# A Critical Analysis of Computational and Imaging Methods for Epileptogenic Lesion Detection

Yusuf Qwareeq*

June 20, 2025

**Abstract**

Detecting the subtle brain lesions responsible for seizures in patients with so-called "MRI negative" epilepsy is a major clinical challenge, leading to poorer surgical outcomes and necessitating complex, often invasive, evaluations. This report presents a critical analysis of three pivotal papers to trace the evolution of computational and imaging techniques developed to address this problem. The analysis follows a three-act narrative, beginning with the definition of a functional biomarker and an early computational solution by Hong and colleagues. It then examines the development of MRF by Ma and colleagues, a revolutionary imaging method for quantitative tissue mapping. The narrative culminates with the work of Ding and colleagues, which synthesizes these threads by applying AI to quantitative imaging data, creating a powerful, non-invasive tool to visualize these elusive lesions. This technological progression illustrates the powerful synergy between targeted clinical need, foundational data acquisition technology, and advanced computational analysis in solving intractable medical challenges.

## Contents

**1 The Clinical Imperative: Localizing Epileptogenic Foci** **4**

**2 The Role of a Narrative: Charting a Path of Scientific Progress** **4**

**3 Report Outline** **5**

**4 Act I: The Foundational Problem: A Critical Analysis of Hong et al. (2014)** **5**
4.1 Research Objectives and Context . . . . . 5
4.2 Methodological Approach . . . . . 6
4.3 Key Findings and Conclusions . . . . . 8
4.4 Critical Appraisal . . . . . 11
4.4.1 Strengths and Contributions . . . . . 11
4.4.2 Statistical Significance and Rigor . . . . . 11
4.4.3 Limitations and Future Directions . . . . . 12

**5 Act II: The Paradigm-Shifting Tool: A Critical Analysis of Ma et al. (2013)** **12**
5.1 Research Objectives and Context . . . . . 12
5.2 Methodological Approach . . . . . 13
5.2.1 The Acquisition: Creating Unique Fingerprints . . . . . 13

*Department of ECE, Temple University, qwareeq@temple.edu

5.2.2 The Dictionary Generation: The Pre-calculated Ground Truth . . . . . 14
5.2.3 The Pattern Matching: From Signal to Quantitation . . . . . 15
5.3 Key Findings and Conclusions . . . . . 15
5.4 Critical Appraisal . . . . . 15
5.4.1 Strengths and Contributions . . . . . 15
5.4.2 Statistical Significance and Rigor . . . . . 15
5.4.3 Limitations and Future Directions . . . . . 19

**6 Act III: The Modern Synthesis: A Critical Analysis of Ding et al. (2025) 19**
6.1 Research Objectives and Context . . . . . 20
6.2 Methodological Approach . . . . . 20
6.2.1 Patient Cohort & Imaging Protocol . . . . . 20
6.2.2 Deep Learning Architecture . . . . . 20
6.2.3 Model Training and Validation . . . . . 23
6.3 Key Findings and Conclusions . . . . . 23
6.4 Critical Appraisal . . . . . 23
6.4.1 Strengths and Contributions . . . . . 23
6.4.2 Statistical Significance and Rigor . . . . . 23
6.4.3 Limitations and Future Directions . . . . . 25

**7 Synthesis, Discussion, and Conclusion 25**
7.1 The Evolutionary Trajectory: From Problem to Tool to Solution . . . . . 25
7.2 Broader Implications and Future Directions . . . . . 26
7.3 Conclusion . . . . . 26

**List of Acronyms 27**

## List of Figures

1 Study flow diagram from Hong et al. (2014), illustrating the patient cohorts for the 3 T and 1.5 T datasets. The diagram shows the final performance metrics, including a sensitivity of 74% and specificity of 100% on the primary 3.0 T dataset. . . . . 8
2 Examples of true positive detections from Hong et al. (2014). For each case, the manual segmentation (green) is shown alongside the initial vertex-wise classifier output (red) and the final, more specific cluster-wise output (blue). . . . . 9
3 Comparison of quantitative features for a true positive detection versus an extralesional finding from Hong et al. (2014). The true lesion (red cluster) is distinguished by its abnormal sulcal depth, while the extralesional cluster (blue cluster) is primarily characterized by abnormal cortical thickness. This example illustrates how the multivariate classifier uses different feature profiles to differentiate true Focal Cortical Dysplasia (FCD) from other morphological anomalies. . . . . 10
4 The Magnetic Resonance Fingerprinting (MRF) acquisition scheme from Ma et al. (2013). (a) A diagram of the Inversion-Recovery Balanced Steady-State Free-Precession (IR-bSSFP) sequence with varying flip angle (FA) and repetition time (TR). (b) The highly undersampled variable-density spiral k-space trajectory. (c, d) Two different pseudorandom patterns of FA and TR used in the experiments. . . . . 13

5 Simulated and measured MRF signals from Ma et al. (2013). (a, b) Simulated signal evolutions (fingerprints) for different tissue types, demonstrating their unique character. (c, d) Validation in phantoms, showing the excellent match between the measured signal (blue) and the corresponding dictionary entry (red). . . . . 14
6 In vivo MRF results from Ma et al. (2013). (a) A single raw image frame, heavily corrupted by aliasing. (b) The measured signal fingerprint from one voxel (blue) and the matched dictionary entry (red). (c-f) The final, co-registered quantitative maps for $T_1$, $\Delta f$, $T_2$, and proton density ($M_0$) respectively, generated from a single rapid scan. . . . . 16
7 Demonstration of MRF's robustness to motion from Ma et al. (2013). (a) Raw image without motion. (b) Raw image with subject motion during the last few seconds. (c, e) $T_1$ and $T_2$ maps from the motion-free scan. (d, f) Nearly identical $T_1$ and $T_2$ maps from the motion-corrupted scan, demonstrating that incoherent motion signals are rejected by the pattern-matching algorithm. . . . . 17
8 Validation of MRF accuracy, efficiency, and precision from Ma et al. (2013). (a, b) Accuracy plots showing high correlation between MRF and standard measurements for $T_1$ and $T_2$. (c, d) Efficiency plots showing MRF outperforms the DESPOT method. (e, f) Precision plots showing high repeatability of MRF measurements over time. . . 18
9 The integrated workflow from Ding et al. (2025). (A) Patient selection flowchart. (B) Data processing pipeline showing the generation of multiple quantitative and derived feature maps from a single MRF acquisition. (C) Deep Learning (DL) model inputs and training scheme. . . . . 21
10 Multiparametric MRF maps from Ding et al. (2025) for four different FCD subtypes (FCD IIa, FCD IIb, Mild Malformations of Cortical Development (mMCD), Mild Malformation of Cortical Development with Oligodendroglial Hyperplasia (MOGHE)). The images demonstrate the rich, quantitative information provided by a single MRF scan, highlighting different lesion characteristics across T1, T2, and tissue fraction maps. . . . . 22
11 The DL framework from Ding et al. (2025). (A) A suite of multiparametric maps derived from a single MRF scan serves as the input to the model. (B) A Convolutional Neural Network (CNN) with a U-Net architecture processes the image data. (C) The model outputs a probabilistic map indicating the location of the predicted lesion. . . 22
12 Examples of successful automated FCD detections from Ding et al. (2025). The ground-truth lesion label is outlined in green, and the model's probabilistic prediction is shown as a yellow heat map. The figure demonstrates the model's ability to accurately localize lesions of varying types and locations. . . . . 24

## List of Tables

1 Comparison of Magnetic Resonance Imaging (MRI) Paradigms. . . . . 19
2 A Comparative Analysis of FCD Detection Methodologies. . . . . 24

## 1 The Clinical Imperative: Localizing Epileptogenic Foci

Epilepsy is a prevalent and burdensome neurological disorder affecting millions of people worldwide. A substantial portion of this population, estimated to be as high as one-third, develops Drug-Resistant Epilepsy (DRE), a condition formally defined by the International League Against Epilepsy (ILAE) as the failure of adequate trials of two tolerated, appropriately chosen, and properly used Anti-Seizure Medication (ASM) schedules to achieve sustained seizure freedom [1]. A systematic review and meta-analysis reported a pooled prevalence of DRE in clinic-based cohorts of 36.3% [2]. For individuals with focal DRE, where seizures originate from a specific brain region, resective neurosurgery offers the most effective path toward seizure control. Multiple meta-analyses have demonstrated that surgery provides a significantly higher chance of seizure freedom compared to continued medical management, with one review reporting seizure-free rates of 57.0% in the surgical group versus 15.3% in the medical treatment group [3].

The success of resective epilepsy surgery is, however, critically predicated on a single, indispensable condition: the accurate presurgical localization and delineation of the Epileptogenic Zone (EZ), the area of cortex indispensable for generating clinical seizures [3]. This requirement introduces a formidable clinical challenge for a significant cohort of patients. An estimated 15% to 40% of patients with focal DRE are classified as "MRI negative" or "non-lesional" [4]. This designation signifies that a conventional radiological review of their structural MRI scans fails to identify a clear epileptogenic lesion. The absence of a visible lesion is a primary negative prognostic indicator, associated with less favorable surgical outcomes compared to MRI-positive cases [5]. This diagnostic impasse is particularly vexing because the most common pathology underlying these cases is often a subtle malformation of cortical development, most frequently FCD Type II, which is highly epileptogenic but can be morphologically inconspicuous on standard imaging [6].

This situation creates a critical and urgent need for more advanced imaging and computational analysis techniques capable of unmasking these occult lesions [7]. Furthermore, it is crucial to recognize that the "MRI negative" label is not a static biological diagnosis but rather a technologically relative one. It represents a moving target, defined by the limitations of a specific detection modality at a given point in time. For instance, studies have repeatedly shown that patients initially diagnosed as MRI negative can have lesions revealed through computational postprocessing of the original scans [8], acquisition at higher magnetic field strengths such as 7 T [9], or the application of alternative imaging modalities like Positron Emission Tomography (PET) or Magnetoencephalography (MEG) [10]. The central clinical problem, therefore, is not merely the detection of lesions in a fixed patient subpopulation, but a continuous technological and analytical pursuit to redefine the very threshold of detectability. The history of progress in this domain is a story of the co-evolution of clinical questions and the technological capabilities developed to answer them.

## 2 The Role of a Narrative: Charting a Path of Scientific Progress

The purpose of this report is to conduct a critical analysis of the scientific and technological advancements that have addressed the challenge of localizing epileptogenic foci in MRI negative epilepsy. This analysis is structured around a central thesis that progress in this field has followed a distinct and cohesive narrative arc, characterized by the identification of a clinical problem, the independent invention of a paradigm-shifting enabling technology, and the ultimate synthesis of these threads into a modern, state-of-the-art solution. This report will trace this narrative through a detailed examination of three pivotal publications that serve as exemplars for each stage of this journey.

The first stage, representing the foundational problem and an early solution, is exemplified by the landmark study from Hong, S., et al. (2014) in *Neurology*. The authors demonstrated that Machine Learning (ML) algorithms could be applied to conventional, qualitative structural MRI data to automatically detect subtle FCD lesions in patients who had been previously classified as MRI negative. This work provided a critical proof-of-concept that clinically relevant information was indeed present, yet hidden, in standard clinical scans and could be extracted computationally [6]. The second act involves a paradigm-shifting tool, MRF, introduced by Ma, D., et al. (2013) in *Nature*. Published in a separate domain and driven by questions of physics and engineering, MRF is a revolutionary, general-purpose imaging technique that fundamentally altered the nature of MRI data acquisition by enabling rapid, simultaneous, and quantitative measurement of multiple intrinsic tissue properties [11]. The narrative culminates in the third stage, a modern synthesis, with the work of Ding, Y., et al. (2025) in *Neurology*. This line of research synthesizes the two preceding threads by applying sophisticated DL models to the rich, quantitative imaging data made possible by the MRF framework, providing a new state-of-the-art solution that demonstrates markedly improved sensitivity for detecting previously invisible lesions [12].

This report will argue that the interplay between these three works provides an intellectual roadmap for understanding scientific progress in this domain. The journey from Hong (Problem Definition) through Ma (Tool Invention) to Ding (Synthesized Solution) is not merely a chronological sequence but a logical and compelling story of how a clinical need drives the adoption and fusion of disparate technological and computational innovations.

## 3 Report Outline

This report is structured to follow the narrative arc established above, providing a comprehensive and critical examination of each component and its contribution to the overall story of scientific progress. Section 4 will provide a detailed critique of the foundational work by Hong et al. (2014), examining its methodology, impact, and inherent limitations, thereby establishing the baseline against which future progress is measured. Section 5 will shift focus to the paradigm-shifting technology of MRF, as introduced by Ma et al. (2013), and will analyze the fundamental principles of MRF and articulate its potential as an enabling technology. Section 6 will then investigate the state-of-the-art work represented by Ding et al. (2025), which constitutes a powerful synthesis of the two preceding research threads. Finally, Section 7 will offer a global synthesis and discussion, revisiting the narrative and reflecting on its broader implications for medical science, including issues of clinical translation, regulatory approval, and the next wave of innovation in Explainable Artificial Intelligence (XAI) and imaging genetics.

## 4 Act I: The Foundational Problem: A Critical Analysis of Hong et al. (2014)

### 4.1 Research Objectives and Context

The 2014 *Neurology* paper by Hong, S., et al., "Automated detection of cortical dysplasia type II in MRI negative epilepsy," emerged from a pressing clinical need and represented a significant inflection point in the application of computational analysis to challenging neuroimaging problems [6]. The study's primary stated objective was to develop and validate a fully automated method capable of detecting FCD Type II lesions in patients with drug-resistant extratemporal epilepsy who were initially classified as "MRI negative" based on routine visual inspection of both 1.5 T and 3.0 T MRI scans. The authors sought to create a classifier that could identify these histologically confirmed, yet radiologically occult, epileptogenic substrates with high sensitivity and specificity.

To appreciate the significance of this objective, it is essential to understand the clinical and technical landscape prior to 2014. FCD is a leading cause of pharmacoresistant epilepsy, and for these patients, surgical resection of the dysplastic tissue offers the best, and often only, chance for seizure freedom. The success of such surgery is, however, critically dependent on the accurate localization of the lesion. A substantial cohort of patients, however, fell into the diagnostically challenging category of "MRI negative" epilepsy. In these cases, even meticulous review of high-field-strength structural MRI by expert neuroradiologists failed to reveal a definitive lesion. This diagnostic impasse was a formidable barrier to surgical treatment. It often forced clinicians to proceed with high-risk, expensive, and invasive diagnostic procedures, such as Stereoelectroencephalography (SEEG), without a clear anatomical target. As the authors note in their discussion, this "blind" approach could lead to sampling errors in up to 40% of cases, where the electrodes miss the true EZ, resulting in failed localization and poor surgical outcomes.

Prior to this study, quantitative methods did exist to aid in FCD detection. Techniques like Voxel-Based Morphometry (VBM) and texture analysis were developed to create statistical maps highlighting features associated with FCD, such as cortical thickening or altered tissue intensity. However, these early-generation tools had a fundamental limitation: they were primarily aids for computer-assisted visualization, not automated detection. They produced a series of parametric maps that still required an expert to visually interpret them, a process the authors correctly identify as being highly dependent on the "reader's familiarity with the algorithm." Furthermore, these methods were largely validated on patients whose lesions were already visible on conventional MRI, leaving their utility in the truly "MRI negative" population uncertain.

The approach taken by Hong et al. was novel and important for two principal reasons. First, it directly confronted the most difficult clinical problem: the MRI negative cohort. Second, it sought to shift the paradigm from computer-aided visualization to fully automated, computer-aided detection. This represented a crucial evolution from creating maps for human interpretation to building a model that could render an objective, probabilistic judgment about the presence and location of a lesion. The decision to employ a surface-based framework, rather than a purely voxel-based one, was also a key methodological choice, designed to more accurately model the sulco-gyral morphology of the cortex and quantify features like cortical thickness and sulcal depth, which are often distorted in FCD.

## 4.2 Methodological Approach

The experimental design detailed by Hong et al. is characterized by its careful patient selection, innovative feature engineering, and a pragmatic ML pipeline tailored to the specific challenges of detecting subtle neuropathology.

The study's primary analysis was conducted on a cohort scanned at 3.0 T. This included 19 patients with drug-resistant extratemporal epilepsy who were initially diagnosed as MRI negative. The "ground truth" location of the FCD lesion for these patients was not established from routine MRI, but rather through a subsequent expert review of advanced quantitative texture maps, a specialized technique in itself. This detail is crucial, as it defines the study's target: lesions that are discoverable only through postprocessing. Of these 19 patients, 15 ultimately underwent surgery, providing definitive histopathological confirmation of FCD Type II (seven with type IIa, eight with type IIb). To robustly assess the classifier's specificity, the authors included two distinct control groups: a cohort of 24 healthy individuals and, critically, a disease control group of 11 patients with an unrelated form of epilepsy (Temporal Lobe Epilepsy (TLE) with hippocampal sclerosis)

who were seizure-free after surgery. This second group ensured the classifier was learning features specific to FCD and not simply detecting any form of brain abnormality. To evaluate the method's generalizability across different scanners and field strengths, an independent validation cohort was assembled. This group, scanned at 1.5 T, consisted of 14 patients with histologically confirmed, MRI negative FCD and 20 matched healthy controls.

The core input for the analysis was high-resolution (1-mm isotropic voxels) T1-weighted 3D structural MRI scans. These images underwent a standard preprocessing pipeline, including correction for intensity non-uniformity, linear registration to a common stereotaxic space, and automated segmentation into Gray Matter (GM), White Matter (WM), and Cerebrospinal Fluid (CSF).

The methodological heart of the paper lies in its feature engineering. Instead of feeding raw voxel intensities to a classifier, the authors extracted a set of five biologically-informed, handcrafted features calculated on a cortical surface model, which preserves the topological structure of the cortex. The five features were cortical thickness, measured as the distance between the GM-WM and pial surfaces to model the cortical thickening commonly observed in FCD; sulcal depth, calculated as the geodesic distance from a vertex to the nearest gyral crown to capture the tendency for small FCDs to reside at the bottom of an abnormally deep sulcus; the absolute mean curvature of the cortical surface, computed to model potential local changes in tissue stiffness and folding patterns; a normalized measure of relative T1-weighted signal intensity within the cortical ribbon to quantify the characteristic hyperintensity of FCD; and the intensity gradient across the GM-WM boundary, which provides a quantitative proxy for the "blurring" of this junction, a pathognomonic sign of FCD Type II.

For classification, the authors implemented a two-stage process using Fisher's Linear Discriminant Analysis (LDA). The first stage operated at the vertex level, training an LDA to distinguish individual "lesional" vertices from "non-lesional" vertices using the five-feature vector. As the authors report, this initial step successfully identified most lesions but at the cost of an extremely high false-positive rate, generating dozens of spurious clusters in both patients and controls. The second stage was the crucial innovation for improving specificity. This second LDA was trained to classify entire clusters of abnormal vertices identified by the first stage. The input to this second classifier was not the original five features, but a set of statistical moments describing each cluster (e.g., mean, standard deviation, skewness, kurtosis of the features within the cluster) and its spatial location. This second classifier effectively learned the morphometric and statistical profile of a true lesion cluster, allowing it to distinguish them from noise-driven false-positive clusters.

The entire model was trained and validated using a Leave-One-Out Cross-Validation (LOOCV) strategy. In this scheme, for each of the 19 patients, a unique classifier was trained on the data from the remaining 18 patients and then tested on the single "left-out" individual. This process was repeated 19 times, ensuring that every patient was used as a test case for a model they were not included in training. This approach is a common and justifiable choice for small datasets, as it maximizes the amount of data used for training in each fold and provides a nearly unbiased estimate of model performance [13].

This two-stage classification architecture can be understood as a pragmatic and intelligent engineering solution to the severe multiple comparisons problem inherent in whole-brain neuroimaging analysis. A typical brain surface contains over 40,000 vertices per hemisphere, meaning a vertex-wise analysis involves tens of thousands of simultaneous statistical tests that dramatically inflate the probability of finding false positives purely by chance [14]. Standard statistical correction procedures, such as the Bonferroni correction, are often overly conservative for spatially correlated neuroimaging

data and risk eliminating true, subtle signals along with the noise [15]. Rather than applying a formal statistical correction like False Discovery Rate (FDR) control [16], the authors developed a data-driven heuristic. The two-stage classifier implicitly operates on the assumption that true lesion clusters possess a different multivariate statistical "shape" and spatial distribution than the clusters generated by random noise. The success of the method hinges on the second classifier's ability to learn this distinction empirically, reflecting a practical approach that prioritizes detection sensitivity over formal statistical error control, a trade-off that defines both the paper's innovative power and its statistical limitations.

### 4.3 Key Findings and Conclusions

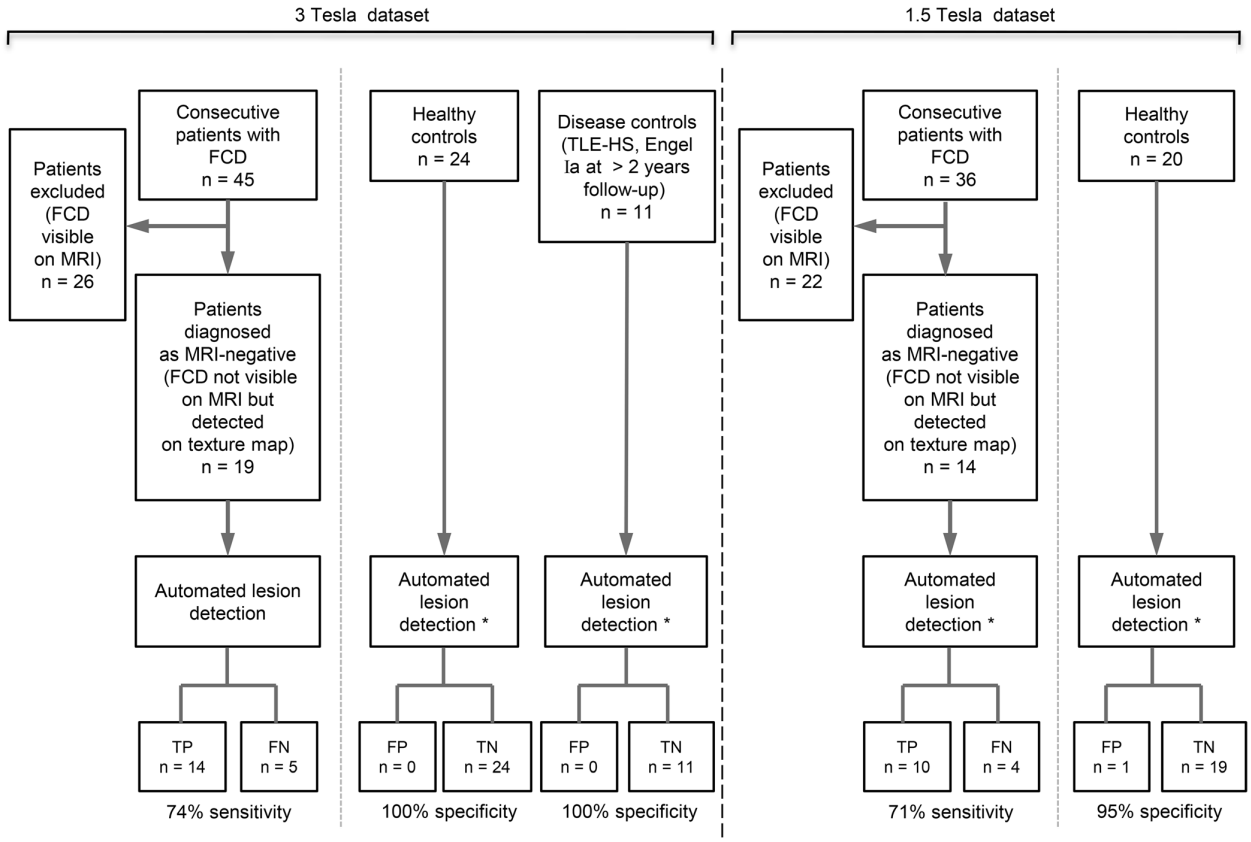


**Figure 1:** Study flow diagram from Hong et al. (2014), illustrating the patient cohorts for the 3 T and 1.5 T datasets. The diagram shows the final performance metrics, including a sensitivity of 74% and specificity of 100% on the primary 3.0 T dataset.

The results presented by Hong et al. provided strong quantitative support for the efficacy of their automated detection pipeline. The primary performance metrics, illustrated in Fig. 1, were compelling. On the primary 3.0 T dataset, the final two-stage classifier achieved a sensitivity of 74%, correctly identifying a lesion cluster that colocalized with the ground-truth manual label in 14 of the 19 MRI negative patients. This result was paired with a specificity of 100%, as the algorithm did not identify any false-positive lesion clusters in the 24 healthy controls or the 11 disease controls with TLE. This perfect specificity in two separate control groups was a particularly strong finding, suggesting the model had learned a signature highly specific to FCD.

The performance on the independent 1.5 T dataset served as a robust test of the method's generalizability. When the classifier, which was trained exclusively on 3.0 T data, was applied to this new cohort, it maintained strong performance. It achieved a sensitivity of 71% (detecting the lesion in 10 of 14 patients) and a specificity of 95%, with the slight decrease in specificity due to a single false-positive detection in a control subject.

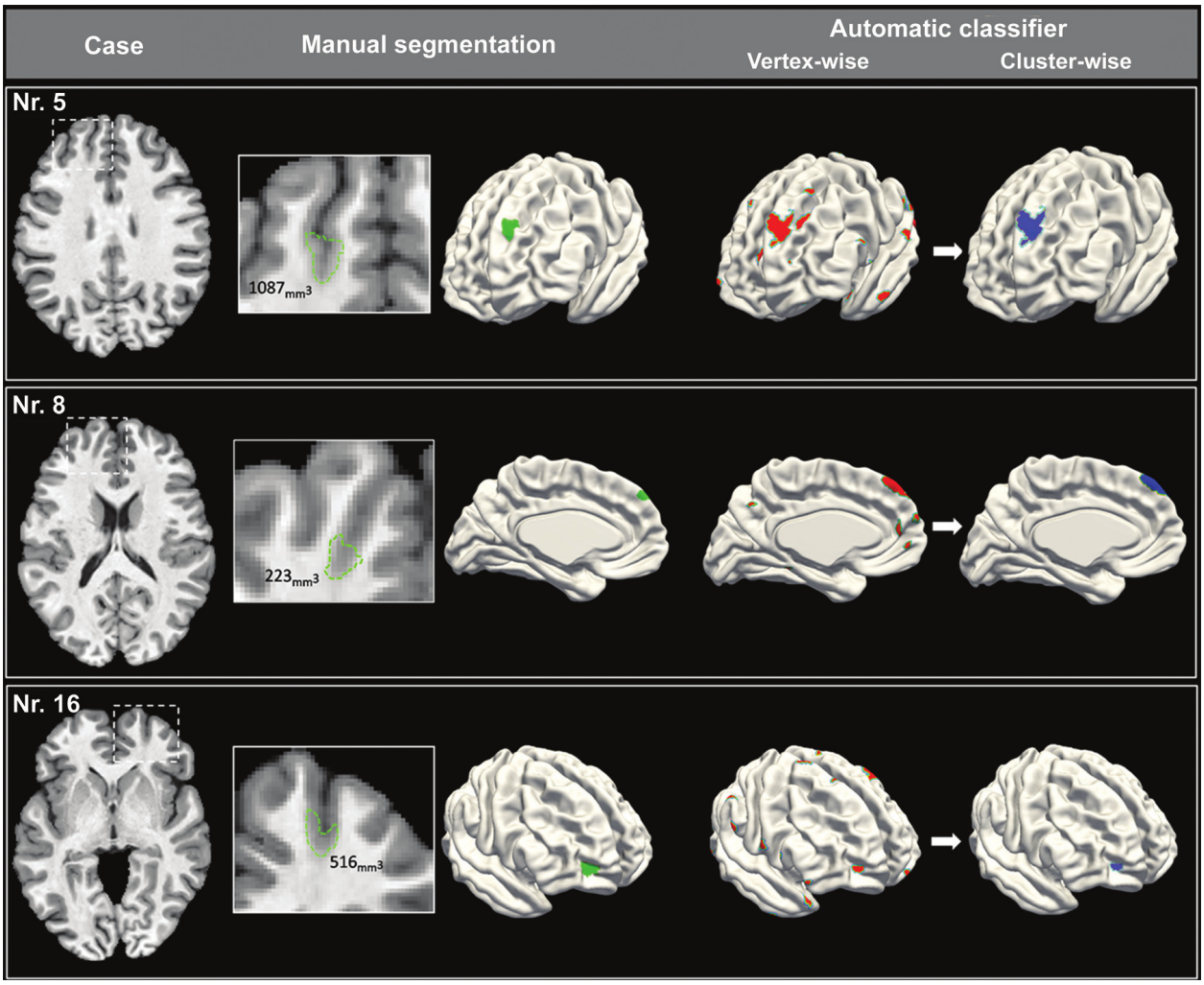


**Figure 2:** Examples of true positive detections from Hong et al. (2014). For each case, the manual segmentation (green) is shown alongside the initial vertex-wise classifier output (red) and the final, more specific cluster-wise output (blue).

Despite the high overall specificity, the classifier's output was not always unambiguous in the patient group. The authors report that of the 14 patients in the 3.0 T cohort where the lesion was correctly identified, only half (7 patients) had a single, clean detection corresponding to the FCD. The other 7 patients also had additional "extralesional" clusters identified elsewhere in the brain, with a median of one such cluster per patient. A retrospective analysis of these extralesional findings, exemplified by the cases shown in Fig. 2, revealed that they were often smaller than the true lesion and were disproportionately characterized by a single abnormal feature. As shown in the feature analysis in Fig. 3, true lesions tended to show abnormalities across multiple features (e.g., thickness, gradient, and depth), while extralesional clusters were often abnormal in only one, most commonly sulcal depth.

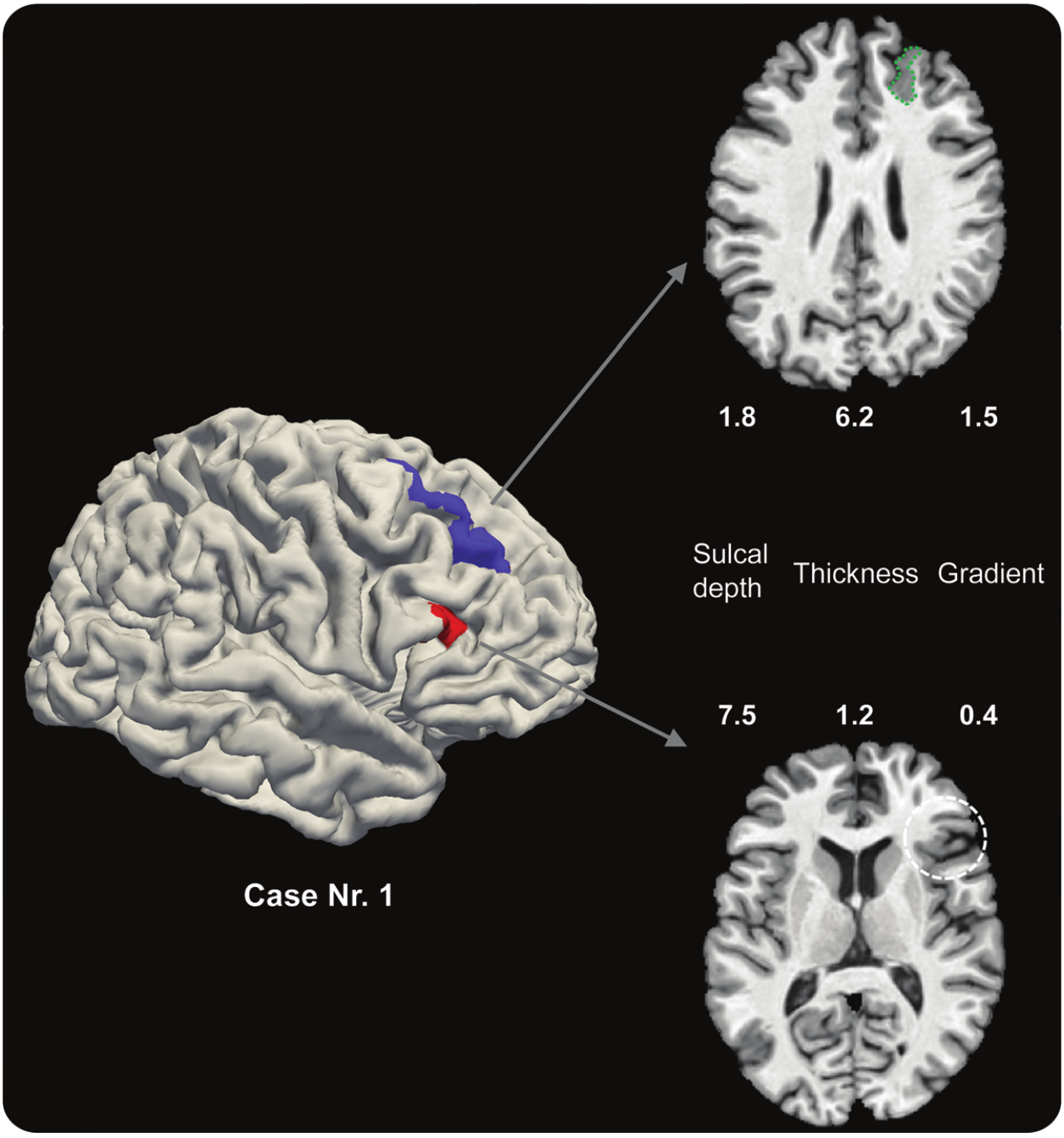


**Figure 3:** Comparison of quantitative features for a true positive detection versus an extralesional finding from Hong et al. (2014). The true lesion (red cluster) is distinguished by its abnormal sulcal depth, while the extralesional cluster (blue cluster) is primarily characterized by abnormal cortical thickness. This example illustrates how the multivariate classifier uses different feature profiles to differentiate true FCD from other morphological anomalies.

Based on these findings, the authors drew several key conclusions. They asserted that their "fully automated multivariate approach offered a substantial gain in sensitivity over standard radiologic assessment," providing a new diagnostic avenue for patients previously considered MRI negative. They highlighted the method's generalizability, supported by the consistent performance on the independent 1.5 T dataset. Ultimately, they proposed that ML tools like theirs could play a significant role in the clinical workflow to "assist presurgical decision-making by facilitating hypothesis formulation about the EZ."

## 4.4 Critical Appraisal

### 4.4.1 Strengths and Contributions

The 2014 paper by Hong et al. stands as a landmark study primarily for its powerful demonstration of a "proof-of-concept." It was among the first studies to show convincingly that an ML algorithm could automatically and objectively identify epileptogenic brain lesions that were occult to conventional, expert-led radiological inspection. This achievement moved the field beyond computer-assisted visualization and into the realm of true automated detection, showing that computational methods could genuinely extend the boundaries of human diagnostic capability.

Its most significant and lasting contribution was the fundamental shift it forced in the conversation around "MRI negative" epilepsy. By providing Class II evidence of lesion detection in this cohort, the paper powerfully argued that "MRI negative" is often a limitation of technology and methodology, not a distinct biological subtype of epilepsy devoid of a structural cause. It demonstrated that the signal for detection is frequently present in the data, but requires a quantitative, multivariate approach to be reliably extracted.

From a clinical standpoint, the development of a fully automated and objective tool for generating hypotheses about lesion location was a major advance. In the complex process of presurgical planning for epilepsy, where diagnostic interpretations can be subjective and variable, this method offered a standardized, data-driven "second opinion." It provided a means to guide the placement of invasive Electroencephalography (EEG) electrodes with a higher degree of confidence, potentially reducing sampling error and improving surgical outcomes.

### 4.4.2 Statistical Significance and Rigor

A focused examination of the study's statistical underpinnings is essential to contextualize its findings. The choices made by the authors reflect a common trade-off in early-stage translational research, prioritizing exploratory power over the stringent error control typical of confirmatory studies. The study's statistical framework is best understood as one optimized for discovery. Adopting strict, confirmatory statistical methods designed to minimize Type I error would have inherently reduced the power to detect the very lesions they were pursuing. Consequently, they chose a high-sensitivity validation strategy and a clever heuristic to manage false positives. This approach is well-suited for a trailblazing study, though the consequence is a lack of robust uncertainty quantification.

Regarding the validation strategy, the use of LOOCV is a defensible choice for the small and rare patient cohort (n=19), as it maximizes the use of available data for training and leads to a performance estimate with very low bias [13]. The primary drawback, however, is that LOOCV estimates can have high variance; because the training sets in each fold are nearly identical, the resulting models are highly correlated, and the final averaged performance metric can be unstable [13]. For multiple comparisons control, the two-stage, cluster-wise classifier is a pragmatic, heuristic

approach that successfully reduced the initial deluge of false alarms without the potentially severe loss of sensitivity that would accompany an overly conservative statistical correction [15]. However, this method lacks any formal statistical guarantee on the error rate across the brain [14]. A significant statistical omission by modern standards is the reporting of uncertainty; given the small sample size and the LOOCV procedure, the 95% confidence interval around the 74% sensitivity is likely to be very wide. Finally, it is important to consider the difference between statistical significance and clinical relevance [17]. A tool with 74% sensitivity will still miss one in four lesions, and the presence of extralesional findings in 50% of positive cases can confuse, rather than guide, surgical planning.

#### 4.4.3 Limitations and Future Directions

The very success of the Hong et al. paper serves to illuminate its own limitations, which in turn provide a compelling rationale for the subsequent evolution of the field. First, the study population may suffer from spectrum bias. The patient cohort was not a random sample of all MRI negative epilepsy patients but a highly selected group known to have lesions that were, at a minimum, detectable by advanced texture analysis. This pre-selection means the reported performance metrics are likely an overestimation of how the algorithm would perform when applied to a more general, unselected clinical population. Second, the problem of "extralesional clusters" is a major unresolved issue. These additional findings create significant clinical ambiguity. Are they simply statistical noise, which would lower the classifier's true specificity, or could they represent regions of genuine but subclinical pathology? Without histopathological correlation, this question remains unanswered.

Arguably the most significant conceptual limitation, however, is the classifier's complete dependence on a small set of manually engineered features. The method's success is fundamentally constrained by the five features the researchers chose based on existing knowledge of FCD pathology. This raises critical questions about whether other, more complex patterns embedded in the MRI data are being missed, and whether the process itself introduces an indelible human bias into the discovery process. These limitations set the stage for the next logical steps. The reliance on a limited set of features from a single MRI contrast highlights a data limitation, creating a need for richer input data from techniques like quantitative, multiparametric imaging. At the same time, the paradigm of manual feature engineering points to a model limitation, motivating the development of more powerful learning architectures, such as Deep Neural Networks (DNNs), which can learn feature representations automatically from the data itself.

## 5 Act II: The Paradigm-Shifting Tool: A Critical Analysis of Ma et al. (2013)

### 5.1 Research Objectives and Context

The 2013 publication of "Magnetic resonance fingerprinting" by Ma et al. in *Nature* represents a watershed moment in the field of MRI [11]. To appreciate its significance, one must consider the prevailing paradigm it sought to disrupt. For decades, clinical MRI relied on producing a series of qualitative, "weighted" images. A radiologist's report would describe a lesion as "hyperintense" on a T2-weighted scan, a term that describes relative signal differences but lacks objective, physical meaning. The signal intensity itself was arbitrary, dependent on scanner hardware and software, rendering it non-comparable across different scanning sessions, let alone different medical centers. The acquisition process was also inherently inefficient. To obtain different image contrasts, such as T1-weighting and T2-weighting, separate scans had to be performed consecutively. Achieving true quantitative maps of physical tissue properties, such as the longitudinal relaxation time ($T_1$) or the

transverse relaxation time ($T_2$), was even more cumbersome and time-prohibitive for routine clinical use.

It is within this context that Ma et al. introduced MRF. The central objective was to introduce a "completely different approach to data acquisition, postprocessing and visualization." The authors aimed to overcome the limitations of conventional MRI by enabling the "simultaneous non-invasive quantification of multiple important properties of a material or tissue" from a single, rapid acquisition. This objective was not an incremental improvement but a fundamental reimagining of the MR experiment itself. Crucially, the paper by Ma et al. must be understood as foundational, "tool-building" research. Its aim was to invent and validate a new method for acquiring and processing MR data, positioning MRF as a general-purpose enabling technology.

## 5.2 Methodological Approach

The core technical innovation of MRF lies in its radical departure from the steady-state paradigm of conventional MRI. It decouples the one-to-one relationship between a pulse sequence and a specific image contrast, instead proposing a single, complex acquisition that encodes information about multiple tissue properties simultaneously into the unique temporal signal evolution of each voxel. This framework can be deconstructed into three key components.

### 5.2.1 The Acquisition: Creating Unique Fingerprints

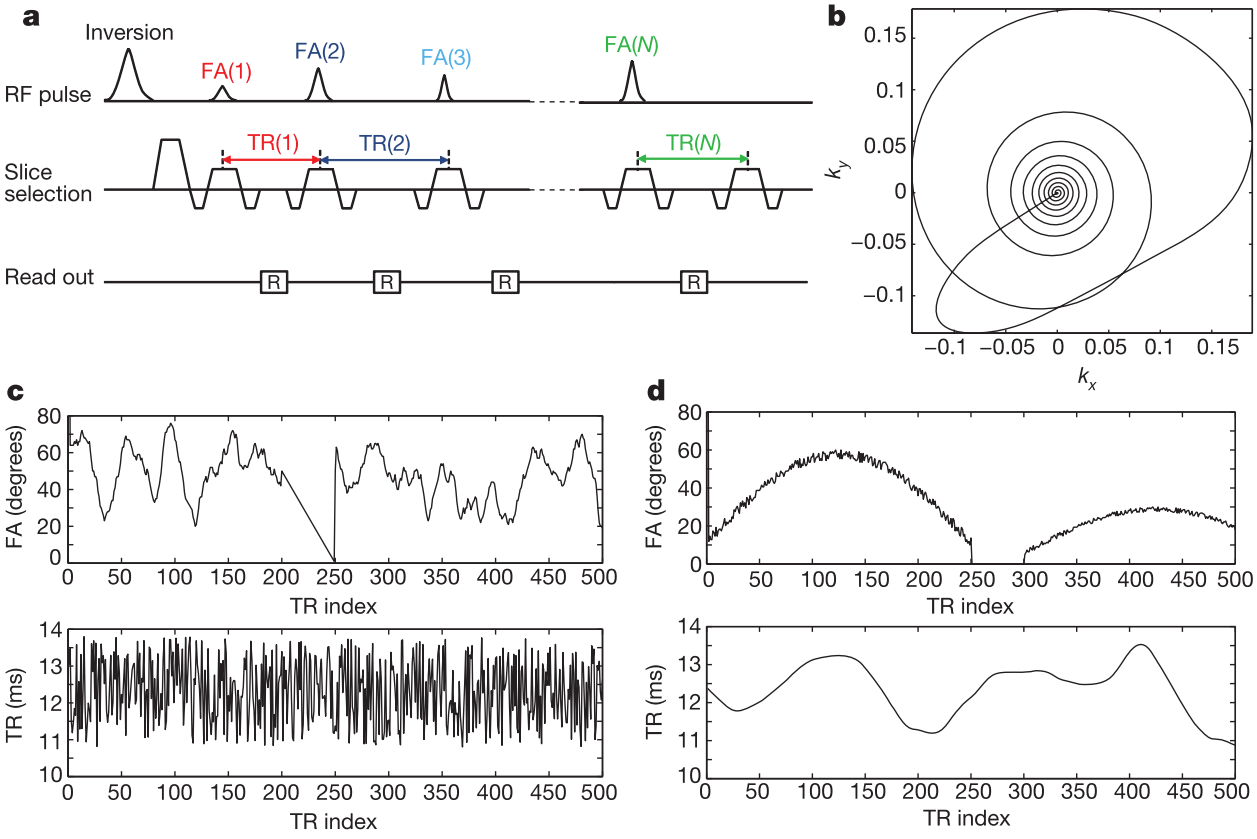


**Figure 4:** The MRF acquisition scheme from Ma et al. (2013). (a) A diagram of the IR-bSSFP sequence with varying FA and TR. (b) The highly undersampled variable-density spiral k-space trajectory. (c, d) Two different pseudorandom patterns of FA and TR used in the experiments.

Unlike conventional pulse sequences that use fixed parameters to achieve a steady-state signal, MRF employs continuously and pseudorandomly varying acquisition parameters. The proof-of-principle implementation was based on an IR-bSSFP sequence, where for each acquisition block, the flip angle and repetition time are changed according to a predetermined, non-repetitive pattern (Fig. 4). This strategy ensures that the magnetization never reaches a steady state; instead, this dynamic signal evolution, the "fingerprint", encodes the quantitative information. To achieve the necessary speed, the acquisition is highly undersampled using a variable-density spiral k-space trajectory that is rotated at each time point.

### 5.2.2 The Dictionary Generation: The Pre-calculated Ground Truth

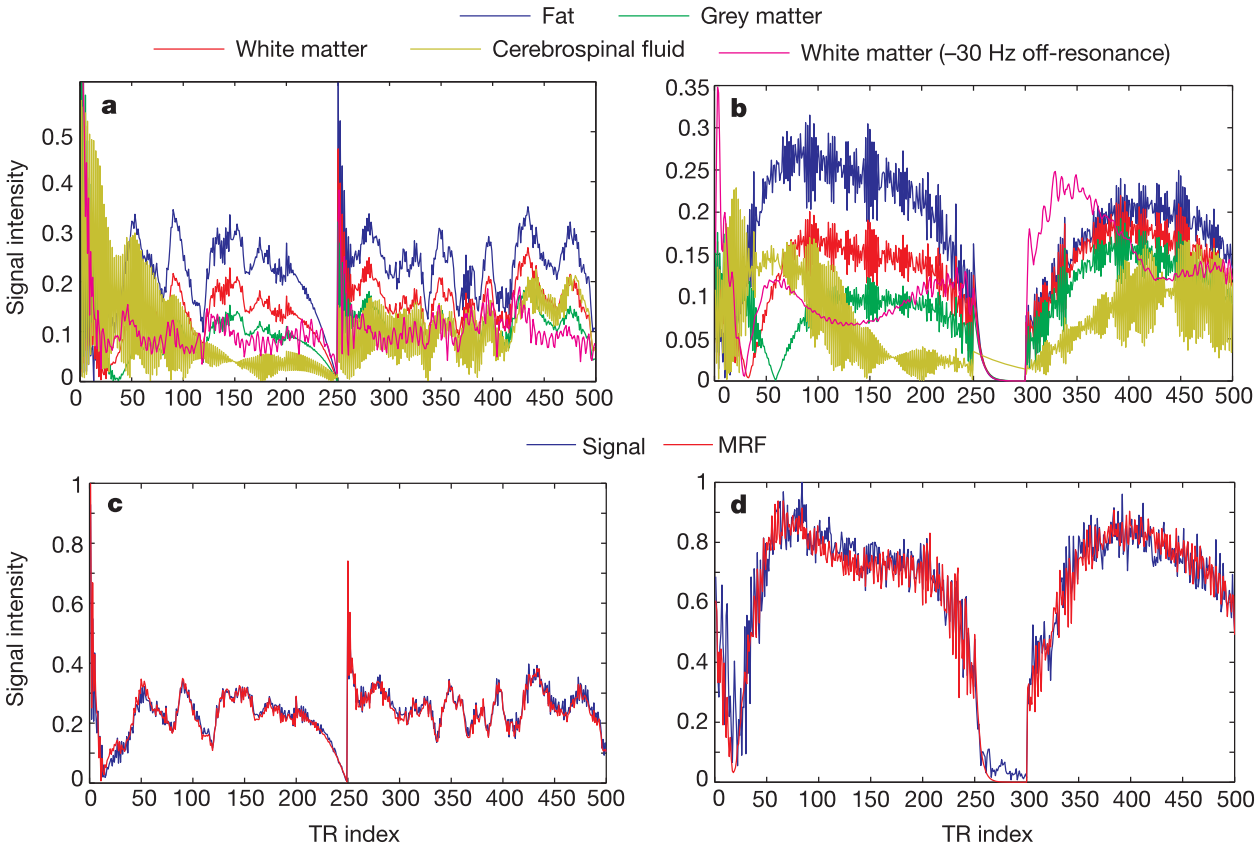


**Figure 5:** Simulated and measured MRF signals from Ma et al. (2013). (a, b) Simulated signal evolutions (fingerprints) for different tissue types, demonstrating their unique character. (c, d) Validation in phantoms, showing the excellent match between the measured signal (blue) and the corresponding dictionary entry (red).

The second component is the creation of a massive, pre-computed lookup table, or "dictionary." Each entry is a theoretical signal evolution, a simulated fingerprint, corresponding to a unique combination of tissue properties ($T_1$, $T_2$, off-resonance $\Delta f$, etc.). This dictionary is generated by simulating the Bloch equations using the exact same pseudorandom sequence parameters from the acquisition. The simulated fingerprints in Fig. 5a and Fig. 5b confirm the fundamental assumption in MRF that different tissues generate unique and distinguishable signal evolutions.

#### 5.2.3 The Pattern Matching: From Signal to Quantitation

The final step is reconstruction. The experimentally measured fingerprint from each voxel is compared to every entry in the dictionary, typically via a normalized dot product. The dictionary entry with the highest dot product is the "best match." The quantitative tissue properties ($T_1$, $T_2$, $\Delta f$) used to generate that dictionary entry are then assigned to the voxel. Because aliasing artifacts are temporally incoherent, they do not correlate well with any physical dictionary entry and are thus largely ignored by the pattern-matching algorithm.

### 5.3 Key Findings and Conclusions

The authors of Ma et al. provided a multi-pronged validation of the MRF concept, progressing from simulation to physical phantoms and finally to in vivo human imaging. The in vivo results were a compelling demonstration of MRF's potential. Fig. 6 showcases the method's extraordinary tolerance to undersampling. A single 2D slice was acquired in just 12.3 s, yet the pattern matching process successfully rejected aliasing artifacts to produce high-quality, co-registered maps of $T_1$, $T_2$, proton density ($M_0$), and off-resonance frequency. The robustness of the technique was tested against bulk subject motion, one of the most common sources of error in clinical MRI. As demonstrated in Fig. 7, even when a scan was corrupted by significant head movement, the resulting quantitative maps were nearly identical to those from a motion-free acquisition.

The quantitative accuracy was rigorously tested against gold-standard spin-echo sequences, as shown in Fig. 8. The MRF-derived $T_1$ and $T_2$ values showed excellent agreement with the standards, achieving concordance correlation coefficients (CCCs) of 0.988 for $T_1$ and 0.974 for $T_2$. Furthermore, MRF was shown to be significantly more efficient (defined as precision per square root of scan time) than contemporary rapid mapping methods. The authors' primary conclusions were that MRF represents a new paradigm for magnetic resonance that enables simultaneous, multiparametric quantitative mapping in a clinically feasible timeframe.

### 5.4 Critical Appraisal

#### 5.4.1 Strengths and Contributions

The revolutionary contribution of MRF is a philosophical shift in experimental design. Conventional MRI is fundamentally reductionist, engineering pulse sequences to isolate one physical contrast. MRF, in contrast, is holistic, creating a complex signal that is an inseparable mixture of multiple underlying physical properties, shifting the burden of separation from acquisition hardware to postprocessing software. This transformed the output of an MRI scanner from qualitative images into maps of objective, reproducible physical constants. This transition to true quantitation is precisely what enables the advanced, data-driven analyses seen in later work [18]. Perhaps MRF's most profound impact is its role as an enabling technology, providing the ideal raw material, rich, quantitative, multiparametric data, for modern computational techniques. The subsequent success of researchers in using MRF data for challenging clinical tasks, such as the detection of subtle FCD [12], is the ultimate validation of this contribution.

#### 5.4.2 Statistical Significance and Rigor

The authors employed robust statistical methods to validate the accuracy and precision of MRF. To assess agreement with gold-standard measurements, they used the CCC ($\rho_c$), which is commendable as it evaluates both correlation and bias. To evaluate the precision and reproducibility, the authors

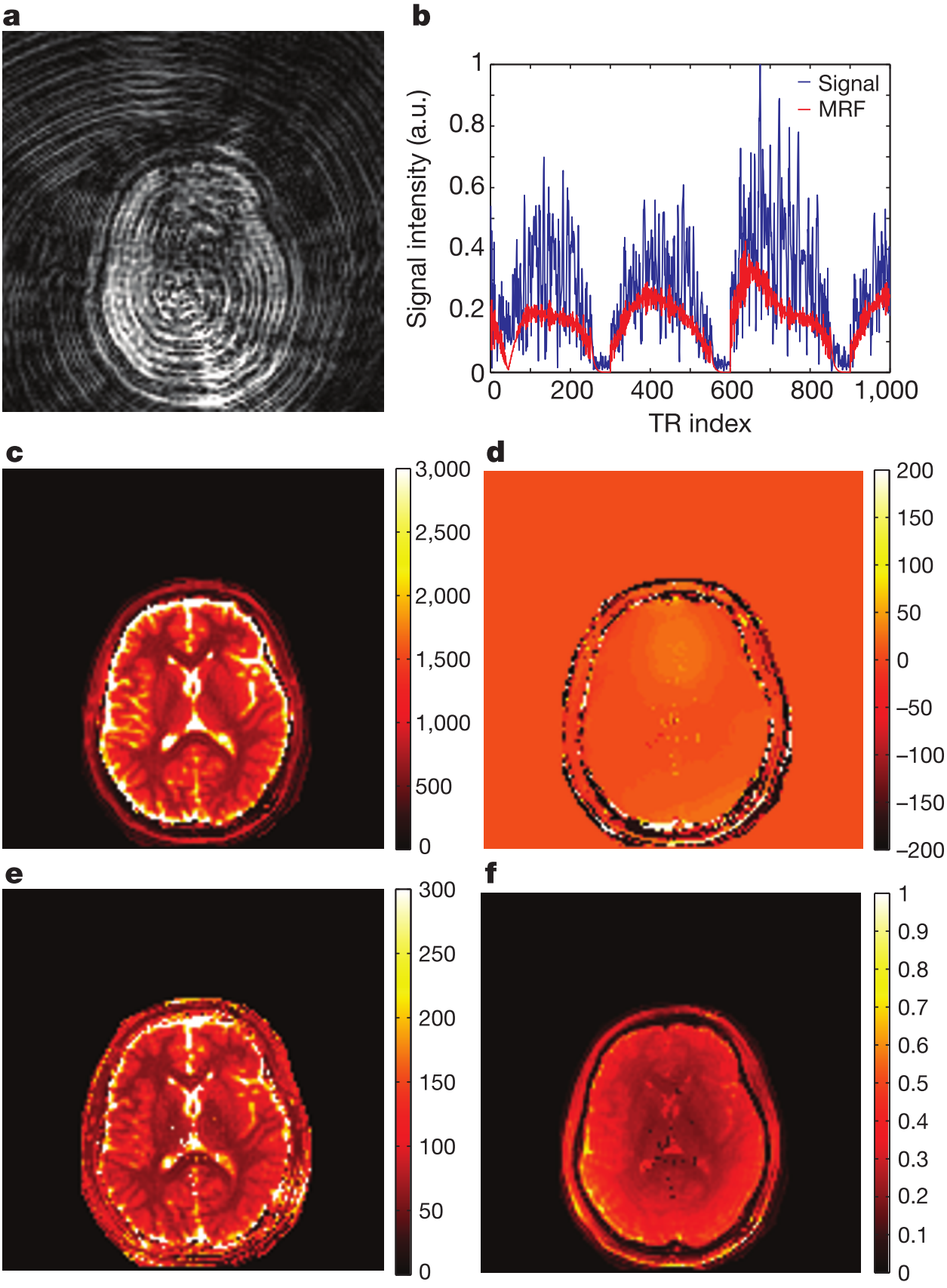


**Figure 6:** In vivo MRF results from Ma et al. (2013). (a) A single raw image frame, heavily corrupted by aliasing. (b) The measured signal fingerprint from one voxel (blue) and the matched dictionary entry (red). (c-f) The final, co-registered quantitative maps for $T_1$, $\Delta f$, $T_2$, and proton density ($M_0$) respectively, generated from a single rapid scan.

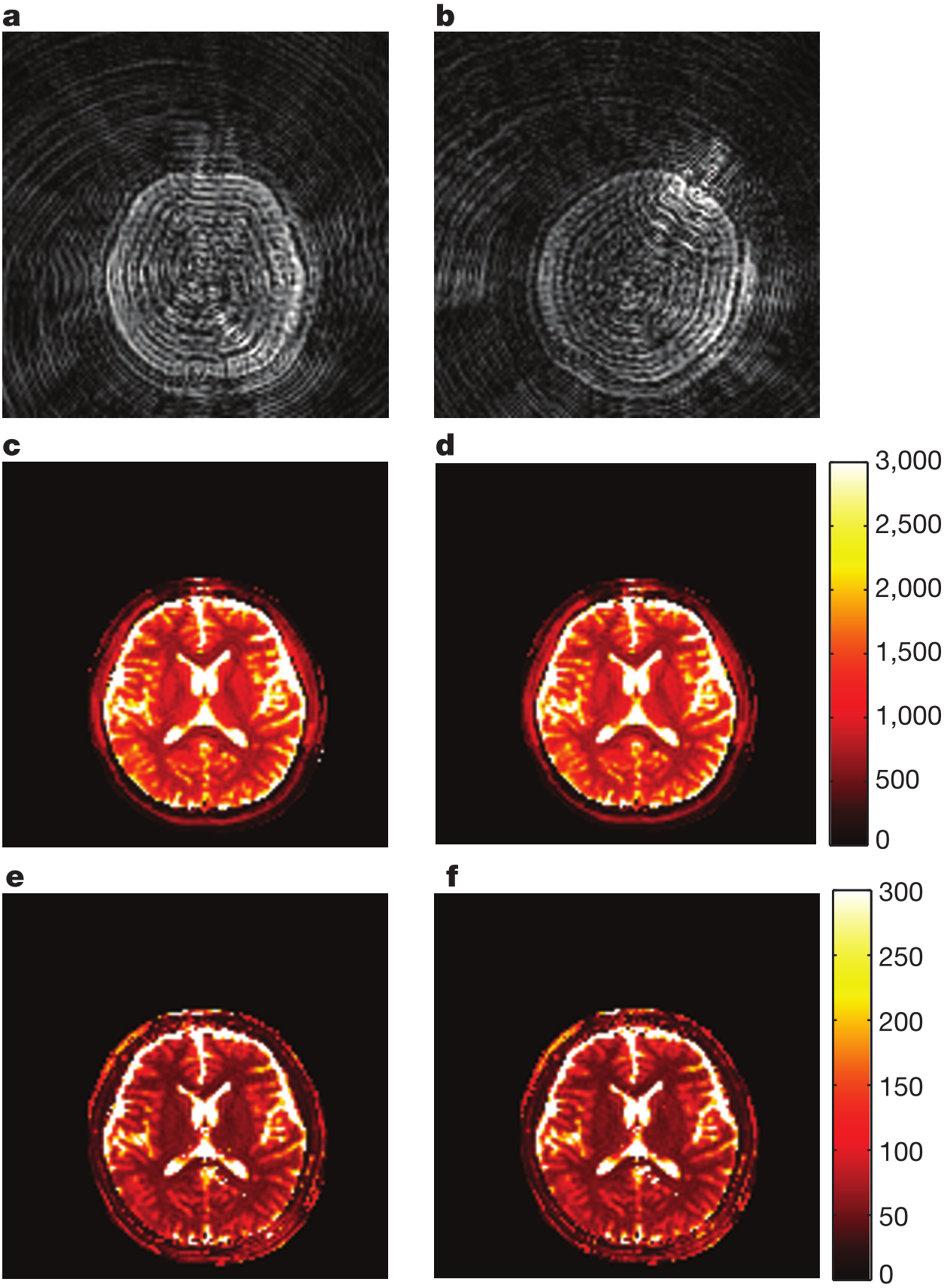


**Figure 7:** Demonstration of MRF's robustness to motion from Ma et al. (2013). (a) Raw image without motion. (b) Raw image with subject motion during the last few seconds. (c, e) $T_1$ and $T_2$ maps from the motion-free scan. (d, f) Nearly identical $T_1$ and $T_2$ maps from the motion-corrupted scan, demonstrating that incoherent motion signals are rejected by the pattern-matching algorithm.

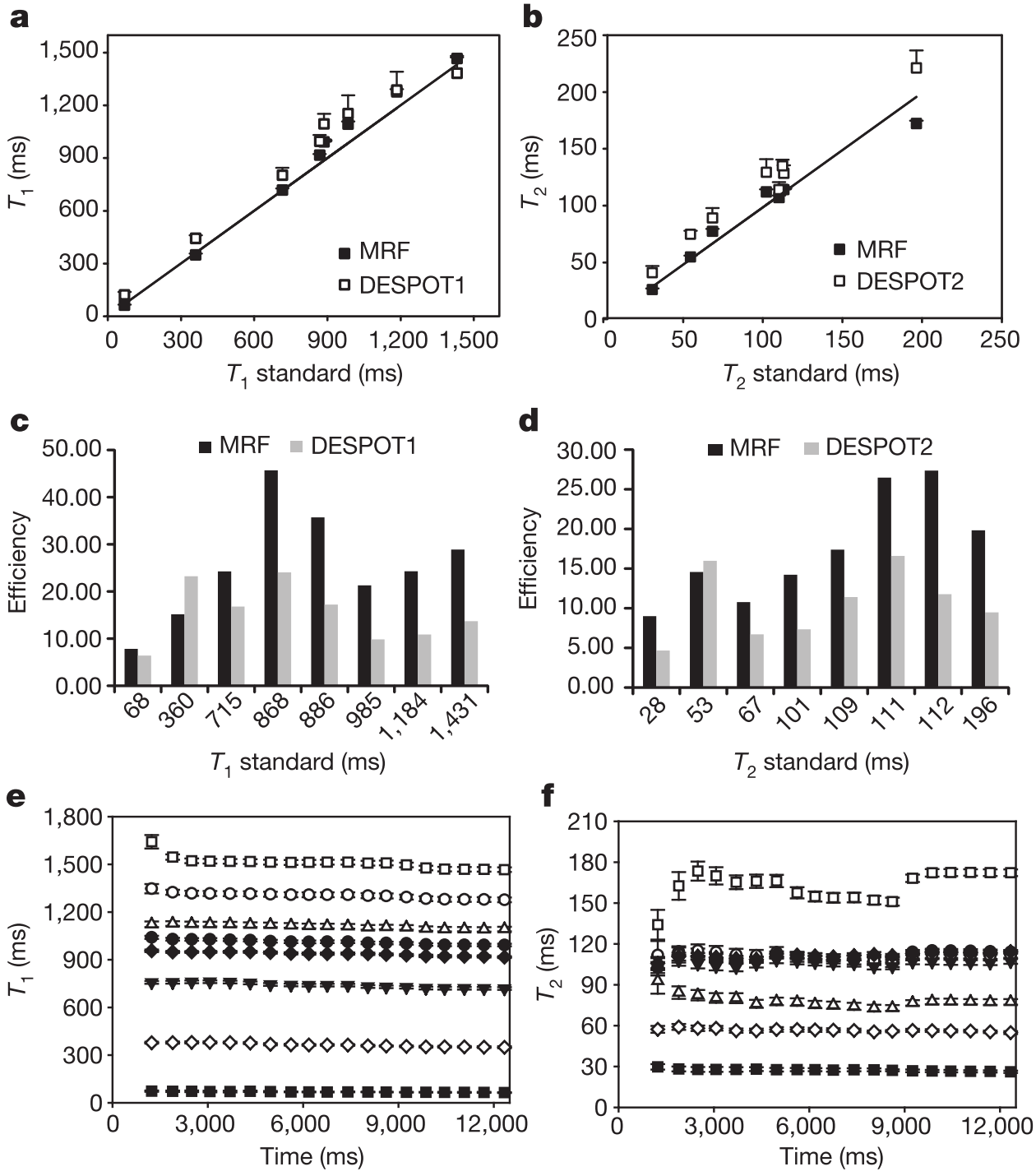


**Figure 8:** Validation of MRF accuracy, efficiency, and precision from Ma et al. (2013). (a, b) Accuracy plots showing high correlation between MRF and standard measurements for $T_1$ and $T_2$. (c, d) Efficiency plots showing MRF outperforms the DESPOT method. (e, f) Precision plots showing high repeatability of MRF measurements over time.

**Table 1:** Comparison of MRI Paradigms.

| Feature | Conventional Qualitative MRI | Conventional Quantitative | MRF |
|---|---|---|---|
| Primary Output | Weighted images (a.u.) | Single quantitative map | Multiple, simultaneous quantitative maps |
| Acquisition Strategy | Steady-state | Multiple steady-state | Single, non-steady-state |
| Parameters per Scan | One (weighted) | One (quantitative) | Multiple (quantitative) |
| Scan Time | Fast per contrast | Slow per parameter | Very fast for all parameters |
| Robustness | Low (to undersampling/motion) | Low (to motion/errors) | High (to undersampling/motion) |

used a bootstrapped Monte Carlo method. Despite this rigor, a notable omission is the use of Bland-Altman plots. While $\rho_c$ provides an excellent summary statistic, a Bland-Altman plot is the gold-standard graphical method for comparing two quantitative measurement techniques and would have visually revealed any systematic bias or whether the magnitude of the error was dependent on the value being measured. This absence represents a minor weakness in an otherwise strong validation.

### 5.4.3 Limitations and Future Directions

The initial implementation of MRF presented in this paper had several acknowledged technical challenges. The first is the computational burden associated with the dictionary, which grows exponentially with each new parameter added. Second, while MRF can model hardware imperfections, it is still sensitive to them. However, the most crucial limitation of this paper, in the context of this report's narrative, is what it did not do: apply this new tool to a specific clinical problem. This limitation is precisely what makes the paper a perfect narrative bridge. Ma et al. provided the scientific community with a powerful new instrument capable of generating rich quantitative data at unprecedented speeds. This created the need for the next layer of innovation: advanced computational algorithms, such as the DL framework presented by Ding et al. [12], which could finally leverage this rich data to solve the difficult clinical problem of detecting MRI negative FCD [6].

## 6 Act III: The Modern Synthesis: A Critical Analysis of Ding et al. (2025)

The 2025 study by Ding et al., "Automated Whole-Brain Focal Cortical Dysplasia Detection Using MR Fingerprinting With Deep Learning," represents the culmination of the research trajectory explored in this report [12]. It serves as a modern synthesis, ingeniously integrating the advanced quantitative imaging paradigm of MRF with the analytical power of DL. This work moves the field beyond hypothesis-driven feature engineering on qualitative images toward a data-driven approach of feature discovery on quantitative, multiparametric data. In doing so, it provides a powerful, state-of-the-art solution to the long-standing clinical challenge of detecting subtle, epileptogenic brain lesions, directly addressing the problems and leveraging the technologies established by the preceding cornerstone papers.

## 6.1 Research Objectives and Context

The research objectives of Ding et al. can be understood as the logical resolution to the narrative arc established by the clinical needs articulated by Hong et al. (2014) and the technological potential introduced by Ma et al. (2013). The paper's central aim was to develop an "MRF-based DL framework for whole-brain FCD detection," a goal that directly confronts the challenge of identifying pharmacoresistant FCD, particularly in cases where conventional MRI fails. This objective is a direct successor to the problem first tackled by Hong et al. Their patient cohort is not only composed of FCD Type II cases but is deliberately heterogeneous, including the subtler subtypes of mMCD and MOGHE. This demonstrates a progression in the field's ambition to push the detection frontier into the most diagnostically challenging populations.

The methodology to achieve this ambitious goal is built upon the quantitative, multiparametric imaging paradigm pioneered by Ma et al. [11]. By acquiring a single 3D whole-brain MRF scan, Ding et al. generate a rich, perfectly co-registered suite of quantitative maps that provide a far deeper, more physically meaningful characterization of tissue than the single, qualitative T1-weighted images used by Hong et al. Perhaps the most profound advance lies in the analytical leap from classical ML to DL. In stark contrast to the pre-defined, "handcrafted" features used by Hong et al., Ding et al. employ a no-new-U-Net (nnU-Net), which is a sophisticated variant of a CNN. The fundamental theoretical advantage of a CNN is its capacity for automatic feature learning, discovering a complex hierarchy of discriminative features directly from the image data itself. This marks a fundamental philosophical shift from "feature engineering" to "representation learning."

## 6.2 Methodological Approach

The methodology presented by Ding et al. represents a state-of-the-art, integrated framework that combines a challenging and clinically relevant patient cohort with an advanced imaging protocol and a powerful DL architecture, summarized in Fig. 9.

### 6.2.1 Patient Cohort & Imaging Protocol

The study's clinical relevance is anchored in its population of 40 patients with pharmacoresistant focal epilepsy and confirmed FCD, alongside a control group of 67 healthy individuals. The inclusion of a heterogeneous set of pathologies is a significant strength [19]. All participants underwent a single 3D whole-brain MRF scan (approx. 12 min) to yield a comprehensive suite of co-registered input maps for the DL model: quantitative $T_1/T_2$ maps, tissue fraction maps (GM, WM, CSF), and derived feature maps like a synthetic T1-weighted image and $T_1/T_2$ z-score maps. The richness of these multiparametric inputs is visualized in Fig. 10.

### 6.2.2 Deep Learning Architecture

The analytical core of the framework is the nnU-Net, a robust, self-configuring implementation of the U-Net architecture. The U-Net, shown schematically in Fig. 11, is a specialized CNN that has become a de facto standard for biomedical image segmentation. It features a symmetric encoder-decoder structure with "skip connections" that enable the network to fuse deep, abstract contextual features with fine-grained spatial details, which is critical for producing precise segmentation masks.

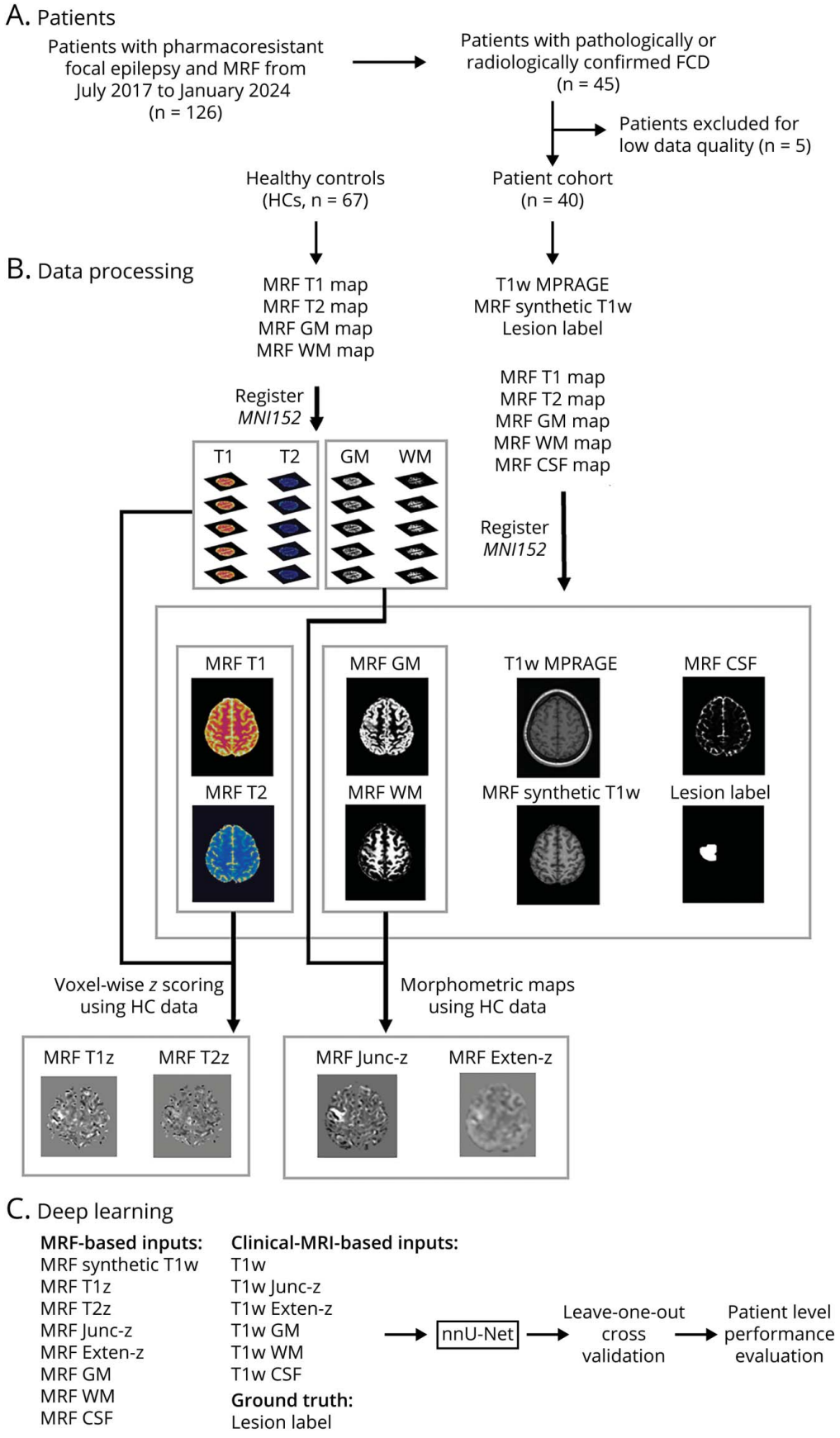


**Figure 9:** The integrated workflow from Ding et al. (2025). (A) Patient selection flowchart. (B) Data processing pipeline showing the generation of multiple quantitative and derived feature maps from a single MRF acquisition. (C) DL model inputs and training scheme.

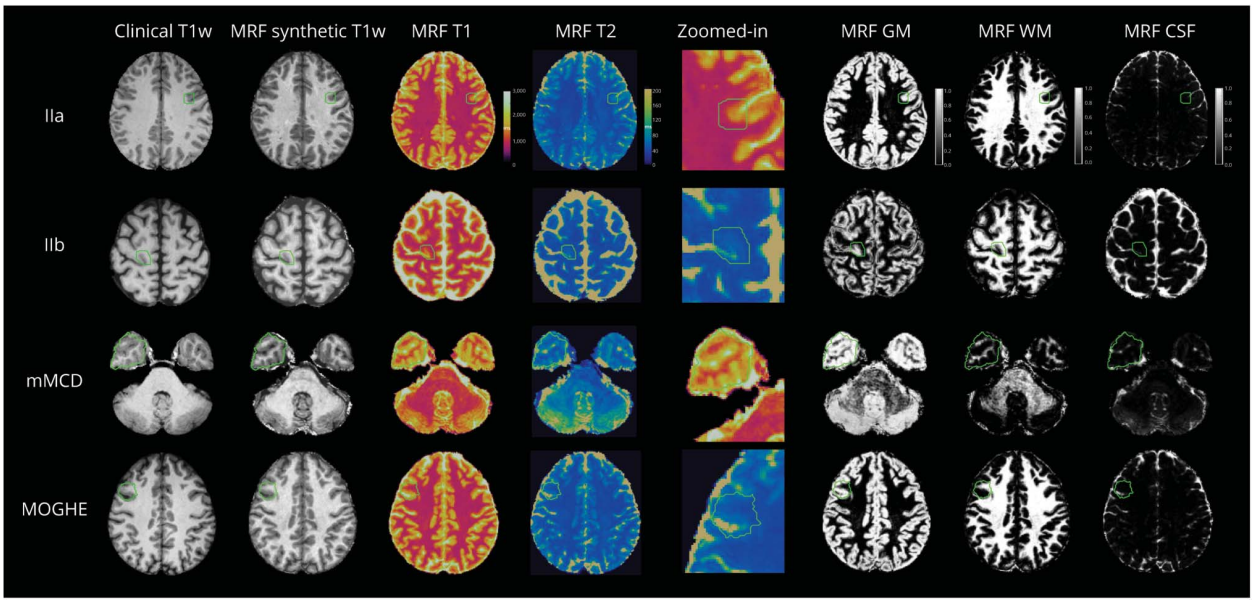


**Figure 10:** Multiparametric MRF maps from Ding et al. (2025) for four different FCD subtypes (FCD IIa, FCD IIb, mMCD, MOGHE). The images demonstrate the rich, quantitative information provided by a single MRF scan, highlighting different lesion characteristics across T1, T2, and tissue fraction maps.

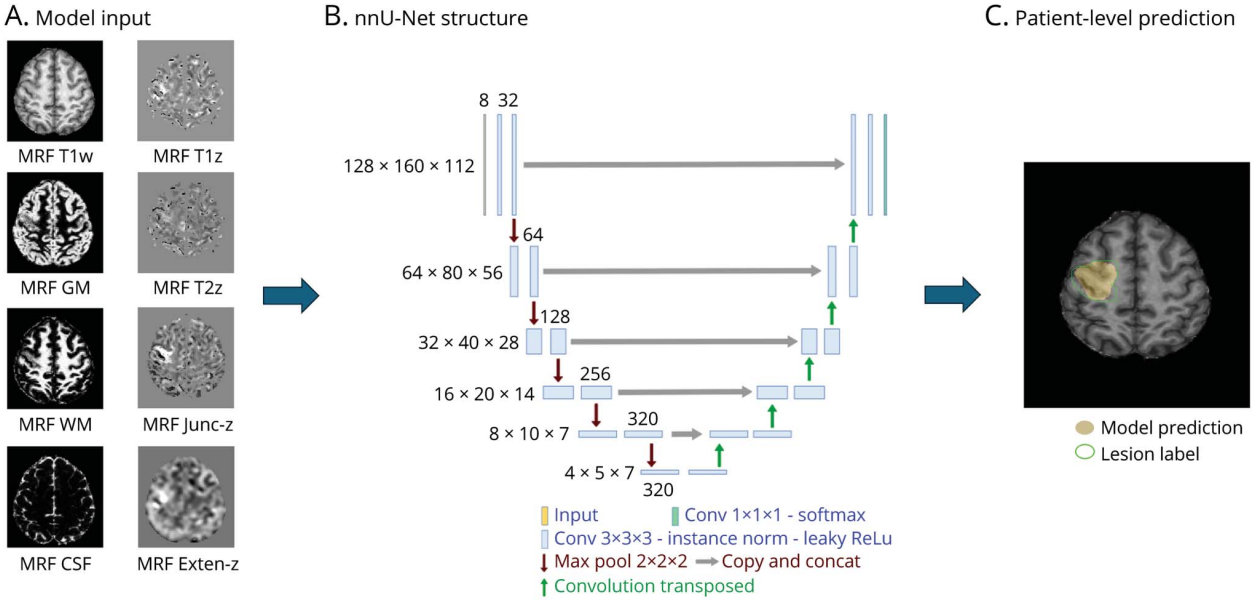


**Figure 11:** The DL framework from Ding et al. (2025). (A) A suite of multiparametric maps derived from a single MRF scan serves as the input to the model. (B) A CNN with a U-Net architecture processes the image data. (C) The model outputs a probabilistic map indicating the location of the predicted lesion.

### 6.2.3 Model Training and Validation

The model was trained in a supervised manner using 3D lesion Regions of Interest (ROIs) manually delineated by experts as the "ground truth." To assess performance rigorously on their cohort of 40 patients, the authors employed LOOCV, a statistically robust choice for smaller datasets as it maximizes the data used for training in each iteration.

## 6.3 Key Findings and Conclusions

The results presented by Ding et al. provide compelling evidence for the efficacy of their synthesized MRF-DL framework, demonstrating a significant leap in performance over existing methods. The optimal model, which used the full suite of MRF-based inputs, achieved a patient-level sensitivity of 80%, successfully identifying the FCD lesion in 32 of the 40 patients. This high sensitivity was accompanied by a clinically acceptable average of 1.7 False Positives (FPs) per patient. Fig. 12 offers a powerful visual summary of these results.

The study's strength is amplified by its direct comparison to alternative approaches. A model trained on conventional clinical T1-weighted images achieved lower sensitivity (70%) with a dramatically higher FP rate (9.7 per patient). This comparison makes it clear that the superior performance is driven primarily by the rich, quantitative information inherent to the MRF inputs. Furthermore, the MRF-DL framework significantly outperformed the well-established Morphometric Analysis Program (MAP18) pipeline, which achieved a sensitivity of only 50%. The authors conclude that their framework demonstrated high efficacy for whole-brain FCD detection, with the key advantage of gathering all necessary data in a single, efficient scan, thereby eliminating inter-scan registration errors.

## 6.4 Critical Appraisal

### 6.4.1 Strengths and Contributions

The primary contribution of this study is its powerful synthesis of two transformative technologies. It surmounts the reliance on handcrafted features in the work of Hong et al. and provides a high-impact clinical application for the MRF technology developed by Ma et al. The model's consistent performance across different FCD subtypes is a major strength, suggesting the quantitative MRF features reveal a fundamental pathological signature. The following table (Table 2) summarizes this technological evolution.

### 6.4.2 Statistical Significance and Rigor

The statistical validation of the framework is generally sound. The choice of LOOCV is appropriate for the study's sample size (n=40). However, while the numerical superiority of the MRF-DL model over benchmarks is clear, the paper does not report formal statistical tests comparing these classifiers (e.g., DeLong's test for Areas Under the Curve (AUCs)) to establish the significance of this superiority. Similarly, the primary sensitivity metric of 80% is presented as a point estimate; reporting a 95% confidence interval would be best practice to convey the estimate's precision. Nonetheless, from a clinical perspective, the low average of 1.7 FPs per patient is a highly promising figure, suggesting a manageable workload for a reviewing neuroradiologist.

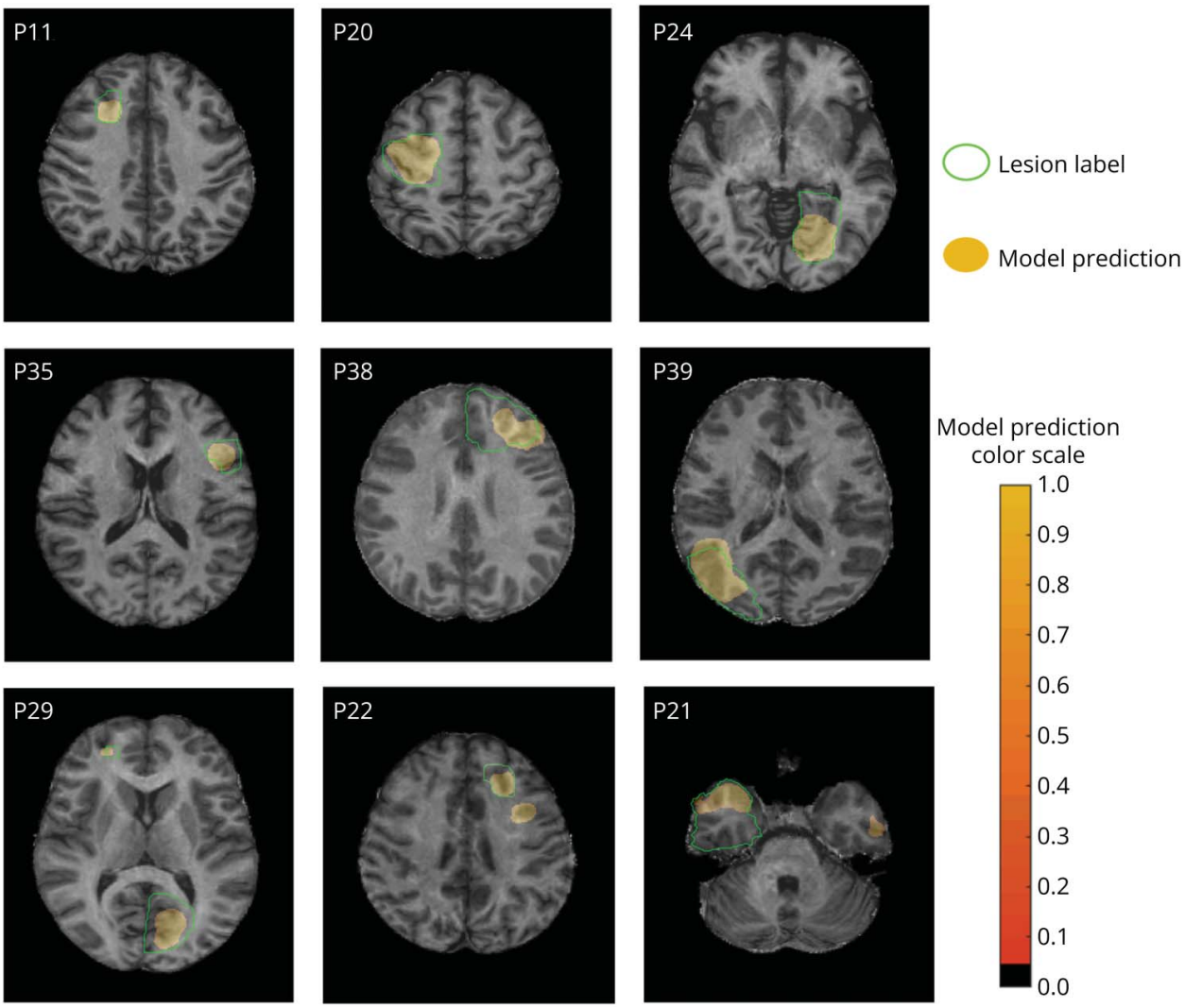


**Figure 12:** Examples of successful automated FCD detections from Ding et al. (2025). The ground-truth lesion label is outlined in green, and the model's probabilistic prediction is shown as a yellow heat map. The figure demonstrates the model's ability to accurately localize lesions of varying types and locations.

**Table 2:** A Comparative Analysis of FCD Detection Methodologies.

| **Methodology** | **Hong et al. (2014)** | **Ma et al. (2013)** | **Ding et al. (2025)** |
|---|---|---|---|
| Primary Goal | Detect MRI negative FCD II | Introduce MRF framework | Detect heterogeneous FCDs |
| Imaging Modality | Qualitative T1-weighted MRI | Quantitative MRF (proof of concept) | Quantitative 3D MRF |
| Feature Engineering | Handcrafted, surface-based | N/A (tech development) | Automated via DL |
| Analysis Model | Classical ML (LDA) | Dictionary matching | DL (nnU-Net) |
| Key Strength | First automated method | Simultaneous quantification | High sensitivity with low FP rate |
| Key Limitation | Brittle features, spectrum bias | No clinical application | Single-center, "black-box" model |

#### 6.4.3 Limitations and Future Directions

Despite its successes, the study's primary limitation is the lack of external validation. The model was developed and tested on data from a single institution. A large-scale, multicenter study is the essential next step to prove the framework's generalizability. This leads to the most critical conceptual challenge for this new generation of Artificial Intelligence (AI) tools: the "black box" problem. The powerful nnU-Net provides a prediction without an explicit explanation of its reasoning. This lack of transparency is a formidable barrier to clinical trust and adoption [20]. The authors acknowledge this by performing a feature importance analysis, but this is only a first step. Answering the question of *why* the model flagged a specific area requires moving toward true XAI, an active area of research aiming to make complex models more transparent and trustworthy [21]. Based on this appraisal, the next logical steps for this research are clear: prospective, multicenter validation to prove generalizability; integration of more sophisticated XAI techniques to build clinical trust; and correlation of model predictions with long-term post-surgical seizure freedom to provide ultimate clinical validation.

## 7 Synthesis, Discussion, and Conclusion

This analysis has synthesized the preceding critiques to construct a cohesive, overarching argument, moving beyond individual papers to demonstrate how these works, in concert, represent a logical and interconnected evolution in the automated detection of FCD. The progression was not merely chronological but represents a powerful and repeatable "Problem-Tool-Algorithm" paradigm for innovation in computational medicine.

### 7.1 The Evolutionary Trajectory: From Problem to Tool to Solution

Scientific progress is often portrayed as a linear march forward, yet the reality frequently involves the convergence of ideas from disparate fields. The central thesis of this report is that the advancement from rudimentary to state-of-the-art FCD detection followed such an evolutionary process, driven by a synergistic interplay of a well-defined clinical problem, the development of a paradigm-shifting measurement tool, and the application of a modern analytical algorithm. A superficial reading of the publication dates is misleading if interpreted as a simple timeline. The true narrative lies in understanding how a foundational tool from a parallel domain was the necessary, albeit chronologically preceding, catalyst that enabled a later synthesis to solve a problem articulated most clearly by a clinical application paper.

The work of Hong et al. began with a clear clinical imperative, implementing a classical ML approach that provided a landmark proof-of-concept: a computational, multivariate approach could identify subtle pathological patterns that eluded expert human visual inspection [6]. However, the paper's greatest contribution was arguably how its success illuminated the profound limitations of the existing toolset, defining a technical void: a clear need for richer, more quantitative raw data. Published a year earlier, Ma et al. introduced MRF, a quintessential "tool in search of a problem" [11]. This work of foundational physics and engineering delivered the critical missing piece needed to fill the technical void. Finally, the paper by Ding and colleagues represents the narrative's climax, masterfully synthesizing the two preceding threads [12]. It directly addresses the clinical problem of FCD detection by applying a modern DL model to the advanced quantitative data generated by MRF, achieving a state-of-the-art lesion detection system that represented a step-change in capability.

### 7.2 Broader Implications and Future Directions

This specific case study reveals a repeatable blueprint for progress in computational medicine and illuminates a clear path forward for research in this field. The "Problem-Tool-Algorithm" synergy observed here is a fundamental pattern of innovation that drives progress across computational medicine [22]. Recognizing this pattern allows for strategic investment in areas where a new tool could unlock a class of intractable problems.

The work of Ding et al. also defines the most critical next steps. There is a clear mandate for rigorous external validation on large, multicenter datasets to prove generalizability [23, 24]. It is also essential to open the "black box" of the AI model. Clinicians are rightly hesitant to base irreversible surgical decisions on an algorithm that cannot explain its reasoning [20, 25]. The application of XAI techniques is a critical step for building clinical trust and may even help discover novel imaging biomarkers [26].

Looking beyond, the true frontier lies in moving from localization to biological characterization. The ultimate clinical question is not just "Where is the lesion?" but "What is the molecular and genetic nature of the epileptogenic network?" This requires the integration of quantitative imaging with other data modalities, such as intracranial Electroencephalography (iEEG) and genomics, creating multi-modal AI models that can learn the relationships between non-invasive imaging features, electrophysiological dysfunction, and the underlying genetic and molecular drivers of the disease [27, 28]. The National Institutes of Health (NIH) grant funding the Ding et al. research already points in this direction by integrating MRF with iEEG data [29]. This is the path toward a true, systems-level understanding and the realization of precision medicine for epilepsy.

### 7.3 Conclusion

In conclusion, this report has confronted one of the most persistent challenges in clinical epileptology: the accurate localization of the EZ in patients with so-called "MRI negative" epilepsy. The failure to identify a discrete lesion in this cohort has historically represented a significant diagnostic and therapeutic impasse [30]. The central argument advanced herein has demonstrated a clear evolutionary trajectory, a progression driven by the co-evolution of imaging technology and data science [31].

This journey commenced with ML establishing a new principle: that lesions, though invisible, could be unmasked computationally. The narrative culminated in demonstrating how advanced computation and superior imaging were powerfully synthesized, shifting the objective from finding a static lesion to modeling a dynamic, pathological system [32]. The ultimate insight gained is a profound lesson in interdisciplinary progress. This evolution was not linear but symbiotic. The initial computational methods highlighted the critical need for higher-fidelity data, creating a clinical pull for technological innovation. In turn, the subsequent explosion of high-resolution imaging data from advanced hardware created both the opportunity and the necessity for more sophisticated data science methodologies. Progress, therefore, is revealed not as the product of isolated breakthroughs, but as a tightly coupled, iterative cycle where advances in imaging hardware and computational modeling perpetually enable and demand more from one another. Ultimately, the evidence affirms that the future of precision epileptology lies in the powerful synergy between our ever-increasing ability to visualize the brain's structure and our growing capacity to model its complex dynamics.

## List of Acronyms

**AI** Artificial Intelligence.

**ASM** Anti-Seizure Medication.

**AUC** Area Under the Curve.

**CCC** concordance correlation coefficient.

**CNN** Convolutional Neural Network.

**CSF** Cerebrospinal Fluid.

**DL** Deep Learning.

**DNN** Deep Neural Network.

**DRE** Drug-Resistant Epilepsy.

**ECE** Electrical & Computer Engineering.

**EEG** Electroencephalography.

**EZ** Epileptogenic Zone.

**FA** flip angle.

**FCD** Focal Cortical Dysplasia.

**FDR** False Discovery Rate.

**FP** False Positive.

**GM** Gray Matter.

**iEEG** intracranial Electroencephalography.

**ILAE** International League Against Epilepsy.

**IR-bSSFP** Inversion-Recovery Balanced Steady-State Free-Precession.

**LDA** Linear Discriminant Analysis.

**LOOCV** Leave-One-Out Cross-Validation.

**MAP18** Morphometric Analysis Program.

**MEG** Magnetoencephalography.

**ML** Machine Learning.

**mMCD** Mild Malformations of Cortical Development.

**MOGHE** Mild Malformation of Cortical Development with Oligodendroglial Hyperplasia.

**MRF** Magnetic Resonance Fingerprinting.

**MRI** Magnetic Resonance Imaging.

**NIH** National Institutes of Health.

**nnU-Net** no-new-U-Net.

**PET** Positron Emission Tomography.

**ROI** Region of Interest.

**SEEG** Stereoelectroencephalography.

**TLE** Temporal Lobe Epilepsy.

**TR** repetition time.

**VBM** Voxel-Based Morphometry.

**WM** White Matter.

**XAI** Explainable Artificial Intelligence.

## References


1. Kwan P, Arzimanoglou A, Berg AT, et al. Definition of Drug Resistant Epilepsy: Consensus Proposal by the Ad Hoc Task Force of the ILAE Commission on Therapeutic Strategies. Epilepsia. 2010 Jun; 51:1069–77. DOI: 10.1111/j.1528-1167.2009.02397.x.
2. Sultana B, Panzini MA, Veilleux Carpentier A, et al. Incidence and Prevalence of Drug-Resistant Epilepsy: A Systematic Review and Meta-analysis. Neurology. 2021 Apr 27; 96:805–17. DOI: 10.1212/WNL.0000000000011839.
3. Wiebe S, Blume WT, Girvin JP, Eliasziw M, and Effectiveness and Efficiency of Surgery for Temporal Lobe Epilepsy Study Group. A Randomized, Controlled Trial of Surgery for Temporal-Lobe Epilepsy. The New England Journal of Medicine. 2001 Aug 2; 345:311–8. DOI: 10.1056/NEJM200108023450501.
4. Sheng J, Liu S, Qin H, Li B, and Zhang X. Drug-Resistant Epilepsy and Surgery. Current Neuropharmacology. 2018 Jan; 16:17–28. DOI: 10.2174/1570159X15666170504123316. PMCID: PMC5771378.
5. Sanders MW, Van der Wolf I, Jansen FE, et al. Outcome of Epilepsy Surgery in MRI-Negative Patients Without Histopathologic Abnormalities in the Resected Tissue. Neurology. 2024 Feb 27; 102:e208007. DOI: 10.1212/WNL.0000000000208007. PMCID: PMC10962914.
6. Hong SJ, Kim H, Schrader D, Bernasconi N, Bernhardt BC, and Bernasconi A. Automated Detection of Cortical Dysplasia Type II in MRI-negative Epilepsy. Neurology. 2014 Jul 1; 83:48–55. DOI: 10.1212/WNL.0000000000000543. PMCID: PMC4114179.
7. Sone D. Making the Invisible Visible: Advanced Neuroimaging Techniques in Focal Epilepsy. Frontiers in Neuroscience. 2021; 15:699176. DOI: 10.3389/fnins.2021.699176. PMCID: PMC8353251.
8. Ahmed B, Brodley CE, Blackmon KE, et al. Cortical Feature Analysis and Machine Learning Improves Detection of "MRI-negative" Focal Cortical Dysplasia. Epilepsy & behavior : E&B. 2015 Jul; 48:21–8. DOI: 10.1016/j.yebeh.2015.04.055. PMCID: PMC4500682.
9. Lee YJ. Advanced Neuroimaging Techniques for Evaluating Pediatric Epilepsy. Clinical and Experimental Pediatrics. 2020 Mar; 63:88–95. DOI: 10.3345/kjp.2019.00871. PMCID: PMC7073377.
10. Wang W, Huang Q, Zhou Q, et al. Multimodal Non-invasive Evaluation in MRI-negative Epilepsy Patients. Epilepsia Open. 2024 Jan 23; 9:765–75. DOI: 10.1002/epi4.12896. PMCID: PMC10984307.
11. Ma D, Gulani V, Seiberlich N, et al. Magnetic Resonance Fingerprinting. Nature. 2013 Mar; 495:187–92. DOI: 10.1038/nature11971. PMCID: PMC3602925.
12. Ding Z, Morris S, Hu S, et al. Automated Whole-Brain Focal Cortical Dysplasia Detection Using MR Fingerprinting With Deep Learning. Neurology. 2025 Jun 10; 104:e213691. DOI: 10.1212/WNL.0000000000213691. PMCID: PMC12089660.
13. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. IJCAI'95. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995 Aug 20:1137–43.
14. Bennett CM, Wolford GL, and Miller MB. The Principled Control of False Positives in Neuroimaging. Social Cognitive and Affective Neuroscience. 2009 Dec; 4:417–22. DOI: 10.1093/scan/nsp053. PMCID: PMC2799957.
15. Nichols T and Hayasaka S. Controlling the Familywise Error Rate in Functional Neuroimaging: A Comparative Review. Statistical Methods in Medical Research. 2003 Oct; 12:419–46. DOI: 10.1191/0962280203sm341ra.

16. Genovese CR, Lazar NA, and Nichols T. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. NeuroImage. 2002 Apr; 15:870–8. DOI: 10.1006/nimg.2001.1037.
17. Sharma H. Statistical Significance or Clinical Significance? A Researcher's Dilemma for Appropriate Interpretation of Research Results. Saudi Journal of Anaesthesia. 2021; 15:431–4. DOI: 10.4103/sja.sja_158_21. PMCID: PMC8477766.
18. Liao C, Wang K, Cao X, et al. Detection of Lesions in Mesial Temporal Lobe Epilepsy by Using MR Fingerprinting. Radiology. 2018 Sep; 288:804–12. DOI: 10.1148/radiol.2018172131.
19. Ganji Z, Hakak MA, Zamanpour SA, and Zare H. Automatic Detection of Focal Cortical Dysplasia Type II in MRI: Is the Application of Surface-Based Morphometry and Machine Learning Promising? Frontiers in Human Neuroscience. 2021; 15:608285. DOI: 10.3389/fnhum.2021.608285. PMCID: PMC7933541.
20. Ghassemi M, Oakden-Rayner L, and Beam AL. The False Hope of Current Approaches to Explainable Artificial Intelligence in Health Care. The Lancet Digital Health. 2021 Nov; 3:e745–e750. DOI: 10.1016/S2589-7500(21)00208-9.
21. Borys K, Schmitt YA, Nauta M, et al. Explainable AI in Medical Imaging: An Overview for Clinical Practitioners - Beyond Saliency-Based XAI Approaches. European Journal of Radiology. 2023 May; 162:110786. DOI: 10.1016/j.ejrad.2023.110786.
22. Yang Y, Xu S, Hong Y, et al. Computational Modeling for Medical Data: From Data Collection to Knowledge Discovery. The Innovation Life. 2024; 2:100079. DOI: 10.59717/j.xinn-life.2024.100079.
23. Collins GS, Dhiman P, Ma J, et al. Evaluation of Clinical Prediction Models (Part 1): From Development to External Validation. BMJ (Clinical research ed). 2024 Jan 8; 384:e074819. DOI: 10.1136/bmj-2023-074819. PMCID: PMC10772854.
24. Cabitza F, Campagner A, Soares F, et al. The Importance of Being External. Methodological Insights for the External Validation of Machine Learning Models in Medicine. Computer Methods and Programs in Biomedicine. 2021 Sep; 208:106288. DOI: 10.1016/j.cmpb.2021.106288.
25. Cross JL, Choma MA, and Onofrey JA. Bias in Medical AI: Implications for Clinical Decision-Making. PLOS digital health. 2024 Nov; 3:e0000651. DOI: 10.1371/journal.pdig.0000651. PMCID: PMC11542778.
26. Dhar T, Dey N, Borra S, and Sherratt RS. Challenges of Deep Learning in Medical Image Analysis—Improving Explainability and Trust. IEEE Transactions on Technology and Society. 2023 Mar; 4:68–75. DOI: 10.1109/TTS.2023.3234203.
27. Nayor M, Shen L, Hunninghake GM, et al. Progress and Research Priorities in Imaging Genomics for Heart and Lung Disease: Summary of an NHLBI Workshop. Circulation Cardiovascular Imaging. 2021 Aug; 14:e012943. DOI: 10.1161/CIRCIMAGING.121.012943. PMCID: PMC8486340.
28. Fan M, Song J, and Qiu Z. Editorial: Biomedical Image or Genomic Data Characterization and Radiogenomic/Image-Omics. Frontiers in Genetics. 2022 Aug 17; 13:994880. DOI: 10.3389/fgene.2022.994880. PMCID: PMC9433120.
29. Ma D, Jones SE, Deshmane A, et al. Development of High-Resolution 3D MR Fingerprinting for Detection and Characterization of Epileptic Lesions. Journal of magnetic resonance imaging: JMRI. 2019 May; 49:1333–46. DOI: 10.1002/jmri.26319.
30. McGonigal A, Bartolomei F, Régis J, et al. Stereoelectroencephalography in Presurgical Assessment of MRI-negative Epilepsy. Brain: A Journal of Neurology. 2007 Dec; 130(Pt 12):3169–83. DOI: 10.1093/brain/awm218.

31. Kini LG, Gee JC, and Litt B. Computational Analysis in Epilepsy Neuroimaging: A Survey of Features and Methods. NeuroImage Clinical. 2016; 11:515–29. DOI: 10.1016/j.nicl.2016.02.013. PMCID: PMC4833048.
32. Jirsa V, Wang H, Triebkorn P, et al. Personalised Virtual Brain Models in Epilepsy. The Lancet Neurology. 2023 May; 22:443–54. DOI: 10.1016/S1474-4422(23)00008-X.
33. Bernstam EV, Hersh WR, Johnson SB, et al. Synergies and Distinctions between Computational Disciplines in Biomedical Research: Perspective from the Clinical andTranslational Science Award Programs. Academic Medicine: Journal of the Association of American Medical Colleges. 2009 Jul; 84:964–70. DOI: 10.1097/ACM.0b013e3181a8144d. PMCID: PMC2884382.
34. Bian Y, Li J, Ye C, Jia X, and Yang Q. Artificial Intelligence in Medical Imaging: From Task-Specific Models to Large-Scale Foundation Models. Chinese Medical Journal. 2025 Mar 20; 138:651–63. DOI: 10.1097/CM9.0000000000003489. PMCID: PMC11925424.
35. Collin CB, Gebhardt T, Golebiewski M, et al. Computational Models for Clinical Applications in Personalized Medicine-Guidelines and Recommendations for Data Integration and Model Validation. Journal of Personalized Medicine. 2022 Jan 26; 12:166. DOI: 10.3390/jpm12020166. PMCID: PMC8879572.
36. Ding Z, Hu S, Su TY, et al. Combining Magnetic Resonance Fingerprinting with Voxel-Based Morphometric Analysis to Reduce False Positives for Focal Cortical Dysplasia Detection. Epilepsia. 2024 Jun; 65:1631–43. DOI: 10.1111/epi.17951. PMCID: PMC11166521.
37. Fuhrman JD, Gorre N, Hu Q, Li H, El Naqa I, and Giger ML. A Review of Explainable and Interpretable AI with Applications in COVID-19 Imaging. Medical Physics. 2022 Jan; 49:1–14. DOI: 10.1002/mp.15359. PMCID: PMC8646613.
38. Gallitto G, Englert R, Kincses B, et al. External Validation of Machine Learning Models—Registered Models and Adaptive Sample Splitting. GigaScience. 2025 Jan 6; 14:giaf036. DOI: 10.1093/gigascience/giaf036.
39. Guo Y, Li T, Gong B, et al. From Images to Genes: Radiogenomics Based on Artificial Intelligence to Achieve Non-Invasive Precision Medicine in Cancer Patients. Advanced Science. 2025 Jan; 12:2408069. DOI: 10.1002/advs.202408069.
40. Hsieh JJL and Svalbe I. Magnetic Resonance Fingerprinting: From Evolution to Clinical Applications. Journal of Medical Radiation Sciences. 2020 Dec; 67:333–44. DOI: 10.1002/jmrs.413. PMCID: PMC7754037.
41. Huo Q, Luo X, Xu ZC, and Yang XY. Machine Learning Applied to Epilepsy: Bibliometric and Visual Analysis from 2004 to 2023. Frontiers in Neurology. 2024; 15:1374443. DOI: 10.3389/fneur.2024.1374443. PMCID: PMC11018949.
42. Jin B, Krishnan B, Adler S, et al. Automated Detection of Focal Cortical Dysplasia Type II with Surface-Based Magnetic Resonance Imaging Postprocessing and Machine Learning. Epilepsia. 2018 May; 59:982–92. DOI: 10.1111/epi.14064. PMCID: PMC5934310.
43. Lee HM, Gill RS, Bernasconi N, and Bernasconi A. Machine Learning in Neuroimaging of Epilepsy. *Machine Learning for Brain Disorders*. Ed. by Colliot O. New York, NY: Humana, 2023. Available from: http://www.ncbi.nlm.nih.gov/books/NBK597482/.
44. Mastoi QUA, Latif S, Brohi S, et al. Explainable AI in Medical Imaging: An Interpretable and Collaborative Federated Learning Model for Brain Tumor Classification. Frontiers in Oncology. 2025; 15:1535478. DOI: 10.3389/fonc.2025.1535478. PMCID: PMC11903279.
45. Monga A, Singh D, de Moura HL, Zhang X, Zibetti MVW, and Regatte RR. Emerging Trends in Magnetic Resonance Fingerprinting for Quantitative Biomedical Imaging Applications: A Review. Bioengineering (Basel, Switzerland). 2024 Feb 28; 11:236. DOI: 10.3390/bioengineering11030236. PMCID: PMC10968015.

46. Panda A, Mehta BB, Coppo S, et al. Magnetic Resonance Fingerprinting-An Overview. Current opinion in biomedical engineering. 2017 Sep; 3:56–66. DOI: 10.1016/j.cobme.2017.11.001. PMCID: PMC5984038.
47. Piffer S, Ubaldi L, Tangaro S, Retico A, and Talamonti C. Tackling the Small Data Problem in Medical Image Classification with Artificial Intelligence: A Systematic Review. Progress in Biomedical Engineering (Bristol, England). 2024 Jun 17; 6. DOI: 10.1088/2516-1091/ad525b.
48. Riley RD, Ensor J, Snell KIE, et al. Importance of Sample Size on the Quality and Utility of AI-based Prediction Models for Healthcare. The Lancet Digital Health. 2025 Jun 2 :100857. DOI: 10.1016/j.landig.2025.01.013.
49. Saroha S. Artificial Intelligence in Medical Education: Promise, Pitfalls, and Practical Pathways. Advances in Medical Education and Practice. 2025 Jun 14; 16:1039–46. DOI: 10.2147/AMEP.S523255. PMCID: PMC12176979.
50. Su TY, Tang Y, Choi JY, et al. Evaluating Whole-brain Tissue-property Changes in MRI-negative Pharmacoresistant Focal Epilepsies Using MR Fingerprinting. Epilepsia. 2023 Feb; 64:430–42. DOI: 10.1111/epi.17488. PMCID: PMC10107443.
51. Van Hees SV, Schilder MB, Keyser A, Sbrizzi A, Kleinloog JP, and Boon WP. Exploring Scenarios for Implementing Fast Quantitative MRI. European Journal of Radiology Open. 2025 May 8; 14:100658. DOI: 10.1016/j.ejro.2025.100658. PMCID: PMC12135375.
52. Van Lanen RHGJ, Colon AJ, Wiggins CJ, et al. Ultra-High Field Magnetic Resonance Imaging in Human Epilepsy: A Systematic Review. NeuroImage Clinical. 2021; 30:102602. DOI: 10.1016/j.nicl.2021.102602. PMCID: PMC7921009.
53. Wang HE, Woodman M, Triebkorn P, et al. Delineating Epileptogenic Networks Using Brain Imaging Data and Personalized Modeling in Drug-Resistant Epilepsy. Science Translational Medicine. 2023 Jan 25; 15:eabp8982. DOI: 10.1126/scitranslmed.abp8982.
54. Wiebe S. Comment: Has Automated Detection of Cortical Dysplasias Come of Age? Neurology. 2014 Jul 1; 83:54. DOI: 10.1212/WNL.0000000000000556.