Clinical Neurophysiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Review

IFCN position statement: use of artificial intelligence in clinical neurophysiology

Aatif M. Husain [a,*], Sandor Beniczky [b], Meriem Bensalem-Owen [c], Jeremy Bland [d], Benjamin H. Brinkman [e], Ritsuko Hanajima [f], Michal Holub [g], Lara Jehi [h], Sanjeev Nandedkar [i], Nortina Shahrizaila [j], Hatice Tankisi [k], Martijn R. Tannemaat [l], M.Brandon Westover [m], Walter Paulus [n]

[a] Department of Neurology, Duke University Medical Center, Durham, NC, USA
[b] Department of Clinical Neurophysiology, Danish Epilepsy Centre Filadelfia, Dianalund and Department of Clinical Medicine/Clinical Neurophysiology, University of Copenhagen & Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark
[c] Department of Neurology, University of Kentucky, Lexington, KY, USA
[d] Department of Clinical Neurophysiology, East Kent Hospitals University NHS Foundation Trust, Kent & Canterbury Hospital, Canterbury, UK
[e] Departments of Neurology, Physiology, and Biomedical Engineering, Mayo Foundation, Rochester, MN, USA
[f] Division of Neurology, Department of Brain and Neurosciences, Faculty of Medicine, Tottori University, Yonago, Japan
[g] Cadwell Industries, Inc., Kennewick, WA, USA
[h] Department of Neurology, Cleveland Clinic and Center for Computational Life Sciences, Lerner Research Institute, Cleveland, OH, USA
[i] Natus Medical Inc., Middleton, Medical College of Wisconsin, Milwaukee, WI, USA
[j] Division of Neurology, Faculty of Medicine, University of Malaya & University Malaya Medical Centre, Kuala Lumpur, Malaysia
[k] Department of Clinical Neurophysiology, Aarhus University Hospital and Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
[l] Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands
[m] Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[n] Department of Neurology, Ludwig Maximilians University, Munich, Germany

ARTICLE INFO

Keywords:
Artificial intelligence
CNP
EEG
EMG
Position statement

ABSTRACT

In recent years, artificial intelligence (AI) has made significant strides, gaining traction across various domains, including clinical medicine. The integration of AI as a decision-support tool introduces complexities, particularly in ensuring patient safety and upholding clinical accountability. This is especially pertinent in clinical neurophysiology (CNP), where AI shows promise in enhancing the interpretation of neurophysiologic data from modalities such as electroencephalography, electromyography, and others. Recognizing the potential and inherent challenges of AI integration, the International Federation of Clinical Neurophysiology (IFCN) has set forth guidelines to steer the responsible development, evaluation, and application of AI technologies in CNP.

The IFCN's core position on AI in CNP emphasizes improving healthcare outcomes, prioritizing patient-centered care, maintaining transparency in AI-generated interpretations, and supporting, rather than replacing, clinical expertise. As AI technology evolves, the IFCN stresses that models implemented in clinical practice, especially when lacking supervision by experts must meet stringent standards for accuracy, safety, and quality control, and that AI cannot substitute expert care without adequate oversight.

The successful integration of AI in CNP hinges on dataset diversity, transparent and ethical training practices, and balancing model complexity with real-world validation. Continuous performance monitoring and clinician feedback are vital to maintaining AI system reliability over time.

The IFCN envisions a future where AI ethically enhances CNP practice through multidisciplinary collaboration, focusing on patient safety, clinical integrity, and scientific rigor. Properly implemented, AI can revolutionize CNP by offering real-time decision support, personalized interventions, and remote monitoring, thus advancing the quality of care in CNP.

* Corresponding author at: Department of Neurology, Duke University Medical Center, Box 2720, 2424 Erwin Road, Suite 1001, Durham, NC 27710, USA.
E-mail address: Aatif.husain@duke.edu (A.M. Husain).

https://doi.org/10.1016/j.clinph.2026.2111893
Accepted 14 April 2026
Available online 16 April 2026
1388-2457/Published by Elsevier B.V. on behalf of International Federation of Clinical Neurophysiology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

In recent years, rapid advances in computational methods broadly categorized as artificial intelligence (AI) have led to the widespread adoption of systems capable of performing tasks traditionally reserved for human expertise. These technologies are increasingly embedded in daily life and across a range of professional domains, including clinical medicine. Applications of AI are already being deployed at scale, often without full consideration of their broader clinical, ethical, and legal implications.

In medicine, the integration of AI as a decision-support tool introduces new complexities. Overreliance on these systems without sufficient oversight risks compromising patient safety and raises pressing ethical concerns about clinical accountability (Ethics and Governance of Artificial Intelligence for Health, 2021; Series and Academy, 2025). These challenges are particularly relevant to the field of clinical neurophysiology (CNP), where AI holds clear promise for enhancing signal and image interpretation, including in modalities such as electroencephalography (EEG), electromyography (EMG) and nerve conduction studies (NCS), evoked potentials (EPs), polysomnography (PSG), neurophysiologic intraoperative monitoring (NIOM), neuromuscular ultrasound, and magnetoencephalography. Beyond data analysis, AI may also support triaging referrals, examination planning, creation of differential diagnoses by integrating neurophysiological data with other clinical information, and generation of CNP test reports.

In CNP, several AI applications are already integrated into routine practice. Some commercial EEG systems incorporate AI-based assessments, and PSG platforms routinely employ automated sleep staging and respiratory event detection algorithms. In contrast, other AI developments, such as predictive models for epilepsy surgery outcomes or AI-based assessment of sleep architecture to assist in identifying patterns suggestive of specific sleep disorders, remain largely investigational and have not yet achieved widespread clinical validation or regulatory integration (Jehi, 2023; Stephansen et al., 2018).

The International Federation of Clinical Neurophysiology (IFCN) acknowledges the urgent need for principles to guide the integration of AI technologies into CNP in a manner that prioritizes patient safety, enhances diagnostic accuracy, and supports professional practice. These principles must clearly define clinical responsibility and ensure that patient welfare remains paramount. Given the pace at which AI is advancing and its growing use in bedside decision-making, it is imperative for the CNP community to articulate shared principles for the safe development, evaluation, and integration of AI in routine clinical care.

To this end, the IFCN convened a multidisciplinary taskforce comprising experts in AI development, neurophysiologists with expertise in biological signal interpretation, and representatives from the medical technology industry. This group was tasked with drafting a position statement to guide the responsible development, evaluation, and implementation of AI tools in CNP. The core principles were developed through a structured expert consensus process that included an initial in-person meeting to define key themes and priorities, followed by a series of virtual meetings and iterative draft revisions. Draft documents were circulated among taskforce members for comment and refinement until consensus was reached. The final version was reviewed and formally approved by the IFCN Executive Committee.

2. Background

AI systems transform large volumes of neurophysiologic data into probabilities or recommendations useful to clinicians. At one end of the spectrum are explicitly programmed, rule-based systems such as manually crafted decision trees; every conclusion can be traced to an "if–then" rule, making these systems transparent and pedagogically valuable (Husain, 2025). More flexible, classical machine-learning models—k-nearest-neighbor classifiers, Bayesian networks, shallow decision trees and related methods—learn patterns from labelled or unlabeled data instead of relying on pre-written rules. Their performance scales with the amount and quality of curated training examples, which in CNP means extensive expert annotations (Nilsson, 2010). At the more complex end lie deep-learning models built from multilayer neural networks. These networks ingest raw waveforms or images and automatically extract hierarchical features, often discovering clinically relevant patterns beyond human perception. Their power, however, often comes with reduced interpretability (Burkart and Huber, 2021).

Bringing an AI model into clinical practice is a cyclic, multidisciplinary endeavor (Fig. 1). It begins with rigorous data collection and ground-truth definition, where high-quality recordings, metadata and expert labels are assembled. Developers then train candidate models, perform internal validation, and document the intended use and technical specifications. Before clinical deployment, the model must undergo external benchmarking against human experts, assessment of safety, bias, and generalizability, and—where applicable—regulatory scrutiny under frameworks such as the Food and Drug Administration's (FDA) Software as a Medical Device (SaMD) guidance. Once an approved system enters daily workflow, real-world performance, and clinician–AI disagreements must be monitored, audited, and fed back to developers under a formal change-management plan, allowing the algorithm to evolve while preserving safety and traceability. Only by attending to every stage of this lifecycle, and by balancing algorithmic power with transparency, can AI augment rather than obscure clinical judgement and improve neurophysiological care.

3. Core position

The IFCN recognizes the transformative potential of AI in CNP and its broader implications for healthcare. Within this field, AI is expected to significantly impact clinical practice, research, and education. AI-assisted tools are already being incorporated into clinical workflows, including automated EEG pattern recognition, EMG analysis, PSG scoring, and structured report generation. Advanced AI-driven analyses of these data are enabling the identification of patterns and biomarkers not readily discernible through conventional visual inspection. As these technologies become more integrated into clinical workflows, they are likely to influence training paradigms and competency expectations for the next generation of clinical neurophysiologists (Husain, 2025).

The core stance of the IFCN on the application of AI is as follows:

- AI should serve to improve healthcare outcomes. Its primary purpose in CNP must be to enhance patient care, in alignment with the IFCN's vision: "To improve healthcare worldwide by understanding the nervous system and optimizing the diagnosis and treatment of its disorders through CNP."
- Integration of AI must prioritize patients. The implementation of AI must consider the needs and perspectives of all stakeholders, most importantly patients, by ensuring equitable access to AI-enabled

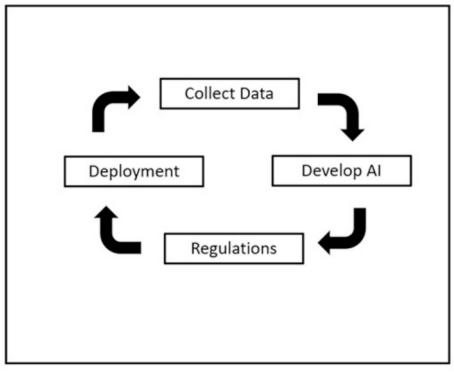


Fig. 1. The cyclic nature of developing an artificial intelligence system.

diagnostics and maintaining a strong focus on patient safety and well-being.

- AI-generated interpretations must be clearly labeled. Transparency is essential. All AI-assisted outputs should be explicitly identified as such to promote clear communication among clinicians, patients, and payors, thereby fostering trust and accountability.
- AI should support—not replace—clinical expertise. In its current state, AI is best used to augment the expertise of clinical neurophysiologists or physician familiar with the interpretation of CNP tests by improving diagnostic accuracy, efficiency, and accessibility. Autonomous interpretation of neurophysiological studies by AI without supervision from a clinical neurophysiologist or physician qualified to interpret CNP tests is not currently recommended.
- Future autonomous AI must meet the highest standards. As AI systems evolve, some applications may achieve sufficient performance for unsupervised use. However, these systems must adhere to the most rigorous standards for accuracy, safety, transparency, and quality control.
- AI is not a substitute for expert care in resource-limited settings. The lack of access to clinical neurophysiologists in certain regions does not justify the use of unsupervised AI-based care, and remote supervision of AI should be considered in these instances. High standards for AI applications must be upheld globally.
- AI research must follow evidence-based standards. Studies reporting on the development and validation of AI applications must conform to rigorous scientific standards and follow Standards for Reporting of Diagnostic Accuracy (STARD) and other similar guidelines for reporting diagnostic accuracy.

4. Training AI models

Training AI models in CNP requires creativity, innovation, and multidisciplinary expertise. The IFCN does not endorse specific regulatory requirements for training practices. Given the diversity of effective training approaches, the IFCN advocates for models to be judged primarily by their demonstrated performance in robust and comprehensive validation processes. The key elements of training AI models are outlined below.

4.1. Dataset diversity

- General training datasets are most effective when they include broad demographic, clinical, and international diversity, accounting for variability in age, sex, ethnicity, and disease type and severity.
- Representative datasets also reflect differences in data acquisition practices, such as the skill levels of recording personnel, environmental noise, and clinical contexts (e.g., EEGs in outpatients, epilepsy monitoring units, intensive care units).

4.2. Specialized datasets

- For models targeting specific clinical scenarios, performance and generalizability are enhanced when the training data captures maximal variability within the defined clinical context.
- Clearly defining the intended scope of use for specialized models is essential to prevent inappropriate application by end users.

4.3. Transparency and ethical considerations

- Transparent documentation of training methods and dataset characteristics supports critical evaluation of biases and limitations, fostering trust and appropriate clinical use.
- Ethical training practices should include proactive efforts to ensure representation from historically disadvantaged or underrepresented patient populations.

4.4. Balancing data Quality, dataset Size, and model complexity

- The IFCN recognizes that effective AI performance may result from diverse training strategies. While some models benefit from large, heterogeneous datasets, others may rely on sophisticated techniques to compensate for limited data.
- There is no one-size-fits-all approach; thus, training practices should remain flexible. The ultimate measure of success is strong validation performance, particularly under challenging or "corner-case" conditions, such as rare pathologies, atypical patient populations, unusual signal artifacts, extreme age groups, or recordings obtained under nonstandard technical conditions, where model performance may degrade.

4.5. Practices likely to enhance model robustness

- Balancing training datasets across key demographic and clinical variables reduces the risk of bias toward specific subgroups.
- Strict separation of training and testing datasets is critical to avoid biased performance estimates. Training procedures must exclude test data from all model development processes, including feature selection and hyperparameter tuning.
- Certain data augmentation techniques, such as flipping EEG channels left-to-right, have been shown to enhance model robustness. However, introducing artificial noise that lacks physiological realism may not improve model generalizability.
- Models trained to exceed individual human performance, by targeting objective clinical outcomes or consensus-based annotations, tend to achieve higher clinical utility. For instance, AI models diagnosing carpal tunnel syndrome may be more effective when trained using NCS results or majority expert consensus rather than individual assessments alone.

4.6. International collaboration

- The IFCN strongly encourages international, multicenter collaborations to develop and share large, standardized, and representative datasets with expert consensus annotations. These collaborative efforts are essential to improving the robustness, generalizability, and clinical relevance of AI models in CNP.

5. Validating AI models

Safely integrating AI into clinical practice requires rigorous, structured validation that ensures not only technical accuracy, but also generalizability, robustness, and interpretability. Inadequate validation can lead to misdiagnosis, clinician mistrust, and diminished credibility of AI tools, ultimately harming patient care. To mitigate these risks, validation must be systematic and comprehensive, with meaningful collaboration among algorithm developers, clinicians, regulatory bodies, and patients. The key elements of rigorous validation are outlined below. Whenever possible, standard reporting guidelines should be followed when publishing validation studies (Ewen et al., 2025).

5.1. Representative data

- As discussed above, representative datasets that reflect the complexity and diversity of real-world CNP practice are essential before validation begins.
- Using representative datasets helps ensure fairness, uncovers algorithmic limitations, and improves generalizability across varied patient populations.

5.2. Structured validation framework

- Validation should progress in a structured, hierarchical manner from internal testing to progressively broader external assessments (de Hond et al., 2023)(Fig. 2):
- Internal validation confirms consistent, reproducible performance within the development dataset using methods such as internal cross-validation, including nested cross-validation for model selection and hyperparameter tuning. This step is critical for algorithm development but insufficient alone for clinical deployment.
- External validation demonstrates generalization beyond the training data and must be blinded to the reference standard (He and Su, 2023). The model architecture and parameters should be pre-specified before testing. Subtypes include:
  - o Holdout Validation: Uses data from patients independent of model development.
  - o Prospective Validation: Assesses model stability over time within the same institution, considering shifts in clinical practice, populations, and equipment.
  - o Geographical Validation: Tests performance across different institutions, systems, equipment types, or regions, ideally through multicenter or leave-one-site-out studies.
  - o Domain Validation: Ensures reliability in different clinical domains (e.g., EEGs performed in pediatric versus adult patients, epilepsy monitoring units (EMU) versus intensive care units (ICU)). Clear documentation of scope is essential to prevent inappropriate use in an unvalidated domain.
  - o Edge Case Validation: Validates models on rare or anomalous clinical scenarios using stress testing or perturbation analysis to ensure robustness in real-world use.

5.3. Reference standard

- The choice of reference standard (or "gold standard") must be explicitly justified and aligned with established best practices for the model's context of use.

5.4. Performance metrics

- Validation metrics must reflect real-world clinical objectives rather than technical accuracy alone. Overreliance on common metrics like accuracy or area under the receiver operating characteristic curve can be misleading, particularly in imbalanced datasets.
- Multiple clinically meaningful metrics should be assessed. These can be grouped into two categories: global metrics, which evaluate the model's probability outputs (such as the area under the receiver operating characteristic curve, area under the precision-recall curve, sensitivity at a given false positive rate, and false detection rate at a given sensitivity level); and threshold-based metrics, which are calculated after applying a specific cutoff to the model's probability output to make a binary (yes/no) decision (including sensitivity, specificity [1 − false positive rate], F1 score, diagnostic odds ratio, and positive and negative predictive values).Criteria for acceptable performance must be clearly enumerated prior to validation, and these criteria must be related to the clinical requirements for the algorithm's intended use case.

5.5. Interpretability vs. Explainability

- Clinicians trust hinges on understanding how AI models generate results. Inherently interpretable models, such as decision trees, sparse linear models, or "this-looks-like-that" frameworks, are preferred for high-stakes CNP applications (Rudin, 2019). This facilitates identifying model biases, logic verification, and trouble-shooting in clinical settings.
- Post-hoc "explanations" of black-box models are often incomplete and may be misleading, representing only approximations of the model's true reasoning. This inherent limitation can lead to false confidence in model outputs, potentially obscuring biases or errors.
- Inherently interpretable models align more closely with IFCN's principles of safety, transparency, and accountability.

5.6. Robustness under Real-World conditions

- AI tools must perform reliably even in suboptimal conditions, such as noisy EEG recordings, missing clinical data, or unexpected patient presentations.
- Systematic robustness testing, including stress testing and perturbation analyses, helps identify vulnerabilities before clinical deployment, ensuring dependable bedside performance (Chen et al., 2024; Erdemir et al., 2021).

5.7. Domain-Specific validation in CNP modalities

- Each CNP modality introduces unique challenges that require tailored validation strategies.
- EEG**:** Models must distinguish normal from abnormal recordings and classify abnormalities into the major categories (epileptiform versus non-epileptiform, focal versus generalized), across varied contexts

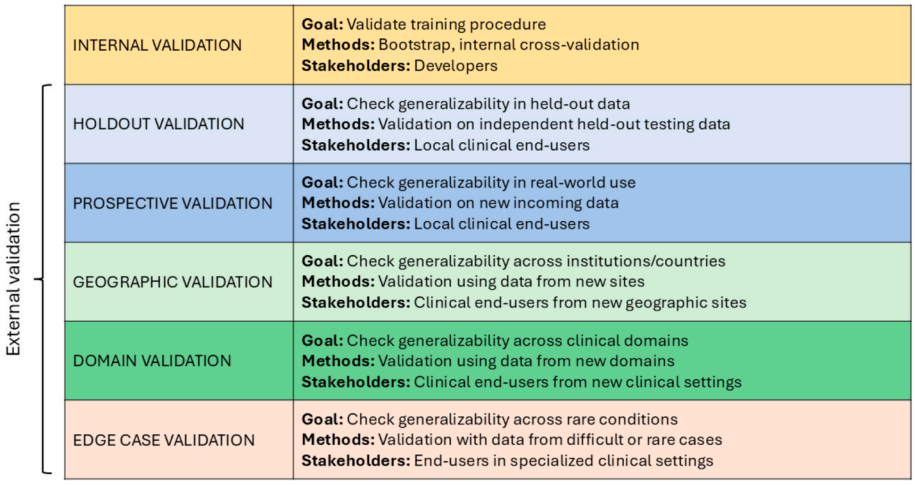


Fig. 2. Structured framework for validation of AI algorithms.

(routine labs, EMUs, ICUs, and point-of-care devices with limited channels.
- EMG/NCS**:** Requires validation in real-time clinical workflows to ensure usability and relevance.
- EP, NIOM, and Neuromuscular Ultrasound**:** Require validation protocols specific to the specialized equipment and clinical scenarios in which they are used.
- These tailored approaches ensure smooth clinical integration, promote user confidence, and reduce disruption to workflow.

5.8. Clinical trials and regulatory Compliance

- Structured clinical trials and regulatory approvals (e.g., FDA, European Medical Device Regulation (CE-mark)) provide vital assurances of safety, efficacy, and clinical utility. These processes build trust among clinicians, patients, payors, and regulators, and are essential before broad clinical adoption.

5.9. Continuous monitoring and re-validation

- Validation must be ongoing rather than a one-time event, given constant changes in clinical practices, technology, and patient demographics.
- Continuous performance monitoring, clinician feedback, periodic re-validation, and retraining with evolving datasets are necessary to ensure long-term reliability and relevance.

5.10. Ethical considerations

- Ethical validation requires fairness and transparency, minimizing biases by using diverse patient datasets for training and validation.
- Transparent reporting of validation methods, datasets, and outcomes upholds IFCN values, ensuring equitable and safe patient care.

6. Benefits of AI in CNP

AI has demonstrated significant potential to enhance diagnostic accuracy, efficiency, and accessibility in multiple domains of CNP.

6.1. Epilepsy and EEG

In epilepsy, AI algorithms can analyze EEG data in real time to detect seizures with high accuracy (Knight et al., 2024). As a foundational step, AI has been used successfully to remove artifacts, filter noise, and identify clean signal segments to optimize performance. A core strength of AI in EEG interpretation lies in pattern recognition, such as differentiating epileptiform activity, localizing abnormalities, and diagnosing encephalopathies. These tasks are often complex, even for epilepsy specialists, yet AI systems have shown performance comparable to expert-level interpretation (Tveit et al., 2023).

6.2. Neuromuscular disorders and EMG

In the neuromuscular domain, ML models can assist in identifying neuropathic or myopathic patterns, reduce inter-rater variability in motor unit potential analysis, and enhance quantitative EMG interpretation (Tannemaat et al., 2023). AI-driven EMG and NCS platforms can support diagnostic decision-making and even help predict prognoses (Martinez-Thompson et al., 2025). By leveraging multichannel recordings, AI can detect correlations and subtle anomalies that might be overlooked by human interpreters. Additionally, AI-powered systems can streamline routine clinical tasks, such as report generation and data entry, improving workflow efficiency and accelerating patient care, particularly in emergency or high-volume settings.

6.3. Sleep disorders and PSG

AI has also been applied to PSG for the diagnosis of sleep disorders. Automated classification of sleep stages continues to improve diagnostic accuracy for conditions such as sleep apnea and narcolepsy, while substantially reducing the time required for manual scoring. This enhances clinical efficiency and throughput. However, currently there is a lack of AI models that can be used in detecting common sleep disorders like insomnia, circadian rhythm disorders, restless legs syndrome and others (Alattar et al., 2024). An emerging area involves AI-driven sleep analysis via smartphones, although diagnostic value is currently limited without concurrent EEG data (Browne et al., 2024).

6.4. Other Aspects of CNP

One of the most promising frontiers for AI in CNP is its role in brain-computer interfaces (BCIs) and neuroprosthetics. These applications utilize neural signals from EEG, electrocorticography (ECoG), or intracortical recordings to control assistive devices. BCIs may enable not only motor control of robotic limbs but also translation of brain signals into speech (Derosiere et al., 2025; Duraivel et al., 2023; Rahman et al., 2024). In addition, BCIs are advancing neuromodulation strategies, including closed-loop neuromodulation, which monitors real-time brain activity to dynamically adjust stimulation parameters for improved therapeutic outcomes (Zrenner and Ziemann, 2024).

Non-invasive and wearable AI-integrated devices are expanding access to neurophysiologic monitoring and personalized treatment. These wearables are increasingly used for seizure detection, sleep evaluation, remote monitoring, and even home-based interventions such as transcranial direct current stimulation (Antal et al., 2022). By offering continuous, real-time data collection and feedback, these tools hold promise for long-term condition management outside traditional clinical environments.

7. Challenges and safety concerns

While the potential value of AI in clinical neurophysiology (CNP) is widely recognized, its successful integration into clinical practice requires addressing several challenges, including regulatory considerations, safety, and system-level barriers.

7.1. Regulatory, ethical, and legal considerations

The integration of AI-assisted technologies into CNP must align with existing regulatory requirements (e.g., FDA, EMA) and data protection laws (e.g., European General Data Protection Regulation). The first step in regulatory approval is specifying the AI tool's clinical purpose, functionality, and target population (Au Yeung et al., 2023). Ethical considerations, including informed patient consent, data security, and the role of AI in clinical decision-making, must also be addressed.

One area of ongoing ambiguity is medical liability (Cestonaro et al., 2023). The degree of legal responsibility often depends on the autonomy of the AI system. In cases involving fully autonomous AI, liability may fall under vicarious liability principles, where institutions or employers could be held accountable. Conversely, with AI-assisted systems, liability may rest with clinicians under the “learned intermediary” doctrine, as they are responsible for reviewing and validating AI-generated outputs before acting upon them.

To foster trust and promote safe implementation, collaboration between all stakeholders, patients, healthcare providers, AI developers, and regulators, should begin early in the development process and continue through deployment. Such collaboration ensures patient autonomy is respected, robust cybersecurity safeguards are established, and consensus is achieved around accountability and liability in clinical use (Gerke et al., 2020; Mesko, 2017).

7.2. Data quality and potential for bias

As discussed above, developing effective AI tools in CNP requires extensive training and validation datasets. The quality and representativeness of these datasets remain critical concerns. Bias may be introduced when datasets fail to reflect the diversity of local patient populations, clinical practices, or diagnostic protocols, an issue particularly pressing in resource-limited settings (Adamson and Smith, 2018). Ways to mitigate these risks have been discussed already. Additionally, active engagement with the global neurophysiology community to define practice standards across various populations is necessary.

7.3. Interoperability and transparency

Healthcare providers are the primary end-users of AI in CNP, and their ability to implement and trust AI tools depends heavily on integration and usability. One key challenge is the lack of standardized data formats and system interoperability. Existing health IT infrastructure may require substantial upgrades to accommodate AI tools, posing logistical and financial challenges, particularly in under-resourced settings. Early involvement of IT, clinical, and administrative stakeholders can help identify infrastructure needs and feasible solutions.

8. Implementation strategy

The integration of AI in CNP practice marks a significant shift in how neurophysiological data are analyzed, interpreted, and applied to patient care. While the potential advantages are clear, the success of AI in clinical environments depends heavily on thoughtful implementation, ongoing validation, and sustained investment in clinician training and infrastructure.

8.1. Training, clinical judgment, and the human element

CNP has traditionally relied on expert visual interpretation and experiential learning. AI introduces opportunities for faster analysis and collaboration, enabling more timely diagnoses and greater operational efficiency. However, to realize these benefits, clinicians and trainees must fully understand AI's capabilities and limitations. It is critical to balance the use of AI with continued emphasis on hands-on training (Au Yeung et al., 2023).

While AI can assist in pattern recognition and data analysis, it cannot replace the practical experience needed to interpret subtle findings or navigate complex clinical scenarios. Clinical judgment remains fundamental. As AI becomes increasingly embedded in practice, there may be pressure to reduce training time or adjust educational standards (Daniello and Weber, 2018; Haneef et al., 2017; Sámano et al., 2018). This must be resisted. Robust clinical training is essential to ensure the next generation of clinical neurophysiologists can independently validate AI outputs and manage cases where AI fails to provide adequate answers. Maintaining high training standards is not optional; it is a safeguard for clinical quality and patient safety (He et al., 2019; Merhi, 2023).

8.2. Ongoing quality control and model validation

AI algorithms must undergo continuous quality assurance and performance monitoring, particularly when retrained with new data. Even minor changes to training datasets or model architecture can introduce bias or class imbalances that degrade performance (Gichoya et al., 2023). Validation is especially important when acquisition hardware changes or software updates are introduced (Feng et al., 2022).

Routine, periodic validation testing ensures that AI tools maintain clinically acceptable accuracy and reliability throughout their lifecycle (Myllyaho et al., 2021; Siebert et al., 2022; Studer et al., 2021). This includes confirming that performance metrics remain stable, clinically relevant, and aligned with the tool's intended use. Without this ongoing scrutiny, AI tools risk drifting away from their original validated performance, potentially compromising patient outcomes.

8.3. Demonstrating value to stakeholders

AI has the potential to reduce costs and streamline workflows, but for widespread adoption, its value must be clearly demonstrated. Health systems and payors must see evidence that AI improves outcomes, facilitates earlier diagnosis, or meaningfully reduces clinician burden. Transparent documentation of when and how AI is used in the diagnostic, interpretive, or reporting process is essential for informed decision-making and accountability.

Clear labeling of AI contributions in clinical workflows helps clinicians, patients, and insurers understand the role AI plays and builds trust in the system. As AI adoption grows, this transparency will be vital to gaining regulatory approval and reimbursement in diverse healthcare systems.

8.4. Reimbursement and regulatory integration

Several countries have begun to incorporate AI-based tools into healthcare regulatory and reimbursement frameworks (Intelligence and Learning, 2025; U.S. Department of Health and Human Services: Strategic Plan for the Use of Artificial Intelligence in Health, Human Services, and Public Health, U.S. 2025; van Kolfschooten and van Oirschot, 2024; Muehlematter et al., 2021). However, widespread clinical adoption still lags in many specialties.

8.5. Collaboration with industry

In the United States adoption of medical AI tools in most use cases has been concentrated in imaging-related areas such as coronary artery disease, diabetic retinopathy, atherosclerosis, liver MRI (Wu et al., 2024). This concentration is largely due to two factors. First, AI-enabled imaging devices allow seamless, trustworthy integration, where data acquisition, analysis, and output all occur within the same system. Second, integrating AI into medical devices at a level acceptable to regulators requires extensive resources, which are typically only available to commercial manufacturers. These investments are justified in high-burden public health areas like diabetes and cardiovascular disease but are more difficult to support for lower-volume fields.

The effective implementation of AI in CNP must involve close partnership with industry. Academic researchers alone may lack the capacity to bring AI tools through the full regulatory pipeline or to embed them into commercial-grade devices. By working with medical device manufacturers, developers can achieve the level of integration and quality assurance needed to satisfy regulatory standards and deliver clinical impact at scale.

Successful deployment of AI in CNP will depend on thoughtful implementation strategies that include strong clinical oversight, robust validation processes, transparent communication with stakeholders, and strategic collaboration with industry partners.

9. Conclusion

The integration of AI into CNP is a transformative development with the potential to improve diagnostic accuracy, increase efficiency, and expand access to specialized care. As this transformation unfolds, the IFCN emphasizes that the successful adoption of AI must be grounded in patient safety, clinical integrity, and scientific rigor. Robust training and validation processes, appropriately representative data, ongoing oversight, and clinician-centered implementation are critical to ensuring that AI complements, rather than compromises, the quality of care.

AI may enable novel developments in CNP including real-time decision support, personalized therapeutic interventions, brain-computer

interfaces, and remote monitoring. Realizing this future will require continued collaboration among clinicians, researchers, industry partners, and regulatory bodies. The IFCN envisions a future where AI is ethically integrated, globally equitable, and continually refined through multidisciplinary collaboration. With robust oversight, shared standards, and a focus on patient-centered outcomes, AI can help advance CNP into a new era, where technology and human judgment work in partnership to deliver high-quality care.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Aatif M. Husain: Consulting for PranaQ; treasurer, IFCN, member, American Academy of Sleep Medicine AI Committee (uncompensated).

Meriem Bensalem-Owen: Grants from Xenon, UCB, Neuropace; member of Council, American Clinical Neurophysiology Society (uncompensated); President of Board of Directors, Epilepsy Foundation of Kentuckiana (uncompensated).

Jeremy Bland: Consulting fees from Phillips Medical Ultrasound; honoraria from King's College London for teaching/lectures.

Benjamin H. Brinkman: Grants from National Institutes of Health, UNEEG Medical, Epilepsy Foundation of America; Royalties from Cadence Neuroscience; honoraria from Eisai Inc; support for attending meetings from Freiburg University; patents for brain impedance, multiscale electrodes, wearable devices; co-chair ILAE Neurotechnology section (uncompensated).

Michal Holub: Employee, Cadwell Industries, Inc.

Lara Jehi: Provisional patent for "predictor of seizure outcome after epilepsy surgery using peri-ictal scalp EEG data; steering committee member of the AI Alliance.

Sanjeev Nandedkar: Employee, Natus Medical, Inc.

Hatice Tankisi: Member, IFCN Executive Committee, President of Europe-Middle East-Africa Chapter of IFCN, European Chapter of Neurology, Teaching Course subcommittee member.

Martijn R. Tannemaat: Grants from ArgenX, NOW, ZonMW, Huma Therapeutics, Cadwell; consulting fees from UCB Pharma, ArgenX, Johnson & Johnson; honorarium from UCB Rare Diseases Connect in Neurology; support for meeting attendance by UCB Rare Disease Connect in Neurology; scientific advisor Spierziekten Nederland (patient advocacy organisation.

M. Brandon Westover: Grants from NIH; co-founder, scientific advisor, consultant to, and with personal equity in Beacon Biosignals.

Sandor Beniczky, Ritsuko Hanajima, Nortina Shahrizaila, Walter Paulus: No relevant conflicts reported.

Acknowledgement

This position statement was developed by the AI Workgroup appointed by the IFCN Executive Committee. The workgroup comprised CNP professionals from around the world representing clinical care, research, and industry.

References

Adamson, A.S., Smith, A., 2018. Machine learning and health care disparities in dermatology. JAMA Dermatol. 154, 1247. https://doi.org/10.1001/jamadermatol.2018.2348.

Alattar, M., Govind, A., Mainali, S., 2024. Artificial intelligence models for the automation of standard diagnostics in sleep medicine—a systematic review. Bioengineering 11, 206. https://doi.org/10.3390/bioengineering11030206.

Antal, A., Luber, B., Brem, A.-K., Bikson, M., Brunoni, A.R., Cohen Kadosh, R., et al., 2022. Non-invasive brain stimulation and neuroenhancement. Clin. Neurophysiol. Pract. 7, 146–165. https://doi.org/10.1016/j.cnp.2022.05.002.

Artificial Intelligence and Machine Learning in Software as a Medical Device. 2025.

Au Yeung, J., Wang, Y.Y., Kraljevic, Z., Teo, J.T.H., 2023. Artificial intelligence (AI) for neurologists: do digital neurones dream of electric sheep? Pract. Neurol. 23, 476–488. https://doi.org/10.1136/pn-2023-003757.

Browne, S.H., Vaida, F., Umlauf, A., Kim, J., DeYoung, P., Owens, R.L., 2024. Performance of a commercial smart watch compared to polysomnography reference for overnight continuous oximetry measurement and sleep apnea evaluation. J. Clin. Sleep Med. 20, 1479–1488. https://doi.org/10.5664/jcsm.11178.

Burkart, N., Huber, M.F., 2021. A survey on the explainability of supervised machine learning. J. Artif. Intell. Res. 70, 245–317.

Cestonaro, C., Delicati, A., Marcante, B., Caenazzo, L., Tozzo, P., 2023. Defining medical liability when artificial intelligence is applied on diagnostic algorithms: a systematic review. Front. Med. 10, 1305756. https://doi.org/10.3389/fmed.2023.1305756.

Chen, X., Huang, W., Peng, Z., Guo, W., Zhang, F., 2024. Diversity supporting robustness: Enhancing adversarial robustness via differentiated ensemble predictions. Comput. Secur. 142, 103861. https://doi.org/10.1016/j.cose.2024.103861.

Daniello, K.M., Weber, D.J., 2018. Education research: the current state of neurophysiology education in selected neurology residency programs. Neurology 90, 708–711.

Derosiere, G., Shokur, S., Vassiliadis, P., 2025. Reward signals in the motor cortex: from biology to neurotechnology. Nat. Commun. 16, 1307. https://doi.org/10.1038/s41467-024-55016-0.

Duraivel, S., Rahimpour, S., Chiang, C.-H., Trumpis, M., Wang, C., Barth, K., et al., 2023. High-resolution neural recordings improve the accuracy of speech decoding. Nat. Commun. 14, 6938. https://doi.org/10.1038/s41467-023-42555-1.

Erdemir, E., Bickford, J., Melis, L., Aydore, S., 2021. Adversarial robustness with non-uniform perturbations. Adv. Neural Inf. Process Syst. 34, 19147–19159.

Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. 1st ed. Geneva: World Health Organization; 2021.

Ewen, J.B., Babiloni, C., Collins, G.S., Ethridge, L.E., Gotman, J., Ikeda, A., et al., 2025. The GREENBEAN checklist for reporting studies evaluating the effectiveness of EEG-based biomarkers. Clin. Neurophysiol. 176, 2110777. https://doi.org/10.1016/j.clinph.2025.2110777.

Feng, J., Phillips, R.V., Malenica, I., Bishara, A., Hubbard, A.E., Celi, L.A., et al., 2022. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. Npj Digit Med 5, 66. https://doi.org/10.1038/s41746-022-00611-y.

Gerke, S., Minssen, T., Cohen, G., 2020. Ethical and legal challenges of artificial intelligence-driven healthcare. Artif. Intell. Healthc., Elsevier 295–336. https://doi.org/10.1016/B978-0-12-818438-7.00012-5.

Gichoya, J.W., Thomas, K., Celi, L.A., Safdar, N., Banerjee, I., Banja, J.D., et al., 2023. AI pitfalls and what not to do: mitigating bias in AI. Br. J. Radiol. 96, 20230023. https://doi.org/10.1259/bjr.20230023.

Haneef, Z., Chiang, S., Rutherford, H.C., Antony, A.R., 2017. A survey of neurophysiology fellows in the United States. J. Clin. Neurophysiol. 34, 179–186.

He, H., Su, W.J., 2023. A law of data separation in deep learning. Proc. Natl. Acad. Sci. 120, e2221704120.

He, J., Baxter, S.L., Xu Jie, Xu., Jiming, Z.X., Zhang, K., 2019. The practical implementation of artificial intelligence technologies in medicine. Nat. Med. 25, 30–36. https://doi.org/10.1038/s41591-018-0307-0.

de Hond, A.A.H., Shah, V.B., Kant, I.M.J., Van Calster, B., Steyerberg, E.W., Hernandez-Boussard, T., 2023. Perspectives on validation of clinical predictive algorithms. NPJ Digit Med 6, 86. https://doi.org/10.1038/s41746-023-00832-9.

Husain, A.M., 2025. Of pilots and copilots: the evolving role of artificial intelligence in clinical neurophysiology. Neurodiagnostic J 65, 2–12.

Jehi, L., 2023. Machine learning for precision epilepsy surgery. Epilepsy Curr 23, 78–83. https://doi.org/10.1177/15357597221150055.

Knight A, Gschwind T, Galer P, Worrell GA, Litt B, Soltesz I, et al. Artificial intelligence in epilepsy phenotyping. Epilepsia 2024:epi.17833. DOI: 10.1111/epi.17833.

van Kolfschooten, H., van Oirschot, J., 2024. The EU Artificial Intelligence Act (2024): Implications for healthcare. Health Policy 149, 105152. https://doi.org/10.1016/j.healthpol.2024.105152.

Martinez-Thompson JM, Mazurek KA, Parra-Cantu C, Naddaf E, Gogineni V, Botha H, et al. Artificial intelligence models using F-wave responses predict amyotrophic lateral sclerosis. Brain 2025:awaf014. DOI: 10.1093/brain/awaf014.

Merhi, M.I., 2023. An evaluation of the critical success factors impacting artificial intelligence implementation. Int. J. Inf. Manag. 69, 102545. https://doi.org/10.1016/j.ijinfomgt.2022.102545.

Mesko, B., 2017. The role of artificial intelligence in precision medicine. Expert Rev. Precis. Med. Drug Dev. 2, 239–241. https://doi.org/10.1080/23808993.2017.1380516.

Muehlematter, U.J., Daniore, P., Vokinger, K.N., 2021. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. Lancet Digit Health 3, e195–e203. https://doi.org/10.1016/S2589-7500(20)30292-2.

Myllyaho, L., Raatikainen, M., Männistö, T., Mikkonen, T., Nurminen, J.K., 2021. Systematic literature review of validation methods for AI systems. J. Syst. Softw. 181, 111050. https://doi.org/10.1016/j.jss.2021.111050.

Nilsson NJ. The quest for artificial intelligence: a history of ideas and achievements. Cambridge New York: Cambridge University press; 2010.

Rahman, N., Khan, D.M., Masroor, K., Arshad, M., Rafiq, A., Fahim, S.M., 2024. Advances in brain-computer interface for decoding speech imagery from EEG signals: a systematic review. Cogn. Neurodyn. 18, 3565–3583. https://doi.org/10.1007/s11571-024-10167-0.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1, 206–215. https://doi.org/10.1038/s42256-019-0048-x.

Sámano, A.G., Mena, J.D.O., Padilla, S.P., Acevedo, G.R., Barraza, J.M.O., San-Juan, D., 2018. Clinical neurophysiology training in a developing country: institutional resources and profiles. J. Clin. Neurophysiol. 35, 242–245.

Siebert, J., Joeckel, L., Heidrich, J., Trendowicz, A., Nakamichi, K., Ohashi, K., et al., 2022. Construction of a quality model for machine learning systems. Softw. Qual. J. 30, 307–335. https://doi.org/10.1007/s11219-021-09557-y.

Stephansen, J.B., Olesen, A.N., Olsen, M., Ambati, A., Leary, E.B., Moore, H.E., et al., 2018. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. Nat. Commun. 9, 5229. https://doi.org/10.1038/s41467-018-07229-3.

Studer, S., Bui, T.B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., et al., 2021. Towards CRISP-ML(Q): a machine learning process model with quality assurance methodology. Mach. Learn. Knowl. Extr. 3, 392–413. https://doi.org/10.3390/make3020020.

Tannemaat, M.R., Kefalas, M., Geraedts, V.J., Remijn-Nelissen, L., Verschuuren, A.J.M., Koch, M., et al., 2023. Distinguishing normal, neuropathic and myopathic EMG with an automated machine learning approach. Clin Neurophysiol Off J Int Fed Clin Neurophysiol 146, 49–54. https://doi.org/10.1016/j.clinph.2022.11.019.

The Learning Health System Series, National Academy of Medicine. An Artificial Intelligence Code of Conduct for Health and Medicine: Essential Guidance for Aligned Action. Washington, D.C.: National Academies Press; 2025. DOI: 10.17226/29087.

Tveit, J., Aurlien, H., Plis, S., Calhoun, V.D., Tatum, W.O., Schomer, D.L., et al., 2023. Automated interpretation of clinical electroencephalograms using artificial intelligence. JAMA Neurol. 80, 805. https://doi.org/10.1001/jamaneurol.2023.1645.

U.S. Department of Health and Human Services: Strategic Plan for the Use of Artificial Intelligence in Health, Human Services, and Public Health. 2025.

Wu, K., Wu, E., Theodorou, B., Liang, W., Mack, C., Glass, L., et al., 2024. Characterizing the clinical adoption of medical AI devices through U.S Insurance Claims. NEJM AI 1. https://doi.org/10.1056/AIoa2300030.

Zrenner, C., Ziemann, U., 2024. Closed-loop brain stimulation. Biol. Psychiatry 95, 545–552. https://doi.org/10.1016/j.biopsych.2023.09.014.