Determination of a Testing Threshold for Lumbar Puncture in the Diagnosis of Subarachnoid Hemorrhage after a Negative Head Computed Tomography: A Decision Analysis

Richard Andrew Taylor, MD, MHS, Harman Singh Gill, MD, Evie G. Marcolini, MD, H. Pendell Meyers, BS, Jeremy Samuel Faust, MD, MS, and David H. Newman, MD

Abstract

Objective: The objective was to determine the testing threshold for lumbar puncture (LP) in the evaluation of aneurysmal subarachnoid hemorrhage (SAH) after a negative head computed tomography (CT). As a secondary aim we sought to identify clinical variables that have the greatest impact on this threshold.

Methods: A decision analytic model was developed to estimate the testing threshold for patients with normal neurologic findings, being evaluated for SAH, after a negative CT of the head. The testing threshold was calculated as the pretest probability of disease where the two strategies (LP or no LP) are balanced in terms of quality-adjusted life-years. Two-way and probabilistic sensitivity analyses (PSAs) were performed.

Results: For the base-case scenario the testing threshold for performing an LP after negative head CT was 4.3%. Results for the two-way sensitivity analyses demonstrated that the test threshold ranged from 1.9% to 15.6%, dominated by the uncertainty in the probability of death from initial missed SAH. In the PSA the mean testing threshold was 4.3% (95% confidence interval = 1.4% to 9.3%). Other significant variables in the model included probability of aneurysmal versus nonaneurysmal SAH after negative head CT, probability of long-term morbidity from initial missed SAH, and probability of renal failure from contrast-induced nephropathy.

Conclusions: Our decision analysis results suggest a testing threshold for LP after negative CT to be approximately 4.3%, with a range of 1.4% to 9.3% on robust PSA. In light of these data, and considering the low probability of aneurysmal SAH after a negative CT, classical teaching and current guidelines addressing testing for SAH should be revisited.
Misdiagnosis and morbidity rates associated with SAH are high, and current clinical practice guidelines recommend lumbar puncture (LP) after a negative noncontrast computed tomography (CT) despite a very low probability of disease.

**Importance**

Lumbar puncture, however, is not without risks including meningitis, neurologic injury, and patient harm from further pursuit of false-positive results. Attempts to balance these complex processes has made the decision of whether to perform LP following CT a decision point both with high clinician variability and with the potential for important impact on patient outcomes.

Decision analysis is a mathematical modeling technique well suited to analyzing complex medical problems with multiple components and determining optimal decision strategies under varying conditions. Among headache patients considered for LP, decision analysis allows for the determination of the comparative impact, in quality-adjusted life-years (QALYs), of performing versus not performing LP at different pretest probabilities of disease. The pretest probability of disease where the two strategies are balanced in terms of QALYs is known as the testing threshold and represents the acceptable miss rate (i.e., if testing was performed at pretest probabilities lower than the threshold the risks of harm to the patient from further testing would outweigh the risk of benefit).

**Goals of This Investigation**

The primary aim of our study was to determine the testing threshold for LP in the evaluation of SAH after a negative noncontrast head CT. As a secondary aim we planned to identify clinical variables that have the greatest impact on this threshold.

**METHODS**

**Study Design**

This study was a decision analysis developed according to published guidelines to estimate the testing threshold for performing LP after a negative noncontrast CT in the evaluation of a patient for SAH. As a decision analysis, the study was only dependent upon data from literature review or expert opinion and was exempt from review by our institutional review board.

**Setting**

The hypothetical base case for our decision analytic model is a 45-year-old patient presenting to the ED with a headache and normal neurologic findings, being evaluated for SAH, after a negative noncontrast CT of the head. Forty-five years represents roughly the mean age of neurologically normal patients enrolled in prospective studies designed to capture patients with SAH and those in whom SAH is an important diagnostic consideration. In our clinical scenario after negative imaging the provider is confronted with two potential diagnostic strategies: perform LP with further testing guided by the results or omit LP and presumably discharge the patient. Subsequent diagnostic and management strategies were chosen to reflect standard practice and accepted guidelines for the evaluation and management of SAH.

**Model Structure**

To model the clinical scenario and diagnostic pathways described, we constructed a decision tree (Figure 1) using decision analysis software (TreeAge Pro 2013, TreeAge Software). The primary node of the decision tree represents the decision to perform or not perform LP. If the provider chooses to perform LP and the findings are positive for xanthochromia or blood then computed tomographic angiography (CTA) is performed. If the LP is negative, the patient is not further evaluated. For those patients with a positive LP, if angiography is positive for a cerebrovascular aneurysm, then the patient potentially undergoes surgical or endovascular aneurysmal repair. Those patients with a positive lumbar and with a negative angiography are not further evaluated for SAH and are treated with standard care.

**Model Inputs and Data Collection**

Data for the model inputs (Tables 1 and 2) were obtained from a methodical literature search and
review, with ancestral search of available evidence for each topic. Using the Integrated Search Interface Web of Knowledge, Google Scholar, and PubMed in 2015, we searched for articles by combining terms subarachnoid hemorrhage, lumbar puncture, CT or computed tomography, contrast-induced nephropathy or acute renal failure, radiation, and cerebral aneurysm in logic-based queries. Two investigators (HM and JF) blinded to study hypothesis reviewed articles or abstracts to determine relevance, extract data using a standardized data form, and grade methodologic quality according to standardized criteria, with disputes resolved by a third reviewer (DN). Credible intervals were constructed using the range of values suggested in the literature, with embedded confidence intervals (CI) where appropriate (typically from systematic reviews with high quality, low heterogeneity meta-analysis).

In modeling the clinical evaluation and treatment of SAH, there are a number of pathophysiologic factors and clinical complications that were considered as

Figure 1. Representative components of the decision tree on whether to perform lumbar puncture after a negative noncontrast head CT. (A) Base of tree with initial decision node (square) and subsequent downstream chance, or probability nodes (circles), and terminal nodes (triangles). Breaks in lines represent further aspects of the decision tree, part of which is demonstrated in (B) with Markov nodes (circles with “M”). CT = computed tomography; LP = lumbar puncture; SAH = subarachnoid hemorrhage.
inputs for the model. Not all SAHs are the result of aneurysmal bleeding, particularly after a negative head CT. This is supported by data from the largest prospective cohort study examining ED headache evaluation for SAH, in which approximately 20% of patients diagnosed with SAH after negative CT evaluation were found to have evidence of an underlying aneurysm.\(^9,20\) There is an estimated 0.4%–7% prevalence of asymptomatic aneurysms in the general population.\(^21\) For false-positive LPs (i.e., false-positive SAH), aneurysms found on CTA will be presumed to be causative and will typically lead to therapeutic procedures (e.g., surgical clipping or endovascular coiling) with the potential for complications including death.\(^22,23\) There are also multiple complications that may arise from LP. In our model post-LP headache and the discomfort of the procedure itself were not considered, as we felt these represented short-term morbidity that would be difficult to reliably model or convert into QALYs.\(^24\) However, we did consider the very small risks of meningitis and paraparesis from LP as these contribute to mortality and long-term morbidity.\(^5,8\) Given the limited amount of information on LP adverse outcomes, within our sensitivity analyses we set the lower bounds of mortality and morbidity rates to 0%. In addition, the model includes complications of contrast administered for CTA (e.g., death from anaphylaxis and renal failure and long-term dialysis dependence).\(^25,29\) Furthermore, the transition of the patient through various states of cancer (i.e., no cancer, cancer, and death) and from the state of having long-term morbidity from SAH to death form SAH were incorporated into Markov nodes.

Data Analysis
To determine a testing threshold for LP for our base case a one-way sensitivity analysis was performed to

### Table 1
List of Input Variables for Decision Analytic Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value, %</th>
<th>Range for Sensitivity Analysis, %</th>
<th>Category in Model</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP sensitivity for SAH</td>
<td>100</td>
<td>94–100</td>
<td>Bayesian</td>
<td>Perry et al.,(^{14}) Claveau and Dankoff(^{40})</td>
</tr>
<tr>
<td>LP specificity for SAH</td>
<td>67</td>
<td>63–71</td>
<td>Bayesian</td>
<td>Perry et al.,(^{14}) Claveau and Dankoff(^{40})</td>
</tr>
<tr>
<td>CTA sensitivity for aneurysm</td>
<td>98</td>
<td>97–99</td>
<td>Bayesian</td>
<td>Carstairs et al.,(^{27}) Westerlaan et al.,(^{41})</td>
</tr>
<tr>
<td>CTA specificity for aneurysm</td>
<td>100</td>
<td>97–100</td>
<td>Bayesian</td>
<td>Carstairs et al.,(^{27}) Westerlaan et al.,(^{41})</td>
</tr>
<tr>
<td>Probability of LP long-term morbidity</td>
<td>0.1</td>
<td>0–0.2</td>
<td>Probability</td>
<td>Thigpen et al.,(^{8}) Evans,(^{24}) Durand et al.,(^{42})</td>
</tr>
<tr>
<td>Probability of death from LP</td>
<td>0.02</td>
<td>0–0.1</td>
<td>Probability</td>
<td>Kooiman et al.,(^{25}) Mitchell et al.,(^{28})</td>
</tr>
<tr>
<td>Probability of death from ARF requiring RRT secondary to CIN</td>
<td>0.1</td>
<td>0–1</td>
<td>Probability</td>
<td>McCullough et al.,(^{26}) Mitchell et al.,(^{27})</td>
</tr>
<tr>
<td>Probability of death from ARF requiring RRT secondary to CIN</td>
<td>35.4</td>
<td>20–100</td>
<td>Probability</td>
<td>Wiebers et al.,(^{36}) Raaymakers et al.,(^{43})</td>
</tr>
<tr>
<td>Probability of death from surgery for asymptomatic aneurysm</td>
<td>2.5</td>
<td>0.8–3.2</td>
<td>Probability</td>
<td>Wiebers et al.,(^{36}) Raaymakers et al.,(^{43})</td>
</tr>
<tr>
<td>Probability of long-term morbidity from surgery for asymptomatic aneurysm</td>
<td>9.2</td>
<td>8.1–10.4</td>
<td>Probability</td>
<td>Rinkel et al.,(^{46})</td>
</tr>
<tr>
<td>Probability of death from anaphylaxis</td>
<td>0.0021</td>
<td>0.0001–0.027</td>
<td>Probability</td>
<td>Cochran et al.,(^{29}) Katayama et al.,(^{44})</td>
</tr>
<tr>
<td>Probability of aneurysmal vs. nonaneurysmal SAH (after negative CT)*</td>
<td>20</td>
<td>10–50</td>
<td>Probability</td>
<td>Perry et al.,(^{29})</td>
</tr>
<tr>
<td>Probability of incidental aneurysm</td>
<td>2</td>
<td>0.4–6</td>
<td>Probability</td>
<td>Li et al.,(^{21}) Vlak et al.,(^{45}) Rinkel et al.,(^{46})</td>
</tr>
<tr>
<td>Probability of death SAH nonaneurysmal</td>
<td>2.6</td>
<td>0.7–9.0</td>
<td>Probability</td>
<td>Rinkel et al.,(^{47})</td>
</tr>
<tr>
<td>Probability of death nonaneurysmal</td>
<td>0</td>
<td>0–4.8</td>
<td>Probability</td>
<td>Rinkel et al.,(^{47})</td>
</tr>
<tr>
<td>Probability of death from SAH (treated/initial correct diagnosis)</td>
<td>5</td>
<td>2–9</td>
<td>Probability</td>
<td>Kowalski et al.,(^{10}) Molyneux et al.,(^{48})</td>
</tr>
<tr>
<td>Probability of long-term morbidity (treated/initial correct diagnosis)</td>
<td>31</td>
<td>24–38</td>
<td>Probability</td>
<td>Kowalski et al.,(^{10}) Molyneux et al.,(^{48})</td>
</tr>
<tr>
<td>Probability of death from initial missed aneurysmal SAH</td>
<td>19</td>
<td>9–35</td>
<td>Probability</td>
<td>Kowalski et al.,(^{10}) Molyneux et al.,(^{48})</td>
</tr>
<tr>
<td>Probability of long-term morbidity from initial missed aneurysmal SAH</td>
<td>31</td>
<td>17–49</td>
<td>Probability</td>
<td>Kowalski et al.,(^{10}) Molyneux et al.,(^{48})</td>
</tr>
<tr>
<td>Annualized long-term mortality rate for SAH morbidity patients†</td>
<td>5</td>
<td>0–10</td>
<td>Markov</td>
<td>Ronkainen et al.,(^{50}) Pyysalo et al.,(^{51})</td>
</tr>
<tr>
<td>Annual cancer rate from CT (head)</td>
<td>0.00035</td>
<td>0–0.001</td>
<td>Markov</td>
<td>Mathews et al.,(^{52}) Smith-Bindman et al.,(^{53}) Salibi et al.,(^{54})</td>
</tr>
<tr>
<td>Annual mortality from cancer (head)</td>
<td>13%</td>
<td>5%–50%</td>
<td>Markov</td>
<td>Siegel et al.,(^{55}) Ostrom et al.,(^{56})</td>
</tr>
<tr>
<td>Annual remission (without symptoms) from cancer</td>
<td>5%</td>
<td>1%–10%</td>
<td>Markov</td>
<td>Østrem et al.,(^{56})</td>
</tr>
<tr>
<td>Patient age (y)</td>
<td>45</td>
<td>18–70</td>
<td>Markov</td>
<td>Perry et al.,(^{9})</td>
</tr>
</tbody>
</table>

Multiple data sources were combined where possible as a weighted means based on study sample size.\(^{57}\)

*Conservative estimate based on available data using positive LP definition of 500 RBCs/hpf.

†Calculated from the lifetime-attributable risk (LAR) of cancer incidence using the linear no-threshold model from the BEIR VII report\(^{13}\) and division by the LAR of cancer incidence by the number of cycles in the model to arrive at a per year risk of cancer.

ARF = acute renal failure; CIN = contrast-induced nephropathy; LP = lumbar puncture; SAH = subarachnoid hemorrhage; RRT = renal replacement therapy.
Two-way sensitivity analyses were performed to evaluate the influence of model variables on the testing threshold and to account for variable uncertainty in the model. When available, a range of values for each variable was obtained from 95% CIs and credible intervals constructed from literature searches as noted above. When these were unavailable a range was derived through assumption and group consensus. Results of the two-way sensitivity analyses are expressed in a tornado diagram.

A limitation of two-way sensitivity analysis is that it is only two-way (i.e., all other variables except two are held constant) and thus unable to examine uncertainty within the model that results from the interaction of more than two variables. To better determine the uncertainty and range of values for the testing threshold, we further analyzed the model through probabilistic sensitivity analysis (PSA) and Monte Carlo simulation. We assumed beta probability distributions for the model variables with distribution parameters determined by data available from literature review or, when not available, assumption and group consensus. Monte Carlo simulation was performed with 500,000 iterations in which each iteration selected random values from the probability distributions of each variable. A testing threshold with 95% CI was determined by analyzing the values of the pretest probability of disease for iterations in which the outcomes for the two decision strategies were equal.

RESULTS

For the base-case scenario (45-year-old presenting to the ED with a headache, normal neurologic status, and negative head CT) based on one-way sensitivity analysis the testing threshold for LP was 4.3%. Adjustments for sex-based life expectancy had no effect. Results of the two-way sensitivity analyses for each variable in the model are demonstrated in Figure 2. In examining all variables in the two-way sensitivity analysis, the range of the test threshold was 1.9%–15.6%, dominated by the uncertainty in the probability of death from initial missed SAH. Other significant drivers of model variation included probability of aneurysmal versus nonaneurysmal SAH after negative head CT, probability of long-term morbidity from initial missed SAH, and probability of renal failure requiring RRT from CIN. In the PSA the mean testing threshold was 4.3% (95% CI = 1.4 to 9.3).

DISCUSSION

The approach to diagnostic testing for specific conditions can be examined and potentially improved by consideration of a threshold for testing at which potential harms and benefits of testing are equal. Within the context of shared decision-making, this information can be used to guide both physicians and patients. In the setting of potentially deadly conditions, however, the utility of this threshold is often overshadowed by barriers including defensive practice due to medicolegal or professional concerns, patient expectation, poor communication, and a focus on diagnostic certainty. We are unaware of prior published literature estimating a testing threshold for LP in the setting of potential SAH after a negative CT.

Our decision analysis calculations suggest that a reasonable test threshold for performance of LP for the detection of SAH in neurologically normal patients with headache and a negative noncontrast head CT is approximately 4.3%, with a range of 1.4% to 9.3% in a robust PSA analysis. These findings contrast with common practice and classical teaching both of which tend to focus on the potential benefits of diagnosis without explicit consideration of harms arising from testing.

The testing threshold after negative imaging in our analysis was raised by a number of factors. First, in...
relevant large studies subjects diagnosed with SAH by LP appear to have mostly nonaneurysmal SAH or false-positive LPs.\textsuperscript{9,14} Because nonaneurysmal atraumatic SAH is associated with nearly universal complete recovery without therapy, detection in such cases yields no improvement in QALYs. Second, data suggest that delays in diagnosis, while undesirable and potentially dangerous, lead to morbidity or mortality in a minority (roughly 10\%) of missed aneurysmal SAH.\textsuperscript{10,33-35} Third, LP includes small but real risks of infection and injury.\textsuperscript{6-8} Finally, patients with positive LP findings typically undergo angiography, incurring risks of anaphylaxis, CIN, and additional radiation exposure. Moreover, incidental aneurysms found during angiography, present in up to 7\% of screening populations,\textsuperscript{22} will commonly be interpreted as culprit lesions and undergo neurosurgical procedures\textsuperscript{22} that include considerable harm rates.\textsuperscript{36}

It is not surprising that the variable with the largest impact on the uncertainty of the model was the probability of death from initial missed aneurysmal SAH, as it is the primary serious outcome from not performing LP and has a wide 95\% CI reported in the literature. Most other important variables were also associated with SAH outcomes. It is also of note that excluding the direct negative effects of LP (morbidity and mortality associated with infection and neurology damage) fails to lower the testing threshold below a level that would favor performing an LP after negative CT in most patients.

Using testing threshold estimates in clinical practice depends upon knowing the probability of SAH after a negative noncontrast head CT. Fortunately, an increasingly high-quality database of prospective studies has begun to fill gaps that have long hampered attempts to examine this issue based on outcomes data.\textsuperscript{2,9,10,14} These investigations, when combined with our analysis, strongly suggest that in most patients with acute headache LP after negative CT with newer-generation scanners, especially when performed under 6 hours of symptom onset, is a more harmful than helpful strategy. This results are also supported by a recent cost-effectiveness study that demonstrated when the CT sensitivity is >99\% (i.e., CT on newer-generation scanner performed less 6 hours from onset of symptoms) no further testing is warranted.\textsuperscript{58} For carefully selected patients (those with a high probability of disease [>20\%] and who present late > 2 days) the likelihood of SAH may exceed testing thresholds in the lowest range of our intervals, suggesting that LP may be a beneficial approach for such patients presuming the most

\begin{figure}
\centering
\includegraphics[width=\textwidth]{tornado.png}
\caption{Tornado diagram of two-way sensitivity analyses of variables in model and their effect on the testing threshold. ARF = acute renal failure; CIN = contrast-induced nephropathy; CT = computed tomography; CTA = CT angiography; LP = lumbar puncture; SAH = subarachnoid hemorrhage; RRT = renal replacement therapy.}
\end{figure}
Because of the declining performance of CT for SAH over time and the complicated aspect of determining a pre-CT probability of disease, we believe that decision aids such as Figure 3 coupled with clinical decision rules that estimate pre-CT probabilities of disease will be helpful in making shared decisions with patients under uncertainty.

LIMITATIONS

The strength of a decision analysis is dependent upon the validity of variable input and the structural assumptions of the model. In our model there are limitations based on the quality and validity of the available literature addressing each input and the inferences that can be reasonably made from observational data. Ideally, there would be randomized trial data to inform outcome predictions based on LP and non-LP approaches following negative CT in such patients. To mitigate these uncertainties we used best available data from a rigorous literature search and review, we employ credible intervals that offer the existing range of published data (rather than 95% CIs) as a means of broadening the potential outputs from our model. In this regard we find it reassuring that varying the statistically most important inputs has a limited impact on decision-making.

As noted above, we made several assumptions regarding the structure of the model to decrease complexity. We chose not to include the short-term effects of LP, anaphylaxis, and SAH (e.g., headache, short-term cognitive deficits) as they are extremely unlikely to contribute to any significant change in the model when compared to more serious long-term effects and death. In addition, we chose not to model cost or other diagnostic testing strategies (e.g., CT/CTA then possible LP or magnetic resonance imaging [MRI]). Cost was not considered as there is no clear accepted standard about the cost per QALY individuals or society would be willing to pay. The strategy of CT/CTA as an initial step is fraught with the consequences of identifying a significant portion of patients with benign headache and incidental aneurysm and previous analysis has shown this strategy to be less effective than CT/LP. A strategy incorporating MRI was not examined because of its reduced availability in acute care settings.

CONCLUSION

In conclusion, our data suggest an explicit threshold approach to lumbar puncture testing for neurologically normal, computed tomography–negative acute headache patients. Our decision analysis calculations suggest this threshold to be approximately 4.3%, with a range of 1.4% to 9.3%. In light of these data, and considering the low probability of aneurysmal subarachnoid hemorrhage after a negative computed tomography, classical teaching and current guidelines addressing testing for subarachnoid hemorrhage should be revisited.

References


