Diagnostic accuracy of interictal source imaging in presurgical epilepsy evaluation: A systematic review from the E-PILEPSY consortium

Brian E Mouthaan[1], Matea Rados[1], Paul Boon[2], Evelien Carrette[2], Beate Diehl[3], Julien Jung[4], Vasilios Kimiskidis[5], Teia Kobulashvili[6], Giorgi Kuchukhidze[6], Pål G Larsson[7], Markus Leitinger[6], Philippe Ryvlin[8], Fergus Rugg-Gunn[3], Margitta Seeck[9], Serge Vulliémoz[9], Geertjan Huiskamp[1], Frans SS Leijten[1], Pieter Van Eijsden[1], Eugen Trinka[6,10], Kees PJ Braun[1], on behalf of the E-PILEPSY consortium

[1]Department of (Child) Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, PO Box 85090, 3508 AB Utrecht, The Netherlands

[2]Reference Center for Refractory Epilepsy, Department of Neurology, Ghent University Hospital, Belgium

[3]National Hospital for Neurology and Neurosurgery, University College London Hospitals, London, United Kingdom and Department of Clinical and Experimental Epilepsy, University College London

[4]Department of Functional Neurology and Epileptology, Institute of Epilepsies (IDEE), Hospices Civils de Lyon, Lyon, France

[5]Laboratory of Clinical Neurophysiology, Medical School, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece

[6]Department of Neurology, Christian-Doppler University Hospital, Paracelsus Medical University, and Centre for Cognitive Neuroscience, Salzburg, Austria

[7]Department of neurosurgery, Clinic of surgery and neuroscience, Oslo University Hospital, Norway

[8]Department of Clinical Neurosciences, Lausanne University Hospital, Lausanne, Switzerland.

[9]EEG and Epilepsy Unit, Department of Clinical Neurosciences, University Hospital of Geneva, Switzerland

[10]Institute of Public Health, Medical Decision Making and HTA, UMIT, Private University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol, Austria

Corresponding author:

Professor K.P.J. Braun, MD, PhD

Department of (Child) Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands

P.O. Box 85090 | 3508 AB UTRECHT

Phone: +31 88 75 543 41

Fax: +31 (0)88 75 55350

E-Mail: k.braun@umcutrecht.nl

Abstract

Objective

Interictal high resolution (HR-) electric source imaging (ESI) and magnetic source imaging (MSI) are non-invasive tools to aid epileptogenic zone localization in epilepsy surgery candidates. We carried out a systematic review on the diagnostic accuracy and quality of evidence of these modalities.

Methods

Embase, Pubmed and the Cochrane database were searched on 13 February 2017. Diagnostic accuracy studies taking post-surgical seizure outcome as reference standard were selected. Quality appraisal was based on the QUADAS-2 framework.

Results

Eleven studies were included: eight MSI (n=267), three HR-ESI (n=127) studies. None was free of bias. This mostly involved: selection of operated patients only, interference of source imaging with surgical decision, and exclusion of indeterminate results. Summary sensitivity and specificity estimates were 82% (95% CI: 75-88%) and 53% (95% CI: 37-68%) for overall source imaging, with no statistical difference between MSI and HR-ESI. Specificity is higher when partially concordant results were included as non-concordant ($p<0.05$). Inclusion of indeterminate test results as non-concordant lowered sensitivity ($p<0.05$).

Conclusions

Source imaging has a relatively high sensitivity but low specificity for identification of the epileptogenic zone.

Significance

We need higher quality studies allowing unbiased test evaluation to determine the added value and diagnostic accuracy of source imaging in the presurgical workup of refractory focal epilepsy.

Highlights

- Interictal source imaging studies are biased and show heterogeneity for population and test method
- Source imaging sensitivity and specificity was 82% (95% CI: 75-88%) and 53% (37-68%) respectively
- Diagnostic accuracy for (extra)temporal and (non)-lesional patients show no statistical differences

Key words

Magnetoencephalography, electroencephalography, source localization, surgery, sensitivity, specificity

1 Introduction

Epilepsy surgery can be a curative treatment option in patients with refractory focal epilepsy. Success of surgery depends on accurate delineation of the epileptogenic zone (EZ). The EZ is a theoretical construct describing the minimum volume of cortical tissue, that is responsible for generation of habitual seizures, and that has to be resected to produce seizure-freedom[52]. Clinical semiology, imaging and electrophysiological data yield important localizing information about the EZ. Video-EEG monitoring (vEEG), magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission tomography (SPECT) and intracranial EEG (iEEG) are frequently used modalities in the presurgical workup[38].

High resolution electric source imaging (HR-ESI) and magnetic source imaging (MSI) are additional electrophysiological techniques to non-invasively localize epileptogenic brain regions. By reconstructing the electric or magnetic potentials as identified by EEG or MEG, locations of underlying source currents are estimated and subsequently combined with structural imaging. Source localization is often based on interictal epileptic discharges (IED), which are frequently observed during EEG or MEG recording. Because IEDs mark the irritative zone and may not always be concordant with the seizure onset zone [2, 35], source localization based on ictal data has been proposed in patients with high seizure frequency [11, 25, 70, 62]. So far most experience has been acquired with interictal ESI and MSI, but complete clinical integration across all epilepsy surgery centers has not been established yet [9, 59, 82]. A number of epilepsy-specific clinical practice guidelines and general recommendations have been published[8, 10, 51, 4]. Yet, disparities in current practice among users on aspects such as the number and positions of sensors, and the selection of inverse and volume conduction models remain [9, 59]. This may be due to the various

technical complex and non-intuitive aspects involved in source localization [14]. MSI and HR-ESI appear to be complementary techniques that differ in their sensitivity for various neural generators: fundamentally MSI is more accurate than ESI in detecting superficial tangentially orientated sources and involves no signal distortion (volume conduction) while EEG allows recording of all source orientations, is more sensitive to deep sources and is less affected by motion artefacts[56]. One previous systematic review on MSI reported that there is insufficient evidence on the use of MSI in the presurgical evaluation[50]. This review did not evaluate HR-ESI and was published nearly a decade ago. In light of the E-PILEPSY network [http://www.e-pilepsy.eu], which aims to harmonize and optimize presurgical diagnostic procedures across European countries, we carried out a systematic review to assess the diagnostic accuracy of interictal HR-ESI and MSI to localize epileptogenic regions of interest in epilepsy surgery candidates.

2 Methods

2.1 Establishment of task force and protocol

As a first step we conducted a broad literature search to allow an orientation on the available literature. Based on this, we established a systematic review protocol containing research questions and study inclusion criteria. A task force was formed of 14 E-PILEPSY members (corresponding co-authors) to allow a broad acceptance of the systematic review protocol and to aid other review tasks (e.g. paper screening, data extraction). Members were familiar with both the field of source imaging and epilepsy surgery, having different educational backgrounds (physicists and physicians) and varying experience (PhD students to

professors). Consensus was reached among task force members on the final systematic review protocol.

2.2 Search strategy

PubMed, Embase and Cochrane were last searched on 13 February 2017 for articles on the diagnostic value of EEG and MEG source localization in epilepsy. We included synonyms and abbreviations for the terms of interest, and used subject headings (i.e. MeSH, EmTree). The search syntaxes are provided in Appendix A. The search strategy was limited to humans, English language, and publication date after 1995. Duplicates were eliminated.

2.3 Study selection

Title and abstract screening of the studies was done by one of two author couples (BM & MR, GH & FL). Discrepancies in eligibility were discussed and final agreement was reached through a consensus meeting. References found in source imaging review papers were screened. Studies were excluded if there was insufficient information to fully assess their eligibility (e.g. full text not available in English, unavailable abstract, unavailable full text). Letters, commentaries, conference abstracts, poster presentations and supplementary materials were also excluded, as were articles focusing on epilepsy not amenable to surgery (i.e. rolandic epilepsy or idiopathic generalized epilepsy), and EEG-fMRI. Other procedures, such as connectivity analysis [16, 87], source volume estimation [13], beta-band activity source imaging [33], slow wave interictal MSI [36] and analysis of high frequency oscillations [88], were not subject of the review due to their limited clinical utilization compared to traditional interictal source imaging.

Studies were then screened on full text by couples of two independent taskforce members. Full text inclusion criteria were: epilepsy surgery candidates, interictal MSI or interictal ESI, diagnostic accuracy based on level of concordance between ESI or MSI source location and the resected area taking seizure outcome as reference standard.

We consulted authors in the case of unavailable full-text. Studies needed to report on sensitivity or specificity including confidence intervals and/or absolute numbers that allow calculation of these statistics. If this was not the case the study was excluded. The full eligibility checklist is provided in Appendix B. Disagreement was discussed and final agreement was reached between the members of each couple before they submitted their full text screening results.

2.4 Critical appraisal and data extraction

An online quality appraisal and data extraction form was created that was first piloted before use. Studies were assessed for methodological quality against modified QUADAS-2 (quality assessment of diagnostic accuracy studies) criteria[90]. Certain aspects of the QUADAS-2 framework were thought to be irrelevant and thus left out from the quality appraisal, such as avoidance of disease progression bias. Results between reviewers of a couple were compared and, if necessary, a web-meeting or email conversation was initiated with a third person to resolve disagreement.

Studies were excluded during data extraction in the case of: 1- sample size less than 10 participants, 2- patients included with less than 6 months follow up, 3- not categorizing

surgical outcome by means of Engel[26] or ILAE classification [92], 4- no classification of concordant and non-concordant ESI or MSI results, 5- low resolution ESI (<64 channels), 6- not presenting results for low resolution (<64 channels) and high resolution (≥64 channels) ESI separately, 7- absence of patients in any of the four groups of the 2x2 contingency table (i.e. zero values).

2.5 Data analysis

We considered diagnostic accuracy as the ability of source localization (HR-ESI or MSI) to detect and localize an epileptogenic source within a brain region that is subsequently validated as epileptogenic based on resection and surgical outcome. Concordance between source location and resected volume was considered as 'test positive' and may represent source localization within resection volume or sublobar co-localization of the source estimate with resection volume. Non-concordance was defined as 'test negative'. Post-surgical outcome was taken as reference standard, discriminating between good and poor surgical outcome (figure 1).

In accordance with most studies, sensitivity was defined as the proportion of good-outcome patients with concordant classification (i.e. test positive)**,** relative to the total number of good-outcome patients. Specificity was defined as the proportion of poor-outcome patients with non-concordant classification (i.e. test negative), relative to the total number of poor-outcome patients.

Positive and negative predictive values (PPV, NPV) were not considered in this study. Since the proportion of patients with either non-localizing source localization results, or those not proceeding to surgery after source localization procedure, was unknown in most studies the calculation of positive and negative predictive values was deemed unreliable. Moreover,

predictive values are expected to vary strongly among studies due to different presurgical workup strategies (comprehensive versus limited workup), surgical strategies (liberal versus conservative resection) and patient characteristics (e.g. lesional versus non-lesional, TLE versus ETLE). Results from patients undergoing hemispherectomy, hemispherotomy, or re-resection (second stage surgery) were excluded from analysis.

To establish homogeneity among studies, we defined “good outcome” as Engel I or ILAE 1-2. Results from studies classifying Engel II or ILAE 3 as good outcome were manually corrected to our proposed definition, if data was provided. If not, the definition as proposed in the study was adopted. For studies quantitatively reporting level of overlap between resection area and source (e.g. 90% of dipoles within resected area) the dichotomization threshold as used in the study was selected- discriminating concordant from non-concordant - to account for the effect of source localization parameters on threshold definition only known by study authors. We selected the concordance definition (i.e. sublobar co-localization or location within resection volume) that the primary study used for sensitivity and specificity calculation and included this in our analysis.

Allocation of partially concordant categories theoretically affects sensitivity and specificity calculations (figure 1). For those studies that provided information, we allocated partially concordant results as normal concordant results to ensure homogeneity across all studies regarding classification. Patients with multiple sources (either within one lobe or across multiple lobes) were also handled as concordant, even when only one source was partially resected. Indeterminate test results (i.e. no source localization possible due to insufficient numbers of IEDs, too many artefacts or scattered sources) were handled as a separate category and reported as a percentage from study sample size for each study [20, 79].

We calculated mean proportions (including 95% confidence intervals) for good-surgical outcome in each concordance group using a weighted binary random-effects model. We calculated a summary estimate of sensitivity and specificity by means of the bivariate linear mixed model by Reitsma to account for the two-dimensional trade-off between sensitivity and specificity[73, 30]. Subgroup analysis was performed for the selection of studies that provided surgical outcome information for indeterminate and partial concordance test results. Subgroup analysis for epilepsy location (TLE versus ETLE) and MRI results (lesional versus non-lesional) was performed. Statistical programming was done using the program Open-Meta Analyst and the mada package used in .R [72, 24].

3 Results

3.1 Study Selection

Figure 2 visualizes the flow of studies through the review process. Our search yielded 1964 papers after removal of duplicates. After title/abstract screening 96 papers were selected for full text assessment. Reference checking of review papers revealed two new studies. Fifty one studies proceeded to data extraction and quality appraisal, of which twelve were prospective. Seven of these were excluded for reason of different study objectives (n=5) [17, 46, 94, 43, 22], not concerning the population of interest (n=1)[5] or for not using a reference standard of interest [91]. In total, forty studies were excluded during data extraction(see Appendix C for list of excluded studies in this phase and their exclusion reasons). This led to a total of 11 studies that were included in the review: eight on MSI, three on HR-ESI [39, 41, 76, 93, 69, 60, 44, 3, 15, 55, 27].

3.2 Study characteristics

All studies were cross-sectional cohort studies, of which five were prospective (table 1). The number of included patients with positive source localization (sample size) ranged from 14 to 52 (median 36). Seven studies reported the proportion of patients with indeterminate test results: one HR-ESI study (16%) and six MSI studies (7-36%, median 17%). Four of these six MSI studies additionally provided surgical outcome results for this group [39, 44, 93, 60].

Three studies reported that indeterminate test results may be resulting from their source localization procedure but did not specify the amount of patients to which this applied [76, 69, 27]. One study did not mention indeterminate test results at all [15].

Regarding population characteristics, one study focused on adults, one on children, and nine included both age groups. Inclusion criteria varied among studies. Indications to perform ESI or MSI were not always explicitly stated, but could be derived from inclusion criteria. Five studies provided information whether resection included single lobe or multiple lobes for each patient.

Seven studies used post-operative MRI to assess the resection volume. Three studies did not mention use of post-op MRI. One study reported that post-op MRI was available, but did not further specify if this was applied to assess resection volume [60]. Concordance with resection volume was defined as sublobar by two and as ‘source estimate within resection volume’ by nine studies.

Concordance definition varied among studies. Six studies specified how partially concordant results were handled. Three studies separately reported surgical outcome for patients with partially concordant results [76, 69, 60]. Three studies did not: two included patients with partial concordance in the complete concordance group [39, 55], and one considered this as

non-concordant [3]. Multifocal source estimates were reported by five MSI studies; their occurrence ranged between 6% and 79% (median, 39%) of the patients[39, 41, 76, 93, 60]. Three studies classified resection of only one source in a multifocal source patient as concordant. One study classified such cases as non-concordant. One study presented data for multi-cluster cases but did not state its classification, and was therefore considered non-concordant by our reviewers.

All but one study used Engel class to define outcome. Definition of good outcome varied between Engel 1a and Engel 1-2. Minimum follow up period ranged from 6 months to one year. Study duration ranged between 2-11 years.

All HR-ESI studies used sensor nets with whole-head coverage consisting of 128-256 EEG electrodes. One study used a realistic head model (FDM)[27] and two a spherical head model (SMAC)[15, 55]. Linear distributed inverse solution based on averaged spikes was used in all studies. One used LORETA[27], one LAURA[15] and one used an unspecified distributed inverse solution [55].

All MSI studies used whole-head MEG. In one study the applied technique varied between double-probe (74 channels) and whole-head (306 channels) MEG [60]. Six out of eight studies used simultaneous EEG to aid IED identification. No averaging was performed in any study. All studies used equivalent current dipole (ECD) as inverse solution. Overall, cluster definition varied among studies.

3.3 Methodological quality

Study quality was generally assessed as “poor” according to QUADAS-2; no study was free from bias (table 2). Studies scored badly on disease spectrum bias, partial verification bias

and inappropriate exclusions from data analysis. All studies enrolled a consecutive sample of patients, none were of case-control design.

Selection bias (i.e. disease spectrum bias) was applicable to eight studies. This was the case for studies selecting specific populations such as patients undergoing or not undergoing intracranial EEG, patients with frontal lobe epilepsy or patients with histologically proven focal cortical dysplasia. A consequence of our inclusion criteria was an additional general disease spectrum bias across all studies. It was believed that exclusion of patients that were considered non-eligible for surgery based on presurgical workup, most likely resulted in an over-estimation of diagnostic accuracy. We did not visualize this in the quality summary to permit between-study difference in selection bias to be noticeable.

One study was biased for the index test based on data-driven threshold selection [93]. Two retrospective studies did not report information on blinding from reference standard information, and were judged as unclear for index test bias.

Reference standard bias was observed in five studies. Good surgical outcome was defined as only Engel 1a by two studies [76, 3] or only Engel 1-2 by one[27]. Three studies included patients with follow up period between 6-12 months[60, 3, 27]. Two studies classified concordance based on sublobar co-localization [44, 27].

Bias regarding study flow was observed in ten studies. In six studies, source localization results were considered in the decision to proceed to surgery (partial verification bias), the decision for coverage/placement of ICEEG, or the area/extend of resection (differential verification bias)[76, 93, 60, 44, 3, 27, 41]. Six studies did not report surgical outcome data for indeterminate test results [39, 76, 69, 3, 27, 41]. All factors possibly led to over-optimistic

diagnostic accuracy. Insufficient data was reported by one study to permit bias judgment [15].

Variations among studies with respect to population, index test specifics, and reference standard specifics were not considered a concern regarding applicability as all these represented part of general clinical practice.

3.4 Diagnostic accuracy

Diagnostic accuracy analysis of MSI and ESI accuracy included a total of 394 patients, 267 on MSI and 127 on HR-ESI, of whom surgical outcome data was available. In 363 patients a localizing source was found: 236 MSI and 127 HR-ESI (figure 3). For all MSI studies, good surgical outcome was reached in 130/236 patients (mean: 54%, 95% CI: 45-63%). For HR-ESI this was 86/127 patients (mean: 67% 95% CI 49-85%). No statistical difference on the probability of good surgical outcome between MSI and HR-ESI studies was observed.

In total, the number of patients with good surgical outcome in the concordant group was higher (172/226 patients, mean: 76%, 95% CI: 67-86%) than the number of patients in the non-concordant group (36/111 patients, mean: 28%, 95% CI: 19-36%). Statistical difference between the concordant and non-concordant group regarding good surgical outcome probability was found for MSI and HR-ESI (table 3).

Surgical outcome data of indeterminate test results was available in four MSI studies; 18/31 patients (mean 56%, 95% CI:33-79%) had good surgical outcome. Sensitivity ranged between 50-96% for MSI and 80-91% for HR-ESI. Specificity ranged between 17-80% for MSI and 56-75% for HR-ESI (Appendix D).

Summary estimates based on the bivariate linear mixed model showed sensitivity and specificity of 82% (95% CI: 75-88%) and 53% (95% CI: 37-68%) for overall source localization. For HR-ESI, summary sensitivity and specificity were 87% (95% CI: 77-93%) and 61% (95% CI: 45-74%) respectively (Figure 4). For MSI, summary sensitivity and specificity were 79% (95% CI: 69-87%) and 46%(95% CI: 25-70%) respectively. HR-ESI and MSI sensitivity/specificity estimates did not show statistical difference ($p>0.05$).

Sensitivity and specificity estimations based on single source locations were based on five MSI studies and were 83% (95% CI: 63-93%) and 22% (95% CI: 20-72%) respectively. No statistical differences regarding sensitivity and specificity estimates were observed when multifocal sources were included in the estimation (Appendix E). Analysis for the allocation of partially concordant results was based on three studies. A statistically significant higher specificity estimate was observed when partially concordant results were categorized as non-concordant, compared to concordant categorization (69% versus 20%, $p<0.05$). Based on four studies reporting surgical outcome of patients with indeterminate source imaging results, statistically significant lower summary sensitivity was observed when indeterminate test results were included in the estimates and considered non-concordant (61% versus 76%, $p<0.05$).

Subgroup analysis showed good surgical outcome in 11/14 lesional patients (mean: 80%, 95% CI: 61-99%) and 46/68 non-lesional patients (mean: 60%, 95% CI: 43-85%) with concordant results. For non-concordant results these were 2/12 in lesional patients (mean: 16%, 95% CI: -1-34%) and 8/18 in non-lesional patients (mean: 42%, 95% CI: 19-66%). Summary sensitivity for lesional and non-lesional patients was similar, and specificity showed now apparent difference ($p=0.059$, Appendix E).

Subgroup analysis for lobar location showed that a good surgical outcome was achieved in 64/96 TLE patients (mean: 61%, 95% CI: 33-89%) and 47/92 ETLE patients (mean: 43%, 95% CI: 21-65%) with concordant results. Good surgical outcome was achieved in 9/24 TLE patients (mean: 37%, 95% CI: 19-54%) and 12/34 in ETLE patients (mean: 42%, 95% CI: 16-68%) of the non-concordant group. Summary sensitivity and specificity for TLE and ETLE subgroup were comparable and no statistically significant differences were observed (Appendix E).

4 Discussion

Electric and magnetic source localization are believed to be valuable techniques in the diagnostic workup of epilepsy surgery candidates. We performed a systematic review and included eight studies on MSI and three on HR-ESI that used seizure outcome after surgery as a reference standard. All studies were highly biased on various aspects, with considerable heterogeneity among studies regarding the included population and test methodology. Bivariate meta-analysis estimated a summary sensitivity and specificity of 82%(95% CI: 75-88%) and 53% (95% CI: 37-68%) for overall source localization and no statistical difference between HR-ESI and MSI was found.

The only previous systematic review on source localization included more studies (17 in total) and reported a higher MSI sensitivity (84% versus 79%) and higher specificity (52% versus 46%) than our study [50]. Separate pooling of sensitivity and specificity permitted authors to include more studies, even when sensitivity or specificity measures were missing in individual studies due to zero values in the 2x2 contingency tables. However, separate pooling of sensitivity and specificity fails to account for the trade-off between these two

measures, the more so when either one is not calculable for all studies . Therefore, we decided to include only studies without zero values in the 2x2 contingency tables and calculated sensitivity and specificity by means of bivariate modelling, at the cost of the number of included studies [23, 53]. Other outcomes also provide information on the clinical value of a test; of which several have been published for MEG. Changes in clinical management after MEG following previous conventional non-invasive presurgical workup, were seen in 21-35% of patients, in whom 11-75% of these changes were considered as crucial or of clear impact[22, 81]. In 23-33% of surgical candidates who required ICEEG, a change in clinical management after MEG was observed, of which 26-39% was classified 'beneficial' according to the authors [84, 46]. The level to which clinical management is changed by HR-ESI remains uncertain as such studies were not discovered by our literature search.

Heterogeneity among studies regarding the included population was observed. It has been proposed that MSI and ESI should preferentially be applied in extratemporal rather than temporal lobe epilepsy, as - in the latter - an epileptic focus may easily propagate through a well-developed and complex limbic network, leading to a more wide-spread irritative zone [51, 86]. None of the individual studies reported statistical differences in test performance between temporal, extratemporal, lesional and non-lesional epilepsy patients, which was also not found by our pooled subgroup analysis[69, 15, 55].

The underreporting of surgical outcome data for patients with indeterminate source imaging results, and the inconsistency in reporting partially concordant results, were an important finding of our study. Indeterminate test results were more frequently reported in MSI studies (six studies) than in HR-ESI studies (one study). Unexpectedly, the majority of

patients with indeterminate MSI test results had good outcome. The importance of these results was highlighted; a statistically significant lower sensitivity was observed when indeterminate test results were included in the analysis of diagnostic accuracy. We further showed that categorizing partially concordant results as non-concordant significantly affects specificity. An explorative analysis showed a statistically significant higher sensitivity for HR-ESI over MSI only when indeterminate test results and partially concordant results were calculated as non-concordant. HR-ESI studies showed, as compared to MSI studies, a very low number of indeterminate and partially concordant results , and they were all prospective and based on distributed inverse methods. Therefore, this result is likely confounded and therefore not reported as a definite result. As addressed by Papanicolaou and colleagues, cases of partial concordance between source estimate and resected area do not have the same significance as cases of complete or non-concordance [68]. In the context of other imaging modalities, partially concordant results may be clinically valuable.

The ability to record epileptic activity from deep midline structures (e.g. mesial temporal regions) is much debated, as such measures are hampered by cortical propagation and relative low signal compared to background brain activity [42, 47, 18, 31]. It often occurs that source localization records only the neocortically (anterio-temporal) propagated hippocampal spikes, not the hippocampal spikes themselves. As surgical strategy aims to resect the underlying hippocampal pathology, the source localization result are left out of the resection volume. From a strict localization perspective, such spatially distinct source solutions do not contribute to identification of the true EZ over other possible EZ's. The high

variation among studies on dealing with such results calls for consensus within the community [77].

Our study has several limitations. First, strict inclusion criteria resulted in few primary studies. A higher number of MSI studies was found compared to HR-ESI and the number of MSI patients outnumbered those with HR-ESI (267 versus 127) reflecting the novelty of HR-ESI relative to MSI.

Second, the quality appraisal was mostly designed for illustrative purposes and the degree to which each quality domain contributes to over- or underestimation of diagnostic accuracy is not quantified. Yet, it is certain that our self-induced patient selection bias, resulting from not taking into account patients who were rejected after presurgical workup, promotes both MSI and HR-ESI diagnostic accuracy over-optimistically. When verification bias (i.e. inclusion of source localization results in the presurgical workup) is present, diagnostic accuracy is corrupted by clinical decision making; surgical resection might easily be expanded after consideration of the source localization results.

Third, inclusion criteria were restricted to ESI studies using 64 EEG electrodes, as this is considered to be the minimum number of channels necessary for accurate localization[51, 48, 80]. Yet, several studies on surgical candidacy have applied ESI based on more widely clinical available long term EEG systems and report sensitivity and specificity ranging between 50-62% and 17-50% [21, 65]. The number of electrodes should not be considered the sole criterion. Adequate coverage of the head, especially inferior temporal regions, is of importance. Although this was not an inclusion criteria, all of our included HR-ESI studies used whole-head electrode coverages, including subtemporal regions therefore improving localization accuracy [75, 78, 7].

We further excluded studies that did not dichotomize their results into concordance categories. A large difference between dichotomization thresholds was observed in the studies by Wilenius et al. and Kim et al. (25% versus 70%) [41, 93]. This proves that a manual dichotomization would have probably disregarded methodological considerations that are often familiar only to those involved in the source localization procedure, and on which threshold selection generally is based [37, 34, 29, 28]. In many of the studies there was underreporting on technical specifics such as artefact handling, spike criteria and selection of the spike interval. Such specifics are important to allow adequate interpretation of study results. Liberal spike criteria and inappropriate artefact handling may be responsible for less accurate localizations. Source localization based on spike peaks have higher SNR compared to spike onset but could possibly be contaminated by propagation effects resulting in different localizations on sublobar level [54]. If such information was available, subgroup analysis could have aided recommendations on these technical aspects.
Fourth, surgical outcome as reference standard – though considered to be the ultimate standard for localization [44] – is not free from uncontrollable variables.

The Engel classification does not allow straightforward comparison between epilepsy surgery centers due to its considerable subjective judgement using terms as "some disabling seizures" and "worthwhile improvement" in seizure frequency [92]. Neither outcome classification (Engel or ILAE classification) includes post-surgical use of anti-epileptic drugs. Absolute proof of removal of the epileptogenic zone might ideally be established by complete seizure freedom off all anti-epileptic drugs following epilepsy surgery. Further, resection is often limited by eloquent cortex, and seizure recurrence after initial postsurgical seizure freedom can occur due to newly evolved epileptogenic tissue [44, 86]. The definition of a true positive is based on the unambiguous proof that resection of a source estimate

results in good surgical outcome. However, an important issue emerges when we attempt to compare studies with different definitions of 'concordant localization'. A first concern is that sublobar regions are defined according to anatomical landmarks and may differ widely in size and shape depending on their location and among patients [55]. In the case of resection volume concordance, the resected area can still be sometimes too large to discriminate between different localizations, especially for multilobar resections [89]. A second aspect is that the size of the source estimate partly depends on the quality of the source solution. A liberal acceptance of weak dipoles with low SNR might result in widespread dipole solutions or excessive large clusters. Yet, they also might just reflect large epileptogenic areas. Also, clusters can be defined as a number of dipoles localized within the same sublobar region [76, 44, 61] or within a region of fixed dimensions [93]. Specificity might be even more unreliable: surgical failure does not necessarily rule out epileptogenicity of resected tissue, as a more widespread epileptogenic network can be present [32]. Also the consideration of a source estimate beyond the resection volume in cases of surgical failure as a 'true' localization is debatable as non-resected areas encompass both epileptogenic and non-epileptogenic regions, [74]. Although surgical outcome as reference standard might not be ideal, different reference standards, such as ICEEG [40, 63, 64, 67, 85, 43, 1, 66], MRI lesion [58, 83, 57] or presumed EZ [45] suffer from limitations as well. It is known that peri-lesional areas are often marked as epileptogenic and good surgical outcome might not always necessitate complete removal of the SOZ [60, 71]. The presumed EZ remains a theoretical construct up to the point of resection [60, 71].

Considering all issues discussed above, future diagnostic accuracy studies require improvements on bias and transparency. Investigators may use the Standards for Reporting of Diagnostic Accuracy checklist for their study [79, 20, 12]**.** Emphasis should be given on prospective study designs, and cohorts should include all patients in whom presurgical source localization procedures are performed and apply alternative reference standards (e.g. seizure onset zone based on ICEEG, lesion location) on those not eventually submitted for surgery. Ideally, decisions to proceed to surgery and the area of resection should be independent from source localization results, but is probably unethical. As an alternative, the presurgical team may be exposed to the source imaging results after they have made the initial decision on surgery so that the influence on clinical decision making can be accounted for to a maximum degree. A normalization algorithm, which accounts for resection size, source estimation size, and differences in procedural approaches should be developed to allow fair comparison between patients and studies. EEG and MEG contain complementary information due to their distinct technical properties. Its combined use is demonstrated to have superior diagnostic accuracy over use of ESI or MSI alone [19]. More studies are needed to further explore the accuracy and feasibility of EEG-MEG fusion source localization[6].

Once an appropriate level of diagnostic accuracy is established, the integrative approach of HR-ESI and MSI within the presurgical workup should be studied by evaluating various combinations with other tests (e.g. MRI with HR-ESI, MSI with PET)[49] and which patient groups (e.g. non-lesional, multilesional) benefit most.

In this systematic review and meta-analysis, diagnostic accuracy of MSI and HR-ESI to localize the epileptogenic regions of interest is strongly affected by poor study quality and likely biased towards an overestimation of diagnostic accuracy. Results from HR-ESI and MSI

should therefore be interpreted with caution and independent support from other diagnostic tools is required to proceed to resective surgery. High quality studies, that allow unbiased MSI and ESI evaluation and judge results in light of source estimate size and resection size are needed to obtain high quality evidence.

5 Conflict of Interest Statement

Paul Boon and Evelien Carrette are members of the European MEG Society (EMEGS) Executive Committee and have received reimbursement for travel expenses in the past. Paul Boon, Margitta Seeck and Serge Vulliémoz have shares in Epilog. All other authors have no conflict of interest to declare.

6 Acknowledgement

This work was supported by the E-PILEPSY network, which has received funding from the European Union, in the framework of the Health Program (2008-2013).

7 References

[1] Agirre-Arrizubieta Z, Huiskamp GJ, Ferrier CH, van Huffelen AC, Leijten FS. Interictal magnetoencephalography and the irritative zone in the electrocorticogram. Brain 2009;132(Pt 11):3060-71.

[2] Alarcon G, Guy CN, Binnie CD, Walker SR, Elwes RD, Polkey CE. Intracerebral propagation of interictal activity in partial epilepsy: implications for source localisation. J Neurol Neurosurg Psychiatry 1994;57(4):435-49.

[3] Almubarak S, Alexopoulos A, Von-Podewils F, Wang ZI, Kakisaka Y, Mosher JC, et al. The correlation of magnetoencephalography to intracranial EEG in localizing the epileptogenic zone: a study of the surgical resection outcome. Epilepsy Res 2014;108(9):1581-90.

[4] Andre-Obadia N, Lamblin MD, Sauleau P. French recommendations on electroencephalography. Neurophysiol Clin 2015;45(1):1-17.

[5] Assaf BA, Karkar KM, Laxer KD, Garcia PA, Austin EJ, Barbaro NM, et al. Magnetoencephalography source localization and surgical outcome in temporal lobe epilepsy. Clin Neurophysiol 2004;115(9):2066-76.

[6] Aydin U, Vorwerk J, Dumpelmann M, Kupper P, Kugel H, Heers M, et al. Combined EEG/MEG can outperform single modality EEG or MEG source reconstruction in presurgical epilepsy diagnosis. PLoS One 2015;10(3):e0118753.

[7] Bach Justesen A, Eskelund Johansen AB, Martinussen NI, Wasserman D, Terney D, Meritam P, et al. Added clinical value of the inferior temporal EEG electrode chain. Clin Neurophysiol 2018;129(1):291-5.
[8] Bagic A, Funke ME, Ebersole J, Committee APS. American Clinical MEG Society (ACMEGS) position statement: the value of magnetoencephalography (MEG)/magnetic source imaging (MSI) in noninvasive presurgical evaluation of patients with medically intractable localization-related epilepsy. J Clin Neurophysiol 2009;26(4):290-3.
[9] Bagic AI. Disparities in clinical magnetoencephalography practice in the United States: a survey-based appraisal. J Clin Neurophysiol 2011;28(4):341-7.
[10] Bagic AI, Knowlton RC, Rose DF, Ebersole JS, Committee ACPG. American Clinical Magnetoencephalography Society Clinical Practice Guideline 1: recording and analysis of spontaneous cerebral activity. J Clin Neurophysiol 2011;28(4):348-54.
[11] Boon P, D'Have M, Vanrumste B, Van Hoey G, Vonck K, Van Walleghem P, et al. Ictal source localization in presurgical patients with refractory epilepsy. J Clin Neurophysiol 2002;19(5):461-8.
[12] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
[13] Bouet R, Jung J, Delpuech C, Ryvlin P, Isnard J, Guenot M, et al. Towards source volume estimation of interictal spikes in focal epilepsy using magnetoencephalography. Neuroimage 2012;59(4):3955-66.
[14] Brodbeck V, Spinelli L, Lascano AM, Pollo C, Schaller K, Vargas MI, et al. Electrical source imaging for presurgical focus localization in epilepsy patients with normal MRI. Epilepsia 2010;51(4):583-91.
[15] Brodbeck V, Spinelli L, Lascano AM, Wissmeier M, Vargas MI, Vulliemoz S, et al. Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. Brain 2011;134(Pt 10):2887-97.
[16] Burgess RC. Evaluation of brain connectivity: the role of magnetoencephalography. Epilepsia 2011;52 Suppl 4:28-31.
[17] Carrette E, De Tiege X, Op De Beeck M, De Herdt V, Meurs A, Legros B, et al. Magnetoencephalography in epilepsy patients carrying a vagus nerve stimulator. Epilepsy Res 2011;93(1):44-52.
[18] Carrette E, Op de Beeck M, Bourguignon M, Boon P, Vonck K, Legros B, et al. Recording temporal lobe epileptic activity with MEG in a light-weight magnetic shield. Seizure 2011;20(5):414-8.
[19] Chowdhury RA, Pellegrino G, Aydin U, Lina JM, Dubeau F, Kobayashi E, et al. Reproducibility of EEG-MEG fusion source analysis of interictal spikes: Relevance in presurgical evaluation of epilepsy. Hum Brain Mapp 2018;39(2):880-901.
[20] Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016;6(11):e012799.
[21] Coutin-Churchman PE, Wu JY, Chen LL, Shattuck K, Dewar S, Nuwer MR. Quantification and localization of EEG interictal spike activity in patients with surgically removed epileptogenic foci. Clin Neurophysiol 2012;123(3):471-85.
[22] De Tiege X, Carrette E, Legros B, Vonck K, Op de Beeck M, Bourguignon M, et al. Clinical added value of magnetic source imaging in the presurgical evaluation of refractory focal epilepsy. J Neurol Neurosurg Psychiatry 2012;83(4):417-23.
[23] Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ 2001;323(7305):157-62.
[24] Doebler P. mada: Meta-Analysis of Diagnostic Accuracy. R package version 0.5.7.; 2015.
[25] Eliashiv DS, Elsas SM, Squires K, Fried I, Engel J, Jr. Ictal magnetic source imaging as a localizing tool in partial epilepsy. Neurology 2002;59(10):1600-10.

[26] Engel J, Jr., Van Ness PC, Rasmussen TB, Ojemann LM. Outcome with respect to epileptic seizures. In: Engel J, Jr., editor Surgical treatment of epilepsies.: Raven Press; 1993, p. 609-21.
[27] Feng R, Hu J, Pan L, Wu J, Lang L, Jiang S, et al. Application of 256-channel dense array electroencephalographic source imaging in presurgical workup of temporal lobe epilepsy. Clin Neurophysiol 2016;127(1):108-16.
[28] Fischer MJ, Rampp S, Scheler G, Stefan H. Comment: A systematic review on MEG and its use in the presurgical evaluation of localization-related epilepsy. Epilepsy Res 2008;82(2-3):238-9; author reply 40-1.
[29] Fischer MJ, Scheler G, Stefan H. Utilization of magnetoencephalography results to obtain favourable outcomes in epilepsy surgery. Brain 2005;128(Pt 1):153-7.
[30] Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. AJR Am J Roentgenol 2006;187(2):271-81.
[31] Gavaret M, Badier JM, Marquis P, Bartolomei F, Chauvel P. Electric source imaging in temporal lobe epilepsy. J Clin Neurophysiol 2004;21(4):267-82.
[32] Hauptman JS, Mathern GW. Surgical treatment of epilepsy associated with cortical dysplasia: 2012 update. Epilepsia 2012;53 Suppl 4:98-104.
[33] Heers M, Hirschmann J, Jacobs J, Dumpelmann M, Butz M, von Lehe M, et al. Frequency domain beamforming of magnetoencephalographic beta band activity in epilepsy patients with focal cortical dysplasia. Epilepsy Res 2014;108(7):1195-203.
[34] Heers M, Rampp S, Kaltenhauser M, Kasper BS, Doelken MT, Stefan H. Monofocal MEG in lesional TLE: does video EEG monitoring add crucial information? Epilepsy Res 2010;92(1):54-62.
[35] Hirsch LJ, Spencer SS, Williamson PD, Spencer DD, Mattson RH. Comparison of bitemporal and unitemporal epilepsy defined by depth electroencephalography. Ann Neurol 1991;30(3):340-6.
[36] Ishibashi H, Simos PG, Castillo EM, Maggio WW, Wheless JW, Kim HL, et al. Detection and significance of focal, interictal, slow-wave activity visualized by magnetoencephalography for localization of a primary epileptogenic region. J Neurosurg 2002;96(4):724-30.
[37] Iwasaki M, Nakasato N, Shamoto H, Nagamatsu K, Kanno A, Hatanaka K, et al. Surgical implications of neuromagnetic spike localization in temporal lobe epilepsy. Epilepsia 2002;43(4):415-24.
[38] Jayakar P, Gaillard WD, Tripathi M, Libenson MH, Mathern GW, Cross JH, et al. Diagnostic test utilization in evaluation for resective epilepsy surgery in children. Epilepsia 2014;55(4):507-18.
[39] Jeong W, Chung CK, Kim JS. Magnetoencephalography interictal spike clustering in relation with surgical outcome of cortical dysplasia. J Korean Neurosurg Soc 2012;52(5):466-71.
[40] Jung J, Bouet R, Delpuech C, Ryvlin P, Isnard J, Guenot M, et al. The value of magnetoencephalography for seizure-onset zone localization in magnetic resonance imaging-negative partial epilepsy. Brain 2013;136(Pt 10):3176-86.
[41] Kim H, Kankirawatana P, Killen J, Harrison A, Oh A, Rozzelle C, et al. Magnetic source imaging (MSI) in children with neocortical epilepsy: surgical outcome association with 3D post-resection analysis. Epilepsy Res 2013;106(1-2):164-72.
[42] Knowlton RC. The role of FDG-PET, ictal SPECT, and MEG in the epilepsy surgery evaluation. Epilepsy Behav 2006;8(1):91-101.
[43] Knowlton RC, Elgavish R, Howell J, Blount J, Burneo JG, Faught E, et al. Magnetic source imaging versus intracranial electroencephalogram in epilepsy surgery: a prospective study. Ann Neurol 2006;59(5):835-42.
[44] Knowlton RC, Elgavish RA, Bartolucci A, Ojha B, Limdi N, Blount J, et al. Functional imaging: II. Prediction of epilepsy surgery outcome. Ann Neurol 2008;64(1):35-41.
[45] Knowlton RC, Laxer KD, Aminoff MJ, Roberts TP, Wong ST, Rowley HA. Magnetoencephalography in partial epilepsy: clinical yield and localization accuracy. Ann Neurol 1997;42(4):622-31.

[46] Knowlton RC, Razdan SN, Limdi N, Elgavish RA, Killen J, Blount J, et al. Effect of epilepsy magnetic source imaging on intracranial electrode placement. Ann Neurol 2009;65(6):716-23.
[47] Koessler L, Cecchin T, Colnat-Coulbois S, Vignal JP, Jonas J, Vespignani H, et al. Catching the invisible: mesial temporal source contribution to simultaneous EEG and SEEG recordings. Brain Topogr 2015;28(1):5-20.
[48] Lantz G, Grave de Peralta R, Spinelli L, Seeck M, Michel CM. Epileptic source localization with high density EEG: how many electrodes are needed? Clin Neurophysiol 2003;114(1):63-9.
[49] Lascano AM, Perneger T, Vulliemoz S, Spinelli L, Garibotto V, Korff CM, et al. Yield of MRI, high-density electric source imaging (HD-ESI), SPECT and PET in epilepsy surgery candidates. Clin Neurophysiol 2016;127(1):150-5.
[50] Lau M, Yam D, Burneo JG. A systematic review on MEG and its use in the presurgical evaluation of localization-related epilepsy. Epilepsy Res 2008;79(2-3):97-104.
[51] Leijten FS, Huiskamp G. Interictal electromagnetic source imaging in focal epilepsy: practices, results and recommendations. Curr Opin Neurol 2008;21(4):437-45.
[52] Luders HO, Najm I, Nair D, Widdess-Walsh P, Bingman W. The epileptogenic zone: general principles. Epileptic Disord 2006;8 Suppl 2:S1-9.
[53] Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and Presenting Results. . In: Deeks JJ, P.M. B, C. G, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.* 1.0 ed. http://srdta.cochrane.org/: The Cochrane Collaboration; 2010.
[54] Maliia MD, Meritam P, Scherg M, Fabricius M, Rubboli G, Mindruta I, et al. Epileptiform discharge propagation: Analyzing spikes from the onset to the peak. Clin Neurophysiol 2016;127(4):2127-33.
[55] Megevand P, Spinelli L, Genetti M, Brodbeck V, Momjian S, Schaller K, et al. Electric source imaging of interictal activity accurately localises the seizure onset zone. J Neurol Neurosurg Psychiatry 2014;85(1):38-43.
[56] Megevand P, Vulliemoz S. Electric and Magnetic Source Imaging of Epileptic Activity. Epileptologie 2013(30):122-31.
[57] Michel CM, Lantz G, Spinelli L, De Peralta RG, Landis T, Seeck M. 128-channel EEG source imaging in epilepsy: clinical yield and localization precision. J Clin Neurophysiol 2004;21(2):71-83.
[58] Morioka T, Nishio S, Ishibashi H, Muraishi M, Hisada K, Shigeto H, et al. Intrinsic epileptogenicity of focal cortical dysplasia as revealed by magnetoencephalography and electrocorticography. Epilepsy Res 1999;33(2-3):177-87.
[59] Mouthaan BE, Rados M, Barsi P, Boon P, Carmichael DW, Carrette E, et al. Current use of imaging and electromagnetic source localization procedures in epilepsy surgery centers across Europe. Epilepsia 2016;57(5):770-6.
[60] Mu J, Rampp S, Carrette E, Roessler K, Sommer B, Schmitt FC, et al. Clinical relevance of source location in frontal lobe epilepsy and prediction of postoperative long-term outcome. Seizure 2014;23(7):553-9.
[61] Murakami H, Wang ZI, Marashly A, Krishnan B, Prayson RA, Kakisaka Y, et al. Correlating magnetoencephalography to stereo-electroencephalography in patients undergoing epilepsy surgery. Brain 2016;139(11):2935–47.
[62] Nemtsas P, Birot G, Pittau F, Michel CM, Schaller K, Vulliemoz S, et al. Source localization of ictal epileptic activity based on high-density scalp EEG data. Epilepsia 2017;58(6):1027-36.
[63] Oishi M, Kameyama S, Masuda H, Tohyama J, Kanazawa O, Sasagawa M, et al. Single and multiple clusters of magnetoencephalographic dipoles in neocortical epilepsy: significance in characterizing the epileptogenic zone. Epilepsia 2006;47(2):355-64.
[64] Oishi M, Otsubo H, Iida K, Suyama Y, Ochi A, Weiss SK, et al. Preoperative simulation of intracerebral epileptiform discharges: synthetic aperture magnetometry virtual sensor analysis of interictal magnetoencephalography data. J Neurosurg 2006;105(1 Suppl):41-9.

[65] Oliva M, Meckes-Ferber S, Roten A, Desmond P, Hicks RJ, O'Brien TJ. EEG dipole source localization of interictal spikes in non-lesional TLE with and without hippocampal sclerosis. Epilepsy Res 2010;92(2-3):183-90.
[66] Ossenblok P, de Munck JC, Colon A, Drolsbach W, Boon P. Magnetoencephalography is more successful for screening and localizing frontal lobe epilepsy than electroencephalography. Epilepsia 2007;48(11):2139-49.
[67] Otsubo H, Ochi A, Elliott I, Chuang SH, Rutka JT, Jay V, et al. MEG predicts epileptic zone in lesional extrahippocampal epilepsy: 12 pediatric surgery cases. Epilepsia 2001;42(12):1523-30.
[68] Papanicolaou AC, Castillo E, Pataraia E, Wheless J, Breier J, Simos P. A response to the article of Lau et al. "A systematic review on MEG and its use in the presurgical evaluation of localization related epilepsy", Epilepsy Research 2008;79(May (2-3)):97-104. Epilepsy Res 2008;82(2-3):237; author reply 40-1.
[69] Papanicolaou AC, Pataraia E, Billingsley-Marshall R, Castillo EM, Wheless JW, Swank P, et al. Toward the substitution of invasive electroencephalography in epilepsy surgery. J Clin Neurophysiol 2005;22(4):231-7.
[70] Pillai J, Sperling MR. Interictal EEG and the diagnosis of epilepsy. Epilepsia 2006;47 Suppl 1:14-22.
[71] Plummer C, Harvey AS, Cook M. EEG source localization in focal epilepsy: where are we now? Epilepsia 2008;49(2):201-18.
[72] R Development Core Team RFfSC, Vienna, Austria. R: A language and environment for statistical computing. 2008.
[73] Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58(10):982-90.
[74] Rikir E, Koessler L, Ramantani G, Maillard LG. Added value and limitations of electrical source localization. Epilepsia 2017;58(1):174-5.
[75] Rosenzweig I, Fogarasi A, Johnsen B, Alving J, Fabricius ME, Scherg M, et al. Beyond the double banana: improved recognition of temporal lobe seizures in long-term EEG. J Clin Neurophysiol 2014;31(1):1-9.
[76] Schneider F, Alexopoulos AV, Wang Z, Almubarak S, Kakisaka Y, Jin K, et al. Magnetic source imaging in non-lesional neocortical epilepsy: additional value and comparison with ICEEG. Epilepsy Behav 2012;24(2):234-40.
[77] Schneider F, Irene Wang Z, Alexopoulos AV, Almubarak S, Kakisaka Y, Jin K, et al. Magnetic source imaging and ictal SPECT in MRI-negative neocortical epilepsies: additional value and comparison with intracranial EEG. Epilepsia 2013;54(2):359-69.
[78] Seeck M, Koessler L, Bast T, Leijten F, Michel C, Baumgartner C, et al. The standardized EEG electrode array of the IFCN. Clin Neurophysiol 2017;128(10):2070-7.
[79] Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. BMJ 2013;346:f2778.
[80] Sohrabpour A, Lu Y, Kankirawatana P, Blount J, Kim H, He B. Effect of EEG electrode number on epileptic source localization in pediatric patients. Clin Neurophysiol 2015;126(3):472-80.
[81] Stefan H, Hummel C, Scheler G, Genow A, Druschky K, Tilz C, et al. Magnetic brain source imaging of focal epileptic activity: a synopsis of 455 cases. Brain 2003;126(Pt 11):2396-405.
[82] Stefan H, Trinka E. Magnetoencephalography (MEG): Past, current and future perspectives for improved differentiation and treatment of epilepsies. Seizure 2017;44:121-4.
[83] Stefan H, Wu X, Buchfelder M, Rampp S, Kasper B, Hopfengartner R, et al. MEG in frontal lobe epilepsies: localization and postoperative outcome. Epilepsia 2011;52(12):2233-8.
[84] Sutherling WW, Mamelak AN, Thyerlei D, Maleeva T, Minazad Y, Philpott L, et al. Influence of magnetic source imaging for planning intracranial EEG in epilepsy. Neurology 2008;71(13):990-6.

[85] Tenney JR, Fujiwara H, Horn PS, Rose DF. Comparison of magnetic source estimation to intracranial EEG, resection area, and seizure outcome. Epilepsia 2014;55(11):1854-63.
[86] Vadera S, Jehi L, Burgess RC, Shea K, Alexopoulos AV, Mosher J, et al. Correlation between magnetoencephalography-based "clusterectomy" and postoperative seizure freedom. Neurosurg Focus 2013;34(6):E9.
[87] van Dellen E, Douw L, Hillebrand A, de Witt Hamer PC, Baayen JC, Heimans JJ, et al. Epilepsy surgery outcome and functional network alterations in longitudinal MEG: a minimum spanning tree analysis. Neuroimage 2014;86:354-63.
[88] van Klink N, Hillebrand A, Zijlmans M. Identification of epileptic high frequency oscillations in the time domain by using MEG beamformer-based virtual sensors. Clin Neurophysiol 2016;127(1):197-208.
[89] Wheless JW, Willmore LJ, Breier JI, Kataki M, Smith JR, King DW, et al. A comparison of magnetoencephalography, MRI, and V-EEG in patients evaluated for epilepsy surgery. Epilepsia 1999;40(7):931-41.
[90] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529-36.
[91] Widjaja E, Shammas A, Vali R, Otsubo H, Ochi A, Snead OC, et al. FDG-PET and magnetoencephalography in presurgical workup of children with localization-related nonlesional epilepsy. Epilepsia 2013;54(4):691-9.
[92] Wieser HG, Blume WT, Fish G, Goldensohn E, Hufnagel A, King D, et al. Proposal for a new classification of outcome with respect to epileptic seizures following epilepsy surgery. Epilepsia 2001;42(2):282-6.
[93] Wilenius J, Medvedovsky M, Gaily E, Metsahonkala L, Makela JP, Paetau A, et al. Interictal MEG reveals focal cortical dysplasias: special focus on patients with no visible MRI lesions. Epilepsy Res 2013;105(3):337-48.
[94] Wu JY, Salamon N, Kirsch HE, Mantle MM, Nagarajan SS, Kurelowech L, et al. Noninvasive testing, early surgery, and seizure freedom in tuberous sclerosis complex. Neurology 2010;74(5):392-8.

8 Legends

Figure 1. Diagnostic accuracy 2x2 contingency table for source localization. Concordant: MSI or ESI source within resection volume or sublobar concordancy with resection volume. Sensitivity: true positives/ (true positives+ false negatives). Good surgical outcome: Engel 1/ILAE 1-2. Poor surgical outcome: Engel ≥2 / ILAE ≥3. Indeterminate test results: e.g. too low number of IEDs, artefacts. Specificity: true negatives / (true negatives+ false positives). Allocation of partially concordant (i.e. partially resected) to concordant or non-concordant group affects sensitivity and specificity. Allocation of indeterminate test results tot the non-concordant group affects sensitivity and specificity.

Figure 2. Flow of studies through review process.

Figure 3. 2x2 contingency table for patient total in all studies (both MSI and HR-ESI).

Figure 4. ~~Summary ROC curve with summary estimates. Left: for HR-EEG and MSI. Middle: HR-EEG and MSII (indeterminate results included as non-concordant), MSII and HR-EEG show statistical difference (p=0.031). Right: HR-EEG and MSII,P (indeterminate and partially concordant results are included as non-concordant) MSII,P and HR-ESI show statistical difference (p=0.001) Individual studies are shown as small symbols. Summary points shown as large symbols, representing sensitivity and specificity estimates pooled using the bivariate linear mixed model. The 95% CI is represented by dotted line (---). Note that HR-EEG is similar in all three figures as partially concordant results and surgical outcome data for indeterminate test results were not reported.~~

Summary ROC curve with summary estimates for HR-ESI (red) and MSI (blue). Individual studies are shown as small symbols. Summary points shown as large symbols, representing sensitivity and specificity estimates pooled using the bivariate linear mixed model. The 95% CI is represented by dotted line (---).

9 Table 1

Type of test	Type of resective surgery	% multiclusters	Resection volume estimation	Study's concordance definition	Minimum follow up (years)	Good outcome definition
MEG (+sim EEG)	SLR (23), MLR (1)	50	Post-op MRI	Complete or partial cluster resection	1	Engel 1
MEG (+sim EEG)	SLR (13), MLR(1)	79	Post-op MRI	>70% of all dipoles within resection volume	1	Engel 1
MEG (+sim EEG)	SLR (16), MLR (2)	39	Post-op MRI	Complete or partial resection of unifocal cluster[c]	2	Engel 1a[d]
MEG	SLR (14), MLR(2)	6	Post-op MRI	>25% of source clusters within resection volume[f]	0,5	Engel 1[g]
MEG (+sim EEG)	Not specified	unclear	Not specified	Complete source estimate resection[c,h,i]	1	ILAE class 1-2
MEG	Not specified	24	Not specified[j]	Complete cluster resection[c]	0,5	Engel 1
MEG (+sim EEG)	Not specified	Unclear	Not specified	Sublobar co-localization of cluster and resection volume	1	Engel 1
MEG (+sim EEG)	Not specified	Unclear	Post-op MRI	Complete cluster resection	0,5	Engel 1a[d]
HR-EEG	SLR (41), MLR(11)	NA	Not specified	Complete source maximum resection	1	Engel 1
HR-EEG	Not specified	NA	Post-op MRI	Complete source maximum resection[m]	1	Engel 1
HR-EEG	Not specified	NA	Post-op MRI	Sublobar co-localization of source maximum and resection volume	0,58	Engel 1-2[d]

Study	Data collection	Sample size[a]	Age group	Population characteristics
Jeong et al., 2012	Retrospective	24[a]	A	FCD (histologically confirmed)
Kim et al., 2013	Retrospective	14[b]	P	Neocortical epilepsy
Schneider et al., 2012	Retrospective	18	A+P	Neocortical MRI negative epilepsy
Wilenius et al., 2013	Retrospective	16[e]	A+P	FCD (histologically confirmed)
Papanicolaou et al., 2005	Prospective	41	A+P	Mixed group
Mu et al., 2014	Retrospective	38	A+P	FLE
Knowlton et al., 2008	Prospective	49[k]	A+P	Inconclusively localizing MRI and vEEG with seizures recorded on ICEEG
Almubarak et al., 2014	Retrospective	36	A+P	Mixed group with localizing ICEEG
Brodbeck et al., 2011	Prospective	52	A+P	Mixed group
Megevand et al., 2014	Prospective	32[l]	A+P	Mixed group, all underwent ICEEG
Feng et al., 2016	Prospective	43	A+P	TLE which did not require ICEEG

Table 1. A: adult; EZ: epileptogenic zone; (F)CD: (focal) cortical dysplasia; MLR: multi-lobar resection; P: paediatric; SLR: single lobe resection; SR: surgical resection sim EEG: simultaneously recorded EEG; a: one patient with scattered source result excluded; b: eight second stage surgery patients excluded; c: surgical outcome for partial concordance category provided; d: insufficient details on Engel classification to reclassify; e: nine patients with scattered source results excluded, 9 patients with repeated surgery excluded f: two thresholds reported by study (10% and 25%), highest threshold selected by reviewers; g: patients with less than 6 months follow-up excluded h: type of source estimate (e.g. single dipoles, clusters) not specified. i: study classifies partially/non-concordant results with poor outcome as indeterminate test result. For data analysis purposes these were considered true negatives by reviewers; j:post-op MRI available, but study did not report its role in resection volume estimation; k: thirteen indeterminate test results excluded; l: 6/38 did not undergo surgery and were excluded; m: When one of the solution points directly neighbouring the source maximum was inside the resection volume, this was considered concordant.

10 Table 2

Type of test	Study	Risk of bias				Applicability concerns		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
MSI	Jeong et al., 2012	-	?	+	-	+	?	+
	Kim et al., 2013	+	?	+	-	+	?	+
	Schneider et al., 2012	-	+	-	-	+	+	+
	Wilenius et al., 2013	-	-	+	-	+	+	+
	Papanicolaou et al., 2005	+	+	+	-	+	+	+
	Mu et al., 2014	-	+	-	-	+	+	+
	Knowlton et al., 2008	-	+	-	-	+	+	+
	Almubarak et al., 2014	-	+	-	-	+	+	+
HR-ESI	Brodbeck et al., 2011	+	+	-	?	+	+	+
	Megevand et al., 2014	-	+	+	-	+	+	+
	Feng et al., 2016	-	+	+	-	+	+	+

Table 2. Quality appraisal of individual studies. + low risk, - high risk, **?** *unclear risk.*

11 Table 3

Concordance category per type of test	Good surgical outcome probability in patients (mean, 95% CI)	Odds ratio	95% Confidence interval	P-value
HR-ESI				
Concordant	75/91 (82%, 67-98%)	9.6	3.8 – 24.0	<0.001
Partially concordant	Not applicable	-	-	-
Non-concordant	11/36 (30%, 15-45%)	Ref	-	-
MSI				
Concordant	97/135 patients (74%, 63-85%)	4.7	1.7- 12.9	0.002
Partially concordant[a]	8/26 patients (30%, 12-47%)	1.7	0.35-8.4	0.512
Non-concordant	25/74 patients (25%, 13-37%)	Ref	-	-

Table 3. Odds ratio of level of concordance for surgical outcome. HR-ESI: High resolution electric source imaging. MSI: magnetic source imaging. a: based on studies reporting surgical outcome for partially concordant cases, three MSI studies in total.