Academic Emergency Medicine
Official Journal of the Society for Academic Emergency Medicine

ORIGINAL RESEARCH CONTRIBUTION

Interpreting Red Blood Cells in Lumbar Puncture: Distinguishing True Subarachnoid Hemorrhage From Traumatic Tap

Amanda D. Czuczman, MD, Lisa E. Thomas, MD, Alyson B. Boulanger, MA, David A. Peak, MD, Emily L. Senecal, MD, David F. Brown, MD, and Keith A. Marill, MD

Abstract

Objectives: The study purpose was to determine the optimal use of lumbar puncture (LP) red blood cell (RBC) counts to identify subarachnoid hemorrhage (SAH) when some blood remains in the final tube.

Methods: A case series was performed at a tertiary emergency department (ED). Records of 4,496 consecutive adult patients billed for LPs between 2001 and 2009 were reviewed. Inclusion criteria were headache (HA), final tube RBCs $\geq$ 5, and neurovascular imaging within 2 weeks of the visit. Demographics, relevant history and physical examination components, LP results, and neuroimaging findings were recorded for 280 patients. True-positive (TP) and true-negative (TN) SAH were strictly defined. Primary outcomes were the areas under the receiver operating characteristic curves (AUC) for final tube RBC count, differential RBC count between the final and initial tubes, and absolute differential RBC count between the final and initial tubes divided by the mean RBC count of the two tubes (also called the percent change in RBC count).

Results: There were 26 TP and 196 TN results; 58 patients were neither. The TP group consisted of 19 patients with visible or possible SAH on imaging (17 on noncontrast head computed tomography [CT; 12 definite and five possible] and two on magnetic resonance imaging), six with xanthochromia and a vascular lesion (aneurysm or arteriovenous malformation [AVM] > 2 mm), and one with xanthochromia and polymerase chain reaction (PCR)-positive meningitis. As a test for SAH, final tube RBC AUC was 0.85 (95% confidence interval [CI] = 0.80 to 0.91). Interval likelihood ratios (LRs) for final tube RBC count were LR 0 (95% CI = 0 to 0.3) for RBCs < 100, LR 1.6 (95% CI = 1.1 to 2.3) for 100 < RBCs < 10,000, and LR 6.3 (95% CI = 4.8 to 23.4) for RBCs > 10,000. Differential RBC count was not associated with SAH, with AUC 0.45 (95% CI = 0.31 to 0.60). However, the percent change in RBC count between the final and initial tubes had an AUC 0.84 (95% CI = 0.78 to 0.90), and the optimal test threshold for SAH was 0.63, with positive LR 3.6 (95% CI = 2.7 to 4.7) and negative LR 0.10 (95% CI = 0.03 to 0.4) for percent change <63% and >63%, respectively. This test added additional independent information to the final tube RBC count based on improved logistic regression model fit and discriminatory ability as measured by the LR test and c statistic, respectively.

Conclusions: Final LP tube RBC count and the percent change in RBC count, but not the simple differential count between the final and initial tubes, were associated with SAH. In this sample, there were no patients with SAH who had RBCs < 100 in the final tube, and RBCs > 10,000 increased the odds of SAH by a factor of 6.

ACADEMIC EMERGENCY MEDICINE 2013; 20:247–256 © 2013 by the Society for Academic Emergency Medicine

From the Department of Emergency Medicine, Beverly Hospital (ADC), Beverly, MA; the Department of Emergency Medicine, Mount Auburn Hospital (LET), Cambridge, MA; and the Department of Emergency Medicine, Massachusetts General Hospital (ABB, DAP, ELS, DFB, KAM), Boston, MA.

Received June 4, 2012; revision received July 30, 2012; accepted September 22, 2012.

Presented at the American College of Emergency Physicians annual meeting, San Francisco, CA, October 2011.

The authors have no relevant financial information or potential conflicts of interest to disclose.

Supervising Editor: Clifton W. Callaway, MD, PhD.

Address for correspondence and reprints: Amanda D. Czuczman, MD; e-mail: amandaczuczman@gmail.com.

The current diagnostic algorithm for the evaluation of subarachnoid hemorrhage (SAH) in the emergency department (ED) is straightforward: noncontrast head computed tomography (CT) and, if negative, lumbar puncture (LP).[1,2] The rationale for this approach is based on the fact that noncontrast head CT is imperfect, with an estimated 93% to 100% sensitivity in detecting acute SAH, which further decreases after 24 hours.[3–5] Performing an LP is especially important in patients being evaluated for SAH who are alert and have normal neurologic exams because these patients are more likely to have a negative head CT scan than are patients with obvious neurologic deficits.[6,7] A neurologically normal patient with a negative head CT

doi: 10.1111/acem.12095
ISSN 1069-6563
PII ISSN 1069-6563583

may still have as high as a 7% to 10% chance of having an SAH, and this is deemed an unacceptable miss rate that mandates further testing.[8,9]

Although the diagnostic algorithm for SAH is clear, the interpretation of the LP data is not. When an LP is performed in the ED, there is often confusion about whether cerebrospinal fluid (CSF) with red blood cells (RBCs) is due to a true SAH versus a traumatic tap, which may be caused by incidental perforation of the epidural venous plexus or cauda equina vessels. Traumatic taps are not uncommon, occurring in an estimated 10% of all ED LPs.[10,11]

It is customary to collect four serial tubes of CSF when performing an LP. SAH is confirmed if xanthochromia is present, and SAH is clearly excluded if fewer than five RBCs are found in the final tube. Xanthochromia is indicative of heme metabolism and takes up to 12 hours to develop after SAH, but it is not expected acutely after traumatic tap. When any number of RBCs greater than five is found in the final tube, there is no reliable method to distinguish traumatic LP from true SAH based on the RBC count.[12] It is generally accepted that in true SAH, there are persistent RBCs from tube 1 to tube 4, usually in the range of 1,000s. However, there is no clear guideline for the number of RBCs required to diagnose SAH.

The purpose of this investigation was to identify the optimal use of LP RBC data in identifying true SAH when some blood remains in the final tube. We hypothesized that the differential RBC count between the final and initial tubes is the best LP test characteristic to identify true SAH, as measured by the area under the receiver operating characteristic curve (AUC).

METHODS

Study Design

This was a retrospective case series of all consecutive adult ED patients billed for LPs between 2001 and 2009. The study was approved by our institutional review board.

Study Setting and Population

This study was conducted at an urban tertiary care emergency center with 90,000 annual visits that serves as a referral receiving center for neurologic emergencies. Computer records of all ED patients older than 16 years billed for LPs between 2001 and 2009 were initially reviewed. To identify the subset of patients with symptoms concerning for spontaneous SAH who had abnormal LP results and had neurovascular imaging performed, we used the following inclusion criteria: presence of headache (HA), $\geq$ 5 RBCs in the final LP tube collected, and CT angiography (CTA) or magnetic resonance angiography (MRA) performed within 2 weeks of the ED visit. Exclusion criteria consisted of presence of ventriculoperitoneal shunt, neurosurgery within 4 weeks preceding the ED visit, CSF sent primarily for cytology, unequivocal history of trauma within 2 weeks preceding the ED visit, failed LP, or no LP performed at our hospital (i.e., no CSF sent).

Study Protocol

Two unblinded emergency physician (EP) reviewers (ADC, LET) abstracted the medical records of patients meeting the inclusion criteria. Differences were resolved by consensus. This comprehensive chart review involved extracting information from the written and computer medical records using a standardized protocol and collection form and entering this data into a Microsoft Excel (Microsoft Corp., Redmond, WA) spreadsheet. Data collected included patient demographics (age, sex), date and time of ED presentation, history of present illness characteristics (duration and quality of HA, worst HA of life, exertion at HA onset), past medical history (hypertension, chronic HA disorder, smoking, history of SAH, warfarin use), social history including current tobacco use and family history of aneurysm, relevant physical exam findings such as neurologic deficits, CSF results (date and time of LP, all LP tube RBC counts), presence of xanthochromia as determined by visual inspection at the laboratory, presence of meningitis proven by culture or polymerase chain reaction (PCR), and all relevant neurovascular imaging results (namely the presence of SAH or a neurovascular lesion such as aneurysm or arteriovenous malformation [AVM]) as interpreted by a neuroradiologist. Two blinded EP reviewers (DAP, ELS) each reviewed a different half of the included records and abstracted these same data points for reliability analysis. When there were discrepancies between the blinded and nonblinded review, a third nonblinded EP (ADC or LET) reviewed the data to determine the final outcome.

Because the aim of our study was to learn how to most effectively interpret CSF RBC data, the best criteria available to use in our criterion standard definition for a true SAH were positive neuroimaging results and the presence of xanthochromia. Therefore, a true-positive (TP) SAH was defined, prior to any analysis, as: 1) presence of SAH on imaging *or* 2) xanthochromia and an aneurysm or AVM > 2 mm on imaging *or* 3) xanthochromia and culture- or PCR-positive meningitis. A true-negative (TN) case was defined as: 1) no SAH on imaging *and* 2) no aneurysm or AVM of any size on imaging *and* 3) no culture- or PCR-positive meningitis *and* 4) no xanthochromia after at least 12 hours of HA (to account for the amount of time it can take for xanthochromia to develop after an SAH). These definitions for TP and TN were selected to be conservative and to ensure that patients included either had definitive SAH or definitely did not have SAH. Cases that were neither TP nor TN (gray zone cases) were excluded in the primary analysis.

More relaxed definitions were created for the purpose of a sensitivity analysis. The sensitivity analysis was performed to gain a sense for the results if we tried to reclassify the gray zone cases (neither TP nor TN) most correctly. Sensitivity analysis A included in its expanded TP definition any aneurysm or AVM > 2 mm first identified on neuroimaging within 2 weeks preceding the ED presentation without requiring the presence of xanthochromia. Sensitivity analysis B incorporated the expanded TP definition of sensitivity analysis A and an expanded TN definition that allowed for negative xanthochromia of any timing (not just $\geq$12 hours after

onset of HA), in addition to the aforementioned TN criteria.

Sensitivity analysis C examined a more refined group of TP cases by excluding those with certain SAH visualized on noncontrast head CT and culture- or PCR-positive meningitis. This allowed a closer look into the challenging population of patients with SAH who have a negative or uncertain head CT.

Outcome Measures

The primary outcome measures in this study were the diagnostic utility of two commonly used methods to differentiate true SAH from traumatic tap on LP RBC analysis: the final tube RBC count and the differential RBC count between the final and initial tubes. Diagnostic utility was assessed with the AUC for each of the two tests. Additional primary outcome measures included three interval likelihood ratios (LRs) computed with 95% confidence intervals (CIs) for empirically determined optimal clinical RBC result intervals for final tube RBC count to distinguish true SAH from traumatic tap. These RBC interval cutoffs were selected based on the LR interval slopes and easy to remember threshold values.

When we inspected the data, it appeared that the TN group had higher RBC count variance between tubes where the final tube might be either much lower or much higher than the first tube. Thus using the AUC, we also analyzed a third test, the percent change in RBC count. This test is defined as the absolute differential between the final and initial tubes divided by the mean RBC count of the two tubes (because the differential is a function of the amplitude of the RBC value). We have used the average of the two RBC values as the denominator instead of simply the first RBC value, because the data suggest symmetry where either the first or the final tube RBC measurement may be closer to the true value. Thus adjustment is made for the average of both. Secondary outcomes measured included a comparison of patient characteristics between the TP and TN groups.

Data Analysis

Receiver operating characteristic (ROC) curves for the final tube RBC count, the differential RBC count between the final and initial tubes, and the percent change in RBC count were constructed using the ROC curve function in SPSS-17 software (IBM, Armonk, NY). The AUC 95% CI was computed assuming a nonparametric data distribution. For the final tube RBC ROC curve, three result intervals were empirically chosen based on optimal test characteristics to distinguish SAH from traumatic tap, and simple threshold values that would be easy to remember. The three intervals chosen were < 100 RBCs, 100 < RBCs < 10,000, and RBCs > 10,000. LRs were calculated for each interval, and LR 95% CIs were estimated using the log method.[13,14]

For the test result interval < 100 RBCs, no patients with SAH had a positive test. Therefore, sensitivity was 0 with 95% CI = 0 to 0.14 calculated with exact statistics using StatXact-3 (Cytel Software Corp., Cambridge, MA). In lieu of using an adjusted log method to calculate the LR 95% CI with sensitivity equal to zero, the upper bound of the LR 95% CI was conservatively estimated by dividing the upper bound of the sensitivity 95% CI by the lower bound of the specificity 95% CI. Using a relatively high estimate for the sensitivity and low estimate for the specificity leads to a relatively higher, less desirable, and thus more conservative estimate for the < 100 RBC LR CI upper bound.[15]

Validation of the estimated LRs in the primary analysis was assessed by also bootstrapping the 95% CIs. Open-source R software, version 2.12 (R Foundation for Statistical Computing, Vienna, Austria), was used to bootstrap 10,000 replicates by sampling with replacement, and the bias corrected and accelerated (BCa) percentile method was used to estimate the 95% CI.[16]

Sensitivity for TP is 0 of 26, and the LR is 0 for RBC < 100. Bootstrapping this sample with zero variance would lead to an unhelpful upper CI bound of zero. The point estimate for the population sensitivity is zero. However, populations with sensitivities slightly greater than zero would also be most likely to yield a sample estimate of zero. The highest population sensitivity that is still most likely to yield a point estimate of 0 of 26 must lie between 0 of 26 and 1 of 26. This population sensitivity can be located using repeated binomial sampling by identifying the median of 50,000 samples of size 26 with varying population sensitivity. A total of 50,000 samples were used to ensure stability and reproducibility. Samples can then be generated for the bootstrap as a conservative replacement for the population sensitivity estimate of zero to yield a useful LR CI.[16]

Comparisons between the TP and TN groups were made using unpaired t-tests for age, Hodges-Lehmann CI for median shift in hours from onset with Statxact-3 software, and chi-square statistics for the remaining nominal variables. To assess the reliability of data extraction of all HPI, imaging, and LP results, unweighted kappa and two-way random effects intraclass correlation coefficients were computed for nominal and interval data, respectively, using SPSS-17 software. The benefit of using both the final tube RBC count and the percent change in RBC count over the final tube RBC count alone was assessed with logistic regression modeling with manual insertion of variables using SAS 9.3 (SAS Institute, Inc., Cary, NC).

RESULTS

Characteristics of Study Subjects

Figure 1 is a flow chart depicting how the initial sample of 4,496 adult ED patients billed for LPs was tapered down to the 280 patients who underwent comprehensive chart review and finally to the TP (n = 26) and TN (n = 196) groups that were used for primary data analysis. However, 10 included patients who were all classified as TN did not have second RBC counts after the first tube, and so these 10 patients could not be included in the parts of the analysis that required repeated measures of the RBC count from two tubes. Table 1 describes the demographic and historical characteristics of the two groups, including relative differences. The TN group included only 194 patients for the nondemographic characteristics because the written medical records for two patients were unavailable. No significant

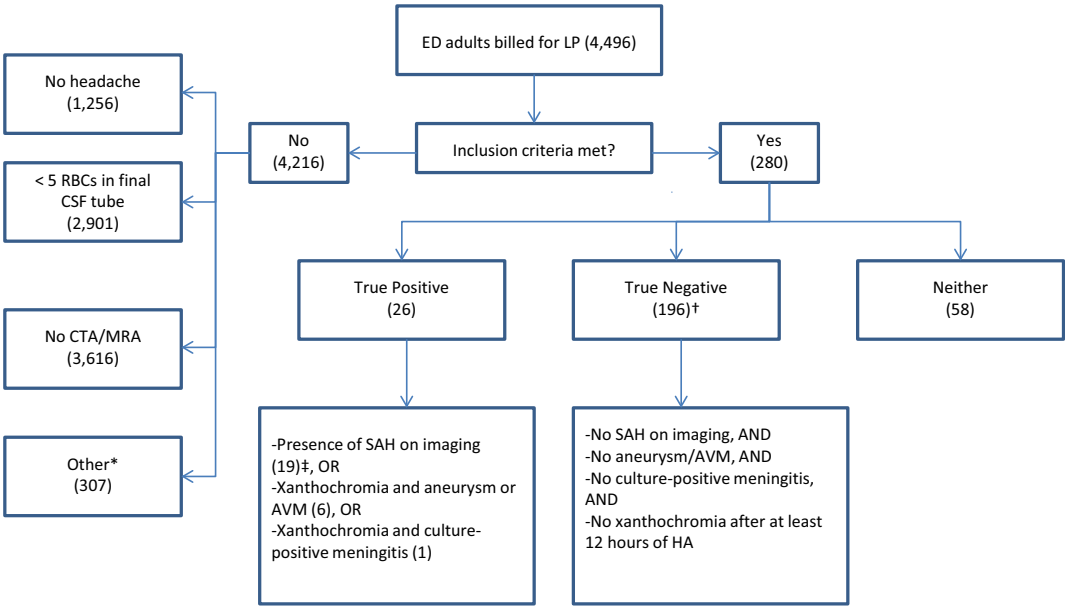

Figure 1. Patient selection. *Other: Failed LP, recent trauma, recent neurosurgery, CSF sent primarily for cytology, VP shunt. †*n* = 186 for analyses requiring RBC counts from two tubes as 10 cases only had one tube. ‡Visible SAH on noncontrast head CT (*n* = 17); visible SAH on brain MRI (*n* = 2). AVM = arteriovenous malformation; CSF = cerebrospinal fluid; HA = headache; LP = lumbar puncture; MRA = magnetic resonance angiography; MRI = magnetic resonance imaging; RBC = red blood cell; SAH = subarachnoid hemorrhage.

Table 1
Patient Demographics and Secondary Outcome Measures

Characteristic	TP (*n* = 26)	TN (*n* = 194)*	Difference† (95% CI)
Age, mean (±SD) (yrs)	50.2 (± 12.6)	42.7 (±15.2)‡	7.5 (1.4 to 13.7)
Female	21 (80.8)	112 (57.1)‡	23.7 (3.1 to 43.0)
Hours from onset, median (IQR)	60 (1.3–120.0)	48 (6.0–120.0)	Median shift 0 (–29.0 to 21.0)
Sudden/thunderclap HA	12 (46.2)	56 (28.9)	17.3 (–3.0 to 41.1)
Worst HA of life	5 (19.2)	40 (20.6)	–1.4 (–17.7 to 19.1)
Exertion at onset	2 (7.7)	22 (11.3)	–3.6 (–15.5 to 15.4)
Hx of HTN	8 (30.8)	39 (20.1)	10.7 (–7.7 to 34.9)
Chronic HA disorder	5 (19.2)	40 (20.4)	–1.2 (–17.6 to 19.2)
Current smoker	6 (23.1)	39 (20.1)	3.0 (–13.9 to 27.3)
Coumadin use	0 (0)	1 (0.5)	–0.5 (–4.8 to 13.9)
Hx of SAH	0 (0)	6 (3.1)	–3.1 (–9.0 to 12.6)
Family Hx of aneurysm	0 (0)	16 (8.2)	–8.2 (–15.7 to 8.6)

Values are reported as *n* (%) unless otherwise noted.
HA = headache; HTN = hypertension; Hx = history; SAH = subarachnoid hemorrhage; TN = true negative; TP = true positive.
**n* = 194 for most data points because we were unable to obtain the written medical record for two patients, so we only had demographic information.
†The final column refers to difference in percentage, except for age and hours from onset.
‡*n* = 196.

differences between these two groups were detected for many characteristics that we often associate with SAH, such as sudden/thunderclap HA, worst HA of life, hypertension, smoking, and family history of aneurysm.

Main Results

Figure 2 shows line plots of the number of RBCs in the initial and final tubes for the TP and TN groups. Note that the lines for the TP group tend to stay relatively constant, while the lines for the TN group often change in either direction (meaning that sometimes the number of RBCs goes down substantially between the initial and final tubes, but sometimes it goes up as well).

Figure 3 illustrates the ROC curve for three tests for SAH: the final tube RBC count, the differential RBC count between the final and initial tubes, and the percent change in the RBC count. As a test for SAH, the number of RBCs in the final tube had an AUC = 0.85 (95% CI = 0.80 to 0.91). In contrast, the differential RBC count was not predictive of SAH, with AUC = 0.45 (95% CI = 0.31 to 0.60). Interestingly, the percent change in RBC count between the final and initial tubes was also a good test for SAH, with an AUC = 0.84 (95% CI = 0.78

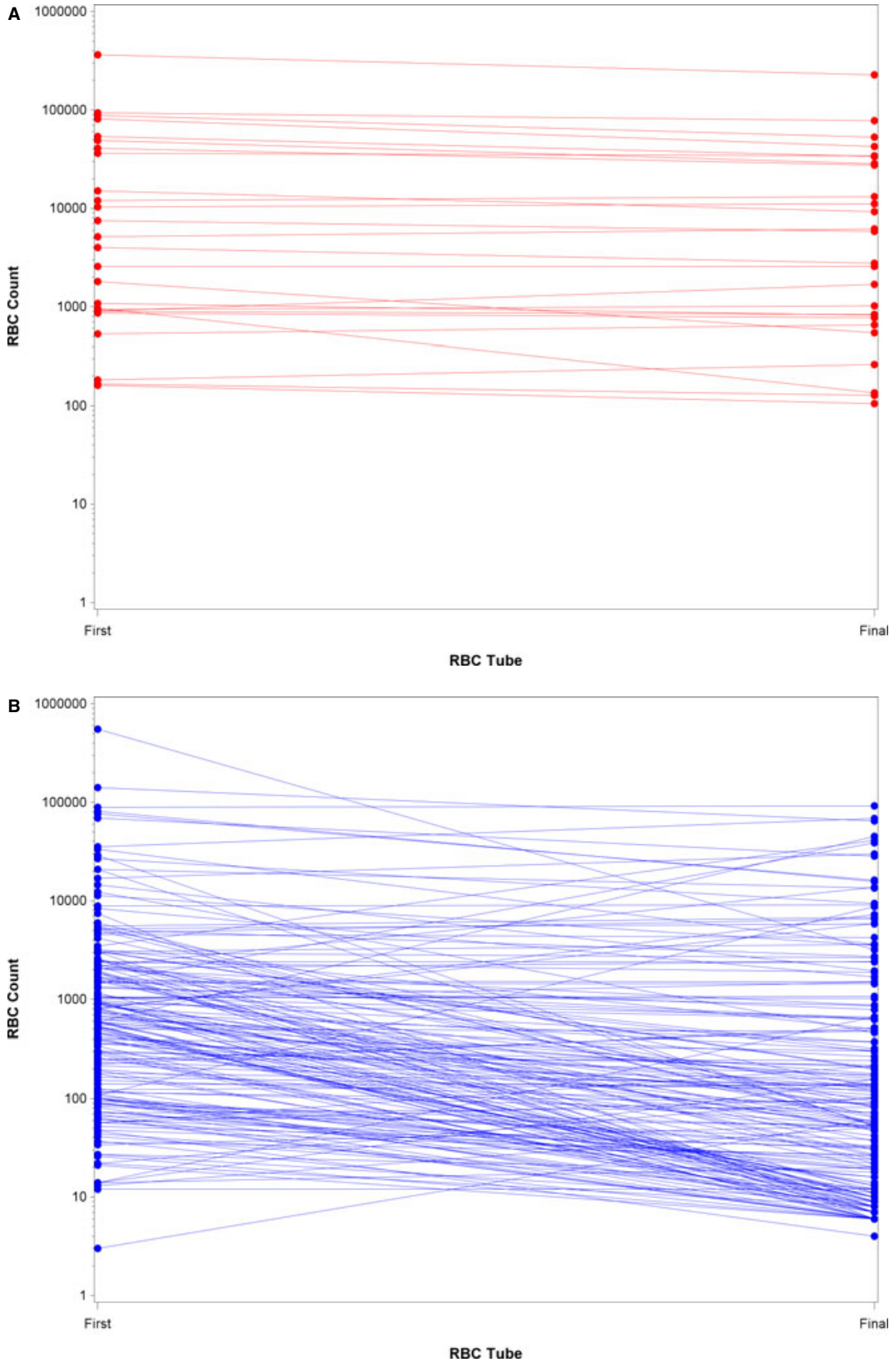

Figure 2. Line plots of the number of RBCs in the initial and final tubes for the TP and TN groups. (A) TP (n = 26). (B) TN (n = 186). RBCs = red blood cells; TN = true-negative; TP = true-positive.

to 0.90). This suggests that if either tube (initial or final) has a much lower or higher RBC count than the other tube, it is less likely to represent a true SAH.

Logistic regression modeling was performed to determine whether adding the percent change in RBC count as a predictor improved the fit and discriminatory

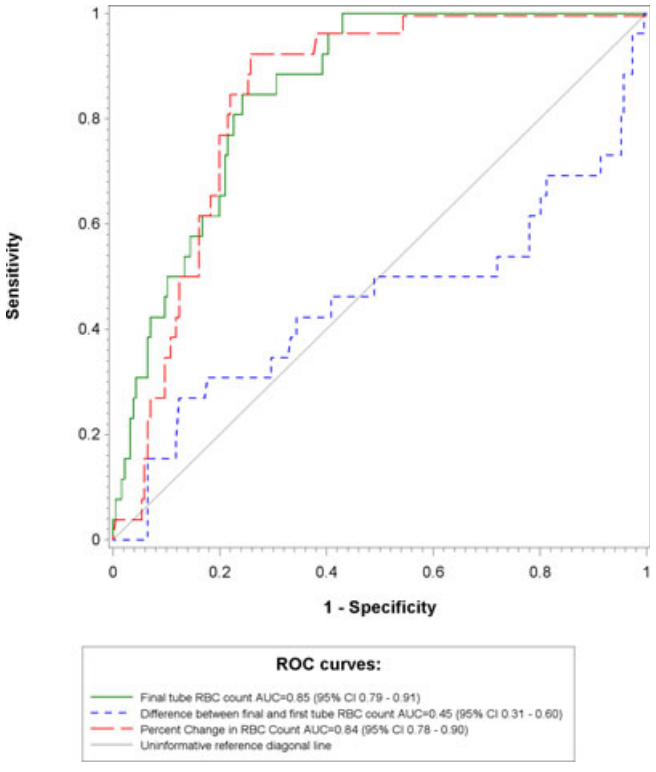

Figure 3. ROC curve for three tests for SAH: the final tube RBC count, the differential RBC count between the final and initial tubes, and the percent change in the RBC count. RBC = red blood cell; ROC = receiver operating characteristic; SAH = subarachnoid hemorrhage.

characteristics of the model that included only the final tube RBC count as a predictor. First, based on the predicted probability diagnostic plots, the two continuous predictors demonstrated reasonable linearity in the logit, and simple linear regression between the two predictors demonstrated essentially no colinearity ($R^2 = 0.03$). The LR test moving from a single predictor variable to the two predictor variables suggested significant improvement in fit ($p < 0.001$). The c statistic, which represents the AUC for the regression models, improved from 0.80 to 0.88, also suggesting improved discrimination.

Figure 4A shows the RBC cutoff values and the three interval LRs for the final tube RBC count. The slopes of the three lines correspond to the interval LR estimates, respectively. Interval LRs for final tube RBC count were 0 for < 100 RBCs, 1.6 for 100 < RBCs < 10,000, and 6.3 for RBCs > 10,000 (Table 2). The numbers of patients in our sample in each of the three interval groups were 108 (49%), 92 (41%), and 22 (10%), respectively. Thus, in 59% of the patients, the final tube RBC test was associated with a relatively large change in the odds of SAH. Alternatively, if one considers a threshold of 100 RBC as the lower limit of a positive test, then the sample sensitivity for true SAH was 26 of 26, or 100% (95% CI = 87% to 100%).

Based on the ROC curve for the percent change in RBC count, we identified that a cutoff of 0.63 for variation could be used to distinguish between true SAH and traumatic tap (Figure 4B). If the percent change in RBC count is under 63%, it is more likely to represent a true SAH (LR 3.6; Table 2). Conversely, if the percent change in RBC count is over 63%, it is more likely to represent

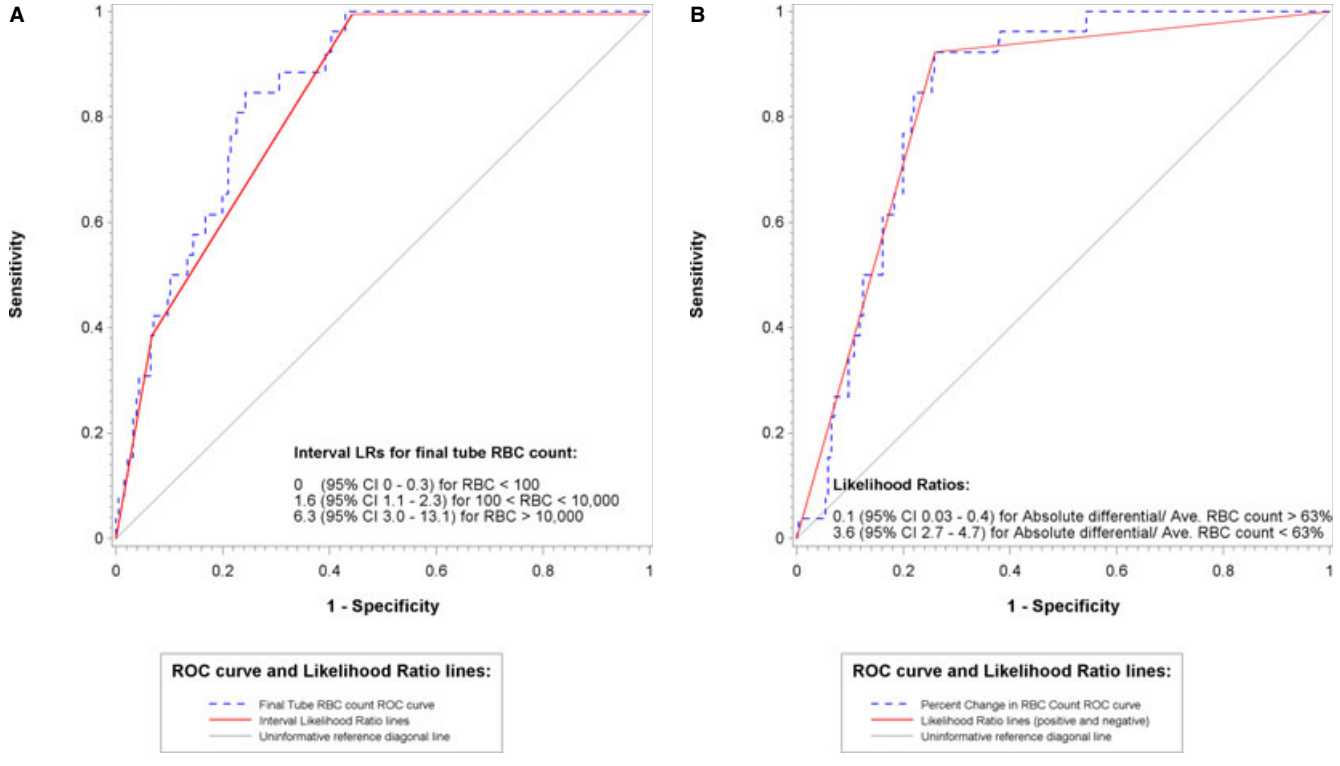

Figure 4. (A) RBC cutoff values and the three interval LRs for the final tube RBC count. (B) Variation cutoff value of 0.63 and positive and negative LRs for the percent change in RBC count. LR = likelihood ratio; RBC = red blood cell; ROC = receiver operating characteristic.

Table 2
Interval LRs for Primary Analysis Versus Sensitivity Analyses

Variable	Primary Analysis	Sensitivity Analysis A	Sensitivity Analysis B	Sensitivity Analysis C
TP	*n* = 26	*n* = 47	*n* = 47	*n* = 13
TN	*n* = 196	*n* = 196	*n* = 211	*n* = 196
Neither ("gray zone")	*n* = 58	*n* = 37	*n* = 22	*n* = 71
LR for RBC < 100	0 (0–0.3) 0 (0–0.2)*	0.5 (0.3–0.8)	0.5 (0.3–0.8)	0 (0–0.5)
LR for 100 < RBC < 10,000	1.6 (1.1–2.3) 1.6 (1.1–2.2)*	1.4 (1.0–1.9)	1.4 (1.0–1.9)	1.8 (1.2–2.7)
LR for RBC > 10,000	6.3 (3.0–13.1) 6.3 (2.8–13.8)*	3.5 (1.6–7.6)	3.5 (1.6–7.4)	5.0 (3.4–7.5)
LR for percent change in RBC count > 63%	0.1 (0.03–0.4) 0.1 (0–0.3)*	0.5 (0.3–0.7)	0.5 (0.3–0.7)	0 (0 0–0.4)
LR for percent change in RBC count < 63%	3.6 (2.7–4.7) 3.6 (2.8–4.8)*	2.5 (1.8–3.5)	2.5 (1.9–3.5)	3.9 (3.0–5.0)

LRs reported with 95% CIs
LR = likelihood ratio; RBC = red blood cell; TN = true-negative; TP = true-positive.
*Bootstrap approach (95% CI).

Table 3
Joint LRs for Primary Analysis

Variable	Percent Change in RBC Count > 63%	Percent Change in RBC Count < 63%
Final tube RBC count < 100	0 (0–0.3); *n* = 90	0 (0–3.1); *n* = 14
100 < final tube RBC count < 10,000	0.4 (0.1–1.4); *n* = 41	3.2 (2.0–5.2); *n* = 45
Final tube RBC count > 10,000	0 (0–6.0); *n* = 9	23.8 (7.0–81.9); *n* = 13

Values reported as LRs with 95% CI; *n* = number of patients in subgroup.
LR = likelihood ratio; RBC = red blood cell.

a false-positive or traumatic tap (LR = 0.1). Table 3 depicts the joint LRs with the combined use of both the final tube interval RBC cutoffs and the variation cutoff of 0.63.

The median number of RBCs in the final tube was 4,309 (interquartile range [IQR] = 745 to 29,838) for the TP group, compared with 69 (IQR = 12 to 498) for the TN group. Within the TP group, those with positive noncontrast head CT findings for SAH (*n* = 17; 12 with certain SAH and five with possible SAH) had a median number of RBCs in the final tube of 5,840 (IQR = 737 to 31,125), compared with 1,690 RBCs (IQR = 519 to 35,000) in those with a negative noncontrast head CT (*n* = 9; Mann-Whitney p = 0.71).

All reliability values using blinded reviewer data were greater than 0.8 except for conventional angiography. In this case, kappa was 0.79 in part because there were few outcomes with aneurysm (*n* = 8). Three cases were correctly identified by both the nonblinded and the blinded reviewers. Four cases (including one performed at an outside hospital) were identified by the nonblinded reviewer, but were classified as not having the study performed by the blinded reviewer, and one case was identified by the blinded reviewer only and was classified as no study performed by the nonblinded reviewer. Presence of xanthochromia had a kappa value of 1.0 (95% CI = 1.0 to 1.0). Intraclass correlation coefficient was also 1.0 for the number of initial (95% CI = 1.0 to 1.0) and final (95% CI = 1.0 to 1.0) tube RBCs. There were no disagreements in any of these three recordings.

Sensitivity Analyses

Sensitivity analysis A used a more relaxed definition for TP SAH so that patients with HA and a newly identified aneurysm or AVM (<2 weeks old, >2 mm in size) but without xanthochromia would be included in the TP SAH group. As a test for SAH, the number of RBCs in the final tube had an AUC = 0.67 (95% CI = 0.58 to 0.76). Again, the differential RBC count was not a good test for SAH, with AUC = 0.55 (95% CI = 0.45 to 0.65).

Sensitivity analysis B incorporated an expanded TN definition (negative xanthochromia at any time, not just $\geq$ 12 hours) in addition to the expanded TP definition from sensitivity analysis A, leading to a gray zone group (neither TP nor TN) of just 22 patients (<10% of study population). As a test for SAH, the number of RBCs in the final tube had an AUC = 0.67 (95% CI = 0.58 to 0.76), compared with AUC = 0.55 (95% CI = 0.45 to 0.65) for the differential RBC count.

Sensitivity analysis C was designed to focus on the true population of interest, those patients with SAH who have a negative noncontrast head CT. For this scenario, 12 patients with certain SAH on head CT and the single patient with certain meningitis based on LP were removed from the TP group, leaving 13 patients with negative or uncertain head CT ("possible SAH") where further evaluation was necessary. For this subgroup, the final tube RBC count had an AUC = 0.87 (95% CI = 0.81 to 0.93), compared with AUC = 0.52 (95% CI = 0.33 to 0.72) for the differential RBC count.

Table 2 illustrates the breakdown of the TP, TN, and gray zone groups for the primary analysis and sensitivity analyses A through C, along with the interval LRs for the final tube RBC count and the percent change in RBC counts. Included are the bootstrapping CIs for the primary analysis to affirm the robustness of the results. For RBCs < 100 and sample sensitivity 0/26, the highest population sensitivity that would still be most likely to

yield this result is 0.026. This estimate was used to compute the LR CI.

DISCUSSION

The purpose of this study was to determine the optimal use of LP RBC counts to identify SAH when some blood remains in the final LP tube. It is current standard practice to evaluate a patient with clinical suspicion for SAH with a noncontrast head CT followed by LP, if the imaging is negative. However, interpreting the LP data remains a diagnostic challenge because it is difficult to distinguish a traumatic tap (RBCs in the CSF caused solely from the procedure) from a true SAH. At this point, there is no clear guideline for the number of RBCs required to diagnose SAH on LP. Xanthochromia is sometimes considered a criterion standard to diagnose SAH, but its reliability is suspect as previously mentioned due to manual assessment, and sensitivity is a function of time from onset of hemorrhage. In our initial sample of 280 patients, 10 patients were excluded as neither TP nor TN because they had xanthochromia but showed no other evidence of SAH on noncontrast head CT, or aneurysm or AVM on CTA (n = 9) or MRA (n = 1). This suggests minimal sample specificity for xanthochromia of approximately 196 of 206 = 95%.

We hypothesized that the differential RBC count between tubes one and four ("clearing of the red cells") is the best LP test characteristic to identify true SAH as measured by the AUC. We found that the number of RBCs in the final LP tube collected, but not the differential count between the final and initial tubes, is associated with true SAH. In our study sample, a 100 RBC threshold in the final tube would effectively rule out SAH. Additionally, RBCs > 10,000 in the final tube was associated with an increase in the odds of SAH by a factor of 6.3.

We also found that the percent change in RBC count between the final and initial tubes is a good test to identify SAH and seems to provide independent information from the final tube RBC count alone. In essence, if the RBC count goes down or up a lot from the initial to final tube, this indicates a traumatic tap rather than a true SAH. This could be explained by the idea that a high number of RBCs in the initial tube may be caused by vascular trauma during needle insertion, while a high number of RBCs in the final tube may be caused by movement or trauma during the procedure itself. However, the limited sample size and resulting CIs in this study suggest that independent confirmation is required before firm predictive conclusions can be drawn.

Our findings are in concurrence with a fairly recent retrospective study of 152 "thunderclap" HA patients with negative noncontrast head CT scans who underwent both LP and conventional four-vessel angiography. Patients with no aneurysm and no xanthochromia had a mean RBC count of 98.7 (n = 93), compared with a mean RBC count of 85,779 in patients with both aneurysm and xanthochromia (n = 13) and RBC count of 20,000 in a single patient with aneurysm but no xanthochromia.[17]

Our negative findings regarding the differential RBC count as a test for SAH are also supported by a retrospective study of 123 patients who underwent both conventional angiography and LP, which showed that a 25% decrease in the number of RBCs between tube 1 and tube 4 can occur even in cases of ruptured aneurysms (n = 2 of total of 8 patients with negative head CT scans and positive aneurysms on angiography).[18] Of note, the authors of this study also concluded that an absolute threshold is impractical in distinguishing traumatic taps from real SAH because they had two cases of positive aneurysm on angiography with a relatively small number of RBCs in tube 1 (69 and 80). However, on closer examination, both of these cases had an increase in the number of RBCs from tube 1 to tube 4, with 100 RBCs and 6,225 RBCs in tube 4, respectively. This coincides with our finding that no patients with < 100 RBCs in the final tube had true SAH.

Although CTA without LP is being evaluated as a potential noninvasive evaluation for SAH in ED patients, and recent data suggest that neither LP nor angiography may be necessary if the HA is acute and head CT is negative,[19] LP remains part of the current work-up. Even if CTA were to supplant LP in the diagnostic algorithm, patients with CTA positive for small aneurysm would likely often require LP to assess for acute SAH. Given that LP also provides helpful information regarding other potential causes of HA (meningitis, encephalitis, idiopathic intracranial hypertension, etc.), it will always remain an important diagnostic tool for the EP. To the best of our knowledge, our study is the first of its kind to directly compare different methods to interpret the RBC count and to provide LRs that can be used to adjust the odds of SAH.

LIMITATIONS

This study is inherently limited by its retrospective study design. The choices of which patients with HA require LP, and which patients with RBCs on LP require neurovascular imaging, are largely subjective based on clinician assessment and judgment. As such, these decisions are subject to many potential biases. Strict methodology, well-defined data points and outcomes, and confirmation of reliable data extraction hopefully limited the additional effect of investigator bias.

Misclassification of TP and TN SAH patients was a major concern, and thus rigorous definitions were used for these two groups. As a result, there was a relatively large group of patients who were classified as neither TP nor TN (gray zone cases) in the primary analysis. Sensitivity analyses were performed to see how the results would change if we attempted to correctly reclassify these gray zone patients into expanded TP and TN definitions. This essentially involves sacrificing accuracy for precision. As gray zone cases are progressively classified as TP or TN in sensitivity analyses A and B and likely some misclassification occurs, the interval LRs tend to contract toward a value of one.

A small TP sample size (TP = 26) leads to relatively wide CIs and limits the generalizability of the study findings without further independent confirmation. Although there is a justified concern for the presence of SAH despite normal noncontrast head CT, a large majority of SAHs are diagnosed by head CT and no LP is required.

Primarily for this reason, patients with SAH who receive LP are rare, and despite data collection over 8 years in a busy referral center, the sample size is limited.

The TP sample obtained is also not precisely equal to the population of interest. Many of the patients with SAH in this study had clearly visible SAH on noncontrast head CT (12 of 26) and also received LP for a variety of reasons. It is uncertain whether the LP RBC count distribution of these patients is the same as those of patients with SAH that is not visible on noncontrast head CT. The latter group is the true population of interest. Sensitivity analysis C was performed to address this issue, and overall the results were similar to the primary analysis. Also, as detailed under Results, there was no major difference in the number of final tube RBCs in our sample between patients with certain or possible SAH visible on head CT and those without visible SAH.

There is an inherent limitation in using xanthochromia as part of the TP and TN definitions because visual inspection of CSF for xanthochromia (the method used at our institution and most others) may be imperfect.[20, 21] Additionally, xanthochromia may no longer be present after 2 weeks of HA duration, which we did not take into account. Finally, no adjustment was made in the 95% CIs for multiple comparisons made in the secondary outcomes analysis.

CONCLUSIONS

In this retrospective analysis, the final lumbar puncture tube red blood cell count was associated with the presence of subarachnoid hemorrhage in patients with headache, but the differential count between the first and last tubes was not. The percent change in red blood cell count between the final and initial tubes was also a good test for subarachnoid hemorrhage. Final lumbar puncture red blood cell counts above 10,000 and below 100 may be useful as important test thresholds to increase or decrease the odds of subarachnoid hemorrhage, respectively, in ED patients who present with headache.

The authors acknowledge the kind assistance of John T. Nagurney, MD.

References

1. Edlow JA, Panagos PD, Godwin SA, Thomas TL, Decker WW. American College of Emergency Physicians Clinical Policies Subcommittee. Clinical policy: critical issues in the evaluation and management of adult patients presenting to the emergency department with acute headache. Ann Emerg Med. 2008; 52:407–36.
2. Bederson JB, Connolly ES Jr, Batjer HH, et al. Guidelines for the management of aneurysmal subarachnoid hemorrhage. A statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association. Stroke 2009; 40:994–1025.
3. Boesiger BM, Shiber JR. Subarachnoid hemorrhage diagnosis by computed tomography and lumbar puncture: are fifth generation CT scanners better at identifying subarachnoid hemorrhage? J Emerg Med. 2005; 29:23–7.
4. Byyny RL, Mower WR, Shum N, Gabayan GZ, Fang S, Baraff LJ. Sensitivity of noncontrast cranial computed for the emergency department diagnosis of subarachnoid hemorrhage. Ann Emerg Med. 2008; 51:697–703.
5. Lourenco AP, Mayo-Smith WW, Tubbs RJ, Sidman R. Does 16-detector computed tomography improve detection of non-traumatic subarachnoid hemorrhage in the emergency department. J Emerg Med. 2009; 36:171–5.
6. Sidman R, Connolly E, Lemke T. Subarachnoid hemorrhage diagnosis: lumbar puncture is still needed when the computed tomography scan is normal. Acad Emerg Med. 1996; 3:827–31.
7. Kassell NF, Torner JC, Haley EC Jr. Jane JA, Adams HP, Kongable GL. The International Cooperative Study on the Timing of Aneursm Surgery. Part 1: Overall management results. J Neurosurg. 1990; 73:18–36.
8. Edlow JA, Wyer PC. Evidence-based emergency medicine clinical question. How good is a negative cranial computed tomographic scan result in excluding subarachnoid hemorrhage? Ann Emerg Med. 2000; 36:507–16.
9. Perry JJ, Spacek A, Forbes M, et al. Is the combination of negative computed tomography result and negative lumbar puncture result sufficient to rule out subarachnoid hemorrhage? Ann Emerg Med. 2008; 51:707–13.
10. Shah KH, Richard KM, Nicholas S, Edlow JA. Incidence of traumatic lumbar puncture. Acad Emerg Med. 2003; 10:151–4.
11. Eskey CJ, Ogilvy CS. Fluoroscopy-guided lumar puncture: decreased frequency of traumatic tap and implications for the assessment of CT-negative acute subarachnoid hemorrhage. AJNR Am J Neuroradiol. 2001; 22:571–6.
12. Shah KH, Edlow JA. Distinguishing traumatic lumbar puncture from true subarachnoid hemorrhage. J Emerg Med. 2002; 23:67–74.
13. Altman DG. Diagnostic Tests. In: Altmand D, Machin D, Bryant T, et al., (Eds). Statistics with Confidence. Oxford, UK: Blackwell Publishing, 2000, pp 108–110.
14. Gart JJ, Nam JM. Approximate interval estimation of the ratio of binomial parameters: a review and corrections for skewness. Biometrics. 1988; 44:323–38.
15. Marill KA, Ingmire TE, Nelson BK. Utility of a single beta HCG measurement in the diagnosis of ectopic pregnancy. J Emerg Med. 1999; 17:419–26.
16. Marill KA. Diagnostic testing and the average absolute likelihood ratio: application to diagnosing wide QRS complex tachycardia and other ED diseases. Am J Emerg Med. 2012; 30:1895–906.
17. Dupont SA, Wijdicks EF, Manno EM, Rabinstein AA. Thunderclap headache and normal computed tomographic results: value of cerebrospinal fluid analysis. Mayo Clin Proc. 2008; 83:1326–31.
18. Heasley DC, Mohamed MA, Yousem DM. Clearing of red blood cells in lumbar puncture does not rule

out ruptured aneurysm in patients with suspected subarachnoid hemorrhage but negative head CT findings. AJNR Am J Neuroradiol. 2005; 26:820–4.
19. Perry JJ, Stiell IG, Sivilotti ML, et al. Sensitivity of computed tomography performed within six hours of onset of headache for diagnosis of subarachnoid haemorrhage: prospective cohort study. BMJ. 2011; 343:d4277.
20. Arora S, Swadron SP, Dissanayake V. Evaluating the sensitivity of visual xanthochromia in patients with subarachnoid hemorrhage. J Emerg Med. 2010; 39:13–6.
21. Soderstrom CE. Diagnostic significance of CSF spectrophotometry and computer tomography in cerebrovascular disease. A comparative study in 231 cases. Stroke. 1997; 8:606–12.