Clinical Neurophysiology Practice 8 (2023) 111–112

Contents lists available at ScienceDirect

Clinical Neurophysiology Practice

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

Letter to the Editor

The future is data-driven: A call to clinical neurophysiology laboratories to standardize your NCS data

Clinical neurophysiology laboratories provide essential diagnostics for neurological diseases. This diagnostic work produces a vast amount of clinical data with great, underutilized potential. Most importantly, these data can be used to create lab-specific reference values that would have an immediate effect on patient diagnostics (Jabre et al., 2015; Nandedkar et al., 2021; Reijntjes et al., 2021). Such reference values could then be further repurposed to calculate the composite Z-scores (Dunker et al., 2022), providing the clinician with a more complete picture of disease severity and progression. In addition, the historical data from neurophysiological laboratories can be used for other purposes including registry studies, laboratory benchmarking and other quality assurance purposes.

Traditionally, reference values for clinical neurophysiology are obtained from healthy subjects. This process is very resource-demanding because it necessitates recruiting healthy subjects, screening them for disease, and employing hospital staff and equipment to collect the necessary data. As a consequence, it is often practically impossible to recruit and test enough healthy subjects to allow for appropriate age stratification. Including healthy children is especially problematic because neurophysiological studies may be unpleasant or painful. Important to note, is that reference values are relatively specific to the laboratory from whence they originated. The use of reference values collected by another laboratory, sometimes going back decades, and obtained with different equipment and methodology, can be inherently inadequate.

New methods, such as e-norms (Jabre et al., 2015), e-ref (Nandedkar et al., 2018), me-ref (Nandedkar et al., 2021) and mixture model clustering (Reijntjes et al., 2021), as well as advances in machine learning, enable each laboratory to create their own reference values based on historical patient data. This is done on the premise that a proportion of patients being examined will have electrodiagnostic findings that are indistinguishable from healthy cohorts (although how you define "healthy" in this context is a discussion in its own right). These methods circumvent most of the challenges inherent in recruiting and testing hundreds or thousands of volunteers, and for NCS, they have yielded promising pilot results that are similar to published reference values (e.g., Punga et al., 2019; Reijntjes et al., 2021).

In the Norwegian "Digital gold mining in neurophysiological data" (DIGMINE) project, we have collected nerve conduction- and electromyography data on over 210,000 Norwegian patients for exactly this purpose. We, and others like us (e.g., working group for a DICOM standard in neurophysiology (Halford et al., 2021)), aim to [further] develop the infrastructure, systems, methods and digital tools that will make it easier for laboratories – alone or in collaboration – to leverage their digital goldmines of historical data.

However, there is no shortcut from collection of raw data from the clinical every-day, to clean data that is easily processed and analyzed. In the real world, data may be stored on inaccessible servers or buried in the local hard drives of electrodiagnostic equipment, sometimes quite literally in the closet. Data retrieved from a hectic clinical practice may be noisy, biased, and lack proper standardization, which can reduce their usefulness. These and other challenges will likely not become apparent before one retrieves the data. Although data cleaning and "post-hoc standardization" is possible to a degree, it is an enormous task that is likely to provide suboptimal results in the end.

Our project has revealed that even in countries with well-standardized methods, similar equipment, settings and training across laboratories, there are pitfalls that threaten how useful the data ultimately becomes. These challenges are likely exacerbated in larger countries with more complex health-care infrastructures. Therefore, directly based on our experience obtaining and compiling NCS data from laboratories across Norway, the DIGMINE project group would like to recommend five ground rules that clinical neurophysiology laboratories should consider moving forward:

1) Do standardize your methods [further]. Data collected on the same machine or on a network of local machines should follow the same protocol and ideally use the same equipment and settings.
2) Do standardize naming. It is important that technicians and clinicians in the same laboratory use the same naming scheme for e.g., nerves, locations and even their own names or initials. Considerable variation in naming of the same procedure can make pooling of data difficult, if not impossible. Common nomenclature across laboratories is challenging, but as a starting example, we append our proposition for nomenclature to be used in Norway, with globally unique identifiers (GUID) that facilitate inter-laboratory pooling of data on smaller scales (Supplemental Files).
3) Do mind the storage. Raw data should be regularly backed up and stored securely where it can be accessed, e.g., on a server or cloud, separate from the local machine, in keeping with legal requirements for data storage in your country.
4) Do strive for complete data. Always record known covariates such as age, height, sex, temperature and measuring distance in a standardized way (e.g., centimeters), but also consider possible covariates such as weight/BMI or dominant limb. Clearly distinguish between studies not performed and actual absent nerve responses.

https://doi.org/10.1016/j.cnp.2023.05.002

5) Do strive for good data quality by minding what is saved in the database, and by avoiding data with low technical quality. As an example, when "improperly" performing a test, e.g. for teaching purposes, either do not save such data, or flag it so it can be easily recognized (for instance by a standardized, identifiable name or code that can be filtered out at a later stage).

By implementing these relatively simple changes, neurophysiological data can be stored with standardized nomenclature and similar structure, perhaps across both laboratories and borders in the future. Such standardization facilitates quality assurance projects and helps to produce reference values that can increase the diagnostic impact of nerve conduction studies. To further this cause, we will suggest an IFCN standardization committee to work towards a more universal model for NCS practices, which includes a standardized nomenclature for NCS recordings. Looking to the future, focusing on standardization basics will allow for more international collaborative studies and pave the way for the development and training of machine-learning based medical decision support systems, in the best interest of neurological patients worldwide.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cnp.2023.05.002.

References

Dunker, Ø., Nilsen, K.B., Olsen, S.E., Åsvold, B.O., Bjørgaas, M.R.R., Sand, T., 2022. Which combined nerve conduction study scores are best suited for polyneuropathy in diabetic patients? Muscle Nerve 65, 171–179.

Halford, J.J., Clunie, D.A., Brinkmann, B.H., Krefting, D., Rémi, J., Rosenow, F., et al, 2021. Standardization of neurophysiology signal data into the DICOM® standard. Clin. Neurophysiol. 132, 993–997.

Jabre, J.F., Pitt, M.C., Deeb, J., Chui, K.K., 2015. E-norms: a method to extrapolate reference values from a laboratory population. J. Clin. Neurophysiol. 32, 265–270.

Nandedkar, S.D., Sanders, D.B., Hobson-Webb, L.D., Billakota, S., Barkhaus, P.E., Stålberg, E.V., 2018. The extrapolated reference values procedure: Theory, algorithm, and results in patients and control subjects. Muscle Nerve 57, 90–95.

Nandedkar, S.D., Stålberg, E.V., Barkhaus, P.E., 2021. MeRef: Multivariable extrapolated reference values in motor nerve conduction studies. Muscle Nerve 63, 737–744.

Punga, A.R., Jabre, J.F., Amandusson, Å., 2019. Facing the challenges of electrodiagnostic studies in the very elderly (>80 years) population. Clin. Neurophysiol. 130, 1091–1097.

Reijntjes, R.H., Potters, W.V., Kerkhof, F.I., van Zwet, E., van Rossum, I.A., Verhamme, C., et al, 2021. Deriving reference values for nerve conduction studies from existing data using mixture model clustering. Clin. Neurophysiol. 132, 1820–1829.

Øystein Dunker *
Oslo University Hospital, Department of Research and Innovation, Norway
* Corresponding author at: Oslo University Hospital, Department of Research and Innovation, Kirkeveien 166, 0450, Norway.
E-mail address: anoeys@ous-hf.no

Marie U. Lie
Aristomo Andries
Kristian Bernhard Nilsen
Oslo University Hospital, Division of Clinical Neuroscience, Norway

Petter Omland
St Olavs Hospital Trondheim University Hospital Department of Neurology and Clinical Neurophysiology, Norway

Andrew Reiner
Oslo Centre for Biostatistics and Epidemiology, Norway

Martijn R. Tannemaat
Robert H. Reijntjes
Leiden University Medical Center Department of Neurology, the Netherlands

Joe F. Jabre
Formerly, University of California Los Angeles David Geffen School of Medicine Department of Neurology, USA

Received 18 April 2023
Received in revised form 3 May 2023
Accepted 8 May 2023

Available online 19 May 2023