



Diagnostic accuracy of ancillary tests for death by neurologic criteria: a systematic review and meta-analysis

Précision diagnostique des examens auxiliaires pour la détermination du décès selon des critères neurologiques : une revue systématique et une méta-analyse

Joel Neves Briard, MD, MSc · Roy Nitulescu, MA · Émile Lemoine, MD, MSc · Polina Titova, BA · Lauralyn McIntyre, MD, MSc · Shane W. English, MD, MSc · Greg Knoll, MD, MSc · Sam D. Shemie, MD · Claudio Martin, MD · Alexis F. Turgeon, MD, MSc · François Lauzier, MD, MSc · Dean A. Fergusson, PhD · Michaël Chassé, MD, PhD

Received: 2 August 2022 / Revised: 20 October 2022 / Accepted: 21 October 2022 / Published online: 8 May 2023
© The Author(s) 2023

Abstract

Purpose Ancillary tests are frequently used in death determination by neurologic criteria (DNC), particularly when the clinical neurologic examination is unreliable. Nevertheless, their diagnostic accuracy has not been extensively studied. Our objective was to synthesize the sensitivity and specificity of commonly used ancillary tests for DNC.

Source We performed a systematic review and meta-analysis by searching MEDLINE, EMBASE, Cochrane databases, and CINAHL Ebsco from their inception to 4 February 2022. We selected cohort and case-control studies including patients with 1) clinically diagnosed death by neurologic criteria or 2) clinically suspected death by neurologic criteria who underwent ancillary testing for DNC. We excluded studies without a priori diagnostic criteria and studies conducted solely on pediatric patients. Accepted reference standards were clinical examination, four-vessel conventional angiography, and radionuclide imaging. Data were

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12630-023-02426-1>.

J. Neves Briard, MD, MSc
Department of Neuroscience, Université de Montréal, Montreal, QC, Canada

Centre de recherche du Centre hospitalier de l'Université de Montréal, 900 rue Saint-Denis, Montreal, QC H2X 3H8, Canada

R. Nitulescu, MA · P. Titova, BA
Centre de recherche du Centre hospitalier de l'Université de Montréal, 900 rue Saint-Denis, Montreal, QC H2X 3H8, Canada

É. Lemoine, MD, MSc
Department of Neuroscience, Université de Montréal, Montreal, QC, Canada

L. McIntyre, MD, MSc · G. Knoll, MD, MSc ·
D. A. Fergusson, PhD
Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

S. W. English, MD, MSc
Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

Department of Medicine, University of Ottawa, Ottawa, ON, Canada

S. D. Shemie, MD
Department of Pediatrics, McGill University, Montreal, QC, Canada

C. Martin, MD
Department of Medicine, Schulich School of Medicine and Dentistry, London, ON, Canada

directly extracted from published reports. We assessed the methodological quality of studies with the QUADAS-2 tool and estimated ancillary test sensitivities and specificities using hierarchical Bayesian models with diffuse priors.

Principal findings Overall, 137 records met the selection criteria. One study (0.7%) had a low risk of bias in all QUADAS-2 domains. Among clinically diagnosed death by neurologic criteria patients ($n = 8,891$), ancillary tests had similar pooled sensitivities (range, 0.82–0.93). Sensitivity heterogeneity was greater within ($\sigma = 0.10$ –0.15) than between ($\sigma = 0.04$) ancillary test types. Among clinically suspected death by neurologic criteria patients ($n = 2,732$), pooled ancillary test sensitivities ranged between 0.81 and 1.00 and specificities between 0.87 and 1.00. Most estimates had high statistical uncertainty.

Conclusion Studies assessing ancillary test diagnostic accuracy have an unclear or high risk of bias. High-quality studies are required to thoroughly validate ancillary tests for DNC.

Study registration PROSPERO (CRD42013005907); registered 7 October 2013.

Résumé

Objectif Les examens auxiliaires sont fréquemment utilisés dans la détermination du décès selon des critères neurologiques (DCN), en particulier lorsque l'examen neurologique clinique n'est pas fiable. Néanmoins, leur précision diagnostique n'a pas été étudiée de manière approfondie. Notre objectif était de synthétiser la sensibilité et la spécificité des examens auxiliaires couramment utilisés pour la DCN.

A. F. Turgeon, MD, MSc
Division of Critical Care Medicine, Department of Anesthesiology and Critical Care Medicine, Université Laval, Quebec City, QC, Canada

Population Health and Optimal Health Practices Research Unit (Trauma-Emergency-Critical Care Medicine), CHU de Québec—Université Laval Research Center, Quebec City, QC, Canada

F. Lauzier, MD, MSc
Division of Critical Care Medicine, Department of Anesthesiology and Critical Care Medicine, Université Laval, Quebec City, QC, Canada

Population Health and Optimal Health Practices Research Unit (Trauma-Emergency-Critical Care Medicine), CHU de Québec—Université Laval Research Center, Quebec City, QC, Canada

Department of Medicine, Université Laval, Quebec City, QC, Canada

Sources Nous avons réalisé une revue systématique et une méta-analyse en effectuant des recherches dans les bases de données MEDLINE, EMBASE, Cochrane et CINAHL Ebsco de leur création jusqu'au 4 février 2022. Nous avons sélectionné des études de cohorte et cas témoins incluant des patients présentant 1) un décès selon des critères neurologiques diagnostiqué cliniquement ou 2) un décès selon des critères neurologiques soupçonné cliniquement qui ont été soumis à des examens auxiliaires pour un DCN. Nous avons exclu les études sans critères diagnostiques a priori et les études menées uniquement auprès de patients pédiatriques. Les normes de référence acceptées étaient l'examen clinique, l'angiographie conventionnelle à quatre vaisseaux et l'imagerie nucléaire. Les données ont été directement extraites de comptes rendus publiés. Nous avons évalué la qualité méthodologique des études avec l'outil QUADAS-2 et estimé les sensibilités et les spécificités des examens auxiliaires à l'aide de modèles hiérarchiques bayésiens avec des distributions préalables diffuses.

Constatations principales Au total, 137 études répondaient aux critères de sélection. Une étude (0,7 %) présentait un faible risque de biais dans tous les domaines de QUADAS-2. Parmi les patients ayant reçu un diagnostic clinique de décès selon des critères neurologiques ($n = 8891$), les examens auxiliaires présentaient des sensibilités combinées similaires (intervalle de 0,82 à 0,93). L'hétérogénéité de sensibilité était plus grande au sein ($\sigma = 0,10$ –0,15) plutôt qu'entre ($\sigma = 0,04$) les types d'examens auxiliaires. Parmi les patients cliniquement soupçonnés de décès selon des critères neurologiques ($n = 2732$), les sensibilités combinées des examens auxiliaires variaient entre 0,81 et 1,00 et les spécificités entre 0,87 et 1,00. La plupart des estimations comportaient une grande incertitude statistique.

Conclusion Les études évaluant la précision diagnostique des examens auxiliaires présentent un risque de biais incertain ou élevé. Des études de haute qualité sont nécessaires pour valider en profondeur les examens auxiliaires pour la DCN.

Enregistrement de l'étude PROSPERO (CRD42013005907); enregistrée le 7 octobre 2013.

Keywords ancillary test · brain death · death by neurologic criteria · meta-analysis · systematic review

M. Chassé, MD, PhD (✉)
Centre de recherche du Centre hospitalier de l'Université de Montréal, 900 rue Saint-Denis, Montreal, QC H2X 3H8, Canada

Department of Medicine, Université de Montréal, Montreal, QC, Canada

Death by neurologic criteria occurs when a catastrophic brain injury causes the permanent loss all cerebral functions essential to life. Accurate death determination by neurologic criteria (DNC) is essential to providing closure to relatives and to ceasing somatic mechanical support in the deceased individual. Patients who are diagnosed with death by neurologic criteria often become organ donors; they are in fact the major source of transplantable organs for individuals with terminal heart, lung, liver, and kidney disease.¹ The cornerstone of DNC is a reliable clinical neurologic examination showing permanent cessation of consciousness and loss of brainstem reflexes, including central apnea as shown by an apnea test.² Perfect specificity in DNC (i.e., absence of false positives) is of paramount importance to ensure that the dead donor rule, which states that organs can only be retrieved from a dead person, is respected.³ In practice, numerous factors commonly known as “clinical confounders” may render the clinical examination unreliable, such as drug intoxication or cervical spinal cord injury. Furthermore, a complete neurologic examination is not always feasible, for instance when apnea testing is not safe because of cardiopulmonary instability. In these scenarios, clinicians often use ancillary tests to assess surrogates of brain function, namely cerebral blood flow (e.g., cerebral four-vessel angiography, computed tomography [CT] angiography), perfusion (e.g., CT perfusion scan), or neurophysiologic function (e.g., electroencephalogram [EEG]-evoked potentials).⁴ In certain jurisdictions, ancillary tests are also compulsory to confirm DNC, even in patients with reliable clinical examinations.⁵

Guidance on the use of ancillary tests for DNC, as well as clinical practice, are heterogeneous both between and within jurisdictions.^{6–8} This may reflect the absence of a comprehensive analysis of the diagnostic validity of ancillary tests. The objective of this study was thus to assess the diagnostic accuracy of commonly used ancillary tests for DNC.

Methods

This study is a systematic review and meta-analysis of diagnostic test accuracy, for which the detailed protocol was published previously.⁹ The review follows strict methodological standards based on the Cochrane Collaboration Diagnostic Accuracy Working Group’s recommendations. Reporting follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.¹⁰

Study selection criteria

The target condition of this review was death by neurologic criteria. Since ancillary tests are applied to both clinically diagnosed DNC patients (in a confirmatory role) and to comatose patients suspected of death by neurologic criteria (in a diagnostic role), studied populations included either 1) patients with clinically diagnosed death by neurologic criteria who underwent confirmatory ancillary testing (only patients with DNC), or 2) comatose patients suspected of death by neurologic criteria who underwent reference standard and ancillary testing for DNC (patients with and without death by neurologic criteria). We included cohort and case–control studies, as well as case series, without restriction by language of publication. As this review principally concerns adult patients, we included study samples composed of at least 80% of adults (18 yr or older). We excluded studies from which we could not obtain or calculate the true and false positive and negative rates from the text, appendices or after contacting the main authors. Studies for which the objective was to determine diagnostic criteria of a specific ancillary test with no *a priori* definition of the diagnostic criteria for death by neurologic criteria were also excluded. Finally, we excluded studies without a valid reference standard, studies conducted on pediatric patients only, case reports (2 or fewer patients), and duplicates or subcohorts of already published cohorts.

Reference standards and index tests

We considered studies that used one of three reference gold standards for DNC: clinical diagnosis (an established cause of brain injury, irreversible coma, absence of brainstem reflexes, and central apnea), conventional four-vessel angiography (no intracranial blood flow), and radionuclide imaging (hollow skull phenomenon).¹¹ In studies where authors included an ancillary test in the reference clinical diagnosis, we considered the combination of the clinical evaluation and this ancillary test as the clinical diagnosis reference standard, with plans to perform subgroup analyses pertaining to this factor. In studies where multiple reference standards were applied to patients, clinical diagnosis was chosen as the preferred reference to allow four-vessel angiography and/or radionuclide imaging to be included in the analysis as index ancillary tests.

We investigated the following ancillary tests: four-vessel angiography, radionuclide imaging (including ^{99m}Tc-pertechnetate angiography, ^{99m}Tc-diethylenetriamine pentaacetate [DTPA] angiography, ^{99m}Tc-hexamethylpropyleneamine oxime [HMPAO] angiography, ^{99m}Tc-HMPAO perfusion with and without single-photon emission computed

tomography [SPECT], or other radionuclide testing), transcranial Doppler ultrasonography (TCD), electroencephalography (EEG; cortical or nasopharyngeal), evoked potentials (brainstem auditory, visual, or somatosensory), CT angiography (CTA; 4-point scale,¹² 7-point scale,¹³ 10-point scale,¹⁴ no intracranial flow criteria, or other criteria), CT perfusion imaging (CTP), magnetic resonance imaging (MRI; time-of-flight angiography, diffusion weighted imaging and apparent diffusion coefficient, arterial spin labeling, or other criteria), magnetic resonance venography, magnetic resonance perfusion imaging, and xenon CT.

Search strategy and study screening

We searched MEDLINE, EMBASE, Cochrane databases, and CINAHL Ebsco from their inception to 4 February 2022, using a comprehensive search strategy developed with an information specialist trained in the conduct of systematic reviews (Electronic Supplementary Material [ESM] eAppendix 1). We also reviewed the reference lists of all published narrative reviews, systematic reviews, and eligible studies for additional references. Two blinded reviewers independently performed study screening at the title/abstract level and then at the full-text level using the same inclusion and exclusion criteria. At each level of the study selection process, disagreements were solved by consensus or by consultation with a third reviewer as needed.

Data collection and methodological quality assessment

Two blinded reviewers independently collected data on study characteristics (study design, location, studied population, inclusion and exclusion criteria, patient characteristics and flow, reference standard and ancillary testing definitions) and results (number of true positives, false positives, true negatives, false negatives, inconclusive results, and patients with missing data). Reviewers used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool to independently assess the methodological quality of each included study.¹⁵ Disagreements in data collection and methodological quality assessment were also solved by consensus or consultation with a third reviewer as needed. When required, reviewers attempted to contact investigators of included studies to clarify extracted data.

Data analysis

Since some studies assessed multiple ancillary test types, descriptive statistics are presented at both the study and assessment levels. Dichotomous variables are reported as

counts and proportions. To estimate ancillary test diagnostic accuracy, we performed two meta-analyses. We conducted the first meta-analysis among studies involving only clinically diagnosed death by neurologic criteria patients (where ancillary tests are used in a confirmatory role). For these studies, only sensitivity could be calculated, since all are either true positives or false negatives. We conducted the second meta-analysis among studies involving comatose patients clinically suspected of death by neurologic criteria (where ancillary tests are used in a diagnostic role). For these studies both sensitivity and specificity could be calculated.

META-ANALYSIS OF CLINICALLY DIAGNOSED DEATH BY NEUROLOGIC CRITERIA PATIENTS

We estimated partially pooled test sensitivities using a hierarchical Bayesian model in which studies were nested within ancillary test types.¹⁶ The three-level (beta-binomial) model was specified as in Kruschke and Vanpaemel, except that we only included one concentration parameter common to each ancillary test type and used a diffuse half-Cauchy prior (with scale parameter of 150) for the concentration parameters rather than a Gamma prior.¹⁷ Results are reported as partially pooled sensitivities and random-effect standard deviations (reported as posterior modes and 95% highest density intervals [HDI]). Partial pooling, achieved through the hierarchical structure of the model, took into account both between-test and between-study variance, ensuring that extreme estimates for assessments with few patients and for ancillary test types with few assessments were moderated.

META-ANALYSIS OF CLINICALLY SUSPECTED DEATH BY NEUROLOGIC CRITERIA PATIENTS

We estimated partially pooled sensitivities and specificities using a different hierarchical Bayesian model in which studies were nested within ancillary test types, as before. The three-level (hierarchical summary receiver operating characteristics [HSROC] curve) model was specified as in Rutter and Gatsonis, except that our model had three levels instead of two, and our priors were slightly more informative (for details, see ESM eAppendix 2).¹⁸ Results are reported as summary receiver operating characteristics (SROC) curves, summary operating points, and partially pooled sensitivities and specificities with random-effect standard deviations (reported as posterior modes and 95% HDI). Detailed model descriptions and Stan codes are provided in the ESM (eAppendix 2).

To explore clinical and statistical heterogeneity, we planned *a priori* to fit separate SROC curves for the

following subgroups, all based on study-level characteristics: 1) demographic group (adult patients only versus mixed children/adult patients), 2) inclusion versus exclusion of an ancillary test in the clinical diagnosis reference standard, 3) delay between clinical DNC and ancillary testing < 24 hr *vs* ≥ 24 hr, and 4) presence or absence of clinical examination confounders. Among these planned subgroup analyses, the following three were deemed feasible because of the low proportions of missing values and sufficient heterogeneity in subgroup composition: 1) patient demographic group, 2) inclusion of an ancillary test in the clinical diagnosis reference standard, and 3) presence of clinical examination confounders. For the latter subgroup analysis, we excluded three studies that used a reference standard different from that in the clinical examination (namely, four-vessel angiography), as clinical confounders do not apply to these reference standards. For the subgroup analysis, we followed the methodology as presented in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (section 10.5.3.3).¹⁹

Although we planned to conduct sensitivity analyses pertaining to the risk of selection bias, risk of interpretation bias for the studied ancillary test, and risk of bias introduced by the interpretation of the reference test, we were unable to pursue these analyses because of the small proportion of studies with a low risk of bias. Nevertheless, for the meta-analysis of clinically confirmed death by neurologic criteria patients, we performed two sensitivity analyses pertaining to our Bayesian models to assess the degree to which partial pooling influenced the meta-analysis results. First, we modified the data in such a way that all assessments had the average sample size, reducing the role of partial pooling. Second, we loosened the priors of the scale parameters. Since estimates from these sensitivity analyses were largely consistent with the main analysis, we consider our results to be robust (sensitivity analysis results not reported). All analyses were performed with R version 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria) and Stan version 2.21.0 via the rstan package in R.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Overall, 137 records met the selection criteria (Fig. 1). Justifications for study exclusion at full-text screening are provided in the ESM (eAppendix 3).

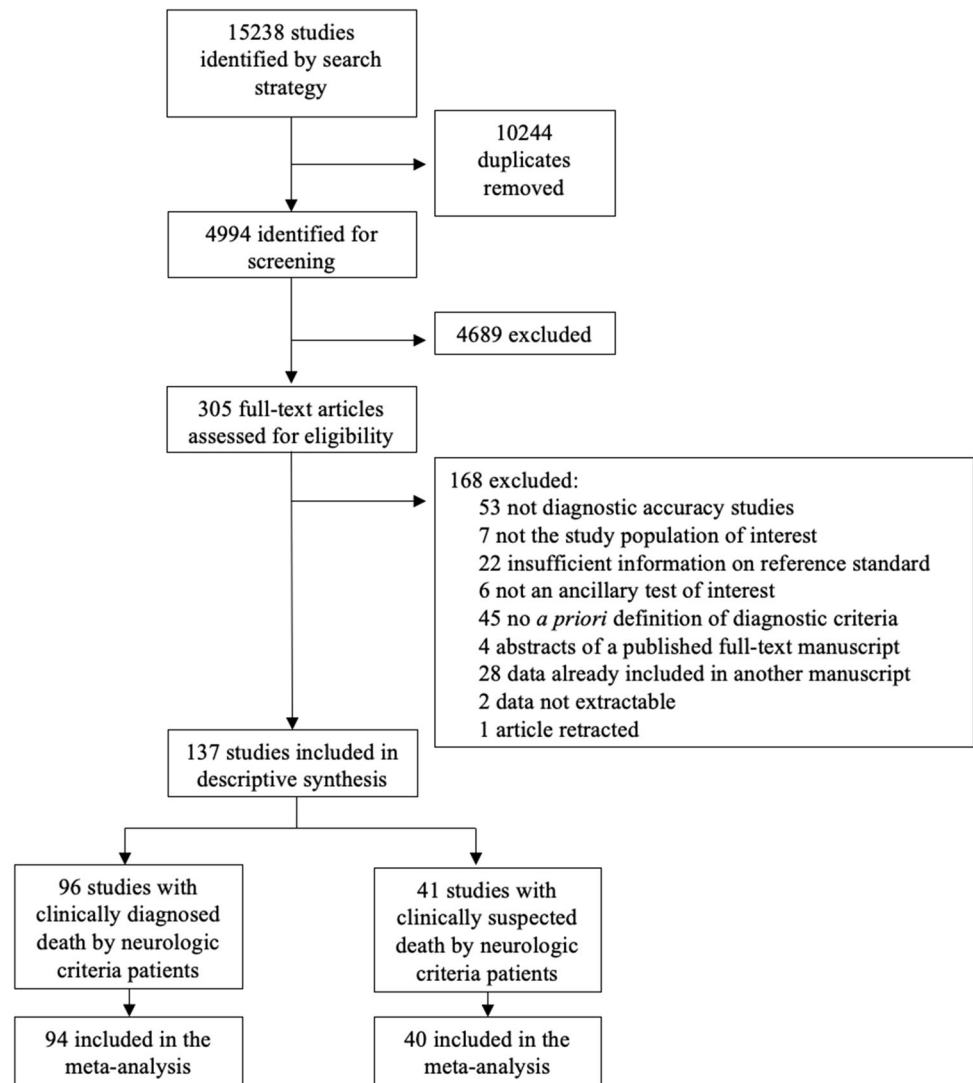
Descriptive analyses

We included 137 studies (Table 1). Ninety-six studies (70%) had been conducted solely on patients clinically diagnosed with death by neurologic criteria and the remaining 41 (30%) included patients with clinically suspected death by neurologic criteria. Studies reported data on a variety of brain injury etiologies: 111 (81%) with traumatic brain injury, 107 (78%) with intracranial hemorrhage, 83 (61%) with hypoxic-ischemic brain injury/cerebral anoxia, 77 (56%) with ischemic stroke, and 95 (69%) with other causes. Fifty-five studies (40%) had been performed without clinical confounders, such as drug intoxication, facial or cervical trauma, hypothermia, and hypotension, whereas 73 (53%) included patients with clinical confounders and 9 (7%) did not specify. Ninety-five studies (69%) did not report the delay between the reference standard and index ancillary test(s), with 28 (20%) clearly indicating a delay < 24 hr and the remaining 14 (10%) with delays ≥ 24 hr.

From the included studies, 230 assessments of ancillary tests were made (Table 2). Reference standards were clinical diagnosis and four-vessel angiography in 99% and 1% of assessments, respectively. Ancillary test types most frequently assessed were TCD (25%), CTA (18%), radionuclide imaging (16%), and EEG (16%). Characteristics of each individual study are detailed in the ESM (eAppendix 4).

Meta-analysis of sensitivity among clinically diagnosed death by neurologic criteria patients

Partially pooled ancillary test sensitivities estimated from the 94 studies comprising only clinically diagnosed death by neurologic criteria patients ($n = 8,891$ ancillary tests applied) were overall similar for all ancillary test types (Fig. 2), and ranged from 0.82 (CTA, 7-point and 10-point scales) to 0.93 (four-vessel angiography). Tests with the highest sensitivity estimates were four-vessel angiography (0.93), ^{99m}Tc-HMPAO perfusion with SPECT (0.90), CTP (0.90), TCD (0.89), MRI using time-of-flight angiography (0.89), MRI using arterial spin labeling (0.89), visual evoked potentials (0.89), and somatosensory evoked potentials (0.89). The standard deviations of the partially pooled sensitivity estimates were larger within each ancillary test type ($\sigma = 0.10$ – 0.15) than the standard deviation between the partially pooled ancillary test

Fig. 1 Flowchart diagram

sensitivities ($\sigma = 0.04$), suggesting heterogeneity within each ancillary test was considerably higher than it was between tests. Data were most abundant for TCD (24% of applied ancillary tests), EEG (30%), and four-vessel angiography (11%) (Table 3). Forest plots for partially pooled sensitivity estimates by ancillary test are provided in detail in the ESM (eAppendix 5; eFigs 1–19).

Meta-analysis of sensitivity and specificity among patients with clinically suspected death by neurologic criteria

Partially pooled ancillary test sensitivities and specificities obtained from 40 studies including patients with clinically suspected death by neurologic criteria ($n = 2,732$ ancillary tests applied) are provided in Fig. 3. Results were mostly driven by TCD data, which represented 41% of applied ancillary tests (Table 3). There were no data for four-vessel angiography. Overall, ancillary test types showed variable

partially pooled sensitivities (0.81–1.00) and specificities (0.87–1.00). The interval estimates (HDI) were considerably wide for the following ancillary test partially pooled specificities: evoked potentials, MRI, ^{99m}Tc -pertechnetate angiography, and ^{99m}Tc -HMPAO angiography. The following ancillary tests had both acceptable partially pooled sensitivity estimates and excellent partially pooled specificity estimates (in terms of both accuracy and precision): CTA (all scales), CTP, ^{99m}Tc -DTPA angiography, ^{99m}Tc -HMPAO perfusion (with or without SPECT), TCD, and EEG. Forest plots for partially pooled sensitivities and specificities by ancillary test, as well as respective SROC curves, are provided in the ESM (eAppendix 5; eFigs 20–53).

Risk of bias assessment

Study risk of bias assessment according to the four QUADAS-2 domains is summarized in Fig. 4. Overall,

Table 1 Descriptive analyses at the study level

Variable	Category	n/total N (%)
Total number of studies		137/137 (100%)
Study population	Clinically diagnosed death by neurologic criteria	96/137 (70%)
	Clinically suspected death by neurologic criteria	41/137 (30%)
Study type	Prospective cohort study	25/137 (18%)
	Retrospective cohort study	34/137 (25%)
	Unclear cohort study	61/137 (45%)
	Prospective case–control study	6/137 (4%)
	Retrospective case–control study	1/137 (0.7%)
	Unclear case–control study	5/137 (4%)
	Case series	5/137 (4%)
Language of publication	English	113/137 (82%)
	Other than English	24/137 (18%)
Demographic group	Adults only	45/137 (33%)
	Adults and children	73/137 (53%)
	Missing	19/137 (14%)
Brain injury causes (not mutually exclusive)	Cerebral anoxia	83/137 (61%)
	Traumatic brain injury	111/137 (81%)
	Intracranial hemorrhage	107/137 (78%)
	Ischemic stroke	77/137 (56%)
	Other causes	95/137 (69%)
Clinical examination confounders	Present	73/137 (53%)
	Absent	55/137 (40%)
	Missing	9/137 (7%)
Delay between reference and index tests	Less than 24 hours	28/137 (20%)
	24 hours or more	14/137 (10%)
	Missing	95/137 (69%)
Risk of bias for patient selection	Low	9/137 (7%)
	Unclear	14/137 (10%)
	High	114/137 (83%)
Risk of bias for patient flow	Low	16/137 (12%)
	Unclear	56/137 (41%)
	High	65/137 (47%)

the proportions of studies with a low risk of bias were 7% for patient selection and 12% for patient flow, whereas the proportion of assessments with a low risk of bias was 44% for the interpretation of the reference standard and 12% for the interpretation of the ancillary test. One study (0.7%) had low risks of bias on all QUADAS-2 elements.

Subgroup analyses

Although data were too sparse to allow reliable subgroup analyses, these did not show significant differences between diagnostic accuracy estimates according to patient demographic group, inclusion of an ancillary test

in the clinical diagnosis reference standard, or presence of clinical examination confounders (ESM eFig. 54).

Discussion

In this systematic review and meta-analysis, we assessed a wide variety of ancillary tests currently used in practice and found that 136/137 eligible studies (99%) had an unclear or high risk of bias on at least one QUADAS-2 domain. Study characteristics support this assessment as most studies were conducted in the presence of clinical confounders that alter the reference standard's validity (53%) or did not specify the delay between reference standards and index tests

Table 2 Descriptive analyses at the assessment level

Variable	Category	n/total N (%)
Total number of assessments		230/230 (100%)
Reference standard	Clinical diagnosis	227/230 (99%)
	With an ancillary test	43/230 (19%)
	Without an ancillary test	184/230 (81%)
Ancillary test type	Four-vessel angiography	3/230 (1%)
	Radionuclide imaging	0/230 (0.0%)
	Four-vessel angiography	22/230 (10%)
	Radionuclide imaging	36/230 (16%)
	CT perfusion imaging	4/230 (2%)
	CT angiography	42/230 (18%)
	Magnetic resonance imaging	5/230 (2%)
	Magnetic resonance perfusion imaging	1/230 (0.4%)
	Magnetic resonance angiography	1/230 (0.4%)
	Transcranial Doppler ultrasound	58/230 (25%)
	Electroencephalography	36/230 (16%)
	Evoked potentials	21/230 (9%)
	Xenon CT	2/230 (0.9%)
	Other	2/230 (0.9%)
Risk of bias for interpretation of reference standard	Low	102/230 (44%)
	Unclear	118/230 (51%)
	High	10/230 (4%)
Risk of bias for interpretation of ancillary test	Low	24/230 (10%)
	Unclear	61/230 (27%)
	High	145/230 (63%)

(69%). Furthermore, most studies only included patients with clinically diagnosed death by neurologic criteria (70%); results from these studies, where the ancillary test is used in a confirmatory role, are not translatable to situations where ancillary tests have a diagnostic role, since they apply to a different patient population and do not assess the trade-off between-test sensitivity and specificity. Finally, we observed significant heterogeneity in sensitivity and specificity estimates across studies. In fact, there was greater heterogeneity in ancillary test accuracy within each ancillary test type than between ancillary test types. These findings likely reflect high variability in the methodological quality of included studies. Since these concerns challenge the internal validity of studies included in this systematic review, caution is mandated in the choice and use of ancillary tests for DNC, as their diagnostic accuracy has not yet been extensively validated in high-quality, rigorous studies.

The methodological shortcomings of the studies included in our meta-analyses call into question the validity of our sensitivity and specificity estimates; however, some general findings remain of interest. First,

current data suggest that ^{99m}Tc -HMPAO perfusion (both with and without SPECT), EEG, and TCD have reasonable diagnostic accuracy. A recent review of national DNC protocols found that these modalities are the most frequently recommended ancillary tests worldwide in addition to four-vessel angiography, for which we did not find data on specificity.⁸ Nevertheless, these tests all have specific limitations in clinical practice. For instance, TCD is not applicable in 10–20% of patients that have a poor acoustic window or in patients with significant structural damage to the cranium; EEG is not appropriate in cases of deep chemical sedation; and nuclear imaging is not universally accessible. Second, there is large uncertainty in the specificity estimates of evoked potentials, MRI, ^{99m}Tc -pertechnetate angiography, and ^{99m}Tc -HMPAO angiography, suggesting that these tests are inappropriate for DNC. Third, in the context of clinically diagnosed death by neurologic criteria, where ancillary tests are used in a confirmatory role, tests have similar sensitivities overall. Some tests have, however, been subject to less investigation, such as CTP and MRI.

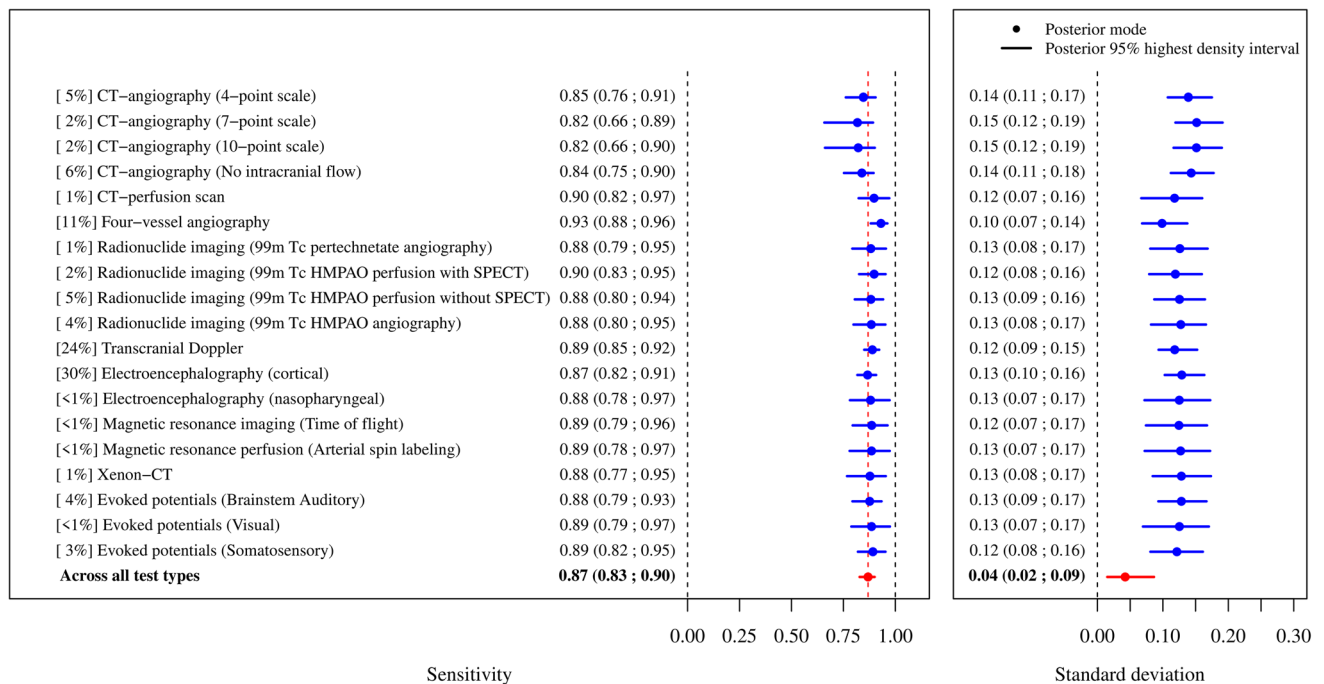


Fig. 2 Ancillary test sensitivities obtained from the studies comprising clinically diagnosed death by neurologic criteria patients. [%] represents the proportion of ancillary tests performed

The recent World Brain Death Project offers guidance on ancillary testing for DNC, some of which is supported by our findings.² First, the project recommendations suggest that four-vessel angiography, TCD, and radionuclide imaging combining diffusible radiopharmaceuticals and SPECT are the three ancillary tests deemed most appropriate for DNC. Our analysis indeed shows that these tests have the most robust diagnostic accuracy based on currently available data, which we reiterate is subject to significant bias. Nevertheless, we did not find data on the accuracy of four-vessel angiography among comatose patients, so it was not possible to estimate this ancillary test's specificity, despite it being historically considered the gold-standard ancillary test for DNC. The World Brain Death Project recommendations also caution against the use of CTA, CTP, and MRA, as they have not been sufficiently studied, which is consistent with our findings. Nevertheless, our results do not provide evidence to support other suggestions made in these guidelines. For instance, there are no data supporting the adjunct role of evoked potentials in patients initially evaluated with EEG. Although there is a physiologic argument to combining evoked potentials, which assess neuronal integrity of the brainstem and cortico-subcortical structures, to electroencephalographic evaluation of supratentorial activity, available data on

in the test type among $n = 8,891$ ancillary tests applied. HMPAO = hexamethylpropyleneamine oxime; SPECT = single-photon emission computed tomography

evoked potential diagnostic accuracy yield specificity estimates with high statistical uncertainty.

Our study has several strengths. Although prior studies have assessed the diagnostic validity of TCD^{20,21} and CTA,^{22,23} our work has assessed the diagnostic accuracy of a wide arsenal of ancillary tests currently used in clinical practice using an exhaustive search strategy. We also excluded studies without an *a priori* definition of ancillary test diagnostic criteria to adhere to strict methodological standards.²⁴ Our findings were robust to statistical model sensitivity analyses, which did not significantly alter estimates. Despite the paucity of data available for several ancillary test types, our analytical model was able to provide clinically useful parameter estimates and heterogeneity estimates for all included ancillary tests. Finally, subgroup analyses allowed us to model and inspect sources of heterogeneity including the presence of clinical confounders to the neurologic examination and the inclusion of an ancillary test in the reference standard. The major limitation of our work is the reliance on data provided by studies with unclear or high risk of bias, which calls into question the validity of our meta-analysis estimates. Importantly, the reference standard, which was a clinical neurologic examination in most included studies, may have been inaccurate in studies where confounders to the examination had not been clearly excluded. The degree to which this influenced our ancillary test diagnostic

Table 3 Number of patients pooled by study population and ancillary test type

Ancillary test	Studies including clinically diagnosed death by neurologic criteria patients		Studies including patients with clinically suspected death by neurologic criteria	
	(<i>n</i> of patients tested)	(%)	(<i>n</i> of patients tested)	(%)
CT angiography (4-point scale)	466	5.2	303	11.1
CT angiography (7-point scale)	191	2.1	79	2.9
CT angiography (10-point scale)	215	2.4	54	2.0
CT angiography (no intracranial flow)	502	5.6	150	5.5
CT perfusion imaging	88	1.0	40	1.5
Four-vessel angiography	951	10.7	0	0.0
Radionuclide imaging (^{99m} Tc-pertechnetate angiography)	64	0.7	254	9.3
Radionuclide imaging (^{99m} Tc-DTPA angiography)	0	0.0	14	0.5
Radionuclide imaging (^{99m} Tc-HMPAO perfusion with SPECT)	135	1.5	76	2.8
Radionuclide imaging (^{99m} Tc-HMPAO perfusion without SPECT)	431	4.8	93	3.4
Radionuclide imaging (^{99m} Tc-HMPAO angiography)	395	4.4	21	0.8
Transcranial Doppler	2,147	24.1	1,108	40.6
Electroencephalography (cortical)	2,643	29.7	264	9.7
Electroencephalography (nasopharyngeal)	6	0.1	0	0.0
Magnetic resonance imaging (time-of-flight angiography)	26	0.3	30	1.1
Magnetic resonance imaging (DWI/ADC)	0	0.0	17	0.6
Magnetic resonance venography	0	0.0	17	0.6
Magnetic resonance perfusion imaging (arterial spin labeling)	5	0.1	0	0.0
Xenon CT	45	0.5	0	0.0
Evoked potentials (brainstem auditory)	324	3.6	98	3.6
Evoked potentials (visual)	10	0.1	0	0.0
Evoked potentials (somatosensory)	247	2.8	114	4.2
Total	8,891	100.0	2,732	100.0

ADC = apparent diffusion coefficient; DWI = diffusion weighted imaging; HMPAO = hexamethylpropyleneamine oxime; SPECT = single-photon emission computed tomography

accuracy estimates is uncertain. Nevertheless, our subgroup analyses comparing the presence or absence of confounders to the clinical examination, and the addition of an ancillary test to the reference standard, did not disclose any significant differences in our point estimates, which suggests that this may not have been a significant source of clinical heterogeneity in our meta-analysis. Studies were also heterogeneous with respect to many other characteristics, such as patient demographics, brain injury etiologies, choice of ancillary tests applied, and ancillary test technology. Moreover, our statistical modeling did not allow us to quantify ancillary tests' areas under the SROC curve, but use of this measure to represent diagnostic accuracy is controversial.²⁵ Our analytic approach also assumed that ancillary tests were exchangeable, although in reality some studies used the same material to interpret different ancillary tests (for instance, some studies examined different CTA scales using the same images), which calls into question whether these tests may have

some dependency unaccounted for by the models. Finally, we did not consider other factors pertaining to the use of ancillary tests, such as cost, reliability, and availability, which are beyond the scope of our study.

In light of our study's findings, we believe high-quality research is warranted to provide accurate and valid measures of DNC ancillary tests' diagnostic accuracy. As technological innovation advances, a growing number of diagnostic tests are being developed in neuroradiology and neurophysiology. Prior to being transposed in clinical practice for DNC, these modalities should undergo thorough accuracy evaluation in rigorous studies applied to appropriate study populations. Hopefully, ancillary tests will eventually evolve to reliably assess cerebral function even in the presence of clinical confounders, such as chemical sedation, instead of relying on other related, but not equivalent, brain physiology parameters (cerebral blood flow or perfusion). As current ancillary tests all assess surrogates for clinical brain function, it is not

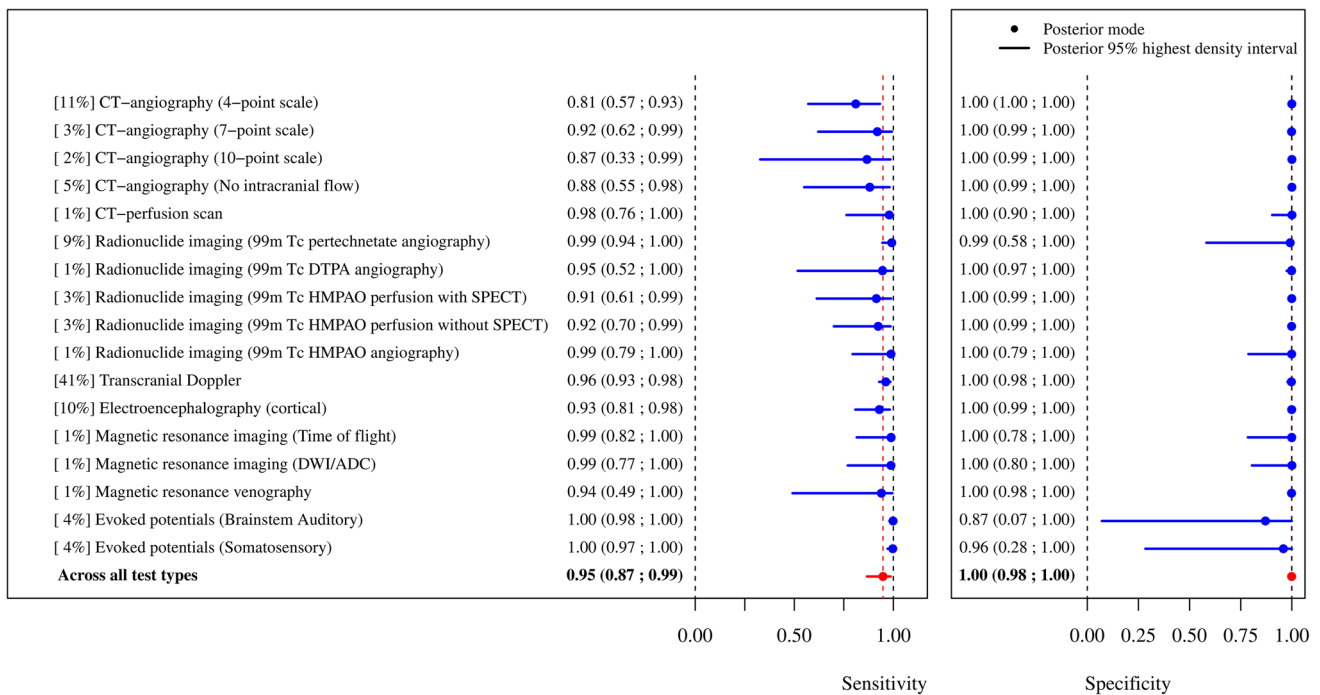


Fig. 3 Ancillary test sensitivities and specificities obtained from the studies comprising clinically suspected death by neurologic criteria patients. [%] represents the proportion of ancillary tests performed in the test type among $n = 2,732$ ancillary tests applied.

ADC = apparent diffusion coefficient; DTPA = diethylenetriamine pentaacetate; DWI = diffused weighted imaging; HMPAO = hexamethylpropyleneamine oxime; SPECT = single-photon emission computed tomography

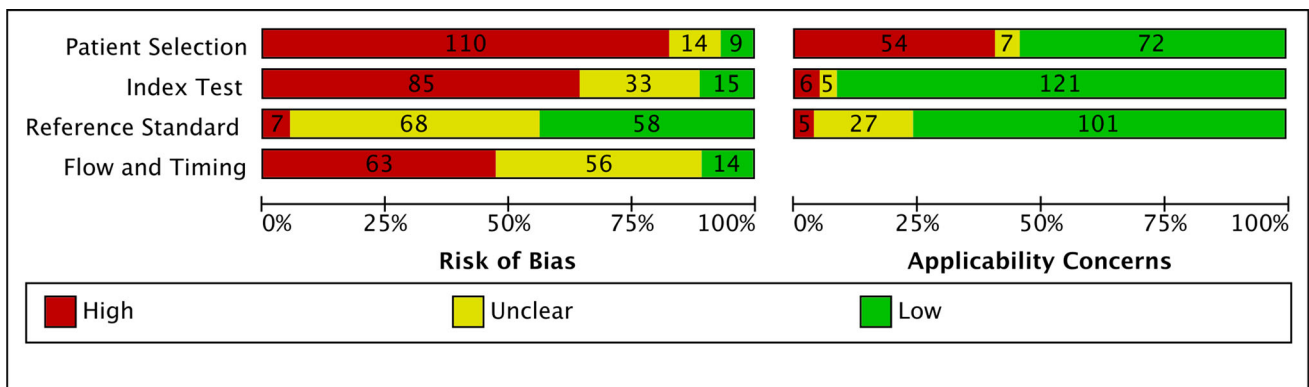


Fig. 4 Quality Assessment of Diagnostic Accuracy Studies-2 study risk of bias summary

surprising that test sensitivities are imperfect, as these tests may often show persistent blood flow, perfusion, or neurophysiologic function that is insufficient to produce pertinent clinical cerebral function, particularly among patients suspected of death by neurologic criteria following a primary infratentorial brain injury.^{26,27} Nevertheless, until significant diagnostic advances in ancillary testing are made, clinical examination should remain the cornerstone of DNC and ancillary testing should retain its role in providing further assurance to the presence of death by

neurologic criteria in situations where the clinical examination may be unreliable or impossible to complete.

In conclusion, clinicians employing ancillary tests in DNC should be aware that the studies assessing their diagnostic accuracy have modest methodological quality and are subject to significant risk of bias. Our findings have implications for clinical practice since patients who require ancillary testing in the process of DNC should be subjected to tests with near-perfect specificity and robustly studied diagnostic accuracy. Further high-quality studies are required to thoroughly validate ancillary tests for DNC.

Author contributions Michaël Chassé designed the study. Joel Neves Briard, Émile Lemoine, Polina Titova, Shane English, and Michaël Chassé performed data collection. Joel Neves Briard, Roy Nitulescu, and Michaël Chassé performed data analysis. Joel Neves Briard and Michaël Chassé drafted the manuscript, which was revised by all other authors for intellectual content. Michaël Chassé is the senior author and supervised the study. All authors reviewed the final version of the manuscript.

Acknowledgments The authors would like to thank Peter Glen and Marie-Helen Doyle for their contribution in the initial screening of titles and abstracts. We would also like to acknowledge Risa Schor and Daniela Ziegler for their contribution to the development of the search strategy.

Disclosures The authors declare no competing interests.

Funding statement This study was funded by the Canadian Institute of Health Research (#KRS-132043). Drs Lauzier and Chassé are recipients of a research salary support award from the Fonds de Recherche du Québec – Santé (FRQS). Dr. Turgeon holds the Canada Research Chair in Critical Care Neurology and Trauma. Dr. English holds a New Investigator Award from the Heart and Stroke Foundation.

Prior conference presentations Authors presented preliminary analyses at the American Academy of Neurology 2021 Annual Meeting and the 2021 Canadian Critical Care Forum.

Editorial responsibility This submission was handled by Dr. David Greer, Guest Editor, *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. Canadian Institute for Health Information. Annual statistics on organ replacement in Canada: Dialysis, Transplantation and Donation, 2010 to 2019. Ottawa, ON: CIHI; 2020.
2. Greer DM, Shemie SD, Lewis A, et al. Determination of brain death/death by neurologic criteria: the World Brain Death Project. *JAMA* 2020; 324: 1078–97. <https://doi.org/10.1001/jama.2020.11586>
3. Truog RD, Robinson WM. Role of brain death and the dead-donor rule in the ethics of organ transplantation. *Crit Care Med* 2003; 31: 2391–6. <https://doi.org/10.1097/01.ccm.0000090869.19410.3c>
4. Plourde G, Neves Briard J, Shemie SD, Shankar J, Chassé M. Flow is not perfusion, and perfusion is not function: ancillary testing in brain death. *Can J Anesth* 2021; 68: 953–61. <https://doi.org/10.1007/s12630-021-01988-2>
5. Lewis A, Liebman J, Kreiger-Benson E, et al. Ancillary testing for determination of death by neurologic criteria around the world. *Neurocrit Care* 2021; 34: 473–84. <https://doi.org/10.1007/s12028-020-01039-6>
6. Braksick SA, Robinson CP, Gronseth GS, Hocker S, Wijdsicks EF, Rabinstein AA. Variability in reported physician practices for brain death determination. *Neurology* 2019; 92: e888–94. <https://doi.org/10.1212/wnl.00000000000007009>
7. Chassé M, Neves Briard J, Yu M, et al. Clinical evaluation and ancillary testing for the diagnosis of death by neurologic criteria: a cross-sectional survey of Canadian intensivists. *Can J Anesth* 2022; 69: 353–63. <https://doi.org/10.1007/s12630-021-02166-0>
8. Lewis A, Bakkar A, Kreiger-Benson E, et al. Determination of death by neurologic criteria around the world. *Neurology* 2020; 95: e299–309. <https://doi.org/10.1212/wnl.00000000000009888>
9. Chassé M, Glen P, Doyle MA, et al. Ancillary testing for diagnosis of brain death: a protocol for a systematic review and meta-analysis. *Syst Rev* 2013; 2: 100. <https://doi.org/10.1186/2046-4053-2-100>
10. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; 339: b2535. <https://doi.org/10.1136/bmj.b2535>
11. Shemie SD, Lee D, Sharpe M, Tampieri D, Young B, Canadian Critical Care Society. Brain blood flow in the neurological determination of death: Canadian expert report. *Can J Neurol Sci* 2008; 35: 140–5. <https://doi.org/10.1017/s0317167100008544>
12. Frampas E, Videcoq M, de Kerviler E, et al. CT angiography for brain death diagnosis. *AJNR Am J Neuroradiol* 2009; 30: 1566–70. <https://doi.org/10.3174/ajnr.a1614>
13. Dupas B, Gayet-Delacroix M, Villers D, Antonioli D, Veccherini MF, Soullou JP. Diagnosis of brain death using two-phase spiral CT. *AJNR Am J Neuroradiol* 1998; 19: 641–7.
14. Combes JC, Chomel A, Ricolfi F, d'Athis P, Freysz M. Reliability of computed tomographic angiography in the diagnosis of brain death. *Transplant Proc* 2007; 39: 16–20. <https://doi.org/10.1016/j.transproceed.2006.10.204>
15. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
16. Rubin DB. Estimation in parallel randomized experiments. *J Educ Stat* 1981; 6: 377–401. <https://doi.org/10.3102/10769986006004377>
17. Kruschke JK, Vanpaemel W. Bayesian estimation in hierarchical models. In: Busemeyer JR, Wang W, Townsend JT, Eidels A (Eds.). *The Oxford Handbook of Computational and Mathematical Psychology*. Oxford: Oxford University Press; 2015: 279–99.
18. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001; 20: 2865–84. <https://doi.org/10.1002/sim.942>
19. The Cochrane Collaboration. Cochrane handbook for systematic reviews of diagnostic test accuracy version 2.0, 2022. Available from URL: <https://training.cochrane.org/handbook-diagnostic-test-accuracy> (accessed November 2022).
20. Chang JJ, Tsivgoulis G, Katsanos AH, Malkoff MD, Alexandrov AV. Diagnostic accuracy of transcranial doppler for brain death confirmation: systematic review and meta-analysis. *AJNR Am J Neuroradiol* 2016; 37: 408–14. <https://doi.org/10.3174/ajnr.a4548>
21. Monteiro LM, Bollen CW, van Huffelen AC, Ackerstaff RG, Jansen NJ, van Vught AJ. Transcranial Doppler ultrasonography to confirm brain death: a meta-analysis. *Intensive Care Med* 2006; 32: 1937–44. <https://doi.org/10.1007/s00134-006-0353-9>

22. *Kramer AH, Roberts DJ.* Computed tomography angiography in the diagnosis of brain death: a systematic review and meta-analysis. *Neurocrit Care* 2014; 21: 539–50. <https://doi.org/10.1007/s12028-014-9997-4>
23. *Taylor T, Dineen RA, Gardiner DC, Buss CH, Howatson A, Pace NL.* Computed tomography (CT) angiography for confirmation of the clinical diagnosis of brain death. *Cochrane Database Syst Rev* 2014; 3: CD009694. <https://doi.org/10.1002/14651858.cd009694.pub2>
24. *Chassé M, Fergusson DA.* Diagnostic accuracy studies. *Semin Nucl Med* 2019; 49: 87–93. <https://doi.org/10.1053/j.semnuclmed.2018.11.005>
25. *Miller SW, Sinha D, Slate EH, Garrow D, Romagnuolo J.* Bayesian adaptation of the summary ROC curve method for meta-analysis of diagnostic test performance. *J Data Sci* 2009; 7: 349–64.
26. *Neves Briard J, Plourde G, Nitulescu R, et al.* Infratentorial brain injury among patients suspected of death by neurologic criteria: a systematic review and meta-analysis. *Neurology* 2023; 100(4): e443–53. <https://doi.org/10.1212/wnl.0000000000201449>
27. *Hoffmann O, Masuhr F.* Use of observational periods or ancillary tests in the determination of brain death in Germany. *Eur Neurol* 2015; 74: 11–7. <https://doi.org/10.1159/000431089>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.