Meta-analysis of diagnostic test accuracy in neurosurgical practice

Julie Dubourg, M.D., Moncef Berhouma, M.D., Michael Cotton, M.D., F.R.C.S., and Mahmoud Messerer, M.D.

Comparative effectiveness research (CER) allows evidence to be evaluated on the effectiveness, benefits, and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition or to improve delivery of care. The purpose of CER is “to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels.” This CER paradigm is not a novel concept as such, and began to take form in the 1850s through the writings of Claude Bernard, a French physiologist. Comparative effectiveness research will no doubt occupy an increasingly important place in clinical practice and health care policy in the future. Consequently, the neurosurgical community should be familiar with CER.

The number of diagnostic options has grown at an increasing rate over the past 2 decades in medicine, and likewise in neurosurgery. Currently, physicians have a large variety of diagnostic tests at their disposal for the same condition. Furthermore, they are faced with patients who require more and more guarantees, as well as the best diagnostic test associated with the fewest complications or side effects. Thus, it is a fundamental task for clinicians to gather data and to use that data to formulate an optimal care plan for their patients.

Meta-analysis represents a high level of evidence-based medicine by summarizing results of well-designed research studies on the same topic, thereby achieving the best estimate of performance of a diagnostic device. Meta-analysis makes it possible to obtain more precise estimates when only small studies are available and can also determine the covariates that may influence results. The process leading to publication of a meta-analysis should be transparent and reproducible. Unlike meta-analyses of randomized controlled treatment trials, the methodology of a meta-analysis of diagnostic test accuracy is less known and understood. Despite the existence for more than 10 years of more and more methodological work, many uncertainties remain, and there is no

C

Comparative effectiveness research is defined by the Institute of Medicine as “the generation and synthesis of evidence that compares the benefit and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition or to improve delivery of care.” The purpose of CER is “to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels.” This CER paradigm is not a novel concept as such, and began to take form in the 1850s through the writings of Claude Bernard, a French physiologist. Comparative effectiveness research will no doubt occupy an increasingly important place in clinical practice and health care policy in the future. Consequently, the neurosurgical community should be familiar with CER.

The number of diagnostic options has grown at an
consensus indicating the best statistical method to synthesize results from studies of diagnostic tests.

The aim of this paper is to improve neurosurgeons’ familiarity with the meta-analysis of diagnostic test accuracy by describing and detailing each stage leading to the publication of such a meta-analysis.

Overview of Meta-Analysis of Diagnostic Test Accuracy

The major steps for conducting a systematic review are as follows: (Fig. 1):32,33 1) identification of the objectives of the meta-analysis; 2) search strategy and study selection; 3) assessment of study quality; 4) data extraction; 5) statistical analysis; and 6) interpretation of results.

Identification of the Objectives of the Meta-Analysis

The diagnostic accuracy of a diagnostic test is its ability to identify or exclude a given disease or target condition. An “index test,” the diagnostic test under evaluation, is measured against a “gold standard,” which is the best currently available test to classify a patient’s disease status. For example, to evaluate the diagnostic accuracy of percutaneous biopsy of the cavernous sinus region through the foramen ovale (index test) in detecting benign or malignant lesions, the reference standard will be the histopathological analysis of the lesion obtained by open craniotomy. This method will correctly and definitively classify patients as having benign or malignant status. According to the gold standard, results of the index test will be classified as true positives, false positives, false negatives, and true negatives, from which the calculation of accuracy indices is possible (Table 1).

Before issuing the study protocol, investigators must ask certain questions about the possible role of the diagnostic test. Three types of questions arise. The first question is: is this test reliable? In other words, is this test accurate? That is, if we say the sodium level is 140 mmol/L, is that really what it is? Also, is the test result precise? That is, if one measures the same patient’s blood 10 times, will the sodium level continue to be measured as 140 mmol/L? This question describes the test’s analytical validity.

The second question is: is this test significant? This question describes the test’s clinical validity. And the final question is: is this test helpful? In other words, does the realization of the test modify patient outcomes? This question describes the test’s clinical utility.

Most physicians are probably satisfied with a test with only clinical validity. However, we believe that it is fundamental that the test also has clinical utility. When defining clinical utility, 3 other questions are raised to define the objectives of the diagnostic test under evaluation.10

Fig. 1. Overview of the complete process of performing a meta-analysis of diagnostic test accuracy.
The first question is known as the replacement test (Fig. 2A): does the index test intend to replace the gold standard test? This is an important question, for it will be impossible to demonstrate the superiority of the index test regarding its diagnostic accuracy because the target condition is defined in terms of the reference standard. Therefore, the index test will be, at best, equal in diagnostic accuracy. The aim will be to demonstrate that the index test is as good as the reference test and to highlight the benefits of the index test in terms of cost, lack of complications, ease of implementation and interpretation, and availability.

The second question is known as the triage test (Fig. 2B): is the index test intended to serve as a triage test? This test should have a 100% sensitivity or specificity depending on whether it is meant to confirm or exclude a diagnosis compared with the reference test.

And the last question is known as the add-on test (Fig. 2C): is the index test intended as an additional test?
This would be the case if there is a reference test with high specificity but poor sensitivity, and there are many false negatives. In this case, it would be useful to have a test with high sensitivity. Both tests will therefore be complementary. Figure 2 shows examples of diagnostic tests in these 3 situations in neurosurgical settings.

Thus, a meta-analysis should evaluate an index test (for replacement, triage, or add-on) for a target condition and compare it to a reference test in a more or less specific population. Once the clinical problem and the clinical role of the index test are identified, a protocol must be written. It should define the context of the study (index test, pathology, reference test), the intended purpose of the index test (replacement, triage, or add-on), the search strategy, study selection, data extraction, validation of the methodological quality of selected studies, analysis and presentation of results, and the expected final interpretation.

Search Strategy and Study Selection

Search Strategy. The search strategy must be determined in the study protocol. Identification of eligible studies is a key step. If improperly performed, potentially eligible articles will not be considered, and this will lead to publication bias and therefore potentially erroneous results. It is important to note that a search to find articles on diagnostic test accuracy is more difficult than a search to find randomized controlled trials.26

The primary sources of data are the electronic databases. There are general medical databases (Medline Pubmed, Embase, Biosis, Cochrane), specific topic databases (Neuroscience Information Framework), and national and regional databases (PASCAL [Europe], IndMed [India], African Index Medicus [Africa], Australasian Medical Index [Australia], and LILACS [Latin America and Caribbean]). These databases may have free or paid access. The ideal strategy is to search in as many databases as possible.

The search strategy is essential. It must be the most sensitive and least restrictive possible; thus, it will use the “free text” and the Medical Subject Heading (MeSH) terms.34 Many search filters have been developed,2,26 but these filters involve an increased risk of not finding all eligible studies, and therefore they are not recommended.36 The search strategy must therefore have no restrictions on language, publication date, or publication type (such as review of the literature, reported cases, and others). The development of key words is an important step; they should be “index test,” “target condition,” “gold standard,” and “target population.” The search structure should not be too specific because of the risk of missing potentially eligible studies, as noted above. The MeSH term “sensitivity and specificity” may appear suitable but is inconsistently applied in most databases. Table 2 shows examples of search strategies. For example, a search strategy for articles about the diagnostic accuracy of ONSD ultrasonography in detecting raised ICP may include the terms “brain injuries,” “ultrasonography,” “intracranial hypertension,” “intracranial pressure,” “optic nerve,” “raised intracranial pressure,” “myelin sheath,” “intraventricular catheter,” “cerebrospinal fluid pressure,” and “diagnosis.”

<table>
<thead>
<tr>
<th>TABLE 2: Examples of search strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>« index test » AND « target condition »</td>
</tr>
<tr>
<td>« gold standard » AND « target condition » AND « specific population »</td>
</tr>
</tbody>
</table>

These different search strategies should be used in the same database and then repeated in all the databases used. It is probably best to import the retrieved references into bibliographic management software (such as Reference Manager, ProCite, EndNote, RefWorks, or others). These software engines save time, and also reduce transcription errors and eliminate duplicate references with 1 click; they can also sort references within subgroups (such as included and excluded).

The search strategy must be kept in mind and explained in the protocol and publication, to further the goal of transparency and reproducibility. A major step in the publication of the meta-analysis will be to also indicate the number of retrieved references, as a result of the search strategy. Thus, for example, the search strategy for an article on ONSD ultrasonography retrieved 699 articles.59 This number constitutes the first element of the flow chart that must be reported in the meta-analysis.

However, this search strategy is not sufficient, especially for studies of diagnostic tests. If performed alone, there is a high risk of publication bias. It has been shown that studies with statistically significant results are more likely to be published than those that show no difference.50 Consequently, it is important that all studies on the same topic must be included to better estimate its true diagnostic accuracy. Other search methods include checking references of retrieved articles, hand-searching clinical meeting abstracts, contacting specialists involved by email, and searching in clinical trial registries such as www.clinicaltrials.gov.

A search is complex and needs to be clearly explained in the methodology section of the article, as do the results of the search. This search will be more effective if there is close contact between clinicians and methodologists and/or librarians.

Study Selection. The eligibility criteria of studies in the meta-analysis should be clearly defined in the protocol. These criteria will vary depending on the clinical problem and may concern the type of design, the population, the index test, the reference test, and the pathology target.

Many designs exist for diagnostic tests, such as cases reported, cross-sectional studies, case-control studies, randomized trials, and cohort studies. In general, each study in which the index test results are compared with gold standard test results will be eligible. However, randomized trials are rarely used to assess diagnostic accuracy because patients will benefit from either the reference test or the index test; thus it is impossible to relate the index test re-
Meta-analysis of diagnostic test accuracy in neurosurgery

results as a function of the gold standard. It is only feasible to assess the impact on patient outcomes. For this reason, cross-sectional studies are widely included.

In an article, this process of analysis must be detailed in the flow chart by indicating the number of studies included and/or excluded and the reasons why. A minimum of 2 authors should independently review the complete list of references found by the search strategy and decide on their eligibility. Once this process is finished, they should meet to check the compatibility of their results and to resolve differences by consensus, or with the help of an additional mentor.

Assessment of Study Quality

Meta-analysis of diagnostic accuracy studies is often characterized by large heterogeneity due to differences in study design. Therefore, it is essential to carefully assess the quality of included studies. Guidelines such as the Standards for Reporting of Diagnostic Accuracy (STARD) initiative are requested by several journals to promote transparency of study design and thereby improve the quality of diagnostic studies. Many tools are available, but the most efficient and most recommended is the Quality Assessment of Diagnostic Accuracy Studies (QUADAS), developed by Whiting et al. in 2003. Initially this tool included 14 items for the assessment of risk bias and sources of variation. Each item is rated “yes,” “no,” or “unclear.” Some authors use an overall quality score regrouping all QUADAS items, but this way of reporting quality is not recommended, because some items could be more important than others. Depending on the subject of the literature review, variables differ in importance. It is therefore up to each author to define the main criteria for his or her study and when there is a need to answer “yes,” “no,” or “unclear.” Close collaboration with a clinician expert in the specialty is thus required.

As mentioned above, it is rarely justified to present the results in an overall score form. The most representative method is to select several or all of the variables and represent them individually using graphics (Fig. 3). Recently, the QUADAS-2 tool was developed to offer additional and improved features, including distinguishing between bias and applicability, and help with assessing the risk of bias.

The 3 main biases of which one must be aware in meta-analyses of diagnostic tests are the spectrum of patients, verification of the “condition” of patients, and the blind comparison.

It is essential that the studies included are composed of the spectrum of patients representative of the target condition in clinical practice. For example, case-control studies are very poor studies to include because the test is very sensitive in the “case” population and the specificity is also very important in the “control” population. Moreover, this spectrum will never be representative of the spectrum of patients encountered in clinical practice.

For verification of the “condition” of patients, the choice of the gold standard is very important. If the gold standard that is used is not optimal, it may incorrectly classify patients and therefore underestimate or overestimate the diagnostic accuracy of the index test. On the other hand, the timing of performing the index test can also result in bias. It is important to avoid using an index test that is a part of the gold standard. It also requires that the lead time between the 2 tests is as short as possible to avoid any change in patient condition status. This verification bias can also be observed if the reference test is not performed in the entire population studied.

The test results must always be interpreted in a double-blind fashion. Otherwise it might influence the interpretation of results by the clinician.

The major problems with this part of the meta-analysis (assessment of quality) mainly concern the number of authors validating the methodological quality of studies and their background, the subjectivity of the exercise, and how to resolve disagreements.

Data Extraction

The data extraction stage should also involve at least 2 authors. They will extract data independently of each study selected through a data extraction sheet, which has already been defined by the protocol. This extraction will allow them to build a 2 × 2 contingency table (Table 1). They will calculate the following indices of diagnostic accuracy: sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and diagnostic odds ratio (Appendix). Again, the authors will meet and agree on a possible consensus on the differences in results observed.

Statistical Analysis

Exploring the Publication Bias. In published studies, favorable findings are more likely to be presented than inconclusive results, which leads to reporting bias. The validity of a meta-analysis of diagnostic test accuracy depends on minimizing bias in the identification of studies. If included studies present results that differ
from relevant studies that are not included, the conclusions of the meta-analysis will be compromised by publication bias. The search for the presence of publication bias can be achieved by constructing a funnel plot and considering its symmetry. The classical methods to achieve this are the Begg method and the Egger test. These methods investigate publication bias by plotting a measure of effect size against a measure of study precision, which appears symmetric if no bias is present. However, assessing this symmetry is very subjective, so nonparametric and parametric linear regression methods have been developed to test for plot asymmetry reflecting publication bias. This funnel plot, often used in meta-analyses of randomized clinical trials, has demonstrated many flaws when used in diagnostic studies. However, alternatives to assess the presence of publication bias have not been very effective. Investigating publication bias in diagnostic accuracy studies is a matter of debate, because the determinants for publication are not the same as for randomized controlled trials, and diagnostic accuracy studies do not routinely use the report of p values, but instead use estimates of sensitivity and specificity with 95% CIs. Therefore, publication bias will not be associated with statistical nonsignificance. Finally, the best choice to avoid publication bias is to perform an exhaustive search strategy.

Exploring for Heterogeneity. In general, there is more heterogeneity between studies of diagnostic tests compared with randomized controlled trials. The search for heterogeneity can be achieved through several tests, but the most powerful is the calculation of the I² statistic (see Appendix). This statistic measures the percentage of variability between summary indices that is due to heterogeneity rather than chance. Generally, a study with an I² > 50% is considered to have substantial heterogeneity. Another widely used test is the Cochran Q test, which indicates heterogeneity when the p value is < 0.05 (Fig. 4). However, searching for heterogeneity is not enough; it is important to consider its causes. Part of the heterogeneity may be due to chance because the number of these studies is usually low. There are other common sources of heterogeneity between studies, such as the use of different thresholds of positivity, differences in methodological quality, various techniques for realization or subjectivity of the test index, various techniques for realization or subjectivity of the gold standard, and variable sources of recruitment involving varied prevalence of the target condition. If heterogeneity is found, the reasons may be explored by relating study-level covariables to an accuracy measure (such as the diagnostic odds ratio), and using meta-regression techniques such as a maximum-likelihood estimation method.

Fig. 4. Forest plot showing pooled sensitivity and specificity of an index test compared with a reference test. Each point indicates sensitivity and specificity estimates for each individual study. Horizontal lines indicate the 95% CI. This graph also shows the combined sensitivity and specificity with 95% CI at the bottom. In these contrived data, the sensitivity of the index test compared with the reference standard is 96% (95% CI 92–98%) and the pooled specificity is 98% (95% CI 95–99%). The higher the sensitivity and specificity are, the more accurate is the index test. In this graph, we can also view the results of total variation across studies that are attributable to heterogeneity rather than to chance. For sensitivity, I² is < 50%, and the p value of the Cochrane Q test is > 0.05, so there is no substantial heterogeneity. The reverse is true for specificity; I² is > 50% and the p value is < 0.05, indicating heterogeneity. When readers observe heterogeneity, they should interpret results more carefully.
Performances Indices. Sensitivity, also known as the true-positive rate, is the ability of the index test to identify a patient with the target condition (Appendix). Conversely, specificity, also known as the false-positive rate, is the ability of the index test to identify healthy patients. When a test has a high sensitivity, a negative result rules out the diagnosis, and when a test has a high specificity, a positive result rules in the diagnosis. It is important to note that there is a negative correlation between sensitivity and specificity. The choice of threshold of the index test determines the values of sensitivity and specificity.

Negative or positive predictive value of the index test is the probability of a negative or positive result being correct.

Clinical usefulness of the index test could also be evaluated by calculating positive and negative likelihood ratios. With a positive likelihood ratio > 1, the probability of the target condition being present increases. When this ratio is < 1, the probability of it being present decreases. When the ratio is 1, the probability is unchanged.

The diagnostic odds ratio is an indicator of overall diagnostic accuracy of the index test (Appendix). With this ratio, the odds of positivity of the index test in patients with the target condition can be calculated, compared with the odds of positivity of the index test in those without the target condition.

Statistical Pooling. In any meta-analysis, the first step is to graph the results of the individual selected studies. The SROC curve consists of representing the paired results for sensitivity and specificity (Fig. 5). In the receiver operating characteristic space, the specificity is displayed on the x-axis and the sensitivity on the y-axis. In this graph, there is a diagonal that represents the value of sensitivity of the target condition being present increases. When this ratio is < 1, the probability of it being present decreases. When the ratio is 1, the probability is unchanged.

The diagnostic odds ratio is an indicator of overall diagnostic accuracy of the index test (Appendix). With this ratio, the odds of positivity of the index test in patients with the target condition can be calculated, compared with the odds of positivity of the index test in those without the target condition.

Statistical Pooling. In any meta-analysis, the first step is to graph the results of the individual selected studies. The SROC curve consists of representing the paired results for sensitivity and specificity (Fig. 5). In the receiver operating characteristic space, the specificity is displayed on the x-axis and the sensitivity on the y-axis. In this graph, there is a diagonal that represents the value of sensitivity and specificity.
and specificity of a “noninformative” index test. The optimal index test will be positioned in the upper right corner (sensitivity and specificity = 1). The area under the curve serves as a global measure of performance. The following guidelines have been suggested for interpreting area under the curve values: low accuracy for values between 0.5 and 0.7, moderate accuracy for values between 0.7 and 0.9, and high accuracy for values greater than 0.9.

The forest plot (Fig. 4) gives an overall estimate of the sensitivity and specificity of the index test and enables sensitivity and specificity to be displayed separately with CIs. However, the forest plot does not display the covariance between sensitivity and specificity. Others possibilities for representing the results graphically are the creation of a Fagan nomogram (Fig. 6) or a conditional probability plot (Fig. 7).

The choice of statistical method is between using a fixed effect model or a random effects model. In a fixed effect model, it is assumed that there is only 1 overall performance of the index test. Therefore, the result is an estimate of this performance, whereas in a random effects model, it is assumed that there are several possible performances and the result is the average of estimates of performance. The random effects model can thus better manage heterogeneity, but is less efficient in terms of estimating the performance in cases involving consistent results. In practice, it is customary to check the presence of heterogeneity and to choose a fixed effect model in the absence of variation. However, it is not essential to ask this question for meta-analyses of studies testing diagnostic indices. In practice, in most cases, a random effects model is used because these studies are often heterogeneous. The random effects model also offers several options: the traditional method, such as the method of Littenberg and Moses, or more sophisticated statistical methods, such as the bivariate random approach or the hierarchical approach.

Studies of diagnostic tests have many indicators of diagnostic performance. Most studies represent estimates of sensitivity and specificity, either alone or combined with other performance indicators. Summing pairwise sensitivity and specificity is not easy because of the frequent presence of a negative correlation of these 2 indicators in studies; for example, this negative correlation may be due to the use of different thresholds of positivity in studies; for example, this negative correlation may be due to the use of different thresholds of positivity between studies.

The principle of the method of Littenberg and Moses is to construct an SROC curve. This approach converts each pair of sensitivity/specificity values into a single performance measure, the diagnostic odds ratio. This method is widely used in meta-analyses published in the literature, but this method has major drawbacks. This method does not distinguish between the detection capability of the disease (sensitivity) and “no disease” (specificity) and ignores the heterogeneity between studies that might possibly be present. These drawbacks result in an inaccurate estimation of CIs using this method.

Two approaches can circumvent the limitations of the method of Littenberg and Moses: the hierarchical SROC model and the bivariate random model. The bivariate random model can directly sum the sensitivities and specificities in the same model using a random approach, taking into account the negative correlation between sensitivity and specificity across studies. This model can examine the effect of covariates. The hierarchical SROC model focuses on estimating an average of the diagnostic performance and the average threshold of positivity of the test used to build an asymmetrical curve. Both models provide valid estimates of the SROC and their CIs. Both models can be combined with statistical software, a method called “mixed models.”

![Image](https://example.com/image.png)

**Fig. 6.** Example of a Fagan nomogram, a very useful tool that allows clinicians without strong statistical knowledge to determine a diagnostic accuracy of the test under evaluation. The clinical utility of the index test is evaluated using the likelihood ratios (LRs) to calculate posttest probability based on the Bayes theorem. For instance, a clinician believes that a patient has a 20% risk (Prior Prob, Pre-test Probability) of having the target condition based on the radiographic clinical features. By drawing a line from the pretest probability (vertical axis on the left) through the likelihood ratios (center vertical axis) and extended to the posttest probability (right vertical axis), you obtain the posttest probability, so you can determine if the realization of the test adds substantial information. In this example, with positive and negative likelihood ratios of 50 and 0.04, respectively, the probability of having the target condition increases from 20% to 93% if the index test is positive (Post_Prob_Pos; red line) and decreases from 20% to 1% (dashed line) if the index test is negative (Post_Prob_Neg). In this case the index test is extremely accurate.
Interpretation of Results

Interpretation of results is the most important step of the meta-analysis, but is often omitted. Nonetheless, this step is crucial in helping clinicians incorporate the results into their medical decision-making. This is a difficult step because it requires a link to be established with clinical practice. The interpretation of results must be structured as follows: 1) reminders of the main results (recall the main issue, number of studies and patients included, characteristics and methodological quality of studies, and the average results of diagnostic performance indices); 2) strengths and weaknesses of the meta-analysis (lack of research literature, low methodological quality of included studies); 3) applicability of the results in clinical practice; and 4) conclusions (implications for clinicians and public health policy, and implications for future research).

Conclusions

Comparative effectiveness research is becoming an important public policy issue. Many efforts have been made by the US government to create a national CER program to improve clinical practice. Meta-analysis of diagnostic test accuracy is a tool that can be used in the CER paradigm. Neurosurgeons need to be familiar with this tool to improve research in their specialty. The ability to assess the validity of results and apply them appropriately to neurosurgical practice is of increasingly paramount importance. The principles of interpretation given in this paper will help neurosurgeons in this difficult task.

Appendix

Sensitivity = proportion of patients with the target condition who have a positive index test result:
True Positives/(True Positives + False Negatives)

Specificity = proportion of patients without the target condition who have a negative index test result:
True Negatives/(True Negatives + False Positives)

Positive Predictive Value = probability of target condition among patients with a positive index test result:
True Positives/(True Positives + False Positives)

Negative Predictive Value = probability of nontarget condition among patients with a negative index test result:
True Negatives/(True Negatives + False Negatives)

Positive Likelihood Ratio (LR) = indicates how much the odds of the target condition increase when the test index is positive:
Sensitivity/(1−Specificity)

Negative LR = indicates how much the odds of the target condition decrease when the test index is negative:
(1−Sensitivity)/Specificity

Diagnostic Odds Ratio = overall indicator of diagnostic accuracy:
(True Positives × True Negatives)/(False Positives × False Negatives)

F Measure for Heterogeneity:
((Q−df) / Q) × 100%
df = degrees of freedom, number of studies − 1; Q = Cochran

Q Statistic
Cochran Q Statistic for Heterogeneity:
∑ wi(θi−0)²
wi = weight of each study; θ = log mean LR; θi = estimate of the log LR for each study

Disclosure

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author contributions to the study and manuscript preparation include the following. Conception and design: Dubourg, Messerer. Drafting the article: Dubourg, Messerer. Critically revising the article: all authors. Approved the final version of the manuscript on behalf of all authors: Dubourg. Study supervision: Dubourg, Messerer.

References

Meta-analysis of diagnostic test accuracy in neurosurgery


Manuscript submitted March 15, 2012. Accepted May 8, 2012. Please include this information when citing this paper: DOI: 10.3171/2012.5.FOCUS1295.
Address correspondence to: Julie Dubourg, M.D., Université de Lyon 1, Lyon, Centre d’Investigation Clinique CIC 201, EPICIME, Hôpital Louis Pradel, Hospices Civils de Lyon, 59 Boulevard Pinel, F-69677 Bron, France. email: julie.dubourg@gmx.fr.