ORIGINAL RESEARCH

Open Access

Artificial intelligence in Emergency Medical Services dispatching: assessing the potential impact of an automatic speech recognition software on stroke detection taking the Capital Region of Denmark as case in point

Mirjam Lisa Scholz[1,2]* , Helle Collatz-Christensen[1], Stig Nikolaj Fasmer Blomberg[1], Simone Boebel[1,2], Jeske Verhoeven[1,2] and Thomas Krafft[2]

Abstract

Background and purpose: Stroke recognition at the Emergency Medical Services (EMS) impacts the stroke treatment and thus the related health outcome. At the EMS Copenhagen 66.2% of strokes are detected by the Emergency Medical Dispatcher (EMD) and in Denmark approximately 50% of stroke patients arrive at the hospital within the time-to-treatment. An automatic speech recognition software (ASR) can increase the recognition of Out-of-Hospital cardiac arrest (OHCA) at the EMS by 16%. This research aims to analyse the potential impact an ASR could have on stroke recognition at the EMS Copenhagen and the related treatment.

Methods: Stroke patient data ($n = 9049$) from the years 2016–2018 were analysed retrospectively, regarding correlations between stroke detection at the EMS and stroke specific, as well as personal characteristics such as stroke type, sex, age, weekday, time of day, year, EMS number contacted, and treatment. The possible increase in stroke detection through an ASR and the effect on stroke treatment was calculated based on the impact of an existing ASR to detect OHCA from CORTI AI.

Results: The Chi-Square test with the respective post-hoc test identified a negative correlation between stroke detection and females, the 1813-Medical Helpline, as well as weekends, and a positive correlation between stroke detection and treatment and thrombolysis. While the association analysis showed a moderate correlation between stroke detection and treatment the correlation to the other treatment options was weak or very weak. A potential increase in stroke detection to 61.19% with an ASR and hence an increase of thrombolysis by 5% in stroke patients calling within time-to-treatment was predicted.

Conclusions: An ASR can potentially improve stroke recognition by EMDs and subsequent stroke treatment at the EMS Copenhagen. Based on the analysis results improvement of stroke recognition is particularly relevant for females, younger stroke patients, calls received through the 1813-Medical Helpline, and on weekends.

*Correspondence: scholz.mirjam@web.de

[1] Emergency Medical Services, Capital Region of Denmark, Telegrafvej 5, 2750 Ballerup, Denmark
Full list of author information is available at the end of the article

Background

According to the World Health Organization (WHO), strokes were the second leading cause of death and the third leading cause of disability-adjusted life years (DALYs) globally, in 2019 [1]. With 63.5 deaths per 100,000 population, stroke is among the top ten causes of death in Denmark [2]. Additionally, with 1024.1 DALYs per 100,000 population, strokes are within the top ten causes of DALYs in Denmark [3]. This is due to stroke patients irretrievably losing approximately 1.9 million neurons, every untreated minute after stroke onset, leading to 1.8 DALYs per minute [4, 5]. Several studies have determined that patients with a time-to-treatment of 90 min showed the best health outcome [6–9]. However, a benefit of intravenous thrombolysis with alteplase for patients with acute ischaemic stroke can be achieved within a time-to-treatment of 4.5 h [6, 9, 10]. Thus, it is crucial to minimize the time between stroke onset and treatment to reduce the mortality as well as the DALYs caused by strokes [11].

To initially get access to hospital treatment in Denmark, patients need to be referred to the hospital by a general practitioner or through the medical helplines (1-1-2 or 1813) of the Emergency Medical Services (EMS) [12]. While the 1-1-2 is the emergency number, the 1813-Medical Helpline (1813) serves as an out-of-hours number providing direct contact to specially trained nurses and physicians within the same emergency dispatch centre of the Capital Region of Denmark [13]. Previous research has shown, that the accurate and early stroke detection by the EMS plays an important role in the timely hospital admission of stroke patients, through dispatching a high priority ambulance ("A" response) [14–17]. Hsieh et al. [18] and Oostema et al. [19] have found that stroke detection by Emergency Medical Dispatchers (EMDs) leads to an improved stroke care and accordingly a better outcome for stroke patients [20]. However, several studies have shown that the accuracy of stroke detection among EMDs are highly variable, between 30 and 83% [8, 21, 22]. In Denmark, an observational study from 2012–2014 found a sensitivity of 66.2% in stroke recognition at the EMS Copenhagen [15]. Additionally, Amtoft et al. [23] identified, that approximately 50% of stroke patients in Denmark did not arrive at the hospital within the stated 4.5 h window of revascularization.

Previous research by Blomberg et al. [24] and Cleve et al. [25] has shown, that artificial intelligence (AI) in emergency medicine increases the accuracy as well as efficiency and reduces the time-to-treatment. Previously, an automatic speech recognition software (ASR) for the detection of Out-of-Hospital cardiac arrests (OHCA) by CORTI AI has proven to increase the sensitivity of OHCA from 72.5 to 84.1% and reduce the median time-to-recognition from 54 to 44 s at the EMS Copenhagen [24]. This software "listens" to the emergency call, processes the audio, transforms it into a textual representation, analyses it and outputs a prediction on the potential presence of cardiac arrest. Based on advanced speech analysis through AI, the technology structures and analyses all sounds and spoken information during a live EMS conversation and converts this data into a valid prediction [25]. The software continuously learns from previous patient consultations and published medical papers in the specific field of cardiac arrest [24, 25].

In line with the EU's values to "become a global leader in innovation in the data economy and its application" [26] and considering the success of the ASR for OHCA by CORTI AI, the question arises, whether an ASR could improve the accuracy and speed of detection, and thus reduce the burden of disease, of other time-critical medical issues, like strokes [27]. Accordingly, this research aims to determine, *how an ASR, at the EMS Copenhagen, could contribute to a more accurate stroke detection and impact the stroke related treatment*. The following research questions address this aim:

(a) How many strokes are detected at the EMS Copenhagen and independently in the EMS numbers 1-1-2 and 1813, throughout 2016–2018?

(b) Is there a difference in stroke characteristics (e.g., stroke type) and patient specific characteristics (e.g., age, sex, time of day, and weekday) between strokes detected and strokes not detected at the EMS Copenhagen, throughout 2016–2018?

(c) Is there a correlation between stroke detection at the EMS Copenhagen and the treatment a stroke patient received, throughout 2016–2018?

(d) Which additional number of strokes could potentially be recognized at the EMS Copenhagen using an ASR?

(e) How could this additional number of strokes detected effect stroke treatment?

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 3 of 17

These analyses were performed to determine necessity of improving the stroke detection rate at the EMS Copenhagen and predict the possible impact of an ASR on stroke recognition as well as the influence on the stroke related treatment.

Methods

This research is a descriptive retrospective quantitative single case study [28, 29].

Setting

The research has been performed with data from the Capital Region of Denmark, with a population of 1.85 million (2020) in an area of 2563 km^2 [30, 31]. In the Capital Region of Denmark the 1-1-2 emergency number and the 1813-Medical Helpline serve as contact points of the EMS [12]. The 1-1-2 and 1813 are part of one emergency medical dispatch centre and allow the assessment of severity of the callers medical condition and an according response independent of the number dialled, thus providing a single point of contact for patients seeking help for emergency and/or acute conditions [12]. While the 1-1-2 serves as an immediate emergency contact, the 1813 is considered an alternative for the GP operated by nurses during the out-of-hours times, between 4 p.m. and 8 a.m. as well as on the weekends. The internal and external validity for our study is ensured by including all the stroke data of the Capital Region of Denmark within the respected timeframe [32, 33].

Data collection

Retrospective research is performed on existing data from 2016–2018 [34]. For 2016–2018, data of 15,258 stroke patients, with the Capital Region of Denmark as emergency site, were extracted from the Danish Stroke Registry, a nationwide clinical database [35]. Stroke patients within this study were distinguished by the types ischaemic and haemorrhagic stroke. This data was joined with the EMS contacts of the stroke patients based on the Danish Det Centrale Personregister (CPR), a personal identification number of Danish citizens, extracted from the EMS database. EMS contacts in the context are considered all contacts to the EMS (1-1-2/1813) by or on behalf of a patient. Only stroke patients with an EMS contact seven days prior or seven days post the onset of the stroke were included in this research, since first stroke symptoms can already occur up to seven days prior to the stroke [36]. For stroke patients with several EMS contacts (based on the CPR), only the contact closest to the onset of stroke was included in this analysis, since this contact is most likely to be the stroke related contact. Stroke patients that did not contact the 1-1-2 or the 1813 were excluded. Stroke contacts to the EMS were coded in "stroke relevant criteria", "stroke nonrelevant criteria", and "missing criteria" based on the criteria of the Danish Index for 1-1-2 and the 1813-Index for the 1813. The Danish Index and the 1813-Index guide the EMDs in assessing the urgency of the emergency situation [13]. The EMS contacts that had an indication of chapter A.26.03. (Suspected stroke, hemiparesis) and A.26.04. (Suspected stroke, reduced consciousness or dizziness) within the Danish Index were coded as "stroke relevant criteria" within this research. All other chapters were considered "stroke nonrelevant criteria". The stroke contacts were coded as "missing criteria" when no criteria based on the Danish Index or the 1813-Index were indicated by the EMD. Additionally, to the variables mentioned before, the following characteristics were included: response plan priority of the EMS, age, sex, year, treatment options thrombolysis, reperfusion, thrombectomy, endovascular, or surgical treatment, incident occurrence on a weekday or the weekend, the EMS number called, and the time-to-call within the time-to-treatment for thrombolysis of 4.5 h.

Outcome measures

The outcomes measured are the number of strokes detected at the EMS and respectively at the 1-1-2 and 1813, as well as the difference between detection when using the two EMS access phone numbers. Additionally, the change of stroke detection throughout 2016–2018, the difference in stroke detection among age, sex, stroke type, year, weekday or weekend, and time of day was determined. Furthermore, correlation between treatment of a stroke patient and detection of stroke through the EMD was analysed. Lastly, a prediction on the presumable number of additional strokes detected at the EMS with CORTI AI and the presumable related change in stroke treatment was made. For this research, strokes are considered detected by the EMS if the criteria of the Danish Index or the 1813-Index were stroke relevant and if a high priority ambulance ("A") was dispatched, as this is the assigned stroke response in Denmark [14]. Accordingly, a stroke is considered as "not detected", if the criteria of the Danish Index or the 1813-Index were stroke relevant, but no "A" response was dispatched or if the criteria of the respective Index were not stroke relevant. For the analysis within this study, the beforenamed outcomes were analysed for "strokes detected" compared to strokes with "missing criteria" and "strokes not detected", separated into "stroke relevant criteria but no "A" response" and "stroke nonrelevant criteria".

Data analysis

To analyse the correlation between two categorical variables within this research, the Chi-Square test was used

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 4 of 17

[37, 38]. This applies to the variables: EMS number (1-1-2/1813), stroke type (ischaemic/haemorrhagic), year (2016–2018), sex (male/female), weekday (Monday-Friday/Saturday-Sunday), treatment (yes/no), and the treatment options, thrombolysis (yes/no), reperfusion (yes/no), thrombectomy (yes/no), endovascular (yes/no), and surgical treatment (yes/no), in correlation with the "strokes detected", "strokes nonrelevant criteria", "stroke relevant criteria but no "A" response", and strokes with "missing criteria". To avoid type I error, the Fisher's Exact test was used for the analysis of characteristics with a cell frequency of < 5, such as the correlation between stroke recognition by the EMD and reperfusion, surgical as well as endovascular treatment [38, 39]. For the further analysis of the correlation analysed with the Fisher's Exact test, the hereafter described analyses for the Chi-Square test were performed. To determine the goodness of fit a Log Likelihood Ratio was performed [40]. The goodness of fit was determined for all the named variables, with a statistically significant p-value of $p \leq 0.05$ (95% CI) [40]. The strength of determined correlations was tested through an association analysis [38]. Within this, the Cramer's V was interpreted, since it is considered a robust test for strength of association within multiple group studies [41]. Lastly, a post-hoc test for independence with adjusted residuals was performed based on a pairwise comparison with a Bonferroni-Holm correction, to adjust the significance level for variables with more than two characteristics [38, 42, 43].

An analysis of variance was performed for the interval scaled variables age and time of day. To respectively choose the suitable statistical test, the normality, using the Shapiro–Wilk test, and the homogeneity of variance, using the Levene's test, were tested for the named variables among the four groups, "stroke detected", "stroke nonrelevant criteria", "stroke relevant criteria but no "A" response", and "missing criteria" [44–46]. Since the Shapiro–Wilk test showed no normal distribution in any of the groups for age or time of day, the Kruskal–Wallis test was chosen to analyse the beforenamed correlations. The Kruskal–Wallis test was solely chosen due to no normal distribution of the stroke patients, however not according to the size of the groups. Additionally, the Wilcoxon–Mann–Whitney test with a Bonferroni-Holm correction was conducted as post-hoc test to determine the correlation between the individual groups through a pairwise comparison [47]. This post-hoc test was respectively chosen, due to the statistically significant Levene's test ($p < 0.05$, 95% CI) determining no homogeneity of variance for the age of the stroke patients and time of call throughout the groups analysed [46].

The results for all analyses were considered statistically significant, when $p \leq 0.05$ (95% CI), or based on the adjusted significance level within the post-hoc test of independence. The statistical analysis was performed with the statistical software R 3.6.3 [48].

Under the condition that the rise of the detection rate for strokes would be the same as the increase in OHCA detection rate through CORTI AI, the presumable increase of the detection rate of strokes using an ASR was calculated. This calculation was conducted based on the results of the analysis of strokes detected by the EMS. The following calculations were performed:

> *Detection rate of stroke with an ASR = (Detection rate of OHCA with an ASR/Detection rate of OHCA without an ASR) * Detection rate of strokes at EMS*
> *Strokes detected with an ASR = Stroke patients calling EMS * Detection rate of strokes with an ASR*
> *Additional strokes detected with an ASR = Strokes detected with an ASR − Strokes detected without an ASR*

Based on the results of this calculation, the potential change in treatment of stroke patients affected by stroke detection through the ASR was determined, under the condition that the number of strokes with "missing criteria" will not be influenced by the ASR. This condition is necessary, due to a lack of information on "missing criteria". The following calculations were performed for the total amount of treatment as well as for each individual treatment:

> Strokes *with treatment$_x$ with an ASR = Strokes not detected through an ASR with treatment$_x$ + Strokes detected through an ASR with treatment$_x$ + Strokes "missing criteria" with treatment$_x$*
> Additional *Strokes with treatment$_x$ with an ASR = Strokes with treatment$_x$ with an ASR − Strokes with treatment$_x$ without an ASR*
> *Treatment$_x$ Rate with an ASR = Strokes with* treatment$_x$ *with an ASR/Strokes*
> Change *in treatment$_x$ Rate with an ASR = (Treatment$_x$ Rate with an ASR/Treatment$_x$ Rate without an ASR) − 1*

The same analysis was performed for the subgroup of stroke patients calling within the time-to-treatment (4.5 h) of thrombolysis ($n = 6013$), as a sensitivity analysis to enable more precise predictions [49, 50]. Within the subgroup analysis the Fisher's Exact test was additionally used for the analysis of thrombectomy due to cell frequencies < 5 [38, 39].

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 5 of 17

Results

For the timeframe 2016–2018, 15,258 stroke patients from the Danish Stroke Registry within the Capital Region of Denmark were included in this research (Fig. 1) [35]. Based on the EMS database, the number of stroke related EMS contacts prior or post seven days of stroke within the respected timeframe were 13,941. 3399 duplicate EMS contacts for the same stroke patient and 1493 stroke patients without contact to the EMS were excluded. Finally, this resulted in the inclusion of 9049 stroke related EMS contacts. Baseline characteristics of the stroke patients included in this research, such as stroke type, sex, age, year, time of day, weekday, stroke relevant criteria, EMS response, EMS number, received treatment, and time-to-call can be found in Table 1.

Outcomes

The results of the Chi-Square and Fisher's Exact test showed a correlation to minimum one of the four groups of stroke recognition within all the considered variables, based on the determined level of significance $p < 0.05$ (Table 2). The standardised residuals show the direction of correlation and were generated within the pairwise comparison and interpreted in relation to the critical z-value, calculated based on the adjusted significance level [37, 51].

A positive correlation between ischaemic stroke and "stroke relevant criteria but no "A" response" and between haemorrhagic stroke and "stroke nonrelevant criteria" have been determined (Fig. 2A). Within the subgroup
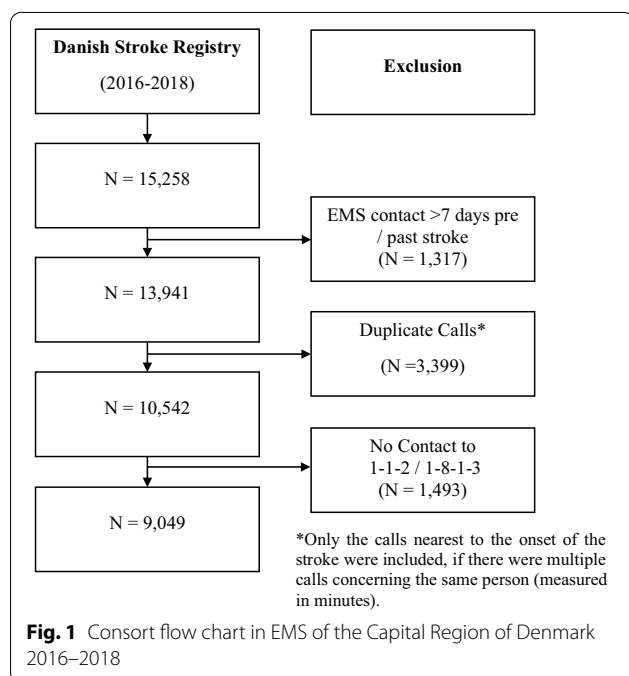
Table 1 Baseline characteristics stroke patients in EMS of the Capital Region of Denmark 2016–2018

Characteristic	N	%
Stroke patients	9049	100.00
Type		
Ischaemic stroke	8215	90.78
Haemorrhagic stroke	834	9.22
Sex		
Male	4827	53.34
Female	4222	46.66
Age—mean (Min.|Q1|Median|Q3|Max.)	71.28 (18|63|73|81|104)	
EMS Number		
1-1-2	5180	57.24
1813-Medical Helpline	3869	42.76
Stroke criteria		
Stroke relevant criteria	5973	66.01
Stroke nonrelevant criteria	1861	20.57
Missing criteria	1215	13.43
Response [53]		
A (life threatening)	6128	67.72
B (urgent, no immediate risk to life)	1028	11.36
C (non-urgent/planned transport)	24	0.27
D (no treatment/medical care required)	4	0.04
F (referral/advice)	45	0.50
Non-Urgent Response*	1820	20.11
Year		
2016	2807	31.02
2017	3021	33.38
2018	3221	35.59
Weekday (Monday–Friday)	5854	64.69
Weekend (Saturday–Sunday)	3195	35.31
Incident time of day		
1–6	677	7.48
7–12	3542	39.14
13–18	3096	34.21
19–0	1734	19.16
Treatment†	1652	18.26
Thrombolysis	1449	16.01
Reperfusion	217	2.40
Endovascular	47	0.52
Surgical	43	0.48
Thrombectomy	239	2.64
Time-to-call		
≤ 4.5 h	6013	66.45
≤ 24 h	8095	89.46

*A patient with a "Non-Urgent Response" typically received a referral to a physician or was advised to drive to the hospital

† A patient who received multiple treatments was considered individually for the individual treatments but cumulated for the total number of treatments


```
Danish Stroke Registry          Exclusion
  (2016-2018)

  N = 15,258

                        EMS contact >7 days pre
                        / past stroke
                        (N = 1,317)
  N = 13,941

                        Duplicate Calls*
                        (N = 3,399)
  N = 10,542

                        No Contact to
                        1-1-2 / 1-8-1-3
                        (N = 1,493)
  N = 9,049
```

*Only the calls nearest to the onset of the stroke were included, if there were multiple calls concerning the same person (measured in minutes).

Fig. 1 Consort flow chart in EMS of the Capital Region of Denmark 2016–2018

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 6 of 17

Table 2 Outcome Chi-Square test

Characteristic	Stroke detected			Stroke not detected					
	Stroke relevant criteria + "A" response			Stroke relevant criteria + no "A" response			Stroke nonrelevant criteria		
	N	%	Residuals	N	%	Residuals	N	%	Residuals
Total	4773	52.75%	–	1200	13.26%	–	1861	20.57%	–
	3613	*60.09%*		*588*	*9.78%*		*1134*	*18.86%*	
Type									
Ischaemic	4345	48.02%	0.8664729	1117	12.34%	2.9572745	1629	18.00%	− 5.4382222
	3288	*54.68%*	*1.730522*	*543*	*9.03%*	***1.631358***	*981*	*16.31%*	*− 5.045376*
Haemorrhagic	428	4.73%	− 0.8664728	83	0.92%	− 2.9572745	232	2.56%	5.4382222
	325	*5.40%*	*− 1.730522*	*45*	*0.75%*	*− **1.631358***	*153*	*2.54%*	*5.045376*
Sex									
Male	2663	29.43%	4.9358538	628	6.94%	− 0.7527075	950	10.50%	− 2.2267499
	2008	*33.39%*	*3.834136*	*294*	*4.89%*	*− 1.826033*	*586*	*9.75%*	*− 1.418067*
Female	2110	23.32%	− 4.9358538	572	6.32%	0.7527075	911	10.07%	2.2267499
	1605	*26.68%*	*− 3.834136*	*294*	*4.89%*	*1.826033*	*548*	*9.11%*	*1.418067*
EMS number									
1-1-2	3566	39.41%	35.486025	427	4.72%	− 16.285076	1106	12.22%	2.139238
	2761	*45.92%*	*23.923099*	*279*	*4.64%*	*− 9.024334*	*765*	*12.72%*	*2.406850*
1813-Medical Helpline	1207	13.33%	− 35.486025	773	8.54%	16.285076	755	8.34%	− 2.139238
	852	*12.17%*	*− 23.923099*	*309*	*5.14%*	*9.024334*	*369*	*6.14%*	*− 2.406850*
Year									
2016	1401	14.48%	− 3.6227039	451	4.98%	5.2774675	615	6.80%	2.1207352
	1062	*17.66%*	*− **2.4569773***	*232*	*3.86%*	*4.9156658*	*369*	*6.14%*	*1.5869673*
2017	1605	17.74%	0.5151980	401	4.43%	0.0250515	627	6.93%	0.3147488
	1208	*20.09%*	*0.9909348*	*185*	*3.08%*	*− 0.8053347*	*369*	*6.14%*	*− 0.3226088*
2018	1767	19.53%	2.9925059	348	3.85%	− 5.1233044	619	6.84%	− 2.3588767
	1343	*22.33%*	***1.3842896***	*171*	*2.84%*	*− 3.9190487*	*396*	*6.59%*	*− 1.2040795*
Weekday									
Week	3278	36.22%	8.381580	719	7.95%	− 3.716612	1275	14.09%	3.868071
	2506	*41.68%*	*4.336712*	*384*	*6.39%*	*− **1.041430***	*805*	*13.39%*	*2.999783*
Weekend	1495	16.52%	− 8.381580	481	5.32%	3.716612	586	6.48%	− 3.868071
	1107	*18.41%*	*− 4.336712*	*204*	*3.39%*	***1.041430***	*329*	*5.48%*	*− 2.999783*
Treatment*	1152	12.73%	15.023994	94	1.04%	− 10.276163	293	3.24%	− 3.282012
	1091	*18.14%*	*9.230893*	*82*	*1.36%*	*− 6.988008*	*269*	*4.47%*	*− **1.895487***
Thrombolysis	1041	11.51%	15.887928	75	0.83%	− 9.901901	222	2.45%	− 5.389991
	1016	*16.89%*	*10.410536*	*73*	*1.21%*	*− 6.665477*	*213*	*3.54%*	*− 4.144533*
Reperfusion	168	1.86%	7.3690592	8	0.09%	− 4.209392	39	0.43%	− 0.9567605
	146	*2.43%*	*4.7924906*	*8*	*0.13%*	*− **2.6258701***	*34*	*0.57%*	*− 0.3453236*
Endovascular	14	0.15%	− 3.1609477	5	0.06%	− 0.5315617	25	0.28%	5.5483350
	13	*0.22%*	*− 3.108757*	*3*	*0.05%*	*− 0.343197*	*19*	*0.32%*	*5.068036*
Surgical	14	0.15%	− 2.6579592	4	0.04%	− 0.7672481	19	0.21%	3.8412820
	8	*0.13%*	*− **3.9074732***	*2*	*0.03%*	*− 0.6252947*	*17*	*0.28%*	*5.1342675*
Thrombectomy	178	1.97%	6.8198305	13	0.14%	− 3.6134411	43	0.48%	− 0.9978678
	158	*2.63%*	*4.6608318*	*10*	*0.17%*	*− **2.4740986***	*37*	*0.62%*	*− 0.4347634*

Characteristic	Missing criteria			Chi-Square test			Cramer's V	Pairwise comparison	
	N	%	Residuals	X-Squared	Df	*p*-value	Estimate [Low. CI.; Upr. CI]	Adj. sig. level	Critical *z*-value
Total	1215	13.43%	–	–	–	–	–	–	–
	678	*11.28%*							
Type									
Ischaemic	1129	12.48%	2.2363844	35.762	3	= 8.406e−08	0.0629 [0.0401; 0.0819]	0.00625	− 2.734269
	628	*10.44%*	*2.028612*	*27.903*		*= 3.807e−06*	*0.0681 [0.0397; 0.0913]*		
Haemorrhagic	91	1.01%	− 2.2363844						
	50	*0.83%*	*− 2.028612*						

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 7 of 17

Table 2 (continued)

Characteristic	Missing criteria			Chi-Square test			Cramer's V	Pairwise comparison	
	N	%	Residuals	X-Squared	Df	p-value	Estimate [Low. CI.; Upr. CI]	Adj. sig. level	Critical z-value
Sex									
Male	586 *333*	6.48% *5.54%*	− 3.8390984 *− **2.467792***	28.702 *15.911*	3	= 2.587e−06 *= 0.01183*	0.0563 [0.0332; 0.0752] *0.0514 [0.0212; 0.0740]*	0.00625	− 2.734269
Female	629 *345*	6.95% *5.74%*	3.8390984 ***2.467792***						
EMS number									
1-1-2	81 *66*	0.90% *1.10%*	− 38.299051 *− 31.542421*	2098.6 *1189.3*	3	< 2.2e−16 *< 2.2e−16*	0.4816 [0.4607; 0.5020] *0.4447 [0.4191; 0.4697]*	0.00625	− 2.734269
1813-Medical Helpline	1134 *612*	12.53% *10.18%*	38.299051 *31.542421*						
Year									
2016	340 *176*	3.76% *2.93%*	− 2.4591338 *− 2.7748033*	55.028 *38.962*	6	= 4.575e−10 *7.281e−07*	0.0551 [0.0375; 0.0674] *0.0569 [0.0344; 0.0715]*	0.0416667	− 2.86526
2017	388 *219*	4.29% *3.64%*	− 1.1524637 *− 0.3789986*						
2018	487 *283*	5.38% *4.71%*	3.5109012 *3.0261982*						
Weekday									
Week	582 *347*	6.43% *5.77%*	− 13.161698 *− 9.446623*	207.03 *94.963*	3	< 2.2e−16 *< 2.2e−16*	0.1513 [0.1299; 0.1712] *0.1257 [0.0989; 0.1497]*	0.00625	− 2.734269
Weekend	633 *331*	6.99% *5.50%*	13.161698 *9.446623*						
Treatment*	123 *118*	1.36% *1.96%*	− 7.886774 *− 5.385627*	260.66 *106.72*	3	< 2.23−16 *< 2.2e−16*	0.1697 [0.1484; 0.1897] *0.1332 [0.1065; 0.1573]*	0.00625	− 2.734269
Thrombolysis	111 *110*	1.23% *1.83%*	− 7.025162 *− 4.733407*	270.13 *117.16*	3	< 2.2e−16 *< 2.2e−16*	0.1728 [0.1515; 0.1928] ***0.1396 [0.1130; 0.1638]***	0.00625	− 2.734269
Reperfusion	2 *2*	0.02% *0.03%*	− 5.4690676 *− 4.5272438*	67.651 *33.67*	3	< 2.2e−16† *= 1.589e−09†*	0.0865 [0.0644; 0.1059] *0.0748 [0.0467; 0.0982]*	0.00625	− 2.734269
Endovascular	3 *2*	0.03% *0.03%*	− 1.4200864 *− 1.132402*	31.165 *25.942*	3	= 1.056e−05† *= 8.998e−05†*	0.0587 [0.0357; 0.0777] *0.0657 [0.0370; 0.0888]*	0.00625	− 2.734269
Surgical	6 *4*	0.07% *0.07%*	0.1015225 *0.2872599*	15.579 *27.909*	3	= 0.003068† *= 2.472e−05†*	0.0415 [0.0168; 0.0599] *0.0681 [0.0397; 0.0913]*	0.00625	− 2.734269
Thrombectomy	5 *4*	0.06% *0.07%*	− 5.2089195 *− 4.3553056*	57.584 *31.176*	3	= 1.929e−12 ***= 2.142e−08†***	0.0798 [0.0575; 0.0991] *0.0720 [0.0438; 0.0953]*	0.00625	− 2.734269

"Italic" numbers indicate the subgroup analysis of stroke patients with a time-to-call < 4.5 h. "Bold" numbers indicate the difference to the initial analysis
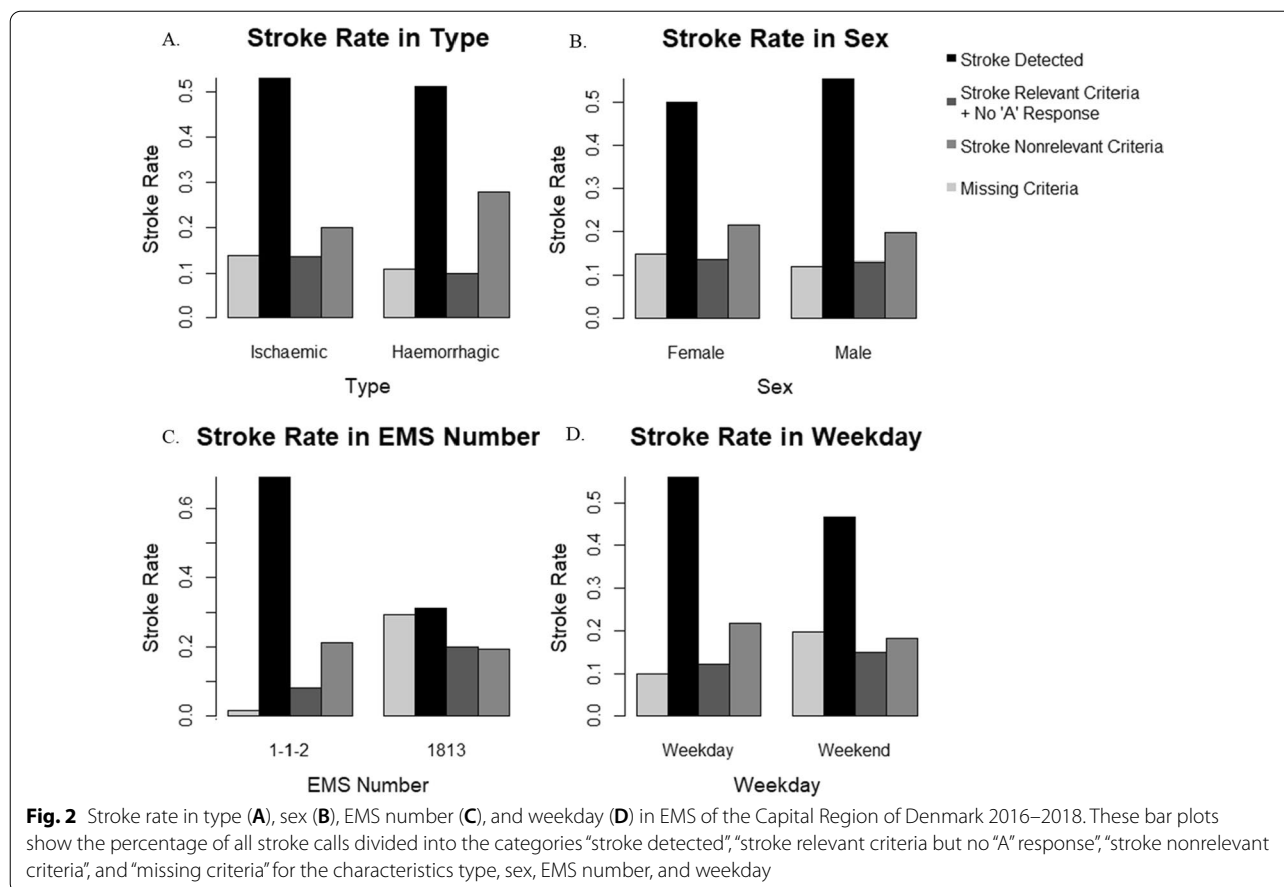
*A patient that received multiple treatments was considered individually for the individual treatments, but cumulated for the total number of treatments

† Results from the Fisher's Exact test, due to the small number of < 5 within the group

analysis no positive correlation between ischaemic stroke and "stroke relevant criteria but no "A" response" was observed. However, the Cramer's V indicates, that based on the degree of freedom, the identified correlation is a weak association effect [38, 52].

Furthermore, a positive correlation exists between male and "stroke detection". Contrarily, a negative association is recorded between female and "stroke detection" (Fig. 2B). However, based on the Cramer's V, the strength of the correlation is very weak for the variable sex in relation to the degree of freedom [38, 52].

A change in correlation throughout the years 2016–2018 was reported. While the "stroke detection" in 2016 has a negative association, 2018 shows a positive association. Conversely, 2016 indicates a positive correlation to "stroke relevant criteria but no "A" response", while a negative correlation is identified in 2018, The correlations reported for the category "stroke detection" could however not be determined within the subgroup analysis. The association reported is considered weak, based on the results of the Cramer's V and under consideration of the respected degree of freedom [38, 52].

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 8 of 17

Fig. 2 Stroke rate in type (**A**), sex (**B**), EMS number (**C**), and weekday (**D**) in EMS of the Capital Region of Denmark 2016–2018. These bar plots show the percentage of all stroke calls divided into the categories "stroke detected", "stroke relevant criteria but no "A" response", "stroke nonrelevant criteria", and "missing criteria" for the characteristics type, sex, EMS number, and weekday

Furthermore, the weekdays (Monday–Friday) are in positive correlation with "stroke detection" and "stroke nonrelevant criteria", while the weekend (Saturday–Sunday) is in positive correlation with "stroke relevant criteria but no "A" response" and "missing criteria" (Fig. 2D). Conversely, in the subgroup analysis no correlation between "stroke relevant criteria but no "A" response" and weekday or weekend could be detected. The reported association is weak based on the degree of freedom [38, 52]. On the weekend 53.11% of all EMS stroke contacts were through 1813, compared to 37.1% within the week.

For the overall treatment, the analysis identified a positive association with regards to "stroke detection". Additionally, for thrombolysis, a positive correlation with "stroke detection" was determined (Fig. 4A). The Cramer's V indicates that, based on the degree of freedom, the strength of the identified correlation is weak for the overall treatment and moderate for thrombolysis [38, 52]. Within the subgroup analysis similar results could be observed, however for thrombolysis only a weak strength of association was identified.

Considering the time-to-call, 75.7% of all "stroke detected" calls were within 4.5 h after stroke onset.

Comparably within the category "stroke nonrelevant criteria" 60.93%, "stroke relevant criteria but no "A" response" 49% and "missing criteria" 55.8% of the calls were within 4.5 h after stroke onset (Table 3).

The Kruskal–Wallis test indicates a statistically significant difference in stroke detection with regards to age (Table 4). While the Wilcoxon–Mann–Whitney post-hoc test with a Bonferroni-Holm correction additionally determines that a statistically significant difference in age between "stroke detected" and "stroke relevant criteria but no "A" response" as well as "missing criteria" and between "stroke relevant criteria but no "A" response" and "stroke nonrelevant criteria" as well as "missing criteria" exists. When considering the mean age of "missing criteria" (72.53), "stroke detection" (71.4), "stroke nonrelevant criteria" (71.08), and "stroke relevant criteria but no "A" response" (69.87), the latter group is statistically significantly younger than the three previously named groups. Additionally, the stroke patients with "missing criteria" are statistically significantly older than the stroke patients detected by the EMD. Comparably, within the subgroup analysis only a statistically significant difference in age was determined between "missing criteria"

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 9 of 17

Table 3 Time-to-call in EMS of the Capital Region of Denmark 2016–2018

Category	Time-to-call		
	Within 4.5 h (%)	Within 24 h (%)	Mean (Min.; Q1; Median; Q3; Max.)
Stroke detected	75.70	94.93	5.65 (0.00045; 0.312; 1.233; 4.296; 165.861)
Stroke relevant criteria + no "A" response	49.00	80.42	16.22 (0.0053; 1.019; 4.721; 18.914; 167.646)
Stroke non relevant criteria	60.93	85.06	14.07 (0.0016; 0.665; 2.451; 11.289; 167.377)
Missing criteria	55.80	83.62	14.16 (0.011; 0.995; 3.391; 12.514; 167.456)

Time of onset is based on the patients recall of symptom onset and thus might be no exact time of onset

Table 4 Outcome variance-analysis

Characteristics	Stroke detected		Stroke not detected				Missing criteria	
	Stroke relevant criteria + "A" response		Stroke relevant criteria + no "A" response		Stroke nonrelevant criteria			
	N	Mean [Min.|Q1|Median|Q3|Max.]	N	Mean [Min.|Q1|Median|Q3|Max.]	N	Mean [Min.|Q1|Median|Q3|Max.]	N	Mean [Min.|Q1|Median|Q3|Max.]
Age	4773 *3613*	71.4 [18|63|73|81|104] *71.06 [19|62|72|81|104]*	1200 *588*	69.87 [21|61|71|80|101] *69.94 [21|61|71|80|101]*	1816 *1134*	71.08 [21|63|73|82|101] *69.96 [21|60|72|81|101]*	1215 *678*	72.53 [19|63|74|83|102] *71.90 [19|62|73|83|102]*
Time of day	4773 *3613*	12.82 [0|9|12|17|23] *12.93 [0|9|13|17|23]*	1200 *588*	13.55 [0|10|14|17|23] *13.36 [0|10|13|17|23]*	1816 *1134*	12.97 [0|9|13|17|23] *12.60 [0|9|12|16|23]*	1215 *678*	13.83 [0|10|15|18|23] *13.56 [0|10|15|18|23]*

Characteristics	Kruskal–Wallis test			Wilcoxon–Mann–Whitney (+ Bonferroni–Holm Correction)			
	Kruskal–Wallis-X^2	Df	p-value				
Age	29.226 *13.944*	3	2.00e−06 *0.002983*		Stroke Detected	Stroke relevant criteria + no "A" response	Stroke non-relevant criteria
				Stroke relevant criteria + no "A" response	0.0017 **0.1368**	–	–
				Stroke non-relevant criteria	1.000 *0.8197*	0.0055 *1.0000*	–
				Missing criteria	0.0068 **0.1551**	8.7e−07 *0.0052*	0.0543 **0.0273**
Time of day	63.455 *24.8*	3	1.073e−13 *1.7e−05*		Stroke Detected	Stroke relevant criteria + no "A" response	Stroke non-relevant criteria
				Stroke relevant criteria + no "A" response	2.8e−05 **0.39660**	–	–
				Stroke non-relevant criteria	0.934 *0.68081*	0.019 *0.04429*	–
				Missing criteria	7.1e−12 *0.00037*	0.055 *0.62118*	3.5e−07 *2.5e−05*

"Italic" numbers indicate the subgroup analysis of stroke patients with a time-to-call < 4.5 h. "Bold" numbers indicate the difference to the initial analysis

and "stroke relevant criteria but no "A" response" as well as "stroke nonrelevant criteria" (Table 4). However, the mean age decreases in following direction "missing criteria" (71.9), "stroke detection" (71.06), "stroke non relevant criteria" (69.96), and "stroke relevant criteria but no "A" response" (69.94).
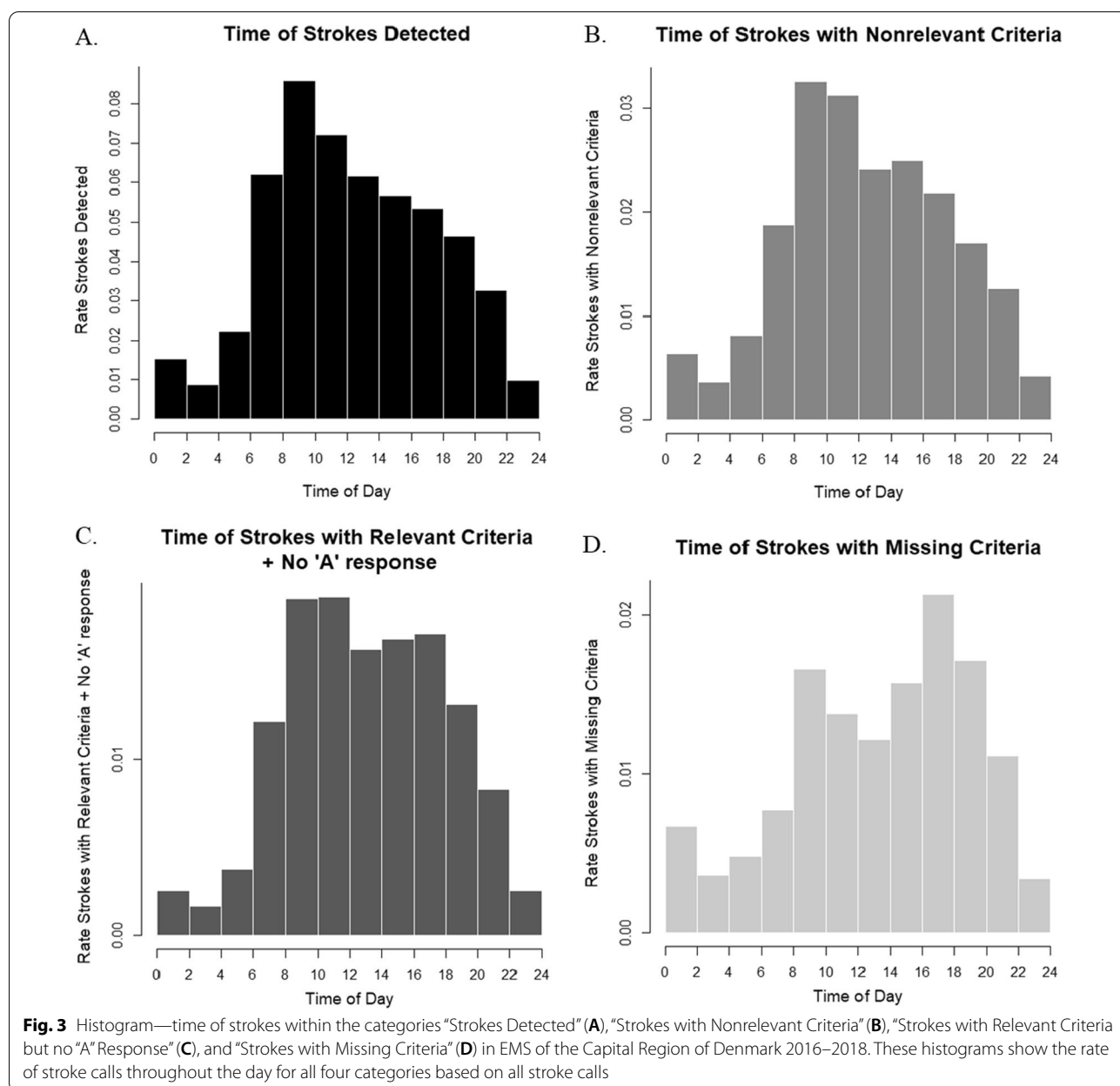
For the time of day, a statistically significant difference in stroke detection was identified (Table 4). A statistically significant difference in the time of the call between "stroke relevant criteria but no "A" response" and "stroke detection" as well as "stroke nonrelevant criteria" has been determined through the Wilcoxon–Mann–Whitney post-hoc test with a Bonferroni-Holm correction. Additionally, a statistically significant difference in time of the EMS call between "missing criteria", "stroke detected", and "stroke nonrelevant criteria" has been detected.

Comparably, no statistically significant difference in time of day between "stroke detection" and "stroke relevant criteria but no "A" response" could be seen in the subgroup analysis (Table 4). It can be observed that the groups "stroke detected" and "stroke nonrelevant criteria" have their peak before 10 a.m. and then steadily decrease (Fig. 3A + B). Comparatively, the group "stroke nonrelevant criteria, but no "A" response" decreases only slightly after 12 p.m. but stays on a relatively high level until 6 p.m. after which the number of strokes with "stroke relevant criteria but no "A" response" decrease (Fig. 3C). When considering the histogram of "time of strokes with

missing criteria", two peaks can be observed, one in the morning and one in the afternoon (Fig. 3D).

Based on the results of the statistical analyses, calculations on how an ASR in the EMS could have potentially impacted the stroke detection in the years 2016–2018 have been performed. This is under the condition that an ASR would improve stroke detection similarly as has been shown for the detection of OHCA by the research of Blomberg et al. For this calculation the strokes with "missing criteria" will be treated as if they are not influenced by the ASR. Presumably, the stroke detection rate in the EMS Copenhagen could

Fig. 3 Histogram—time of strokes within the categories "Strokes Detected" (**A**), "Strokes with Nonrelevant Criteria" (**B**), "Strokes with Relevant Criteria but no "A" Response" (**C**), and "Strokes with Missing Criteria" (**D**) in EMS of the Capital Region of Denmark 2016–2018. These histograms show the rate of stroke calls throughout the day for all four categories based on all stroke calls
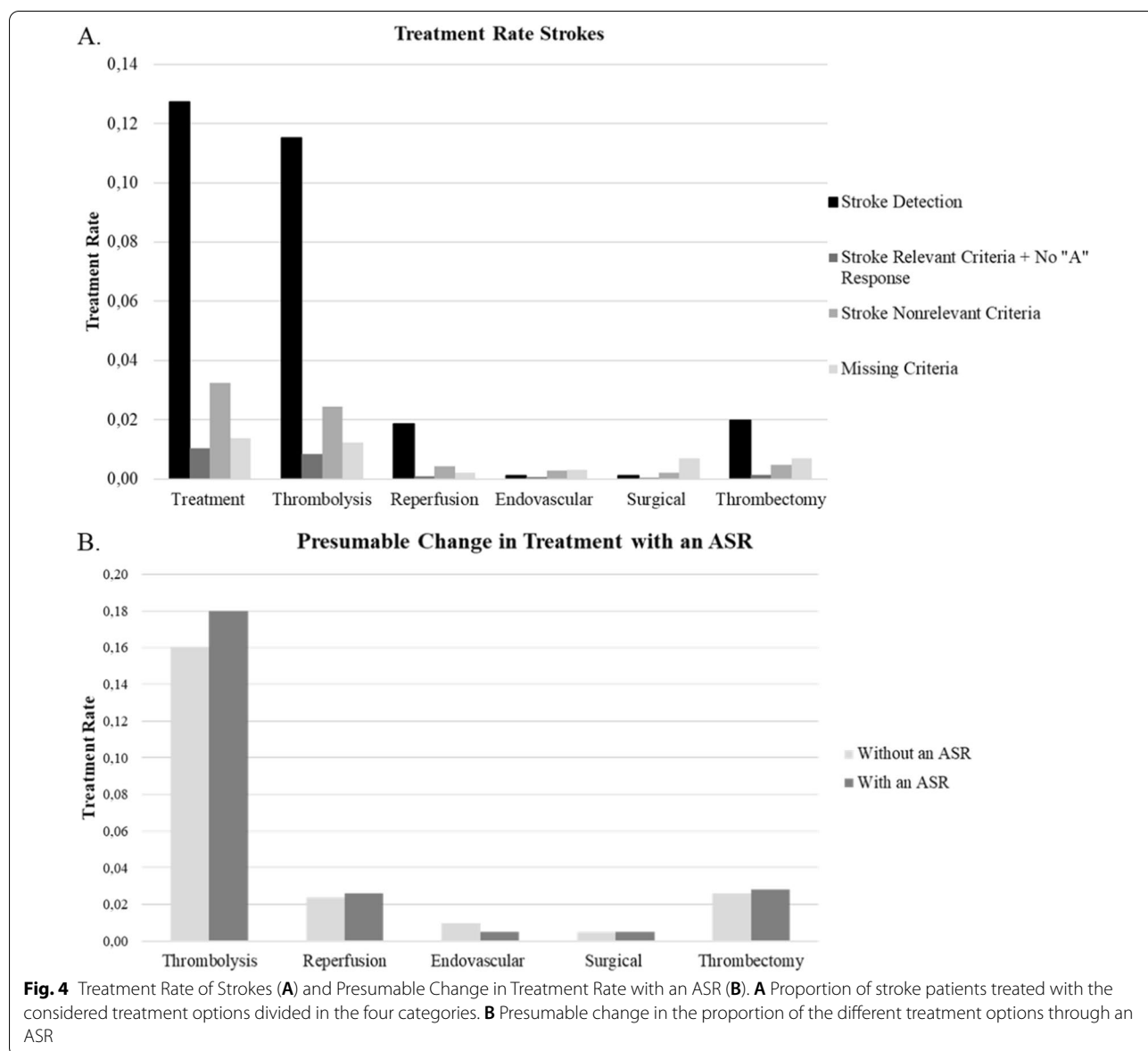
rise to 61.19% [24]. Therefore, a supporting ASR tool could assumingly have increased the amount of strokes detected by 764 (16%) from 4773 to 5537 ($n = 9049$) in the years 2016–2018. Additionally, assuming that the EMS contact was within the appropriate time-to-treatment, the thrombolysis rate among stroke patients could increase from 16 to 18%. Comparatively, the reperfusion rate could increase from 2.4 to 2.6%, the thrombectomy rate from 2.6 to 2.8%, and the surgical treatment rate from 0.48 to 0.49%. However, based on the data analysis and under the named conditions, the endovascular treatment rate would decrease from 0.52 to 0.49% (Fig. 4B). Under consideration of the time-to-call (4.5 h), the subgroup analysis indicated that

the stroke detection rate within this subpopulation the stroke detection rate with an ASR could increase to 69.7% and thus increase the thrombolysis rate within the stroke patients calling within time-to-treatment of thrombolysis by 5%. Contrarily, the amount of endovascular treatment would have presumably decreased by 14%, while surgical treatment would have decreased by 16% if an ASR would have been used for stroke recognition within the years 2016–2018.

Discussion

The analysis suggests that a significant number of stroke calls are not detected as strokes (33.83%) within the 1-1-2 and 1813 emergency medical contact points. Considering

Fig. 4 Treatment Rate of Strokes (**A**) and Presumable Change in Treatment Rate with an ASR (**B**). **A** Proportion of stroke patients treated with the considered treatment options divided in the four categories. **B** Presumable change in the proportion of the different treatment options through an ASR

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 12 of 17

the positive effects stroke recognition at the EMS takes on the stroke related outcome, the improvement of stroke detection at the EMS is crucial [14–16, 18, 19]. This research suggests the usage of an ASR, based on the model of CORTI AI for OHCA, to increase stroke recognition at the EMS from 52.75 to 61.19%. This increased detection rate through an ASR might decrease the number of multiple EMS calls for stroke patients, due to an earlier detection of the stroke and an accurate response within the first call. However, further research to determine the reason for multiple EMS calls would be necessary. Based on the condition that the stroke detection rate would increase by the same amount as the OHCA detection rate increased through CORTI AI, the rate of stroke patients treated with thrombolysis will rise by 5% within the group of stroke patients calling within time-to-treatment for thrombolysis [24, 54]. Additionally, the ASR might lead to an increase in thrombectomy of 8%, reperfusion of 8%, and surgical treatment of 2%. However, these increasing rates for thrombectomy, reperfusion and surgical treatment are to be viewed with caution. While the strength of the identified correlation between stroke recognition and treatment is moderate for thrombolysis, it is weak or very weak for the other treatment options. Additionally, the calculations have been made based on theoretical background and under the condition that the patients call the EMS within the treatment specific time-to-treatment. While 66.45% of all EMS contacts are within time-to-treatment of thrombolysis (4.5 h), 89.46% are within 24 h after stroke onset. Mosley et al. [55] confirm these findings by reporting that less than 50% of the stroke related calls were within 60 min after stroke onset [55]. In the future, a prospective study on the change in treatment through an increase in stroke detection would be interesting.

The described results suggest, that with an increased stroke detection at the EMS, the rate of stroke patients receiving endovascular treatment might decrease. The subgroup analysis of stroke patients with a time-to-call < 4.5 h showed a decrease in endovascular and surgical treatment. Nonetheless, this must be considered with caution, since endovascular treatment is regarded as an alternative for unsuccessful thrombolysis or patients not eligible for thrombolysis and surgical treatment is only carried out occasionally and under selected circumstances [56, 57]. While the time-to-treatment for thrombolysis is 4.5 h, endovascular treatment can be received within six to eight hours after stroke onset [56, 58]. Thus, patients who are not eligible for thrombolysis due to the closure of the window of time-to-treatment might receive endovascular treatment. In contrast, other reasons influence the choice of endovascular treatment [56]. Additionally, due to the low number of endovascular ($n=47$)

and surgical ($n=43$) treatment, the results of these categories cannot be emphasized, but further research with a larger number of stroke patients treated with endovascular and surgical treatment would be necessary to draw conclusions [59].

Moreover, several additional factors, like stroke detection by the caller, recognition by the paramedic on scene, pre-conditions, and personal characteristics impact the stroke patients eligibility for treatment [14, 60]. Jones et al. [61] determined that symptoms like speech problems as well as posterior circulation symptoms were least likely to be recognised as stroke related. Further research on the beforementioned connections as well as on mortality and on the score of the modified ranking scale, which defines a patients clinically discrete disability caused by a stroke on a scale of seven levels, would be helpful in order to draw precise and grounded conclusions on the effect of EMS stroke detection on patient outcome [62, 63]. Nonetheless, stroke detection by the EMS might impact the treatment, specifically thrombolysis. The relevance of an ASR for stroke detection at the EMS is underlined by the substantial amount (49% of "stroke relevant criteria but no "A" response" and 60.93% of "stroke nonrelevant criteria") of calls within time-to-treatment for thrombolysis in the categories "strokes not detected".

Based on the results of the analysis, it can be argued that the ASR could specifically impact the detection of those characteristics with a negative correlation to "stroke detected" or a positive correlation to one of the categories within "strokes not detected".

The analysis indicates that an improvement of stroke detection is particularly important for calls to the 1813 Medical Helpline, due to the observed negative correlation of stroke detection within 1813-calls. Thus, the ASR should be used for both access numbers 1813 and 1-1-2. The negative correlation may be influenced by non-recognition of atypical stroke symptoms by the caller, thus the 1813 instead of the 1-1-2 is called [64]. However, for validation further research is needed.

When training the ASR specific attention should be placed on haemorrhagic strokes, due to the positive correlation between haemorrhagic strokes and "stroke nonrelevant criteria" and the small representation of haemorrhagic strokes (9.22% of all strokes). Several authors argue, that it is particularly important to take into account underrepresented groups, e.g. haemorrhagic stroke patients ($n=834$) and patients with "stroke nonrelevant criteria" ($n=1215$), when training an ASR, in order to avoid a bias, that could possibly cause an erroneous stroke detection algorithm [65, 66]. An ASR could also positively influence the stroke detection rate of females, due to the negative correlation to stroke

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 13 of 17

detection. Lisabeth et al. [67] and Rathore et al. [68] support this finding by describing, that women reported a larger amount of non-traditional stroke symptoms.

In the data analysis a negative correlation between stroke detection and weekends was determined, hence the ASR for stroke detection could particularly improve the stroke detection on weekends. A possible explanation could be, that on weekends 53.11% of all stroke related EMS calls are to the 1813, while within the week only 37.1% are to the 1813. As argued before, 1813-calls might entail more atypical stroke symptoms not detected compared to the 1-1-2, resulting in a decline of the detection rate on the weekends [64].

Interestingly, stroke patients within the group "stroke relevant criteria but no "A" response" are significantly younger than the patients within the other groups. Considering the research by Singhal et al. [69], detection of stroke among younger patients, is challenging due to infrequency in comparison to stroke mimics and missing awareness among the general population as well as the EMDs. This might result in an EMS contact outside of the window of time-to-treatment or missing recognition of severity and thus in no "A" response. This is strengthened by "stroke relevant criteria but no "A" response" having the lowest proportion of calls (49%) within the time-to-treatment for thrombolysis. This reasoning is also supported by the subgroup analysis showing no statistically significant difference in age between the categories "stroke detection" and "strokes not detected".

The category "stroke relevant criteria but no "A" response", has a different distribution throughout the time of day, compared to strokes detected and strokes with non-relevant criteria. While the latter have a peak between 8 a.m. and 10 a.m. and thereafter steadily decrease, the beforenamed category is comparatively steady between 8 a.m. and 6 p.m. The peak of stroke relevant criteria in the morning might be due to the so called "wake-up stroke", for which EMDs have a high awareness, since one out of five strokes is a "wake-up stroke" [70]. Comparably, in the afternoon a greater diversity among emergency calls occurs, which might result in a higher difficulty to detect strokes [71]. For these calls an ASR supporting the EMD in the stroke detection would be useful to detect and send the correct response. Due to the subgroup analysis identifying no statistically significant difference in time of day between "stroke detection" and "strokes not detected" an ASR would be relevant for increasing stroke detection throughout the whole day. The steady number of calls with "stroke relevant criteria but no "A" response" might be caused by a delayed emergency call and despite the stroke detection by the dispatcher, but due to the closure of the window of time-to-treatment, no "A" response. This argument

is supported by the outcome of the subgroup analysis, showing no statistically significant difference in time of day between "stroke detection" and "stroke relevant criteria but no "A" response". Further research to conduct the reason for stroke relevant criteria but no "A" response is necessary. The "missing criteria" show two peaks throughout the day, between 8 a.m. and 10 a.m. as well as between 4 p.m. and 6 p.m. These peaks can be explained by the majority of "missing criteria" within 1813 calls, and the increased amount of 1813 calls between 8 a.m. and 10 a.m. on the weekends and between 4 p.m. and 6 p.m. during the week, due to its mission as out-of-hours general practitioner [13]. However, additional factors that might be influenced by an ASR, were not considered in this study.

Due to a significant increase of stroke detection throughout the years 2016–2018, as shown in our analysis, it might be argued that no further technical support might be necessary to improve stroke detection. However, since a significant decrease has only been seen within the group "stroke relevant criteria but no "A" response", but not within the group "stroke nonrelevant criteria", this argument can be discarded due to the ASR presumably impacting the recognition of strokes with currently "stroke nonrelevant criteria", by increasing the detection of stroke symptoms and thus indicating "stroke relevant criteria". Additionally, the subgroup analysis indicating no statistically significant increase in stroke detection throughout the years 2016–2018 supports the need of an ASR for improving stroke recognition by the EMDs. The increase in stroke detection seen for 2016–2018 might be influenced by the publication by Viereck et al. [15] in 2016 on the recognition of strokes through EMDs, after which small changes have been made in the algorithm of the 1-1-2. Another reason for the change in stroke detection throughout the years 2016–2018, could be the results of a research conducted at the University of Kentucky Stroke Center impacting the stroke recognition campaign, "FAST" (Face, Arm, Speech, Time) to "BE-FAST" (Balance, Eyes, Face, Arm, Speech, Time) in 2017, through including visual symptoms on stroke [72]. This revision might have led to an increasing sensibility for strokes within the population possibly resulting in a clearer expression of the symptoms to the EMS and an increasing sensibility of EMDs for stroke related symptoms [72].

The question arises, whether other options could increase stroke detection by EMS call-takers. Past research analysed the influence of educational training modules as well as stroke recognition scales and protocols, such as the "FAST"-Tool [17, 73–75]. However, Oostema et al. [73] reported, that the increase in stroke recognition after an educational intervention was

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 14 of 17

limited to three months and might increase the rate of false positive stroke detection due to a higher sensibility to symptoms related to stroke [73, 76]. Additionally, the systematic review by Oostema et al. [17] discovered, that the correct usage of the scales and protocols has not been analysed in the included studies, resulting in lacking security of the right usage. It is to be mentioned, that educational programmes for EMDs might increase the rate of false positive stroke detection due to a higher sensibility to symptoms related to stroke [76].

Like the correct use and acceptance of scales and protocols, the acceptance and adoption of the ASR into the EMS call by the EMD, is relevant for its effect on stroke detection. Blomberg et al. [77] reported a lack of compliance with the suggestions of CORTI AI by the EMDs, which resulted in no increase of OHCA detection within the EMS Copenhagen. Considering the results of educational interventions, the introduction of an ASR for strokes at the EMS could be accompanied by, for example educational interventions addressing challenges in the uptake of the ASR, in order to ensure the effect of the ASR [73–75, 77]. The European Institute of Innovation and Technology (EIT) Health states that to improve the uptake and effect of AI in healthcare, investments in the education of healthcare workers to ensure digital literacy, the exchange of best practice in the field of AI in healthcare throughout the EU and improvement of collaboration is essential [78].

Despite the lack in compliance with the ASR and thus the limitation of the effect, no sole usage of an ASR should be aimed for, due to possible input and algorithm bias as well as the missing consideration of the emotional component [79, 80]. In summary, the combination of an ASR with a well-trained human professional can substantially increase the number of correctly detected strokes [24, 25].

Limitations
The definition of stroke detection as "stroke relevant criteria" and an "A" response, might not represent all the strokes detected within the EMS. Possibly, strokes were detected within the category "stroke relevant criteria but no "A" response", and still, due to the closure of the window of time-to-treatment no "A" response was sent. For those cases obviously, an ASR would not impact the stroke detection. Contrarily, strokes might have been detected within the category "missing criteria", but no criteria were indicated within the system, yet an "A" response had been sent as the correct stroke response. Likewise, possibly "stroke nonrelevant" criteria were indicated within the system, but the EMD recognised the stroke and sent an "A" response. Due to the definition of stroke response within the EMS Copenhagen, the proxy of "stroke relevant criteria" and "A" response was considered the most accurate to define stroke detection for this research.

Another limitation of the stroke related emergency calls is, that all EMS calls seven days prior and post stroke were included within this study, even if the emergency call was not related to the stroke of the patient but was due to another medical issue. However, research has shown that strokes typically impact the health of the patient significantly, through post-stroke and pre-stroke symptoms, thus the number EMS contacts of stroke patients not related to the stroke might be comparably small [36, 67]. The choice to include stroke calls seven days prior and post stroke could be affecting the response made by the medical dispatcher, depending on the time of symptom onset named within the call and thereby diminish the effect of the outcome. Unfortunately, the data on time of symptom onset is not documented and thus not available and must therefore be considered a blind spot within this research. Additionally, the subgroup and time-to-call analysis is limited due to the determination of stroke onset within the patient-doctor consultation based on the patients recall of time of symptom onset. Thus, the possibility of recall bias needs to be considered in the interpretation of the results [81, 82].

The internal validity, which is described as to which extent the study accurately measures the concept, might be also limited due to the assumption, that an ASR for stroke has the same effect on the increase of detection at the EMS, as CORTI AI on OHCA, since OHCA symptoms are more specific compared to stroke symptoms [33, 83–86]. Thus, the possibility of stroke mimics, which are defined as disorders showing stroke symptoms, such as for example brain tumours, metabolic disorders, or migraines, and are diagnosed as strokes are likelier than false positive OHCA [85]. This is supported by the research by Watkins et al. [84] detecting a specificity of 99.4% within OHCA, while according to Hatzitolios et al. [85] 5% and to Hosseininezhad and Sohrabnejad [86] 14.9% of all stroke-like symptoms are stroke mimics. However, since CORTI AI for OHCA is, to the researcher's knowledge, the only ASR within an EMS context, the presumable increase of 16% based on CORTI AI was chosen. Hence, this limitation must be considered when referring to the presumable increase in stroke detection, especially since Blomberg et al. [24] reported a decrease in specificity for OHCA detection with the ASR from 98.8 to 97.3% ($p < 0.001$). Under consideration of the beforenamed rate of stroke mimics, the decrease in specificity of stroke detection might be higher with an ASR compared to the decrease in specificity of OHCA. Further research on the topic of specificity in stroke detection through ASR should be performed to elaborately address this point and to discuss possible mitigation strategies.

Scholz *et al. Scand J Trauma Resusc Emerg Med*   (2022) 30:36

Page 15 of 17

The transferability to the population of Denmark would need further research, since the results conducted for the Capital Region of Denmark, with the specialty of the 1813, might not be transferable to the entirety of Denmark [87]. Additionally, the transferability to other countries might be limited, due to country specific EMS and population characteristics. Thus, the assessment of transferability on superordinate level using for example the PIET-T Model might be helpful [88]. Because of the mentioned limiting factors, the results of this study should be interpreted with caution and considered as directing and indicating further research fields.

Conclusion

An ASR can presumably improve the recognition of stroke. Based on the results of this research, an intervention to increase stroke recognition is important for the EMS Copenhagen, specifically among females, younger stroke patients, within the 1813-Medical Helpline, and on weekends. Under consideration of the beforenamed conditions and limitations, an ASR could have a positive effect on stroke detection, and thereafter on stroke treatment, specifically on thrombolysis.

Abbreviations
1813: 1813-Medical Helpline; AI: Artificial intelligence; ASR: Automatic speech recognition software; CI: Confidence interval; DALY: Disability-adjusted life year; EMD: Emergency Medical Dispatcher; EMS: Emergency Medical Service; OHCA: Out-of-Hospital cardiac arrest.

Author contributions
MLS analysed and interpreted the data and was a major contributor in writing the manuscript. SNFB performed the data collection and preparation and supported in the data interpretation. HCC and TK supported in the data analyses and interpretation. HCC, TK, JV, and SB provided critical revisions to the manuscript. All authors read and approved the final manuscript.

Availability of data and materials
The data that support the findings of this study are available from Emergency Medical Services, Capital Region of Denmark, Denmark but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Emergency Medical Services, Capital Region of Denmark, Denmark.

Declarations

Ethics approval and consent to participate
This research was approved by the ethics committee of Maastricht University. In compliance with the general data protection regulation, this study was registered at the Danish Data Protection Agency (PVH-2014-002). Additionally, the study was registered at the Danish Patient Safety Authority (R-21013122). Since no patients were involved in the study, no consent of individual patients was required. Only anonymized data was used for the analysis.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Emergency Medical Services, Capital Region of Denmark, Telegrafvej 5, 2750 Ballerup, Denmark. [2]Department of Health, Ethics and Society, Care and Public Health Research Institute (CAPHRI), Faculty of Health, Medicine and Life Sciences, Maastricht University, P.O. Box 616, 6200 MD Maastricht, Netherlands.

References
1. World Health Organization. Global Health Estimates: life expectancy and leading causes of death and disability [Internet]. 2020 [cited 2021 Jan 6]. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates.
2. World Health Organization. Global health estimates: leading causes of death [Internet]. 2020 [cited 2021 Jan 6]. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death.
3. World Health Organization. Global health estimates: leading causes of DALYs [Internet]. 2020 [cited 2021 Jan 6]. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys.
4. Saver JL. Time is brain—quantified. Stroke. 2006;37(1):263–6.
5. Meretoja A, Keshtkaran M, Saver JL, Tatlisumak T, Parsons MW, Kaste M, et al. Stroke thrombolysis: save a minute, save a day. Stroke. 2014;45(4):1053–8.
6. Saver JL, Levine SR. Alteplase for ischaemic stroke—much sooner is much better. Lancet. 2010;375(9727):1667–8.
7. Marler JR, Tilley BC, Lu M, Brott TG, Lyden PC, Grotta JC, et al. Early stroke treatment associated with better outcome. Neurology [Internet]. 2000;55:1649–55.
8. Ragoschke-Schumm A, Walter S, Haass A, Balucani C, Lesmeister M, Nasreldein A, et al. Translation of the "time is brain" concept into clinical practice: focus on prehospital stroke management. Int J Stroke [Internet]. 2014;9(3):333–40.
9. Bluhmki E, Albers GW, Hamilton SA, Kennedy P, Lees KR, Bluhmki E, et al. Time to treatment with intravenous alteplase and outcome in stroke: an updated pooled analysis of ECASS, ATLANTIS, NINDS, and EPITHET trials. Lancet. 2010;375:1695–703.
10. Berge E, Whiteley W, Audebert H, De Marchis GM, Fonseca AC, Padiglioni C, et al. European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke. Eur Stroke J. 2020;6:1–62.
11. Gupta S, Sharme DK, Gupta MK. Artificial intelligence in diagnosis and management of ischemic stroke. Biomed J Sci Tech Res. 2019;13(3):9964–7.
12. The Capital Region of Denmark. Emergency Medical Services [Internet]. 2020 [cited 2021 Jan 7]. https://www.regionh.dk/english/Healthcare-Services/Emergency-Medical-Services/Pages/default.aspx.
13. Lindskou TA, Mikkelsen S, Christensen EF, Hansen PA, Jørgensen G, Hendriksen OM, et al. The Danish prehospital emergency healthcare system and research possibilities. Scand J Trauma Resusc Emerg Med [Internet]. 2019;27(1):100–7.
14. Rudd AG, Bladin C, Carli P, De Silva DA, Field TS, Jauch EC, et al. Utstein recommendation for emergency stroke care. Int J Stroke. 2020;15(5):555–64.
15. Viereck S, Møller TP, Klingenberg Iversen H, Christensen H, Lippert F. Medical dispatchers recognise substantial amount of acute stroke during emergency calls. Scand J Trauma Resusc Emerg Med. 2016;24:89–95.

Scholz *et al. Scand J Trauma Resusc Emerg Med* (2022) 30:36

Page 16 of 17

16. Rajajee V, Saver J. Prehospital care of the acute stroke patient. Tech Vasc Interv Radiol. 2005;8(2):74–80.

17. Oostema JA, Carle T, Talia N, Reeves M. Dispatcher stroke recognition using a stroke screening tool: a systematic review. Cerebrovasc Dis. 2016;42:370–7.

18. Hsieh M-J, Chien K-L, Sun J-T, Tang S-C, Tsai L-K, Chiang W-C, et al. The effect and associated factors of dispatcher recognition of stroke: a retrospective observational study. J Formos Med Assoc. 2018;117:902–8.

19. Oostema JA, Chassee T, Reeves M. Prehospital emergency care emergency dispatcher stroke recognition: associations with downstream care. Prehospital Emerg Care. 2018;22:466–71.

20. Owens Johnson C, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol [Internet]. 2019;18:439–58.

21. Abboud ME, Band R, Jia J, Pajerowski W, David G, Guo M, et al. Prehospital emergency care recognition of stroke by EMS is associated with improvement in emergency department quality measures. Prehospital Emerg Care. 2016;20(6):729–36.

22. Adams HP, Gregory Del Zoppo C, Alberts MJ, Bhatt DL, Brass L, Furlan A, et al. Guidelines for the early management of adults with ischemic stroke A guideline from the American Heart Association/American Stroke Association Stroke Council, Clinical Cardiology Council, Cardiovascular Radiology and Intervention Council, and the Atheros. Circulation. 2007;116(18):e515.

23. Amtoft AC, Danielsen AK, Hornnes N, Kruuse C. A qualitative inquiry into patient reported factors that influence time from stroke symptom onset to hospitalization. J Neurosci Nurs. 2021;53(1):5–10.

24. Blomberg SN, Folke F, Kjaer Ersbøll A, Christensen HC, Torp-Pedersen C, Sayre MR, et al. Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. Resuscitation [Internet]. 2019;138:322–9. https://doi.org/10.1016/j.resuscitation.2019.01.015.

25. Cleve A, Devillers D, Palladini M, Paris J, Michael R, Faure E, et al. Detecting Out-of-Hospital cardiac arrest using artificial intelligence. Brussels: European Emergency Number Association; 2020.

26. European Commission. White paper on artificial intelligence-A European approach to excellence and trust White Paper on artificial intelligence A European approach to excellence and trust. COM(2020) 65. Brussels; 2020.

27. European Commission. A European strategy for data. COM(2020) 66. Brussels; 2020.

28. Baxter PE, Jack SM. Qualitative case study methodology: study design and implementation for novice researchers. Qual Rep [Internet]. 2008;13(4):544–59.

29. Crowe S, Cresswell K, Robertson A, Huby G, Avery A, Sheikh A. The case study approach. BMC Med Res Methodol [Internet]. 2011;11:100–9.

30. Statistics Denmark. Area 1. Janurary by region and time. StatBank Denmark. 2021.

31. Statistics Denmark. Population at the first day of the quarter by age, sex, region and time. StatBank Denmark. 2020.

32. Das S, Mitra K, Mandal M. Sample size calculation: basic principles. Indian J Anaesth [Internet]. 2016;60(9):652–6.

33. Frambach JM, van der Vleuten CPM, Durning SJ. AM last page. Quality criteria in qualitative and quantitative research. Acad Med. 2013;88(4):552.

34. Hess D. Retrospective studies and chart reviews. Respir Care [Internet]. 2004;49(10):1171–4.

35. Johnsen S, Ingeman A, Holmager Hunborg H, Zielke Schaarup S, Gyllenborg J. The Danish stroke registry. Clin Epidemiol. 2016;8:697–702.

36. Rothwell PM, Warlow CP. Timing of TIAs preceding stroke: time window for prevention is very short. Neurology. 2005;64(5):817–20.

37. Agresti A. An introduction to categorical data analysis. New Jersey: Wiley; 2001.

38. Kim H-Y. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod. 2017;42(2):155.

39. Jung SH. Stratified Fisher's exact test and its sample size calculation. Biom J. 2014;56(1):129–40.

40. Cochran WG. The χ2 test of goodness of fit. Ann Math Stat. 1952;23(3):315–45.

41. Mchugh ML. The Chi-square test of independence Lessons in biostatistics. Biochem Medica. 2013;23(2):143–52.

42. Cangur S, Ankarali H. Examining the probabilities of type i error for unadjusted all pairwise comparisons and Bonferroni adjustment approaches in hypothesis testing for proportions. Int J Stat Med Res. 2014;3(4):404–11.

43. Kim H-Y. Statistical notes for clinical researchers: post-hoc multiple comparisons. Restor Dent Endod. 2015;40(2):172–6.

44. Stoline MR. The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs. Am Stat. 1981;35(3):134–41.

45. Harchavanich D. A comparison of type I error and power of Bartlett's test, Levene's test and O'Brien's test for homogneity of variance tests. Southeast Asian J Sci. 2014;3(2):181–94.

46. Glass GV. Testing homogeneity of variances. Am Educ Res J. 1966;3(3):187–90.

47. Fay MP, Proschan MA. Wilcoxon–Mann–Whitney or T-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. Stat Surv. 2010;4:1–39.

48. Myers JL, Well AD, Lorch RF Jr. Research design and statistical analysis. 3rd ed. New York: Routledge; 2010.

49. Gorris LGM, Yoe C. Risk analysis: risk assessment: principles, methods, and applications. In: Motarjemi Y, editor. Encyclopedia of food safety. Amsterdam: Elsevier; 2014. p. 65–72.

50. Pichery C. Sensitivity analysis. In: Wexler P, editor. Encyclopedia of toxicology. 3rd ed. Amsterdam: Elsevier; 2014. p. 236–7.

51. Haberman SJ. The analysis of residuals in cross-classified tables. Biometrics. 1973;29(1):205–20.

52. Cohen J. The t test for means. Stat Power Anal Behav Sci. 1988;2:20–6.

53. Hovestaden R. AKUTBEREDSKABET ÅRSRAPPORT 2019. Copenhagen: Region Hovestaden; 2019.

54. Waller J, Kaur P, Tucker A, Amer R, Bae S, Kogler A, et al. The benefit of intravenous thrombolysis prior to mechanical thrombectomy within the therapeutic window for acute ischemic stroke. Clin Imaging. 2021;79:3–7.

55. Mosley I, Nicol M, Donnan G, Patrick I, Dewey H. Stroke symptoms and the decision to call for an ambulance. Stroke. 2007;38(2):361–6.

56. DeAugustinis K. Acute Ischemic Stroke: The Role for Endovascular Therapy [Internet]. Evidence-Based Medicine Consult. 2015. https://www.ebmconsult.com/articles/acute-ischemic-stroke-endovascular-therapy.

57. Sykora M, Diedler J, Jü Ttler E, Steiner T, Zweckberger K, Hacke W, et al. Intensive care management of acute stroke: surgical treatment. Int J Stroke. 2010;5:170–7.

58. Demaerschalk BM, Cheng NT, Kim AS. Intravenous thrombolysis for acute ischemic stroke within 3 hours versus between 3 and 4.5 hours of symptom onset. Neurohospitalist. 2015;5(3):101–9.

59. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ. 2015;351:h5651.

60. Haghani A, Yang S. Real-time emergency response fleet deployment: concepts, systems, simulation & case studies. In: Zeimpekis V, Trantilis CD, Giaglis GM, Minis I, editors. Dynamic fleet management concepts, systsmy, algorithms & case studies. New York: Springer; 2007. p. 133–62.

61. Jones SP, Bray JE, Gibson JME, McClelland G, Miller C, Price CI, et al. Characteristics of patients who had a stroke not initially identified during emergency prehospital assessment: a systematic review. Emerg Med J. 2021;38(5):387–93.

62. Broderick JP, Adeoye O, Elm J. Evolution of the modified Rankin scale and its use in future stroke trials. Stroke [Internet]. 2017;48(7):2007–12.

63. Kwon S, Hartzema AG, Duncan PW, Lai SM. Disability measures in stroke: relationship among the Barthel Index, the functional independence measure, and the modified Rankin Scale. Stroke [Internet]. 2004;35(4):918–23.

64. Handschu R, Poppe R, Rauß J, Neundörfer B, Erbguth F. Emergency calls in acute stroke. Stroke. 2003;34(4):1005–9.

65. Kuner C, Svantesson DJB, Cate FH, Lynskey O, Millard C. Machine learning with personal data: Is data protection law smart enough to meet the challenge? Int Data Priv Law. 2017;7(1):1–2.

66. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science (80-). 2019;366:447–53.

67. Lisabeth LD, Brown DL, Hughes R, Majersik JJ, Morgenstern LB. Acute stroke symptoms: comparing women and men. Stroke. 2009;40(6):2031–6.

68. Rathore SS, Hinn AR, Cooper LS, Tyroler HA, Rosamond WD. Characterization of incident stroke signs and symptoms findings from the atherosclerosis risk in communities study. Stroke. 2002;33(11):2718–21.

69. Singhal AB, Biller J, Elkind MS, Fullerton HJ, Jauch EC, Kittner SJ, et al. Recognition and management of stroke in young adults and adolescents. Neurology. 2013;81(12):1089–97.

70. Biggs D, Silverman ME, Chen F, Walsh B, Wynne P. How should we treat patients who wake up with a stroke? A review of recent advances in management of acute ischemic stroke. Am J Emerg Med. 2019;37(5):954–9.

71. Møller TP, Ersbøll AK, Tolstrup JS, Østergaard D, Viereck S, Overton J, et al. Why and when citizens call for emergency help: An observational study of 211,193 medical emergency calls. Scand J Trauma Resusc Emerg Med. 2015;23(1):1–10.

72. Aroor S, Singh R, Goldstein LB. BE-FAST (Balance, eyes, face, arm, speech, time) reducing the proportion of strokes missed using the FAST mnemonic. Stroke. 2017;48:479–81.

73. Oostema JA, Chassee T, Baer W, Edberg A, Reeves MJ. Brief educational intervention improves emergency medical services stroke recognition. Stroke. 2019;50(5):1193–200.

74. Krebes S, Ebinger M, Baumann AM, Kellner PA, Rozanski M, Doepp F, et al. Development and validation of a dispatcher identification algorithm for stroke emergencies. Stroke. 2012;43(3):776–81.

75. Mattila OS, Puolakka T, Ritvonen J, Pihlasviita S, Harve H, Alanen A, et al. Targets for improving dispatcher identification of acute stroke. J Stroke. 2019;14(4):409–16.

76. Watkins CL, Leathley MJ, Jones SP, Ford GA, Quinn T, Sutton CJ. Training emergency services' dispatchers to recognise stroke: an interrupted time-series analysis on behalf of the Emergency Stroke Calls: Obtaining Rapid Telephone Triage (ESCORTT) Group. BMC Health Serv Res. 2013;13:318.

77. Blomberg SN, Christensen HC, Lippert F, Ersbøll AK, Torp-Petersen C, Sayre MR, et al. Effect of machine learning on dispatcher recognition of Out-of-Hospital cardiac arrest during calls to emergency medical services. JAMA Netw Open [Internet]. 2021;4(1):e2032320.

78. EIT Health. Healthcare workforce and organisational transformation with AI-enacting change. Riund Tabke Series 2020. Summary Report. Munich; 2021.

79. Bolander T, Dk T. What do we loose when machines take the decisions? J Manag Gov [Internet]. 2019;23:849–67. https://doi.org/10.1007/s10997-019-09493-x.

80. Spanglerid D, Hermansson T, Smekal D, Blomberg H. A validation of machine learning-based risk scores in the prehospital setting. PLoS ONE [Internet]. 2019. https://doi.org/10.1371/journal.pone.0226518.

81. Hassan E. Recall bias can be a threat to retrospective and prospective research designs. Internet J Epidemiol. 2005;3(2):339–412.

82. Coughlin SS. Recall bias in epidemiologic studies. J Clin Epidemiol. 1990;43(1):87–91.

83. Heale R, Twycross A. Validity and reliability in quantitative studies. Evid Based Nurs [Internet]. 2015;18(3):66–7.

84. Watkins CL, Jones SP, Hurley MA, Benedetto V, Price CI, Sutton CJ, et al. Predictors of recognition of out of hospital cardiac arrest by emergency medical services call handlers in England: a mixed methods diagnostic accuracy study. Scand J Trauma Resusc Ermeg Med. 2021;29:7.

85. Hatzitolios A, Savopoulos C, Hippokratia GN. Stroke and conditions that mimic it: a protocol secures a safe early recognition. Hippokratia. 2008;12(2):98–102.

86. Hosseininezhad M, Sohrabnejad R. Stroke mimics in patients with clinical signs of stroke. Casp J Intern Med. 2017;8(3):213–6.

87. Alanazy ARM, Wark S, Fraser J, Nagle A. Factors impacting patient outcomes associated with use of emergency medical services operating in urban versus rural areas: a systematic review. Int J Environ Res Public Health [Internet]. 2019;16(10):1728–44.

88. Schloemer T, Schröder-Bäck P. Criteria for evaluating transferability of health interventions: a systematic review and thematic synthesis. Implement Sci. 2018;13(1):1–17.

Publisher's Note