et al. How accurate is the reporting of stroke in hospital discharge data? A pilot validation study using a population to based stroke registry as control. *J Neurol* 2013; 260(2):605-613.
13. Karpati T, Cohen-Stavi CJ, Leibowitz M, Hoshen M, Feldman BS, Balicer

- RD. Towards a subsiding diabetes epidemic: trends from a large population to based study in Israel. *Popul Health Metr* 2014;12:32.
14. Woodfield R, Grant I; UK Biobank Stroke Outcomes Group; UK Biobank Follow-Up and Outcomes Working Group, Sudlow CL. Accuracy of Electronic Health Record Data for Identifying Stroke Cases in Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes Group. *PLoS One*. 2015;23;10(10):e0140533.
 15. Bousser MG, Chabriat H. Vascular dementia. *Bull Acad Natl Med*. 2012;196(2):409-28;
 16. Norrving B. Long to term prognosis after lacunar infarction. *Lancet Neurol* 2003;2:238-245.
 17. Andersen SD, Skjøth F, Yavarian Y, Bach FW, Lip GY, Larsen TB. Multiple Silent Lacunes Are Associated with Recurrent Ischemic Stroke. *Cerebrovasc Dis*. 2016;42(1-2):73-80.
 18. Arboix A, Martí-Vilalta JL. *Expert Rev Neurother*. Lacunar stroke. 2009;9(2):179-96.
 19. Smith EE, Saposnik G, Biessels GJ, Doubal FN, Fornage M, Gorelick PB. Prevention of Stroke in Patients with silent cerebrovascular disease: A Scientific Statement for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke* 2017;48(2):e44-e71.
 20. Kong T, Chen J, Sun K, Zhang W, Wang J, Song L et al. Aspirin reduced recurrent stroke risk in patients with lacunar stroke. *Acta Neurol Scand*. 2019;140(1):78-83.
 21. Townsend H. Natural language processing and clinical outcomes: the promise and progress of NLP for improved care. *J AHIMA* 2013;84:44-5.
 22. Murphy DR, Meyer AN, Sittig DF, Meeks DW, Thomas EJ, Singh H3. Application of electronic trigger tools to identify targets for improving diagnostic safety. *BMJ Qual Saf* 2019;28(2):151-9.

Open Access Declaration

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source of content.