Original Research

Development of an algorithm for identifying paraneoplastic ischemic stroke in association with lung, pancreatic, and colorectal cancer

Rebecca Kassubek [ID], Marc-Andre G. R. Winter, Jens Dreyhaupt, Mona Laible, Jan Kassubek [ID], Albert C. Ludolph and Jan Lewerenz

Abstract

Background: Paraneoplastic ischemic stroke has a poor prognosis. We have recently reported an algorithm based on the number of ischemic territories, C-reactive protein (CRP), lactate dehydrogenase (LDH), and granulocytosis to predict the underlying active cancer in a case-control setting. However, co-occurrence of cancer and stroke might also be merely incidental.
Objective: To detect cancer-associated ischemic stroke in a large, unselected cohort of consecutive stroke patients by detailed analysis of ischemic stroke associated with specific cancer subtypes and comparison to patients with bacterial endocarditis.
Methods: Retrospective single-center cohort study of consecutive 1612 ischemic strokes with magnetic resonance imaging, CRP, LDH, and relative granulocytosis data was performed, including identification of active cancers, history of now inactive cancers, and the diagnosis of endocarditis. The previously developed algorithm to detect paraneoplastic cancer was applied. Tumor types associated with paraneoplastic stroke were used to optimize the diagnostic algorithm.
Results: Ischemic strokes associated with active cancer, but also endocarditis, were associated with more ischemic territories as well as higher CRP and LDH levels. Our previous algorithm identified active cancer-associated strokes with a specificity of 83% and sensitivity of 52%. Ischemic strokes associated with lung, pancreatic, and colorectal (LPC) cancers but not with breast and prostate cancers showed more frequent and prominent characteristics of paraneoplastic stroke. A multiple logistic regression model optimized to identify LPC cancers detected active cancer with a sensitivity of 77.8% and specificity of 81.4%. The positive predictive value (PPV) for all active cancers was 13.1%.
Conclusion: Standard clinical examinations can be employed to identify suspect paraneoplastic stroke with an adequate sensitivity, specificity, and PPV when it is considered that the association of ischemic stroke with breast and prostate cancers in the stroke-prone elderly population might be largely incidental.

Keywords: cancer, C-reactive protein, granulocytes, ischemic stroke, lactate dehydrogenase, magnetic resonance imaging, paraneoplastic, receiver operating characteristics

Correspondence to:
Rebecca Kassubek
Department of Neurology, University of Ulm, Oberer Eselsberg 45, Ulm 89081, Germany
rebecca.kassubek@uni-ulm.de

**Marc-Andre G. R. Winter
Mona Laible
Jan Lewerenz**
Department of Neurology, University of Ulm, Ulm, Germany

Jens Dreyhaupt
Institute of Epidemiology and Medical Biometry, University of Ulm, Ulm, Germany

**Jan Kassubek
Albert C. Ludolph**
Department of Neurology, University of Ulm, Ulm, Germany

German Center for Neurodegenerative Diseases (DZNE) Ulm, Ulm, Germany

Introduction

Cancer-associated paraneoplastic hypercoagulability frequently results in arterial thromboembolism into the brain,[1] which thus can be referred to as paraneoplastic ischemic stroke. Upon autopsy, cerebral post-ischemic lesions can be detected in up to 15% of cancer patients.[2] Similar to the risk of venous thromboembolism, the risk for

1

paraneoplastic ischemic stroke depends on the type of the underlying tumor with lung and pancreatic cancers reportedly harboring the highest risk.[3] Paraneoplastic ischemic stroke may even be the first symptom of occult cancer.[4] In known cancer, the frequency of paraneoplastic venous as well as arterial thromboembolism[1,5] such as ischemic stroke[6–8] peaks early after diagnosis of cancer and decreases thereafter, indicating that cancer treatment prevents additional events. Compared to cerebral ischemia with conventional risk factors, paraneoplastic ischemic stroke often involves multiple territories and is associated with a high stroke-associated morbidity and mortality.[9–12] However, unfractionated subcutaneous heparin[13,14] may reduce the risk of stroke recurrence in these patients. Low-molecular-weight heparin also corrects cancer-associated hypercoagulability,[15] and this correction was found to be associated with a better prognosis.[16] Thus, although the knowledge regarding optimal treatment of paraneoplastic ischemic stroke is still rudimentary, early identification of an underlying cancer associated with hypercoagulability is the prerequisite for any therapeutic decision that may improve prognosis.

The challenge for the clinician is to diagnose whether an ischemic stroke is of paraneoplastic origin, whether a cancer diagnosis is known or not, as many cancer patients share similar demographic characteristics and vascular risk factors with stroke patients, such as older age and smoking.[17] Thus, many studies sought to identify the typical characteristics of cancer-associated, presumably paraneoplastic stroke in comparison to ischemic stroke not associated with cancer. These characteristics have been found to comprise the simultaneous involvement of two or more vascular territories,[18–20] elevated lactate dehydrogenase (LDH),[6] elevated D-dimers as marker for hypercoagulability,[19–22] as well as a systemic inflammatory response with increased C-reactive protein (CRP)[18,22–24] and relative granulocytosis.[6,20] All these laboratory markers are also elevated in cancer patients compared to controls;[25,26] however, CRP and D-dimer elevations are more pronounced in patients with cancer and concomitant ischemic stroke compared to patients with cancer only.[27,28] Likewise, higher LDH is associated with a higher risk for venous thromboembolism in cancer patients.[29,30] To date, no definite biomarker for paraneoplastic ischemic stroke exists, as ischemic stroke in cancer patients caused by conventional vascular risk factors might also involve multiple territories. In addition, systemic inflammation is well known to increase the risk of cerebral ischemia.[31] Finally, the laboratory and imaging characteristics of paraneoplastic stroke are similar to ischemic strokes associated with bacterial endocarditis.[32]

In a case-control setting, we previously generated a multiple model with four variables, that is, the number of territories, LDH and CRP levels, as well as relative granulocytosis, that identified ischemic stroke associated with active solid cancer with a sensitivity and specificity of 78% and 95%, respectively.[6] The aim of the current study was to test the performance of our previous model to detect cancer-associated ischemic stroke in a large, unselected cohort of consecutive stroke patients. In addition, we sought to improve the detection of paraneoplastic stroke by detailed analysis of ischemic stroke associated with specific cancer subtypes and comparison to patients with the clinical diagnosis of bacterial endocarditis.

Methods

Recruitment of the consecutive ischemic stroke cohort

Acute ischemic strokes admitted to the Stroke Unit of the Department of Neurology, University Hospital Ulm (RKU), Ulm, Germany, between 1st October 2014 and 31st May 2017, were identified by reviewing our prospectively collected database. This period directly followed recruitment of our previous study, 1st January 2009 to 30th September 2014.[6] Cases with ischemic stroke associated with primary brain tumors and intracerebral metastases were excluded. Ischemic stroke diagnosed as associated with probable or proven bacterial endocarditis was analyzed separately. This categorization was performed based on the clinical diagnosis made by the treating physicians who cared for the patient during acute treatment.

Data collection for the consecutive ischemic stroke cohort

Only patients with data for magnetic resonance imaging (MRI), CRP, LDH, as well as percentage of granulocytes obtained at admission were enrolled. MRI was performed with a Magnetom TIM Symphony 1.5T MR scanner (Siemens, Erlangen, Germany) equipped with a 12-channel

head coil and transversal diffusion-weighted imaging (DWI) with diffusion-sensitizing gradients in at least three orthogonal directions and b values of 0 and $1000\,mm^2/s$ together with transversal gradient recalled echo, T2-/T1-weighted sequences, and coronal fluid attenuated inversion recovery (FLAIR) sequences. Acute vascular lesions were attributed to four vascular territories: left and right carotid arteries and left and right vertebral arteries. In case that the respective vertebral artery was connected to the basilar artery territory, lesions in this vertebral artery territory and the basilar artery territory were counted as one territory. Data for the percentage of granulocytes among leukocytes, serum LDH activity, and CRP were obtained from the routine laboratory data at the day of admission.

Clinical history, National Institutes of Health Stroke Scale (NIHSS, National Institute of Health, Bethesda, MD, USA; https://www.ninds.nih.gov/health-information/public-education/know-stroke/health-professionals/nih-stroke-scale) score at admission and dismissal and death during hospitalization due to stroke as well as the type of cancer and presence of metastatic disease was extracted from the case files with MRI, CRP, LDH, and granulocyte data available. Patients with additional diagnosis of a malignant cancer were grouped into those with ischemic stroke associated with either active or inactive cancer. The latter was postulated if the diagnosis of malignant cancer was more than 12 months prior stroke and no evidence of active tumor occurred and the patient did not undergo any tumor-specific therapy for at least 12 months prior to ischemic stroke. Conversely, active cancer was assumed if either cancer diagnosis or the last cancer-specific treatment occurred less than 12 months before ischemic stroke. Cancer-associated ischemic stroke was defined simply as ischemic stroke in the presence of active cancer. As both common conventional stroke pathophysiologies, large vessel atherosclerosis and atrial fibrillation, and cancer share common risk factors like smoking and alcohol consumption,[33,34,35] this descriptive category was assigned independently of the presence of conventional stroke risk factors.

Combined cohort for the five most common cancer subtypes

To analyze the subtype-specific characteristics of cancer-associated ischemic stroke, all patients with active cancer from 1st January 2009 to 31st May 2017 were combined. Those identified from 1st January 2009 to 30th September 2014 were reported previously.[6] The characteristics of the five most common cancer subtypes, lung, pancreatic, prostate, colorectal, and breast cancer, were analyzed separately.

Statistical analysis

Statistical analysis was performed by use of the MATLAB-based Statistics Toolbox (The MathWorks, Inc., Natick, MA, USA), GraphPad Prism (GraphPad Software, San Diego, CA, USA), SAS version 9.4 (SAS Institute, Cary NC, USA), RStudio (RStudio, Inc., Boston, MA, USA), and Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). For bivariate statistics, differences in frequency of multiple categorical variables were analyzed using the chi-square test. In case that multiple pairwise *post hoc* Fisher's exact tests were performed on a multigroup dataset as indicated, p values were Bonferroni-corrected. Groupwise difference regarding ordinal and continuous variables was analyzed using Kruskal–Wallis tests with Bonferroni-adjusted Dunn's post-test as indicated. Cutoffs were optimized using receiver operator characteristic (ROC) analysis with Youden's J statistics.

To assess the overrepresentation of certain solid tumor types in the cohort with active cancer and stroke, the mean 5-year prevalence for all solid tumors was calculated using the 5-year prevalence for all tumors after subtracting the 5-year prevalence of Hodgkin and non-Hodgkin lymphomas, multiple myelomas, and leukemias for each year from 2009 to 2017 using data of the Deutsches Krebsregister (German Cancer Registry, https://www.rki.de/EN/Content/Health_Monitoring/Cancer_Registry/cancer_registry_node.html). The average relative 5-year prevalence for lung, pancreatic, colorectal, breast, and prostate cancer was calculated by dividing the average 5-year prevalence for each tumor type by the average 5-year prevalence for all solid tumors. To test for overrepresentation of active cancers of one of these subtypes in the ischemic stroke cohort, the proportion of this cancer in the stroke with active cancer cohort with solid tumors from 1st January 2009 to 31st May 2017 was divided by the proportion this cancer accounts for in the average 5-year prevalence of solid tumors in Germany.

Relationships of two variables were analyzed by Spearman rank correlation and simple linear regression. Nonlinear hypothesis-free relationships between two variables were visualized using a smoothing spline algorithm with three knots.

Probabilities of the current data for cancer-associated stroke were initially calculated using the published multiple logistic regression model, using the following equation:

$$\text{logit}(p) = \ln\left(p/(1-p)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

with p: probability of cancer, X_1: number of territories, X_2: mg/l CRP, X_3: percentage granulocytes, X_4: U/l LDH, and estimates for the regression coefficients β_0, β_1, β_2, β_3, and β_4 of −11.7523, 1.4438, 0.0410, and 0.000926, respectively.[6]

For the new model based on the current cohort only, variable selection was performed using forward and backward variable selection based on Akaike Information Criterion using the step-function in R Version 4.1.3 (Posit PBC, Boston, MA, USA). This resulted in the equation:

$$\text{logit}(p) = \ln\left(p/(1-p)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

with p: probability of cancer, X_1: number of territories, X_2: mg/l CRP, and the estimates for the regression coefficients β_0, β_1, and β_2 were −4.207377, 0.554057, and 0.011009, respectively.

For the final model optimized to identify stroke associated with lung, pancreatic, and colorectal (LPC) cancers, all strokes from 1st January 2009 to 31st May 2017 associated with these three entities were combined with the strokes without active cancer from 1st January 2009 to 31st May 2017. The resulting cohort was split into two separate datasets: one comprising stroke affecting one territory only, and the other affecting all multiterritorial strokes. Input variables were age (years), presence of atrial fibrillation, sex, number of territories (multiterritorial cohort only), Ln(mg/l CRP + 1), LDH (U/l), and relative granulocytes (%). The multiple regression models to detect LPC cancers were optimized by forward and backward variable selection based on Akaike

information criterion using the step-function in R Version 4.1.3.

For strokes with one territory, this resulted in the equation:

$$\text{logit}(p) = \ln\left(p/(1-p)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

with p: probability of cancer, X_1: Ln(CRP + 1), X_2: % granulocytes, and the estimates for the regression coefficients β_0, β_1, and β_2 were −4.67215, 0.92771, and 0.02710, respectively.

For multiterritorial ischemic strokes in the equation:

$$\text{logit}(p) = \ln\left(p/(1-p)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

with p: probability of cancer, X_1: age, X_2: sex, X_3: number of territories, X_4: Ln(CRP + 1), and the estimates for the regression coefficients β_0, β_1, β_2, β_3, and β_4 of −3.25649, −0.07307, −1.12719, and 0.90398, respectively.

Given the exploratory nature of this study, the results from the statistical tests should not be interpreted as confirmatory but in an exploratory way. A two-sided p value <0.05 was considered to be significant.

Results

Characteristics of patients with ischemic stroke and cancer

We identified 1612 ischemic strokes in 1548 patients with a full data set for MRI and values for CRP, LDH, and granulocytes. The main characteristics of the overall cohort are shown in Supplemental Table 1. Sixty-seven of 1612 (4.2%) ischemic strokes were associated with an active malignancy. Thereof, one patient suffered from ischemic stroke twice during the recruitment period. One ischemic stroke was in a patient with two different active malignancies, lung cancer, and sarcoma. For five ischemic strokes associated with active cancer, the patients had a history of another malignancy, now classified as inactive. One hundred and six ischemic strokes (6.8%) in 105 patients with a history of cancer, in 11 cases with 2 cancers, classified as inactive were

identified. Eighteen strokes (1.1%) in 18 patients were documented as associated with probable or proven bacterial endocarditis. One ischemic stroke was not only associated with the diagnosis of bacterial endocarditis but also active bile duct cancer. In total, strokes associated with the documented diagnosis of probable or proven bacterial endocarditis were fourfold less frequent than ischemic strokes associated with active cancer.

Active cancer-associated ischemic strokes comprised 21 different types of solid malignancies, one active cancer of unknown primary, one not further described solid active cancer, and four active hematological malignancies (Table 1). Compared to inactive tumors, distant metastases were present more frequently in active tumors (64% *versus* 6%, $p < 0.0001$) as were lymph node metastases (70% *versus* 41%, $p = 0.04$). Among the active cancers associated with ischemic stroke, the most common were lung (14/67; 20.6%), prostate (10/67; 14.7%), breast (9/67; 13.2%), colorectal (7/67; 10.3%), and pancreatic (5/67; 7.4%) cancers. In line with our previous study[6] of the 63 ischemic strokes for which the time of cancer diagnosis could be determined, 12 (19%) were associated with cancer diagnosed within 4 weeks prior to or even following ischemic stroke, while in another 19 stroke episodes (30%), the diagnosis was made from 2 months up to 1 year prior to ischemic stroke. Taken together, close to 50% of all diagnoses of tumors classified as active occurred within 1 year prior to ischemic stroke.

Stroke associated with active cancer is frequently multiterritorial and associated with elevated CRP and LDH

There were no significant differences among the ischemic events without cancer association, associated with history of cancer now classified as inactive, with active cancer, or with bacterial endocarditis regarding the patients' age, sex distribution, and NIHSS at admission and dismissal (Table 2). However, while 7 out of 67 ischemic strokes associated with active cancer (10.4%) were associated with death during the hospital stay due to cerebral ischemia, the in-hospital mortality of ischemic stroke not associated with cancer was more than threefold lower (44/1420, 3.1%, $p = 0.0066$).

The number of territories, CRP, and LDH levels in ischemic stroke with history of now categorized as inactive cancer was not different from stroke

without known history of cancer, thereby validating our categorization of cancers as inactive or active. In contrast, involvement of three or four territories occurred significantly more often in ischemic stroke associated with active cancer (3.8-fold, $p < 0.001$), and the median CRP and LDH were significantly higher (CRP 22.5-fold, $p < 0.0001$, LDH 1.4-fold, $p < 0.05$) compared to stroke not associated with cancer (Table 2; Figure 1). Surprisingly and in contrast to our previous findings,[6] the percentage of granulocytes was not changed by the presence of active cancer. In contrast, ischemic strokes associated with probable or proven bacterial endocarditis were mostly multiterritorial and associated with very high CRP and LDH levels as well as granulocytosis (Table 2; Figure 1; Supplemental Table 2).

Validation and limitations of the previous multiple model of ischemic stroke associated with active cancer

In the next step, we used the current cohort to confirm the multiple logistic regression model generated in a case-control setting.[6] Probabilities for cancer-associated cerebral ischemia calculated by use of the previously developed algorithm were significantly higher in ischemic stroke associated with active cancer in the current cohort ($p < 0.0001$), while stroke associated with history of inactive cancer yielded identical probabilities when compared to stroke without any cancer association nor endocarditis [Figure 2(a)]. Since events associated with endocarditis might mimic paraneoplastic stroke, our strategy to preemptively exclude these events before generating models to detect paraneoplastic stroke is considered to be valid. ROC analysis resulted in an area under the curve (AUC) of 0.67 [95% CI 0.57–0.75, Figure 2(a), left middle panel] that was not higher than that for CRP alone (0.72, 95% CI 0.65–0.79). A Youden's J-optimized cutoff of >0.76 again detected ischemic stroke with an adequate specificity of 83% but poor sensitivity of 52% [Figure 2(a) and (b)]. Among the stroke subgroup with a probability >0.76, active cancer was 4.8-fold more frequent (11.8%) when compared to strokes with probability below the cutoff [2.5%, Figure 2(a)]. However, 19% of all ischemic events were associated with a probability above the cutoff.

As the performance of the previous model seemed unsatisfactory, we set out to optimize the detection of paraneoplastic stroke using the current dataset

Table 1. Number of different tumor types in the cohort with ischemic stroke associated with cancer.

Strokes associated with tumor	All	One or more active tumors*	Inactive tumor
Strokes (patients)	173 (171)	67 (66)	106 (105)
Strokes categorized by most recently diagnosed tumor entity	*N* (%)	*N* (%)	*N* (%)
Breast	36 (20.8)	9 (13.4)[a]	27 (25.5)[b]
Prostate	31 (17.9)	10 (14.9)[c]	21 (19.8)[d],*
Colorectal	23 (13.3)	7 (10.4)[e],$	16 (15.1)[f]
Lung	16 (9.2)	14 (20.9)[g]	2 (1.9)[h]
Hematologic	11 (6.4)	4 (6.0)	7 (6.6)[h]
Urothelial	9 (5.2)	4 (6.0)	5 (4.7)
Melanoma	9 (5.2)	2 (3.0)	7 (6.6)
Renal	7 (4.0)	1 (1.5)	6 (5.7)[h]
Pancreatic	5 (2.9)	5 (7.5)	0 (0)
Uterus	4 (2.3)	1 (1.5)	3 (2.8)
Gastric	3 (1.7)	1 (1.5)	2 (1.9)
Unknown	3 (1.7)	1 (1.5)[i]	2 (1.9)
Hepatocellular	2 (1.2)	1 (1.5)	1 (0.9)
Testis	2. (1.2)	0 (0)	2 (1.9)
Sarcoma	2 (1.2)	1 (1.5)	1 (0.9)
Ovarian	2 (1.2)	2 (3.0)	0 (0)
Biliary tract	1 (0.6)	1 (1.5)	0 (0)
Carcinoid	1 (0.6)	0 (0)	1 (0.9)
Thyroid	1 (0.6)	0 (0)	1 (0.9)
Urachus	1 (0.6)	0 (0)	1 (0.9)
Larynx/pharynx	1 (0.6)	1 (1.5)	0 (0)
Spinalioma	1 (0.6)	0 (0)	1 (0.9)
CUP	1 (0.6)	1 (1.5)[j]	0 (0)
IPMN	1 (0.6)	1 (1.5)	0 (0)
Known lymph nodes metastases, *N* (%)	24 (14.0)	16 (23.9)	8 (7.5)

(Continued)

Table 1. (Continued)

Strokes associated with tumor	All	One or more active tumors*	Inactive tumor
Known distant metastases, N (%)	30 (17.5)	27 (40.3)	3 (2.8)
Diagnosis prior stroke, months (median, IQR)	54 (13–130)	15 (3–101)	82 (34–155)

aOne patient with preexisting inactive contralateral breast cancer.
bOne each with preexisting inactive non-Hodgkin lymphoma, melanoma, and renal carcinoma.
cOne patient with preexisting inactive colorectal carcinoma.
dOne each with preexisting also inactive colorectal, bladder and larynx carcinoma, and Hodgkin lymphoma.
eOne patient with preexisting inactive prostate cancer.
fOne each with gastric and urachal carcinoma.
gTwo patients with preexisting inactive breast cancer, one patient with active lung cancer and sarcoma, two patients with preexisting (inactive) breast cancer.
hOne each with a preexisting cancer of the same category.
iImaging suggestive for metastatized pancreatic cancer, not biopsy performed due to medical condition.
jOne patient with preexisting inactive spinalioma.
Two patients, one each with an active colorectal carcinoma ($) and an inactive prostate carcinoma (*) were enrolled twice as two episodes with ischemic strokes occurred.
CUP, cancer of unknown primary; IPMN, intraductal papillary mucinous neoplasm; IQR, interquartile range.

Table 2. Characteristics of patients without tumor, inactive tumor, active tumor, or endocarditis within the cohort of one consecutive stroke patients.

1548 patients/1612 strokes	No. tumor (1420 strokes)	inactive tumor (106 strokes)	Active tumor (67 strokes)	Endocarditis (18 strokes)	Statistics
Sex, female/male, N (%)	632/788 (45/55)	41/65 (39/61)	28/39 (42/58)	8/10 (44/56)	p = 0.6823
Age (years), median (IQR)	77 (66–84)	78 (72–82)	77 (69–82)	75 (66–81)	p = 0.3578
NIHSS admission, median (IQR), N	3 (1–6), 1222	3 (1–6), 88	3 (2–8), 57	3 (1–10), 16	p = 0.9098
NIHSS discharge, median (IQR), N	1 (0–4), 1115	1 (0–3), 79	2 (2–5), 52	1 (1–3), 16	p = 0.6792
Atrial fibrillation, N (%)	457 (32)	43 (41)	17 (25)	7 (39)	p = 0.1652
Death, N (%)	44 (3)	4 (4)	7 (10)	1 (6)	**p = 0.0143**
Territories (1/2/3 + 4), N (%) (IQR)	1099/247/74 (77/17/5)	87/12/7 (82/11/7)	39/15/13 (58/22/19)	4/6/8 (22/33/44)	**p < 0.0001**
CRP (mg/l), median (IQR)	4.6 (1.8–11.3)	4.8 (2.1–11.8)	23.2 (5.9–67.8)	99.0 (40.2–146.0)	**p < 0.0001**
LDH (U/l), median (IQR)	227 (195–276)	237 (188–281)	258 (201–327)	324 (249–463)	**p = 0.0002**
Granulocytes (%), median (IQR)	68 (60–76)	70 (61–77)	71 (62–81)	84 (72–88)	**p = 0.0001**

Statistical analysis was performed using Chi-Square and Kruskal–Wallis.
p values < 0.05 in bold.
CRP, C-reactive protein; IQR, interquartile range; LDH, lactate dehydrogenase.

excluding stroke associated with endocarditis as new exploratory cohort. Variable selection among the four variables, number of territories, CRP, LDH, and percentage of granulocytes, led to a multiple logistic regression model with only the number of territories and CRP remaining as predictors. Again, the resulting probabilities were significantly higher for stroke associated with active cancer, as they were when calculated for the 18 strokes associated with endocarditis [both $p < 0.0001$, Figure 2(b), left panel]. Although the AUC for this model to detect ischemic stroke associated with active cancer overlapped with the AUC of the previous model [0.71, 95% CI 0.64–0.78, Figure 2(b), middle left
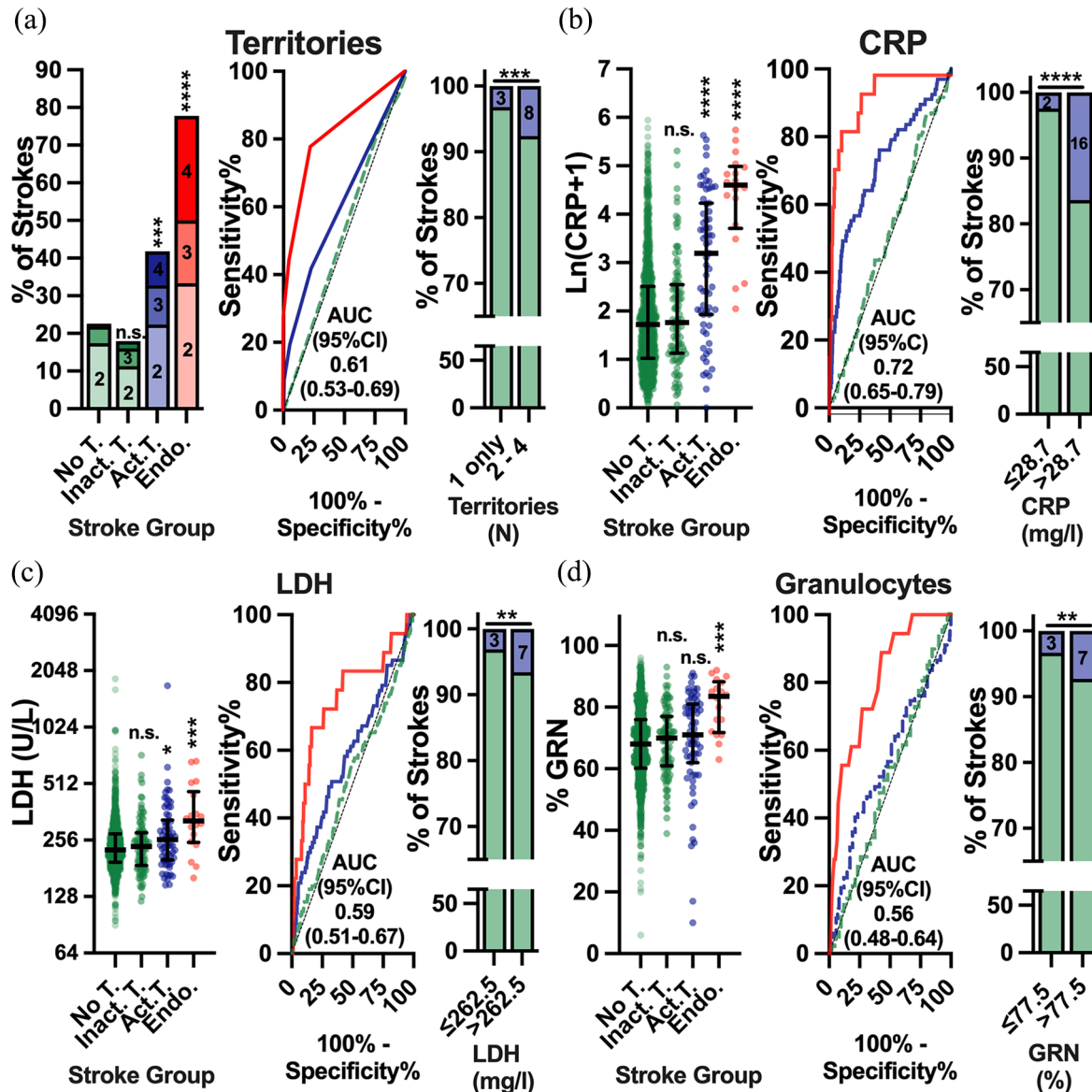
Figure 1. Patients with ischemic stroke associated with active tumor as well as patients with ischemic strokes associated with endocarditis show ischemic strokes affecting multiple territories as well as an increase in inflammatory and cell-lysis indicating blood parameters. The number of large territories affected by ischemic strokes (a), CRP (b), LDH (c), and granulocytes (% GRN, d) in patients without known history of tumor (No. T., green), known neoplastic disease judged as inactive (Inact. T., green), with active tumor (Act. T., blue), or with diagnosis of bacterial endocarditis (Endo, red). Left panels: Stacked box blots with the percentage of patients with ischemic stroke involving >1 territory with the indicated number [progressively darker shades, (a)] or scatter plots (b–d) with CRP in mg/l logarithmized after adding 1 [Ln(CRP + 1), (b)]. Middle panels for (a)–(d): ROC analysis of the number of territories (a), CRP (b), LDH (c), and percent GRN (d). The AUC for active tumors *versus* no tumor indicated (continuous lines: AUC > 0.5). Right panels: Performance of cutoffs (GRN) to distinguish patients with active tumor from those without known tumor optimized for specificity and sensitivity using the Youden index. The number of patients above and below each cutoff was normalized to 100% with the percentage of those with active tumor shown in blue with the percentage indicated. Statistical analysis was performed using Bonferroni-corrected Fisher's exact tests after dichotomization [1–2 territories *versus* 3–4 territories, (a), left panel] or using Kruskal–Wallis tests with Dunn's post-tests (left panels) both comparing against stroke without tumor, ROC analysis (middle panels), and Fisher's exact tests (right panels).

n.s. $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$,***$p < 0.001$, ****$p < 0.0001$.

AUC, area under the curve; CRP, C-reactive protein; GRN, granulocytes; LDH, lactate dehydrogenase; ROC, receiver operator characteristics.
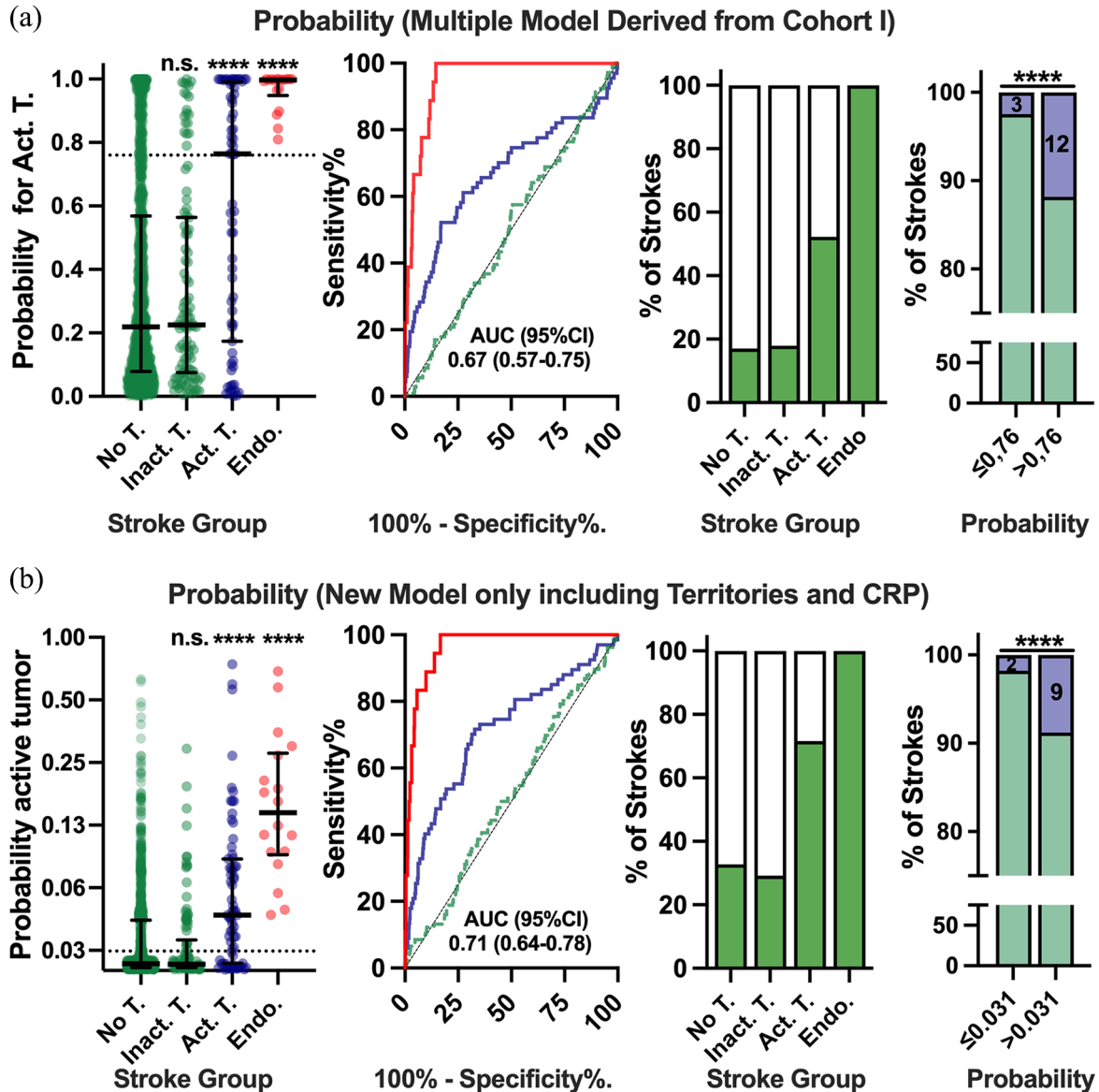
Figure 2. Performance of the algorithm to identify stroke associated with active cancer based on the previous case-control study and generation of a new algorithm. [(a), left graph] Probabilities for stroke associated with active cancer based on the algorithm of Kassubek *et al.*[6] in patients without known tumor (No. T., green), with inactive tumor (Inact. T., green), active tumor (Act. T., blue), and endocarditis (Endo., red). The Youden-optimized cutoff generated by comparing strokes without and with active tumor of the present cohort (>0.76) is indicated by the dotted line. (Left middle graph) ROC analysis of the probabilities generated for strokes without active tumor (green line), with active tumor (blue line), or endocarditis (red line) in comparison to the group without cancer. The AUC with 95% confidence interval (95% CI) for patients with active tumor *versus* no known tumor is indicated. (Middle right panel) Percentage of false-positive (green/red) or correctly positive strokes in each subgroup identified by applying the cutoff of >0.76. (Right panel) Enrichment of stroke associated with active tumor (blue) in the total cohort after exclusion of strokes associated with endocarditis by applying the cutoff of >0.76. The percentage of strokes associated with active cancer in each of the two groups is indicated. (b) Performance of the new algorithm generated by logistic regression with stepwise variable selection with number of territories and CRP as independent predictors shown as in (a). The Youden-optimized cutoff for the probability generated using this model was 0.031. In the left graphs, the median and interquartile range is shown. Statistical analysis was performed using the Kruskal–Wallis test with Dunn's post-test comparing the strokes with inactive tumor, active tumor, or endocarditis to strokes associated without tumor (left graphs) or Fisher's exact test using the number of patients in each group (right graphs). ****$p < 0.0001$.
AUC, area under the curve; CRP, C-reactive protein; n.s., not significant; ROC, receiver operator characteristics.

panel], the sensitivity and specificity were 72% and 67% when applying a Youden's J optimized cutoff of >0.031 [Figure 2(b), middle right panel], with the frequency of active cancer in the strokes with probabilities above the cutoff (8.8%) being 4.9-fold higher than in those with probabilities below the cutoff (1.8%). Of note, the increased sensitivity of this new model was largely due to the fact that in this model, 34% of all strokes (more than a third) had probabilities above the cutoff. Thus, this strategy failed to generate a tool to efficiently select stroke patients for cancer screening.

Differential distribution of characteristics typical for paraneoplastic stroke among different cancer entities

To generate a new strategy to identify 'paraneoplastic' ischemic stroke, we assumed that certain active tumors highly prevalent in the stroke-prone elderly population might only be coincidental findings and not causative for the ischemic event by inducing a paraneoplastic coagulopathy. This contamination with incidental pairings of stroke and active cancer might prevent the development of a sensitive and specific algorithm. We hypothesized that selecting tumor subtypes where an incidental nonparaneoplastic ischemic stroke is least likely might pave the way to generate a more efficient screening tool for paraneoplastic stroke. To this end, we assumed that tumors most prone to be associated with paraneoplastic stroke should be overrepresented in the cohort with stroke associated with active cancer. In addition, we set out to analyze the tumor subtype dependency of the markers known to suggest paraneoplastic etiology, the number of territories, CRP, LDH, and relative granulocytosis, as well as the correlation among these variables.

In a first step, we selected all strokes associated with the five most frequent active tumor entities in the current cohort, that is, lung ($N=14$), prostate ($N=10$), breast ($N=9$), colorectal ($N=7$), and pancreatic ($N=5$) cancer (Table 1). To increase group sizes, we combined those strokes with those identified in the preceding period analyzed in the previous study.[6] Among a total of 138 cerebral ischemic events associated with active solid cancer, this yielded groups for lung, prostate, breast, colorectal, and pancreatic cancers comprising 34 (24.6%), 21 (15.2%), 21 (15.2%), 17 (12.3%), and 10 (7.3%) stroke episodes, respectively (Supplemental Table 3). Compared to the average relative 5-year prevalence of those

entities from 2009 to 2017 according to the German Cancer Registry among all solid cancers (lung 5.8%, prostate 18.8%, colorectal 14.6%, breast 21.6%, and pancreas 1.2%), the relative proportion of strokes associated with active pancreatic or lung cancer were 5.9- and 4.3-fold higher, respectively, while colorectal and prostate cancer were 1.2-fold and breast cancer 1.4-fold less frequent among patients with stroke associated with active cancer [Figure 3(a)].

When the five groups of ischemic strokes associated with different types of cancer were analyzed for characteristics for paraneoplastic stroke, involvement of three or four vascular territories was observed much more often in ischemic strokes associated with active lung (10/32, 31%, $p<0.0001$), pancreatic (4/10, 40%, $p=0.0074$), and colorectal cancer (5/17, 29%, $p=0.0091$), but less so in ischemic strokes associated with prostate cancer (5/21, 24%, $p=0.0247$) and not in breast cancer (3/18, 17%, $p=0.3588$) when compared to strokes not associated with active cancer [81/1526, 5%, Figure 3(b)]. All four vascular territories were involved in only <1% of ischemic strokes not associated with an active tumor (12/1526) and were associated with active breast or prostate cancer in none of the events. In contrast, ischemia in all four territories was present in 24%, 19%, and 10% of the strokes associated with active colorectal, lung, and pancreatic cancer, respectively.

In patients with both active cancer and stroke, conventional stroke risk factors like atrial fibrillation or atherosclerosis might be less prevalent at younger age. Thus, it can be hypothesized that at younger age, stroke due to paraneoplastic hypercoagulopathy might be comparatively more frequent. *Compatible* with this hypothesis, the median age at ischemic stroke in patients with lung cancer was significantly lower [71 years, interquartile range (IQR) 65–75 years] than in patients without active cancer (77 years, IQR 66–83 years, $p=0.0036$), as in patients with colorectal cancer (67 years, IQR 63–80 years, $p=0.2923$) for whom the difference did not reach statistical significance, potentially due to smaller cohort size. However, in patients with stroke and active prostate (77 years, IQR 74–80 years), breast (78 years, IQR 68–85 years) and pancreatic cancer (79 years, IQR 68–83 years), the median age was virtually identical when compared to strokes not associated with active cancer [$p>0.9999$ for all pairs, Figure 3(c)].
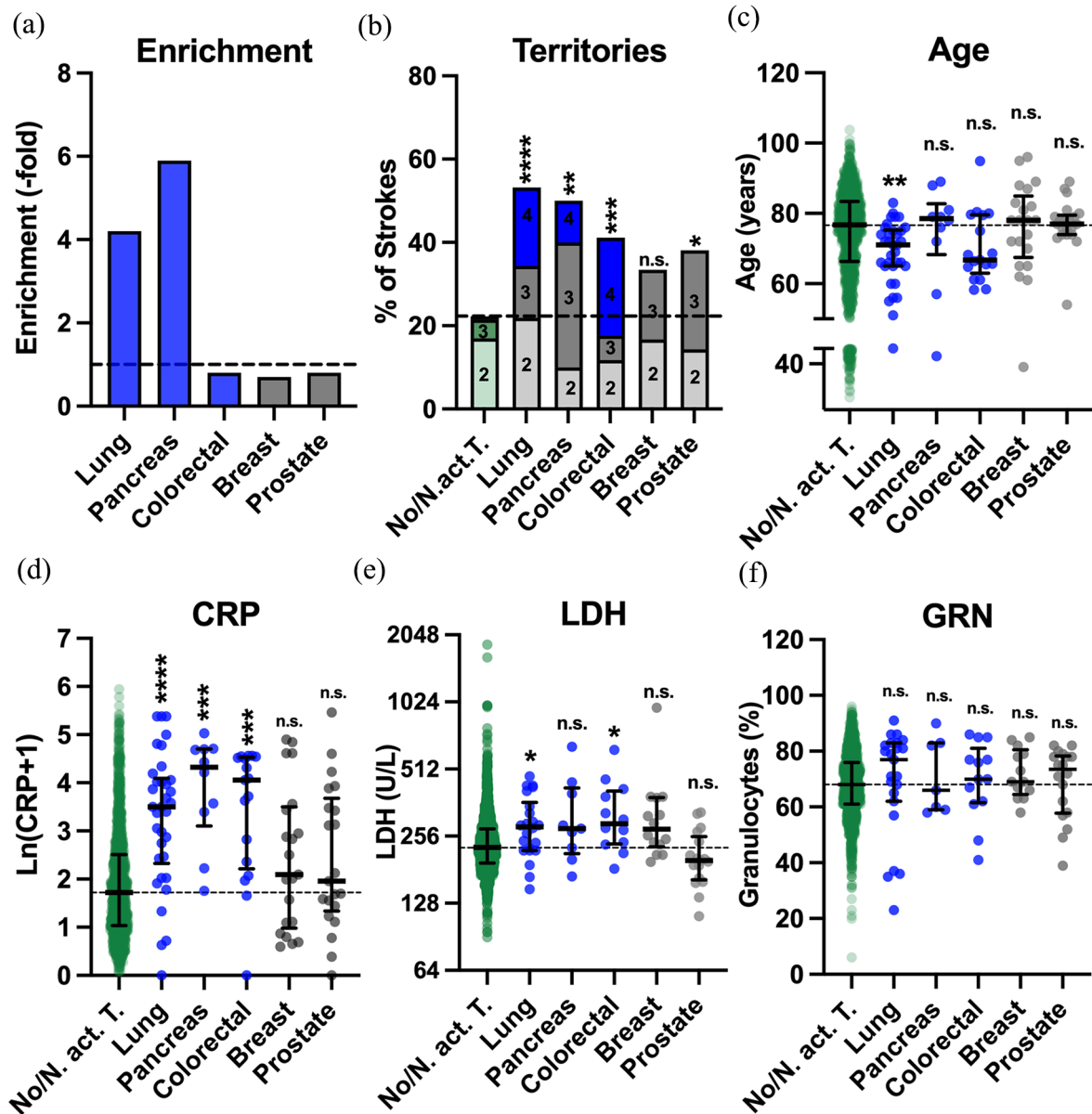
Figure 3. Patients with active lung, colorectal, and pancreatic cancer share some characteristics typical of paraneoplastic stroke. (a) Relative frequency of lung, pancreatic and colorectal (blue), as well as breast and prostate cancers (gray) combined from Kassubek *et al.*[6] and the current cohort compared to the German Cancer Registry. The dotted line indicates no change in frequency. (b) Proportion of multiterritorial strokes in patients without tumors (No. T., green) and with indicated active tumors (lung/colorectal/pancreatic: blue, breast/prostate: gray). Higher numbers of ischemic territories are indicated by darker shades with the number of territories given. (c) Age, (d) blood concentrations of CRP [mg/l, logarithmized after adding the correction factor 1, Ln(CRP + 1)], and (e) LDH and the percentage of GRN among leukocytes in the same groups as in (a). The median and interquartile range in indicated. The median of the cohort without tumors is indicated by dotted lines. Statistical analysis was performed using Bonferroni-corrected Fisher's exact tests after dichotomization (1–4 territories *versus* 4 territories) (b) or using Kruskal–Wallis tests with Dunn's post-tests (c–f) both comparing against stroke without tumor.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.
CRP, C-reactive protein; GRN, granulocytes; LDH, lactate dehydrogenase; n.s., not significant.

Systemic inflammation is known to result in a procoagulant state increasing the risk of cerebral ischemia in general,[31] and CRP levels are higher in cancer-associated stroke compared to active cancer only.[27,28] Compared to strokes not associated with active cancer (median CRP 4.6 mg/l,

IQR 1.8–11.3), median CRP levels were 7-, 16-, and 12-fold higher when stroke occurred in association with active lung cancer (32.1 mg/l, IQR 9.4–58.9, $p < 0.0001$), pancreatic cancer (74.8 mg/l, IQR 23.6–109.0, $p = 0.0001$), and colorectal cancer (56.9 mg/l, IQR 8.3–91.8, $p < 0.0001$), respectively. In contrast, CRP was not higher in the presence of either breast (7.1 mg/l, IQR 1.7–32.2, $p = 0.6881$) or prostate cancer [6.7 mg/l, IQR 3.4–43.5, $p = 0.5657$, Figure 3(d)].

Higher LDH has been identified as a risk factor for cancer-associated venous thromboembolism.[29,30] When compared to median LDH levels in ischemic stroke not associated with active cancer (228 U/l, IQR 194–276 U/l), LDH levels in ischemic stroke associated with active lung (282 U/l, IQR 221–363 U/l, $p = 0.0380$) and colorectal (291 U/l, IQR 236–408 U/l, $p = 0.0281$) cancers were significantly elevated. A nonsignificant tendency for higher LDH levels was found for pancreatic (277 U/l, IQR 214–422 U/l, $p = 0.3673$) and breast cancer (276 U/l, IQR 229–381 U/l, $p = 0.0517$). LDH levels in prostate cancer were nominally lower [199 U/l, IQR 163–255 U/l, Figure 3(e)].

Neutrophil extracellular traps are thought to be involved in cancer-associated thrombogenesis.[36] However, although when compared to ischemic stroke not associated with active cancer (68%, IQR 61–76%), the median relative granulocytosis was only numerically higher in ischemic stroke associated with lung (77%, IQR 62–83%, $p = 0.6130$) and prostate cancer (74%, IQR 58–78%, $p > 0.9999$) and not higher in pancreatic (66%, IQR 59–83%), colorectal (70%, IQR 62–81%), and breast cancer [69%, IQR 65–81%, all $p > 0.9999$, Figure 3(f)].

In summary, our analysis revealed five characteristics indicating that an ischemic stroke in the presence of an active cancer is indeed paraneoplastic due to cancer-induced hypercoagulopathy: (1) enrichment of the tumor type among stroke patients compared to the German Cancer Registry, (2) younger age, (3) more multiterritorial involvement, (4) elevated CRP, and (5) elevated LDH compared to ischemic stroke not associated with cancer, as these were differentially distributed among the five most common cancer subtypes associated with ischemic stroke. Of note, in the total cohort of 1612 ischemic strokes as well as in ischemic strokes not associated with cancer, we found a strong positive correlation of numbers of territories, age, CRP, and LDH with each other (Supplemental Tables 5–8), indicating that they are pathophysiologically linked.

Ischemic strokes associated with active LPC cancers, but not prostate and breast cancer, are most typical for paraneoplastic stroke

The group of ischemic strokes associated with active lung cancer exhibited all five characteristics identified as suggestive for paraneoplastic origin [Figure 3(a)–(e)]. In stroke associated with pancreatic and colorectal cancer, three of five were positive [enrichment in the cancer-associated stroke cohort, higher number of territories, and increased CRP for pancreatic cancer; higher number of territories, higher CRP and LDH for colorectal cancer (Figure 3(a)–(e))]. In contrast, none of the five characteristics was present in strokes associated with breast cancer, and only one, a higher number of ischemic territories, was present in stroke associated with prostate cancer.

Thus, for the purpose of our analysis, we grouped ischemic strokes associated with active lung, pancreatic, and colorectal as LPC-cancer-associated ischemic stroke and those with breast and prostate (BP) cancers as BP-cancer-associated strokes. LPC-cancer-associated strokes were considered to be mostly paraneoplastic, that is, due to paraneoplastic coagulopathy, while in BP-cancer-associated strokes, both diseases were considered to be coincidental in the majority of patients.

This result was confirmed by comparing ischemic strokes associated with either active LPC or BP cancers from the years 2009 to 2017 to those not associated with active cancer from 2014 to 2017 with regard to the numbers of ischemic territories, age, CRP, and LDH levels. The proportion of ischemic strokes involving all territories was significantly higher in strokes associated with LPC cancers (11/59, 19%, $p < 0.0001$) but not BP cancers [0/39, 0%, $p > 0.9999$, Figure 4(a)]. Compared to median age at stroke not associated with active cancer (77 years, IQR 66–83 years), age at stroke associated with active LPC cancer (72 years, IQR 65–79 years, $p = 0.0042$), but not with BP cancers (79 years, IQR 73–82 years, $p = 0.8379$) was significantly younger [Figure 4(a)]. Correspondingly, we observed a tendency for younger age being associated with more
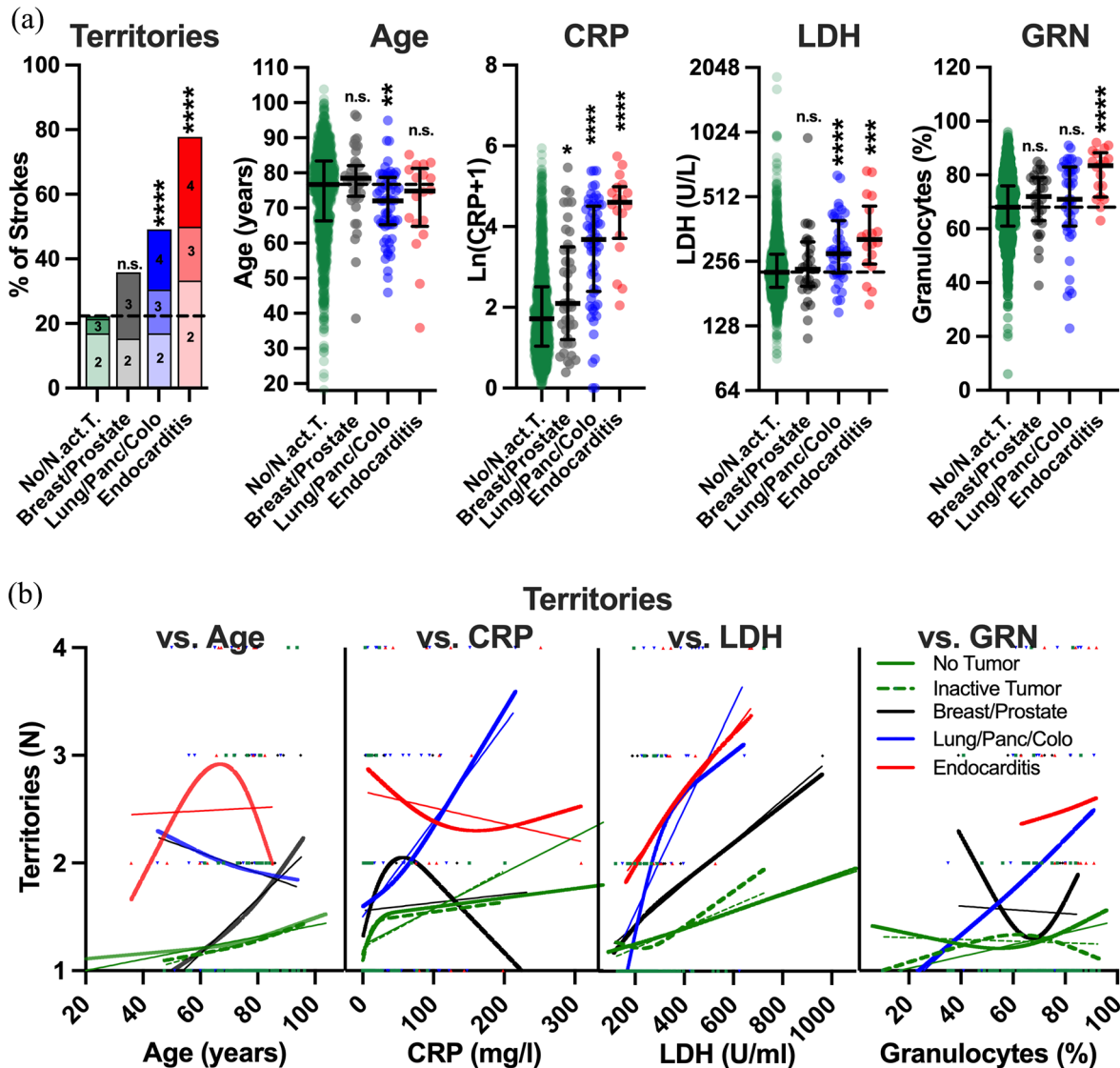
Figure 4. Strokes associated with lung, colorectal, and pancreatic cancer exhibit a unique positive correlation of the number of territories involved with blood CRP levels. The combined cohorts of patients with active lung, colorectal, and pancreatic cancer from Figure 2 were grouped as were the patients with breast and prostate cancers. (a) Proportion of multiterritorial strokes (left panel), age (middle left panel), CRP (middle panel), LDH (middle right panel), as well as the percentage of GRN (right panel) in patients without known tumor and with known inactive tumor (green), breast and prostate cancers combined (breast/prostate, gray), lung, pancreatic and colorectal cancer (Lung/Panc/Colo, blue), and endocarditis (red). The number of territories is indicated by darker shades. For statistical analysis, the exact number of patients with a certain number of territories was used. Statistical analysis was performed using Bonferroni-corrected Fisher's exact tests after dichotomization [1–2 territories *versus* 3–4 territories, (a), left panel] or using Kruskal–Wallis tests with Dunn's post-tests [(a), other panels] both comparing against stroke without tumor. (b) Relationship of the number of territories involved in the cerebral ischemia with the age, CRP, LDH, and percentage of granulocytes. The bold lines (green solid: no tumor; green dotted: inactive tumor; gray: breast and prostate cancers; blue: lung, colorectal, and pancreatic cancer; red: endocarditis) show the results of spline regression, the fine lines of simple linear regression. Details for the correlations including statistical analysis can be found in Supplemental Tables 5–8.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.
CRP, C-reactive protein; GRN, granulocytes; LDH, lactate dehydrogenase; n.s., not significant.

ischemic territories in the LPC cancer group, while the opposite was observed in BP-cancer-associated ischemic stroke [Figure 4(b), left panel,

Supplemental Table 5]. While median CRP levels were only 1.6-fold higher in ischemic stroke associated with BP cancers (7.1 mg/l, IQR

2.2–32.8 mg/l, $p = 0.0445$), an 8.5-fold increase in CRP compared to ischemic strokes not associated with active cancer (4.6 mg/l, IQR 1.8–11.3 mg/l) was observed in cerebral ischemia with active LPC cancer (median CRP: 38.8 mg/l, IQR 10.0–89.8 mg/l, $p < 0.0001$) [Figure 4(a)]. In addition, the number of territories significantly increased with increasing CRP in LPC but not BP cancers [Figure 4(b), middle left panel, Supplemental Table 6]. Median LDH levels were also significantly higher in the LPC-cancer-associated stroke group (278 U/l, IQR 226–397 U/l, $p = 0.0001$) when compared to ischemic strokes not associated with active cancer (228 U/l, IQR 194–276 U/l), a difference not observed in the strokes associated with BP cancers (235 U/l, 196–315 U/l, $p > 0.9999$). Correspondingly, the number of ischemic territories increased by 0.51 with each 100 U/l LDH (95% CI 0.21–0.82) in LPC-cancer-associated strokes, but only by 0.2 (0.0–0.39) in BP-cancer-associated strokes [Figure 4(b), middle right panel, Supplemental Table 7], leading to a significant correlation of the number of ischemic territories with LDH in LPC ($p = 0.0003$) but not BP-cancer-associated strokes ($p = 0.1278$). Relative granulocytosis did not differ between patients with active LPC cancers, BP cancers, and those without active cancer [Figure 4(a)].

In ischemic stroke associated with probable or definite bacterial endocarditis, the pattern of multiterritorial stroke associated with high CRP and high LDH was quite similar, but even more pronounced when compared to strokes associated with LPC cancers. However, the median age of stroke associated with endocarditis (75 years, IQR 65–81 years) was similar to age at strokes without cancer association; in addition, relative granulocytosis was much higher (84%, IQR 72–88%, $p < 0.0001$). Increased granulocytosis was observed neither in strokes associated with LPC cancers (71%, 61–83%, $p = 0.2862$) nor with BP cancers [72%, 63–79%, $p = 0.7077$, Figure 4(a)]. In addition, the number of ischemic territories in bacterial endocarditis appeared to be independent of age and CRP [Figure 4(a), left panels, Supplemental Tables 5/6].

Identification of 'paraneoplastic' ischemic stroke by using association with active LPC cancer as a surrogate parameter

We now had demonstrated that ischemic strokes associated with active LPC cancers, in contrast to those associated with active BP cancers, exhibit not only multiple characteristics assumed to be typical for paraneoplastic stroke but also a unique interdependency between these characteristics. Next, we set out to employ these results to generate an optimized model to identify paraneoplastic stroke using all the ischemic strokes associated with LPC cancers as surrogates for strokes with paraneoplastic etiology. To this end, we combined all ischemic strokes associated with LPC cancers from 2009 to 2017 of which all four variables (numbers of ischemic territories, CRP, LDH, and relative granulocytosis) were available ($N = 37$; 10 ischemic stroke from the previous study, 27%; 27 from the current cohort, 73%), with all ischemic strokes not associated with active cancer from 2014 to 2017 (no cancer + inactive cancer, $N = 1526$) to generate a new training cohort. In contrast to our previous attempts, we now included age and sex as well as the presence of atrial fibrillation to the list of putatively relevant predictors which consequently were subjected to variable selection. However, as a vast majority of all ischemic strokes in this cohort was restricted to one territory (1203/1563, 77%), these models were heavily biased for monoterritorial strokes (not shown). Thus, we decided to calculate two independent models, one for ischemic strokes affecting a single vascular territory only [17 with LPC cancer, 3 of which (18%) from the previous cohort, 1186 without active cancer, Supplemental Tables 9 and 10], the other for multiterritorial strokes (20 with LPC cancer, 7 of which, 35%, of the previous cohort, 340 without active cancer, Supplemental Tables 9 and 11]. As CRP generally led to a better fit when transformed by logarithmic calculus (not shown), we used log-transformed CRP. For the model for single-territory strokes, CRP and relative granulocytosis remained as most relevant predictors for LPC-cancer-associated stroke after variable selection (details: see Methods section). For multiterritorial strokes, age, sex, CRP levels, and numbers of territories remained to predict LPC-cancer-associated stroke. Both models were recalculated using the maximal number of strokes from the previous cohort for which information regarding variables that had been removed upon variable selection were missing (Supplemental Table 12 – 1 territory: 21 strokes, 7 (33%) from previous cohort, 10, 4, and 7 associated with lung, pancreatic, and colorectal cancers, respectively; 2–4 territories: 29 strokes, 16 (55%) from the previous cohort, 17, 5, and 7 associated with

lung, pancreatic, and colorectal cancers, respectively). With this expanded dataset, all remaining variables significantly affected the probability that an ischemic stroke was associated with active LPC cancer (Supplemental Figure 2). Interestingly, lower granulocytes were identified as predictive for monoterritorial LPC-cancer associated stroke. Indeed, at low CRP, a relative low percentage of granulocytes was characteristic for LPC-cancer-associated stroke (Supplemental Figure 3).

To achieve comparable cutoffs for both models adequate for their application as a screening tool and to adjust for the fact that 1.4-fold more strokes were added from the previous cohort to the number of strokes associated with active LPC cancer to the multiterritorial (compared to uniterritorial stroke) group, we chose cutoffs for LPC cancer probabilities that yielded a positive predictive value (PPV) of 10% for uniterritorial (probability >0.0382, sensitivity 57%, specificity 91%) and 14% for multiterritorial stroke (probability >0.0165, sensitivity 93%, specificity 52%).

Next, we calculated the probabilities for LPC-cancer association as a surrogate for paraneoplastic etiology for the current cohort of 1612 consecutive ischemic strokes (one stroke with both active cancer and endocarditis included in the endocarditis group), using the two logistic regression models depending on which number of territories was involved. Probabilities for ischemic strokes associated with the history of now inactive cancer were not different from probabilities in the group of strokes without cancer association [Figure 5(a); Supplemental Figure 4; Supplemental Tables 13/14]. Ischemic strokes associated with active LPC cancer according to both models displayed significantly increased probabilities, while ischemic strokes associated with BP cancer only showed a tendency for higher probabilities for 'paraneoplastic' etiology and those associated with other active tumors had probabilities ranging between both groups. In each of the two models, the highest probabilities for LPC-cancer-associated stroke were found among strokes associated with proven or definite bacterial endocarditis (Supplemental Figure 4; Supplemental Tables 13/14). When applying the cutoffs generated using the training cohort, all probabilities of ischemic strokes associated with the diagnosis of endocarditis were categorized as LPC-cancer-associated. By use of this strategy,

21/27 (78%) of all ischemic strokes associated with either LPC or active cancer in general, respectively, were detected (Supplemental Table 15). When preemptively excluding strokes associated with endocarditis from the analysis, 41/313 ischemic strokes with probabilities above the cutoffs (13%) were indeed associated with active cancer, that is, less than every ninth stroke [Figure 5(b); Supplemental Table 15]. The three different groups of strokes associated with active cancer were differentially enriched among the strokes with probabilities above the cutoffs. While strokes associated with LPC cancer were enriched 14.3-fold among strokes with probabilities above the cutoff (compared those with probabilities equal or below), those associated with other types of active cancer and BP cancer were 4.5- and 3.7-fold enriched, respectively [Figure 5(b)].

Performance of the new approach to identify paraneoplastic ischemic stroke in comparison to previous models

Finally, we compared the sensitivity, specificity as well as PPV of all three consecutive approaches to identify stroke associated with active cancer, especially with LPC cancers which we assumed to be more likely truly paraneoplastic [Figure 5(c)]. In comparison to the original model published by us[6], the new composite territory-dependent model exhibited a specificity to detect both ischemic strokes associated with active tumors in general as well as for those associated with active LPC cancers which was quite similar to the previously published model [Figure 5(c); Supplemental Table 16]. Of note, while the sensitivity to detect ischemic strokes associated with active tumors in general did not change much from 52.2% to 61.2%, the sensitivity to detect ischemic stroke associated with LPC cancers increased from to 59.3% to 77.8%, almost to the level of the second model which lacked specificity. While the PPV to detect ischemia associated with all active cancers remained unchanged (13.1% *versus* 11.8%), it increased 1.24-fold for stroke associated with LPC cancers [Figure 5(c), middle panel]. Consequently, the enrichment of LPC-cancer-associated strokes in the stroke above cutoffs compared to those below was 2.2- and 1.7-fold higher in the LPC-optimized compared to the previous models [Figure 5(c)]. The LPC-optimized model tended to enrich strokes associated with BP cancers slightly less and these associated with other active tumors slightly more
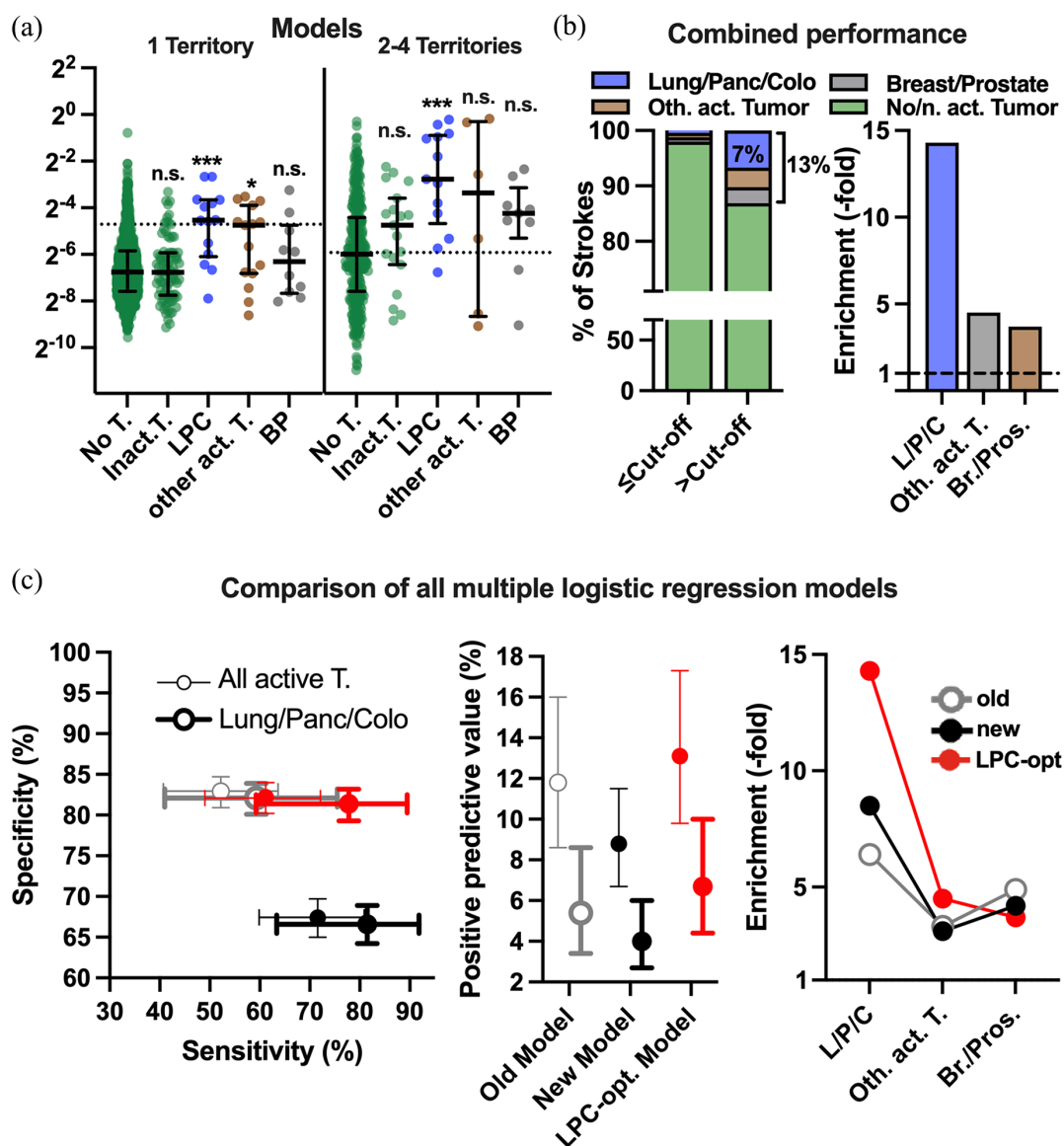
Figure 5. Optimized model taking into account the differences of mono- and multiterritorial stroke associated with cancer using lung, pancreatic, and colorectal cancer as surrogate parameter to select for paraneoplastic stroke. (a) Distribution of probabilities generated by the multiple logistic regression models for strokes with one territory only (left panel) with CRP and percentage of granulocytes as independent variables or for multiterritorial strokes with number of territories, CRP, age, and sex as independent variables (right panel). Strokes not associated with tumor (No T.) and with inactive Tumor (Inact. T.) are shown in green, strokes either associated with lung, pancreatic, or colorectal cancer (L/P/C) in blue, strokes associated with breast or prostate cancer (Br./Pros.) in gray, and all other active tumors (Oth. act. T.) in brown. The respective cutoffs (1 territory: 0.04131, 2–4 territories: 0.01049) are indicated by dotted lines. The mean and interquartile range for each group is indicated. [(b), left panel] Percentage of strokes with active tumors in the group of strokes with probability below or above the respective cutoff. The percentage of correctly positive strokes associated with lung, pancreatic, and colorectal cancer (blue) and all active cancers (bracket) is indicated. (Right panel) Enrichment of the three different groups of cancer-associated strokes by applying the two territory-dependent cutoffs among the strokes with probabilities above cutoff compared to the strokes with probabilities equal and below the cutoffs. (c) Sensitivity and specificity (left panel), positive predictive value (middle panel) as well as enrichment of strokes with different active tumors (thin lines in left and middle panel) or L/P/LC cancer (bold lines in left and middle panel) in the total current cohort after exclusion of stroke associated with endocarditis in the model optimized to detect stroke associated with L/P/C cancer (red circles), the new logistic regression model (black circles), and the model published previously (Kassubek *et al.*[6]; empty circles, gray). The 95% confidence intervals are shown for sensitivity, specificity, and the positive predictive value. Statistical analysis in (a) was performed using Kruskal–Wallis tests with Dunn's post-tests compared to the cohort without tumor.
***p < 0.001.
CRP, C-reactive protein; GRN, granulocytes; LDH, lactate dehydrogenase; n.s., not significant.

than the other two models [Figure 5(c)]. In summary, our analysis suggests that the application of our territory-dependent LPC cancer-optimized models might not only identify stroke patients with a high likelihood of paraneoplastic stroke but might also be informative whether an ischemic stroke is of paraneoplastic origin in a patient with non-LPC cancer.

Discussion

It is well established that the risk of both venous[5] and arterial thromboembolism[1] is increased in newly diagnosed malignancies. Arterial thromboembolism in cancer frequently presents as ischemic stroke.[1,7] Typically, paraneoplastic ischemic stroke affects multiple if not all vascular territories of the brain.[37] Due to often relentlessly recurring ischemic events, it is associated with a bleak prognosis.[10,16]

Paraneoplastic intravascular coagulopathy is assumed to be the underlying cause in most cases of paraneoplastic ischemic stroke.[38] Pathological activation of coagulation is reflected by highly increased D-dimers, repeatedly reported to be present in cancer-associated stroke.[20,39–41] Of note, D-dimers as markers of paraneoplastic intravascular coagulopathy are much higher in patients with assumed paraneoplastic stroke compared to those suffering from the same type of cancer only.[42] Why and how some tumors induce paraneoplastic coagulopathy while others of the same tumor type do not remains to be explored. Overall, adenocarcinomas seem to be associated with the highest risk of paraneoplastic coagulopathy.[28,41] In addition, in paraneoplastic venous thrombosis, D-dimers strongly correlate with inflammatory blood markers.[43] Interestingly, high blood levels of some epithelial glycoproteins commonly used as tumor markers for carcinomas, that is, CA-125 aka Mucin-16, are associated with paraneoplastic ischemic stroke.[44] Experimentally, injecting these proteins into the circulation directly activates coagulation through adhesion-dependent, reciprocal activation of neutrophils and platelets.[45] Of note, mucins (including Mucin-16) are upregulated by inflammatory stimuli in keeping with their biological function as constituents of mucus protecting epithelial membranes upon infection.[46] This regulation was demonstrated to remain active in various carcinoma cells *in vitro*.[47] Thus, it was concluded that the tight correlation of inflammatory cytokines and Mucin-16 expression in carcinoma tissue

results from the induction of Mucin-16 expression due to intratumoral inflammation rather than cytokine induction as a result of mucin expression.[47] Not surprisingly, not only D-dimers but also CRP strongly correlates with CA-125 blood levels in patients with cancer-associated ischemic stroke.[41] Although the value of elevated D-dimer levels has been extensively proven in many studies in this context,[16,20] D-dimers are not always part of routine blood analyses and might only represent the most distal event of paraneoplastic coagulopathy. In our previous study,[6] we demonstrated that a multiple model comprising the basic routine laboratory parameters CRP, LDH, as well as the percentage of granulocytes in addition to the number of ischemic cerebral territories correctly classified stroke associated with solid cancer with a sensitivity of 78% and a specificity of 95% in a case-control setting. We now conducted a population-based confirmatory study comprising consecutive 1612 ischemic strokes including 67 with active cancer. Although applying of our previous model to the current dataset confirmed adequate specificity of 83%, sensitivity was poor with 52%. In addition, a new similar model based on the current data did not yield to relevant improvement.

Of note, ischemic stroke and most common cancers share age as a common risk factor.[17] Especially, cancers with high prevalence in the elderly population that is also at high risk for stroke, BP cancers, might constitute a relevant proportion of cancer patients with ischemic stroke by mere coincidence of both diseases. Moreover, excessive alcohol consumption and smoking are established cancer risk factors, while smoking increases risk for both large artery atherosclerosis and atrial fibrillation and alcohol for atrial fibrillation, both underlying common conventional stroke pathophysiologies.[33–35] In theory, those ischemic strokes coincident with active cancer but not linked to paraneoplastic coagulopathy should resemble those due to conventional stroke risk factors without cancer association.[48] This contamination of the cancer-associated strokes with nonparaneoplastic strokes will limit the discriminatory power of any model that tries to detect paraneoplastic strokes just based on their association with active cancer in general. To circumvent potential circular conclusions, for example, by defining paraneoplastic stroke as cancer-associated stroke with features reportedly typical for paraneoplastic stroke such as multiterritorial involvement[20,49] or selecting cancer-associate

ischemic stroke without conventional stroke risk factors only as done in some previous studies,[20,50] we opted to select ischemic strokes associated with carefully selected active cancers to increase the probability of paraneoplastic stroke among cancer-associated strokes. We hypothesized that in association with some types of cancer, strokes due to paraneoplastic coagulopathy predominate over those due to known or unknown conventional risk factors. We assumed that these types of stroke-associated cancers should meet three of the following criteria: (1) enrichment in our stroke population compared to the general population according to the German Cancer Registry and, when associated with an ischemic stroke, these are characterized by (2) more territories involved, (3) higher CRP, and (4) higher LDH levels than our control population with ischemic stroke without active cancer. In our total 'stroke with cancer' group (combining all cancer-associated strokes of the current and the previous cohort), LPC and BP cancer were the most common cancer diagnoses. Only the first three, LPC cancers, met these criteria. In line with our categorization, a recent study demonstrated that the prevalence of BP cancers was not higher among 85,893 Danish patients with ischemic stroke compared to a 10-fold larger Danish control population, while lung and colorectal cancer were clearly overrepresented in stroke patients.[51] In a study from the United States based on Medicare data, of the three most common types of cancer, lung and breast cancers but not prostate cancer were associated with an increased hazard ratio (HR) for ischemic stroke, with the HR of breast cancer being 1.6 lower than for lung cancer.[8] In addition, stroke associated with lung cancer was reported to be associated with both higher D-dimers and CRP, while in breast cancer-associated ischemic stroke, only CRP was found upregulated and none of the two markers were upregulated in prostate cancer-associated ischemic stroke.[22]

Our more detailed analysis revealed that for strokes associated with LPC cancers, but neither breast and prostate cancers nor bacterial endocarditis, steep increases in the average number of ischemic territories depending on both higher CRP and LDH levels were characteristic. Our finding corresponds quite well to observations that CRP positively correlates with D-dimers in paraneoplastic coagulopathy associated with venous thrombosis[43] and higher D-dimer levels

in turn with a higher number of ischemic territories in cancer-associated stroke.[19,47] In summary, our findings reported here as well as previous reports are compatible with the hypothesis that the extent to which certain tumors, especially LPC cancer, induce inflammation dictates the activation of coagulation and thereby multiple ischemic events typical for paraneoplastic ischemic stroke.

Based on this assumption, we set out to establish and evaluate a strategy to detect LPC-cancer-associated stroke. Taking the much lower frequency of tumor-associated strokes among strokes involving only one territory into account, two separate models for strokes with either one or multiple territories were generated, both with cut-offs adjusted to yield a PPV of 10% for detecting LPC-cancer-associated stroke among all strokes associated with LPC cancer of this and the previous cohort combined with all ischemic strokes not associated with active cancer of the current cohort. When applied to the total current cohort, a PPV for active cancer-associated stroke in general of 13% with a unique enrichment of LPC tumors (PPV 7%) was achieved. Of note, our data demonstrate that bacterial endocarditis has to be excluded before testing whether a certain stroke might be paraneoplastic, as endocarditis-associated strokes will inevitably be categorized as paraneoplastic, as they are characterized by multiterritorial ischemic lesions, high CRP, and high LDH levels. Interestingly, in contrast to paraneoplastic stroke, prominent granulocytosis seems to be characteristic for endocarditis-associated stroke.

Ischemic strokes associated with active LPC cancers as our surrogate marker for paraneoplastic stroke were detected with a fair sensitivity and specificity of 77.8% and 81.4%, respectively. Using D-dimers only to predict ischemic stroke associated with active cancer in a retrospective monocentric cohort study quite similar to ours, a sensitivity of 79.6% and a specificity of 66.7% at a D-dimer cutoff of >0.55 mg/l were found.[39] A predictive score for cancer-associated stroke based on increased serum levels of D-dimer and fibrinogen and the absence of hyperlipidemia yielded a sensitivity of 68% and a specificity of 88%,[40] a quite similar approach combining elevated D-dimers (≥3 mg/l), anemia (Hb ≤ 12.0 g/dl), and history of smoking yielded a sensitivity of 53% and specificity of 83%.[21] Thus, our approach

seems at least equivalent to others, however, with the advantage that basic routine laboratory parameters are used (not D-dimers). This has the important advantage that the incidental finding of elevated D-dimers in clinical routine might result in unnecessary or even harmful investigations, for example, contrast-enhanced CT to exclude pulmonary embolism.[52]

Of note, our study has several limitations. First, the retrospective nature of our exploratory study and the considerable overlap of the training cohort with the cohort used to test our algorithm has to be considered as a limitation. Thus, all results must be categorized as merely hypothesis generating and before any clinical application, our model needs to be tested by a second external validation cohort. Second, upon the mandatory future prospective testing of our model, it may become apparent that our model for detection of uniterritorial paraneoplastic stroke might be less suited to identify cerebral ischemia associated with yet unknown, and therefore untreated, cancer, as low granulocytes in the presence of CRP in the lower range, possibly a result of chemotherapy-induced granulocytopenia, will increase the probability of active cancer according to this model. Third, a PPV of 13% means that only one in eight patients with a cancer probability above the cutoff will indeed suffer from active cancer. Thus, those patients tested likely to suffer from potentially paraneoplastic ischemic stroke upon application of our algorithm will have to undergo additional diagnostic procedures. Based on the data currently available, testing for both D-dimers[19,21] and CA-125[44,53] might be the most promising strategy prior to perform more laborious or even invasive procedures including X-rays, ultrasound, or colonoscopy for cancer detection.

In summary, the proposed combination of levels of basic routine laboratory parameters for inflammation and cell turnover together with the number of vascular territories involved in ischemic stroke seems to be a very promising screening tool to identify patients with high risk of paraneoplastic stroke. The next mandatory steps will be to validate the algorithm described here using a larger multicenter confirmatory cohort and to prospectively test which consecutive diagnostic steps are the most appropriate in those stroke patients with high tumor risk.

Declarations

Ethics approval and consent to participate
The study has been approved by the Ethics Committee of Ulm University, Germany (reference No. 133/17) and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. For this retrospective study on clinical data (without any data from an individual person), a written consent for participation or publication was waived and thus not sought.

Consent for publication
Not applicable.

Author contributions
Rebecca Kassubek: Conceptualization; Data curation; Investigation; Project administration; Supervision; Validation; Writing – original draft.

Marc-Andre G. R. Winter: Investigation; Writing – review & editing.

Jens Dreyhaupt: Formal analysis; Investigation; Methodology; Writing – review & editing.

Mona Laible: Investigation; Writing – review & editing.

Jan Kassubek: Data curation; Investigation; Writing – review & editing.

Albert C. Ludolph: Investigation; Writing – review & editing.

Jan Lewerenz: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Writing – original draft.

Competing interests
The authors declare that there is no conflict of interest.

Availability of data and materials
The data presented in this study are available from the corresponding author on reasonable request.

ORCID iDs
Rebecca Kassubek [iD] https://orcid.org/0000-0001-5384-7334

Jan Kassubek [iD] https://orcid.org/0000-0002-7106-9270

References

1. Naschitz JE. Cancer-associated atherothrombosis: the challenge. *Int J Angiol* 2021; 30: 249–256.

2. Graus F, Rogers LR and Posner JB. Cerebrovascular complications in patients with cancer. *Medicine (Baltimore)* 1985; 64: 16–35.

3. Navi BB, Reiner AS, Kamel H, *et al.* Association between incident cancer and subsequent stroke. *Ann Neurol* 2015; 77: 291–300.

4. Salazar-Camelo RA, Moreno-Vargas EA, Cardona AF, *et al.* Ischemic stroke: a paradoxical manifestation of cancer. *Crit Rev Oncol Hematol* 2021; 157: 103181.

5. Timp JF, Braekkan SK, Versteeg HH, *et al.* Epidemiology of cancer-associated venous thrombosis. *Blood* 2013; 122: 1712–1723.

6. Kassubek R, Bullinger L, Kassubek J, *et al.* Identifying ischemic stroke associated with cancer: a multiple model derived from a case-control analysis. *J Neurol* 2017; 264: 781–791.

7. Lun R, Roy DC, Hao Y, *et al.* Incidence of stroke in the first year after diagnosis of cancer-a systematic review and meta-analysis. *Front Neurol* 2022; 13: 966190.

8. Navi BB, Howard G, Howard VJ, *et al.* New diagnosis of cancer and the risk of subsequent cerebrovascular events. *Neurology* 2018; 90: e2025–e2033.

9. Grazioli S, Paciaroni M, Agnelli G, *et al.* Cancer-associated ischemic stroke: a retrospective multicentre cohort study. *Thromb Res* 2018; 165:33–37.

10. Kneihsl M, Enzinger C, Wunsch G, *et al.* Poor short-term outcome in patients with ischaemic stroke and active cancer. *J Neurol* 2016; 263:150–156.

11. Navi BB, Singer S, Merkler AE, *et al.* Recurrent thromboembolic events after ischemic stroke in patients with cancer. *Neurology* 2014; 83: 26–33.

12. Seystahl K, Hug A, Weber SJ, *et al.* Cancer is associated with inferior outcome in patients with ischemic stroke. *J Neurol* 2021; 268: 4190–4202.

13. Kawano H, Honda Y, Amano T, *et al.* Subcutaneous heparin therapy for patients with cancer-associated stroke. *J Stroke Cerebrovasc Dis* 2019; 28: 399–404.

14. Yamaura G, Ito T, Miyaji Y, *et al.* Therapeutic efficacy of heparin and direct factor Xa inhibitors in cancer-associated cryptogenic ischemic stroke with venous thromboembolism. *Thromb Res* 2021; 206: 99–103.

15. Jang H, Lee JJ, Lee MJ, *et al.* Comparison of enoxaparin and warfarin for secondary prevention of cancer-associated stroke. *J Oncol* 2015; 2015: 502089.

16. Lee MJ, Chung JW, Ahn MJ, *et al.* Hypercoagulability and mortality of patients with stroke and active cancer: the OASIS-CANCER study. *J Stroke* 2017; 19: 77–87.

17. Woock M, Martinez-Majander N, Seiffge DJ, *et al.* Cancer and stroke: commonly encountered by clinicians, but little evidence to guide clinical approach. *Ther Adv Neurol Disord* 2022; 15: 17562864221106362.

18. Kikuno M, Ueno Y, Takekawa H, *et al.* Distinction in prevalence of atherosclerotic embolic sources in cryptogenic stroke with cancer status. *J Am Heart Assoc* 2021; 10: e021375.

19. Wang JY, Zhang GJ, Zhuo SX, *et al.* D-dimer >2.785 μg/ml and multiple infarcts ≥3 vascular territories are two characteristics of identifying cancer-associated ischemic stroke patients. *Neurol Res* 2018; 40: 948–954.

20. Xu Y, Wu Z and Xu H. Cancer-related cryptogenic stroke involving the bilateral anterior and the posterior circulations: diagnostic value of clinical and imaging characteristics. *Front Neurol* 2022; 13: 1032984.

21. Selvik HA, Bjerkreim AT, Thomassen L, *et al.* When to screen ischaemic stroke patients for cancer. *Cerebrovasc Dis* 2018; 45: 42–47.

22. Selvik HA, Thomassen L, Bjerkreim AT, *et al.* Cancer-associated stroke: the Bergen NORSTROKE study. *Cerebrovasc Dis Extra* 2015; 5: 107–113.

23. Cocho D, Gendre J, Boltes A, *et al.* Predictors of occult cancer in acute ischemic stroke patients. *J Stroke Cerebrovasc Dis* 2015; 24: 1324–1328.

24. Park H, Kim J, Ha J, *et al.* Histological features of intracranial thrombi in stroke patients with cancer. *Ann Neurol* 2019; 86: 143–149.

25. Gouin-Thibault I and Samama MM. Laboratory diagnosis of the thrombophilic state in cancer patients. *Semin Thromb Hemost* 1999; 25: 167–172.

26. Heikkila K, Ebrahim S and Lawlor DA. A systematic review of the association between circulating concentrations of C-reactive protein and cancer. *J Epidemiol Community Health* 2007; 61: 824–833.

27. Anai M, Saruwatari K, Ikeda T, *et al*. Clinical impact of cerebral infarction in patients with non-small cell lung cancer. *Int J Clin Oncol* 2022; 27: 863–870.

28. Xie X, Chen L, Zeng J, *et al*. Clinical features and biological markers of lung cancer-associated stroke. *J Int Med Res* 2016; 44: 1483–1491.

29. Conteduca V, Scarpi E, Wetterskog D, *et al*. Plasma tumor DNA is associated with increased risk of venous thromboembolism in metastatic castration-resistant cancer patients. *Int J Cancer* 2022; 150: 1166–1173.

30. Nitta S, Kawai K, Kimura T, *et al*. Predictors of venous thromboembolism development before and during chemotherapy for advanced germ cell tumor. *Jpn J Clin Oncol* 2020; 50: 338–343.

31. Muir KW, Tyrrell P, Sattar N, *et al*. Inflammation and ischaemic stroke. *Curr Opin Neurol* 2007; 20: 334–342.

32. Hu W, Wang X and Su G Infective endocarditis complicated by embolic events: pathogenesis and predictors. *Clin Cardiol* 2021; 44: 307–315.

33. Belbasis L, Mavrogiannis MC, Emfietzoglou M, *et al*. Environmental factors, serum biomarkers and risk of atrial fibrillation: an exposure-wide umbrella review of meta-analyses. *Eur J Epidemiol* 2020; 35:223–239.

34. Mast H, Thompson JL, Lin IF, *et al*. Cigarette smoking as a determinant of high-grade carotid artery stenosis in Hispanic, black, and white patients with stroke or transient ischemic attack. *Stroke* 1998; 29: 908–912.

35. Jun S, Park H, Kim UJ, *et al*. Cancer risk based on alcohol consumption levels: a comprehensive systematic review and meta-analysis. *Epidemiol Health* 2023; 45: e2023092.

36. Rosell A, Martinod K, Mackman N, *et al*. Neutrophil extracellular traps and cancer-associated thrombosis. *Thromb Res* 2022; 213(Suppl. 1): S35–S41.

37. Nouh AM, Staff I and Finelli PF. Three territory sign: an MRI marker of malignancy-related

38. Bang OY, Chung JW, Lee MJ, *et al*. Cancer-related stroke: an emerging subtype of ischemic stroke with unique pathomechanisms. *J Stroke* 2020; 22: 1–10.

39. Guo YJ, Chang MH, Chen PL, *et al*. Predictive value of plasma (D)-dimer levels for cancer-related stroke: a 3-year retrospective study. *J Stroke Cerebrovasc Dis* 2014; 23: e249–e254.

40. Jiang J, Shang X, Zhao J, *et al*. Score for predicting active cancer in patients with ischemic stroke: a retrospective study. *Biomed Res Int* 2021; 2021: 5585206.

41. Nezu T, Hosomi N, Naito H, *et al*. Clinical characteristics and tumor markers in ischemic stroke patients with active cancer. *Intern Emerg Med* 2022; 17: 735–741.

42. Navi BB, Sherman CP, Genova R, *et al*. Mechanisms of ischemic stroke in patients with cancer: a prospective study. *Ann Neurol* 2021; 90: 159–169.

43. Setiawan B, Budianto W, Sukarnowati TW, *et al*. Correlation of inflammation and coagulation markers with the incidence of deep vein thrombosis in cancer patients with high risk of thrombosis. *Int J Gen Med* 2022; 15: 6215–6226.

44. Jovin TG, Boosupalli V, Zivkovic SA, *et al*. High titers of CA-125 may be associated with recurrent ischemic strokes in patients with cancer. *Neurology* 2005; 64: 1944–1945.

45. Shao B, Wahrenbrock MG, Yao L, *et al*. Carcinoma mucins trigger reciprocal activation of platelets and neutrophils in a murine model of Trousseau syndrome. *Blood* 2011; 118: 4015–4023.

46. van Putten JPM and Strijbis K. Transmembrane mucins: signaling receptors at the intersection of inflammation and cancer. *J Innate Immun* 2017; 9: 281–299.

47. Morgado M, Sutton MN, Simmons M, *et al*. Tumor necrosis factor-alpha and interferon-gamma stimulate MUC16 (CA125) expression in breast, endometrial and ovarian cancers through NFκB. *Oncotarget* 2016; 7: 14871–14884.

48. Kono T, Ohtsuki T, Hosomi N, *et al*. Cancer-associated ischemic stroke is associated with elevated D-dimer and fibrin degradation product levels in acute ischemic stroke with advanced cancer. *Geriatr Gerontol Int* 2012; 12: 468–474.

49. Finelli PF and Nouh A. Three-territory DWI acute infarcts: diagnostic value in

ischemic stroke (Trousseau syndrome). *Neurol Clin Pract* 2019; 9: 124–128.

cancer-associated hypercoagulation stroke (Trousseau syndrome). *AJNR Am J Neuroradiol* 2016; 37: 2033–2036.

50. Liu M, Ellis D, Duncan A, *et al*. The utility of the markers of coagulation and hemostatic activation profile in the management of embolic strokes of undetermined source. *J Stroke Cerebrovasc Dis* 2021; 30: 105592.

51. Tybjerg AJ, Babore AD, Olsen TS, *et al*. Types of occult cancer in stroke and the relation to

smoking. *Acta Neurol Scand* 2020; 142: 486–492.

52. Moore AJE, Wachsmann J, Chamarthy MR, *et al*. Imaging of acute pulmonary embolism: an update. *Cardiovasc Diagn Ther* 2018; 8: 225–243.

53. Maezono-Kandori K, Ohara T, Fujinami J, *et al*. Elevated CA125 is related to stroke due to cancer-associated hypercoagulation. *J Stroke Cerebrovasc Dis* 2021; 30: 106126.