scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A comprehensive dataset of pattern electroretinograms for ocular electrophysiology research

Itziar Fernández[1,2,3]✉, Rubén Cuadrado-Asensio[2,4], Yolanda Larriba[1,5]✉, Cristina Rueda[1,5] & Rosa M. Coco-Martín[2,4]

The Pattern Electroretinogram (PERG) is an essential tool in ophthalmic electrophysiology, providing an objective assessment of the central retinal function. It quantifies the activity of cells in the macula and the ganglion cells of the retina, assisting in the differentiation of macular and optic nerve conditions. In this study, we present the IOBA-PERG dataset, an extensive collection of 1354 transient PERG responses accessible on the PhysioNet repository. These recordings were conducted at the Institute of Applied Ophthalmobiology (IOBA) at University of Valladolid, over an extended period spanning nearly two decades, from 2003 to 2022. The dataset includes 336 records, ensuring at least one PERG signal per eye. The dataset thoughtfully includes demographic and clinical data, comprising information such as age, gender, visual acuity measurements, and expert diagnoses. This comprehensive dataset fills a gap in ocular electrophysiological repositories, enhancing ophthalmology research. Researchers can explore a broad range of eye-related conditions and diseases, leading to enhanced diagnostic accuracy, innovative treatment strategies, methodological advancements, and a deeper understanding of ocular electrophysiology.

Background & Summary

Visual electrophysiology involves a variety of non-invasive techniques to quantify electrical signals throughout the visual pathway, from the retina to the optic nerve and the primary visual cortex. These tests are progressively becoming more accessible and invaluable in both clinical and research contexts, aiding our comprehension of ocular diseases[1]. They complement other diagnostic tools and support innovative therapeutic approaches[2]. Electrophysiological measurements play a key role in assessing the functioning of the optic nerve, diagnosing hereditary retinal diseases, and monitoring traumatic injuries. They assist in managing inflammatory disease outbreaks such as Birdshot chorioretinopathy and autoimmune retinopathies[3]. Beyond ocular applications, they find utility in neurological conditions[4–6], monitoring drug-induced toxicity[7], and as biomarkers for tracking disease progression and treatment responses in neurological disorders[8].

Among the various visual electrophysiology tests, the electroretinogram (ERG) is a diagnostic test utilized in ophthalmology to measure the electrical activity of the retina in response to light stimuli[2]. Various types of ERG tests serve distinct purposes in assessing retinal function and identifying a range of retinal conditions. Among these, the pattern ERG (PERG) stands out as a specialized test designed to thoroughly assess macular and ganglion cell functionality[9]. This specific area of the retina is responsible for central vision and fine-detail visual tasks, therefore it plays a critical role in our daily lives. The importance of PERG lies in its ability to identify early signs of macular disorders and conditions that affects the optic nerve, even before substantial loss of visual field occurs. Additionally, PERG serves as a valuable tool for long-term monitoring of the progress of these disorders, allowing healthcare professionals to assess the effectiveness of treatment strategies and make necessary adjustments when needed. Beyond its role in early detection and monitoring, PERG can be employed for differential

[1]Department of Statistics and Operations Research, University of Valladolid, Valladolid, 47011, Spain. [2]Institute of Applied Ophthalmobiology (IOBA), University of Valladolid, Valladolid, 47011, Spain. [3]Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Carlos III National Institute of Health, Valladolid, 47011, Spain. [4]Department of Surgery, Ophthalmology, Otorhinolaryngology and Physiotherapy, University of Valladolid, Valladolid, 47005, Spain. [5]Mathematics Research Institute of the University of Valladolid (IMUVA), University of Valladolid, Valladolid, 47011, Spain. ✉e-mail: itziar.fernandez@uva.es; yolanda.larriba@uva.es

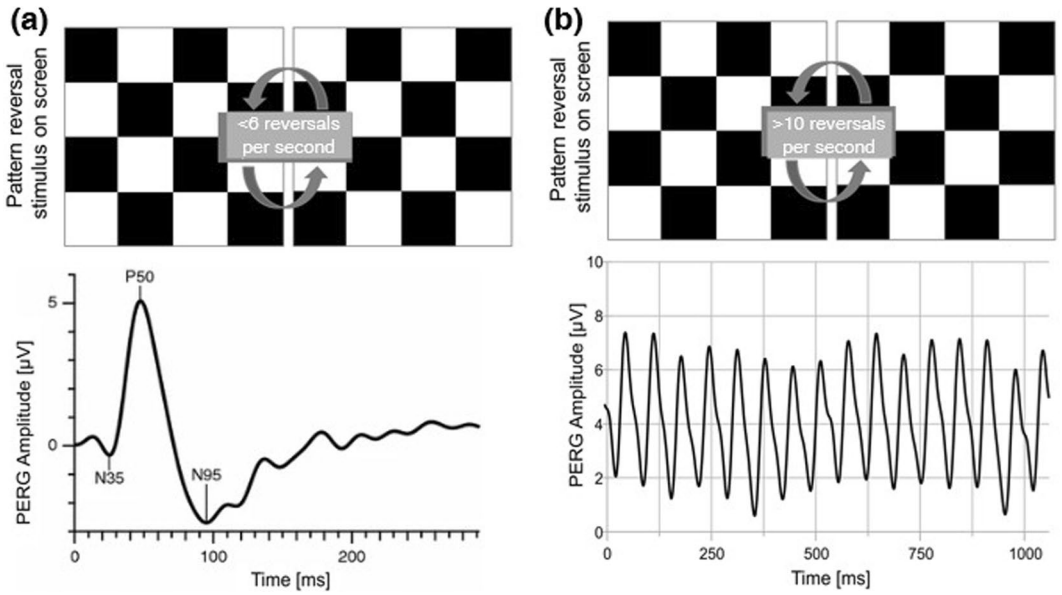

Fig. 1 Typical PERG waveforms. (**a**) Transient PERG, generated at temporal rates less than 6 pattern reversals per second (rps). (**b**) Steady-state PERG, generated at temporal rates higher than 10 rps.

diagnoses. In situations where retinal disorders present similar symptoms but originate from diverse underlying causes, specific focus on macular function helps provide more accurate diagnoses. Finally, PERG extends beyond the boundaries of clinical practice and is frequently used in clinical research and trials related to retinal and optic nerve diseases. Its ability to offer precise data on macular function makes it a useful tool for evaluating the efficacy of new treatments and interventions.

The PERG response is generated through a checkerboard stimulus that includes alternating black and white squares, reversing multiple times per second to stimulate the central 15 degrees of the retina. PERG is a relatively small signal, typically falling within the amplitude range of 2-8 microvolts (μV). This limited signal amplitude requires a high degree of technical precision during the recording process. Therefore, it is strongly advised to closely adhere to the well-established guidelines specified by the International Society for Clinical Electrophysiology of Vision (ISCEV)[10–12].

The PERG waveform varies depending on the temporal frequency of the stimulus, either transient or steady-state PERG (ssPERG). On the one hand, a standard PERG captures a transient response at low temporal frequencies (below 6 reversals per second, rps, or equivalent to less than 3 Hertz, Hz). In individuals with normal vision, this transient PERG consists of an initial small negative component around 35 milliseconds (ms), followed by a larger positive component at 45–60 ms, and a substantial negative component at 90-100 ms. Conventionally, these components are designated as N35, P50, and N95, respectively, based on their polarity and approximate latency. N95 is linked to the function of the retinal ganglion cells, while P50 indicates the activity of the macular photoreceptors and serves as an indicator of macular function[13]. On the other hand, a steady-state waveform is attainable with a rapid stimulus rate (exceeding 3.5 Hz). This results in a periodic, roughly triangular response whose amplitude and phase features remain constant over time (see Fig. 1). ssPERG does not allow for the measurement of individual components within the PERG pattern and, although it provides valuable information about overall retinal function, it may lack the diagnostic precision required to differentiate specific retinal diseases when compared to the transient PERG. In this dataset, recorded in a clinical context, we focused on transient PERG, the test recommended in such settings by ISCEV standards[10,14], as it offers more precise information about the specific site and extent of retinal dysfunction compared to ssPERG.

As far as we are aware, PERG signal datasets are not currently available in public repositories. Access to datasets generated and analyzed in various studies can often be obtained from the corresponding author upon request. However, it is crucial to acknowledge that, in many cases, the raw data is not accessible. Instead, it is frequently presented in the form of summarized component wave amplitudes and their corresponding implicit times.

In this study, we present a extensive transiet PERG dataset. It is comprised of 1354 signals gathered from 304 participants enrolled at the Institute of Applied Ophthalmobiology (IOBA), a University of Valladolid-affiliated institution in Spain, over the period from 2003 to 2022. Throughout this elongated recruitment period, 23 individuals participated in multiple visits, resulting in a total of 336 records. As a part of the routine clinical assessment, subjects provided detailed clinical information, including their age, gender, visual acuity, and between 1 and 3 diagnoses by ophthalmology specialists. Notably, this dataset ensures access to at least one PERG signal for each eye, facilitating research into patterns and variations in ocular responses.

The value of such an exhaustive and well-curated dataset cannot be overstated. It provides researchers and scientists with unprecedented access to an extensive collection of PERG data, facilitating comprehensive studies, identifying emerging patterns, and gaining deep insights into the complexities of the visual system. These insights possess the potential to substantially propel our comprehension of various ocular disorders and

conditions, including, but not limited to, optic neuropathies or retinal disorders. This dataset offers invaluable knowledge regarding the progression of these conditions and holds the capacity to contribute to the development of innovative treatments. Moreover, by making this dataset accessible to a wider scientific community, it not only fosters collaboration but also promotes advancements within the field of ophthalmology. It serves as an indispensable resource for validating research findings and refining methodological approaches, driving innovation and progress within the field. The comprehensive and high-quality nature of this dataset presents a great opportunity for the scientific community, particularly due to its ability to fill a significant gap in ocular electrophysiological signal repositories. Illustrating its utility, the dataset has already been partially utilized in a study focused on developing physiologically plausible statistical models for the analysis and prediction of PERG signals using the innovative FMM (Frequency Modulated Möbius) approach[15].

Methods

The dataset described in this paper was compiled for a project focused on the automated analysis of electrical signals obtained through ocular electrophysiology tests, which received approval from the IOBA research committee (approval 2021/47). The rigorous approval process guarantees strict adherence to ethical and research standards. Following IOBA's standard clinical practice, all patients were informed about the procedures, risks, and benefits of the ocular electrophysiological tests. Additionally, they were made aware of the potential use of their de-identified data for research purposes. All participants signed an informed consent form, indicating their agreement and acknowledging their understanding of the information provided.

Data collecting. From 2003 to 2022, a total of 336 ocular electrophysiology visits were conducted at IOBA, a research institute affiliated with the University of Valladolid in Spain. These visits involved the measurement of transiet PERG signals from a diverse group of 304 subjects, representing a cross-section of patients, volunteers, or participants involved in eye-related research. As a result of these visits a substantial collection of 1354 transient PERG signals was collected. The dataset encompasses all individuals who underwent transient PERG testing in the Retina Unit at IOBA during this period, ensuring comprehensive inclusion without selective exclusion criteria. Transient PERG was primarily used to assess visual prognosis in patients with retinal diseases who retained some vision. It is generally not indicated for patients with a logMAR visual acuity exceeding 1.0, which corresponds to the threshold for legal blindness, as its results are often expected to be flat. Additionally, PERG may be useful in cases where the cause of vision loss is unclear in patients with normal ocular fundus, although such scenarios are relatively uncommon.

Data were compiled during routine checkups. The results were reviewed and analyzed by a specialized ophthalmologist belonging to the Retina Unit, who conducted comprehensive clinical evaluations and diagnoses.

Recording protocol. All PERG signals were recorded by highly trained optometrist using the computerized Optoelectronic Stimulator Vision Monitor MonPack 120 (Metrovision, Pérenchies, France), strictly adhering to the ISCEV guidelines[10–12]. The ISCEV guidelines involves a standardized and well-defined procedure to guarantee a high degree of consistency and reliability in measurements. During the data collection period, there was only one hardware and software upgrade, which occurred midway through. This upgrade is unlikely to have had a major impact on data consistency. Additionally, from 2003 to 2022, the ISCEV PERG standards have been updated to reflect advancements in technology and a deeper understanding of electrophysiological testing. Key changes include more flexible stimulus parameters, updated recording conditions with a preference for binocular recordings, refined signal processing techniques, and stricter quality control measures. These adjustments improve the accuracy and consistency of PERG recordings without fundamentally altering the basic principles of the protocol described bellow.

A binocular recording was carried out using single-use, conveniently sterilized electrodes with their integrity verified before insertion. A recording gold electrode was accurately positioned on the corneal surface, while a separate reference electrode was placed on the skin, near the outer canthus of each eye on the same side (ipsilateral). Additionally, a surface electrode was placed on he forehead and connected to the amplifier to "ground input".

The subjects were meticulously prepared for the examination, ensuring they were in a comfortable and relaxed state throughout the process, with their heads in a stable position against a head-rest. To preserve accommodation and, consequently, retinal image quality, the PERG signals were recorded without dilatation of the pupils and with the necessary optical correction for an optimal visual acuity. This correction was initially based on the participants' current refractive prescriptions. However, if these prescription did not provide 20/20 vision, the refraction was manually reassessed during the visit, and the optical correction was adjusted accordingly. Participants were seated in a quiet, dark room at a distance of 1 m from the monitor, with the screen occupying 12° vertically and 16° horizontally of the visual field, at a photopic luminance of 4 lux. Explicit instructions were provided to the participants, directing them to fixate on a central target in the stimulator (0.5°, 3.32 cd/m^2), with an emphasis on minimizing any unnecessary eye and/or face movements. The lighting conditions in the testing room were thoughtfully controlled, maintaining a subdued ambient light environment before presenting the visual stimuli to the subjects. These conditions remained constant throughout all recordings.

A black and white reversing checkerboard pattern was employed, featuring a reversal rate of 4 rps, equivalent to 2 Hz. The stimulus was displayed on a cathode-ray tube (CRT) monitor to mitigate flash artifacts that can occur during pattern reversals. The checkerboard pattern featured a check size of 1.0°, with white areas exhibited a photopic luminance exceeding 80 candela per square meter, and the contrast between the black and white squares was maximized, nearly reaching 100%. Furthermore, a frame rate of 75 Hz was utilized to present the stimuli with precision. Each eye was subjected to an average of 230 pattern reversal stimuli. The analysis period,

# PERG signals	2	4	6	8	10
# Records	49	240	41	5	1

Table 1. Overview of number of PERG signals per record.

Variable	Data Type	Description
`id_record`	string	four-digit unique PERG record identifier
`date`	date	PERG recording date encoded as `YYYY-MM-DD`
`age`	integer	age at recording in years
`sex`	categorical	sex (`Male` or `Female`)
`diagnosis1-3`	string	up to three different ocular diagnoses per record
`va_re`	double	visual acuity right eye on logMAR scale
`va_le`	double	visual acuity left eye on logMAR scale
`unilateral`	categorical	eye affected by a unilateral condition (right eye, `RE`, or left eye, `LE`)
`rep_record`	string	identifier of the record matches the subjects (in the format `id:XXXX`)
`comments`	string	additional information

Table 2. Columns provided in the metadata file `participants_info.csv`.

or sweep time, was set at 150 milliseconds with 250 milliseconds intervals between reversals. For optimal data accuracy, a higher sampling rate of 1700 Hz was employed during PERG recording.

Data processing. PERG signals were recorded using amplification systems and electrodes, and the raw data was collected in digital form. The signal processing adhered to the clinical standards integrated into the used devices. This encompassed a series of essential steps, including preprocessing, artifact detection and correction, and signal averaging. These steps collectively aimed to elevate data quality while eliminating any unwanted noise.

To initiate the enhancement process, a series of preprocessing steps were carried out, encompassing crucial procedures such as filtering and baseline correction. The recorded signals were filtered within a 1–100 Hz band to isolate the relevant frequency components and remove noise outside this range.

Subsequently, the data was segmented into epochs, with each segment corresponding to a single stimulus presentation. These segments underwent meticulous artifact detection procedures to maintain data integrity. Computerized algorithms analyzed the data to detect abnormal data points. These anomalies, often stemming from sources like eye blinks, saccades, muscle interference, or electrical noise, were identified by the algorithms through predefined threshold values. A fast rejection threshold of 8 μV was applied to eliminate data points with high-amplitude fluctuations, while a slow rejection threshold of 50 μV was used to filter out slower artifacts that could affect the signal. Any data point exceeding these established thresholds was flagged as a potential artifact. Flagged data points were replaced with interpolated values, effectively eradicating their influence on the final data.

Signal averaging emerges as a fundamental step in PERG signal processing due to their typically low amplitude. This process is instrumental in augmenting the signal-to-noise ratio, thereby enabling the extraction of meaningful insights from the data. A minimum of 100 artifact-free sweeps were acquired and then subjected to averaging. In cases where the PERG response exhibited small amplitude, was undetectable, or was overshadowed by significant background noise, a higher number of sweeps became imperative to yield reliable results.

Data de-identification. To safeguard the confidentiality and privacy of the subjects involved, all protected health information has been meticulously removed from the dataset, and a comprehensive de-identification process was applied. The anonymization process began with the identification of personal information, which was categorized into direct and indirect identifiers. Direct identifiers, such as names, addresses, phone numbers, and specific clinical history references, were removed from the dataset. Clinical history references were replaced with a unique four-digit codes generated randomly to preserve data utility while ensuring confidentiality. Indirect identifiers, including dates of birth and visit dates, were handled through generalization and random offset techniques. Specifically, birthdates were replaced with only the age of participants (in years), and, for PERG records, acquisition dates were subtly shifted by a random offset, preserving chronological order but obscuring specific dates. These measures ensure the protection of individual privacy while maintaining the value of the dataset for research purposes.

Data Records

The dataset, known as PERG-IOBA[16], is accessible through the PhysioNet repository[17]. The dataset contains 336 records, with each record corresponding to a single visit, and it encompasses a total of 1354 PERG signals. Each record in the dataset guarantees the presence of at least one PERG signal for each eye.

PERG signal data. The PERG signal data are presented in comma-separated value (CSV) format, with a dedicated file for each record. These files adhere to a standardized naming convention, featuring a four-digit unique identifier that has been exclusively designed for this collection. Importantly, this unique identifier is entirely independent of any information found in the participants' medical records.

All CSV files include a `TIME` column and at least one PERG signal data for each eye, identified as `RE_1` and `LE_1` for the right and left eye, respectively. The time is encoded as `YYYY-MM-DD hh:mm:ss.ms`. To accommodate cases where the test is repeated during the same visit, additional columns labeled as `RE_2`, `RE_3`, and so forth, along with `LE_2`, `LE_3`, and so on, are included to encompass multiple signals collected for each eye. Furthermore, to provide temporal information for the repeated tests, columns `TIME_2`, `TIME_3`, and so on, are incorporated into the CSV records whenever applicable.

In total, there are 1354 PERG signals distributed across 336 records, with a number of signals per record ranging from 2 to 10. A detailed breakdown regarding the number of signals per record is presented in Table 1. Note that the number of PERG signals is always a multiple of 2, as data from both eyes are consistently incorporated.

Metadata. Metadata for all PERG records are provided in CSV format, organized within the file `participants_info.csv` containing 12 columns. Table 2 gives an overview of the variables included in this table.

A total of 69 different diagnoses have been recorded across three variables (`diagnosis1-3`), and each record in the dataset has been allocated at least one diagnosis. One hundred and six out of 336 records (31.5%) are categorized as "normal", indicating the absence of ocular pathology. The overwhelming majority of diagnoses affect both eyes, as only 2.4% of the 336 records display unilateral involvement, evenly distributed between the right and left eye. The distribution of all diagnoses is depicted in Fig. 2.

Among the diagnoses, 188 (56.0%) records indicate retinal involvement, 34 (10.1%) denote a neuro-optic disorder, 8 (2.4%) reveal retinal toxicity, and 2 (0.6%) signify amblyopia. Within the category of retinal involvement, several diagnostic subclasses can be distinguished, including cone disorder, and rod-cone, cone-rod, central or inner retina affectation. Meanwhile, central neuro-ophthalmological and optic nerve affectation are subclasses of neuro-ophthalmological disorders. Figure 3 provides a summary of the distribution of diagnostic classes and their corresponding subclasses in the PERG-IOBA dataset.

Demographic data comprises information on `age` and `sex`, with 47.6% being male and 52.4% female. The age refers to the subjects's age at the time of the PERG recording. The age distribution for the complete dataset and segregated by gender is presented in Table 3.

Out of the 304 participants, 23 (7.6%) have multiple records, each corresponding to a follow-up visit. In particular, 19 of them has two records, one has three, another has four, and two more have five follow-up visits. The variable `rep_record` is designed to aggregate record identifiers belonging to the same individual. Each entry in this variable follows the format `id:XXXX`, where `XXXX` represents the specific record identifier. Different records are separated by hyphens (`-`).

Regarding the clinical data, both right and left eye visual acuity measurements are recorded (variables `va_re` and `va_le`, respectively) with a missing rate of 5.7%. Visual acuity plays a critical role in the evaluation of PERG signals, as the precision and reliability of PERG measurements rely on the participant's ability to perceive visual stimuli clearly. We employ the logMAR (logarithm of the Minimum Angle of Resolution) scale for the assessment of visual acuity. This scale quantifies visual acuity by assigning a value based on a person's ability to discern progressively smaller optotypes, such as letters or symbols, on an eye chart. Within the logMAR scale, lower values indicate better visual acuity, with 0 representing perfect vision, and negative values indicating even sharper vision. This is a standardized scale that ensures accurate and consistent evaluation of visual acuity in clinical and research settings. The distributions of visual acuity are very similar between eyes and also between genders (see Table 3). Figure 4 illustrates the distribution of average visual acuity across diagnostic subclasses.

Technical Validation

As mentioned above, the PERG is a small signal, demanding a high level of technical precision during the acquisition process. The PERG-IOBA records were obtained in a clinical environment, guided by a specialized optometrist and later interpreted by an ophthalmologist specializing in ocular electrophysiology. Adhering to clinical protocols, the personnel involved in the tests continuously monitored the signals and equipment, performing periodic evaluations and secure fixation of the electrodes. Ensuring the precision of PERG signals during the recording procedure involved a series of precautions. These measures included comprehensive patient preparation, meticulous electrode placement for optimal contact without discomfort, and maintaining consistent lighting conditions. Stimuli calibration adhered to established standards for luminance, contrast, and pattern size. Electrode management and patient fixation were prioritized to prevent contamination and minimize eye movement artifacts. Signal amplification levels were carefully adjusted to capture responses within the optimal range without saturation. Another aspect to highlight is the implementation of regular equipment calibration checks, initially conducted annually as per the manufacturer's recommendations and ISCEV guidelines[18,19]. However, more frequent evaluations were carried out when necessary to promptly address any deviations and maintain consistent performance. Specific equipment calibration parameters were meticulously maintained. These parameters include adjusting stimulus luminance of white areas within the range of 100 to 500 cd/m^2 80 to 100 cd/m^2 and a mean luminance of 45–50 cd/m^2, according with the ISCEV standards. Also, constant overall screen luminance during checkerboard reversals was verified. Other parameters include maintaining a sampling frequency of 1700 Hz, calibrating stimulus voltage levels to prevent saturation while ensuring adequate response, setting system response time typically between 1 to 10 ms post-stimulus, precise alignment of visual stimuli on the monitor, maintaining background noise levels below 5 μV, and ensuring electrode impedance remains below 5 kΩ. Calibration also involves setting amplifier gain to amplify signals without distortion, typically around 10,000x.

The quality of biomedical signals and, by extension, the subsequent analysis, can be greatly affected by background noise. This noise can originate from various sources such as electronic components, environmental factors, or inherent limitations in the measurement or transmission process. In the case of PERG, emphasizing the crucial role of signal averaging in the acquisition process is essential for mitigating the impact of background

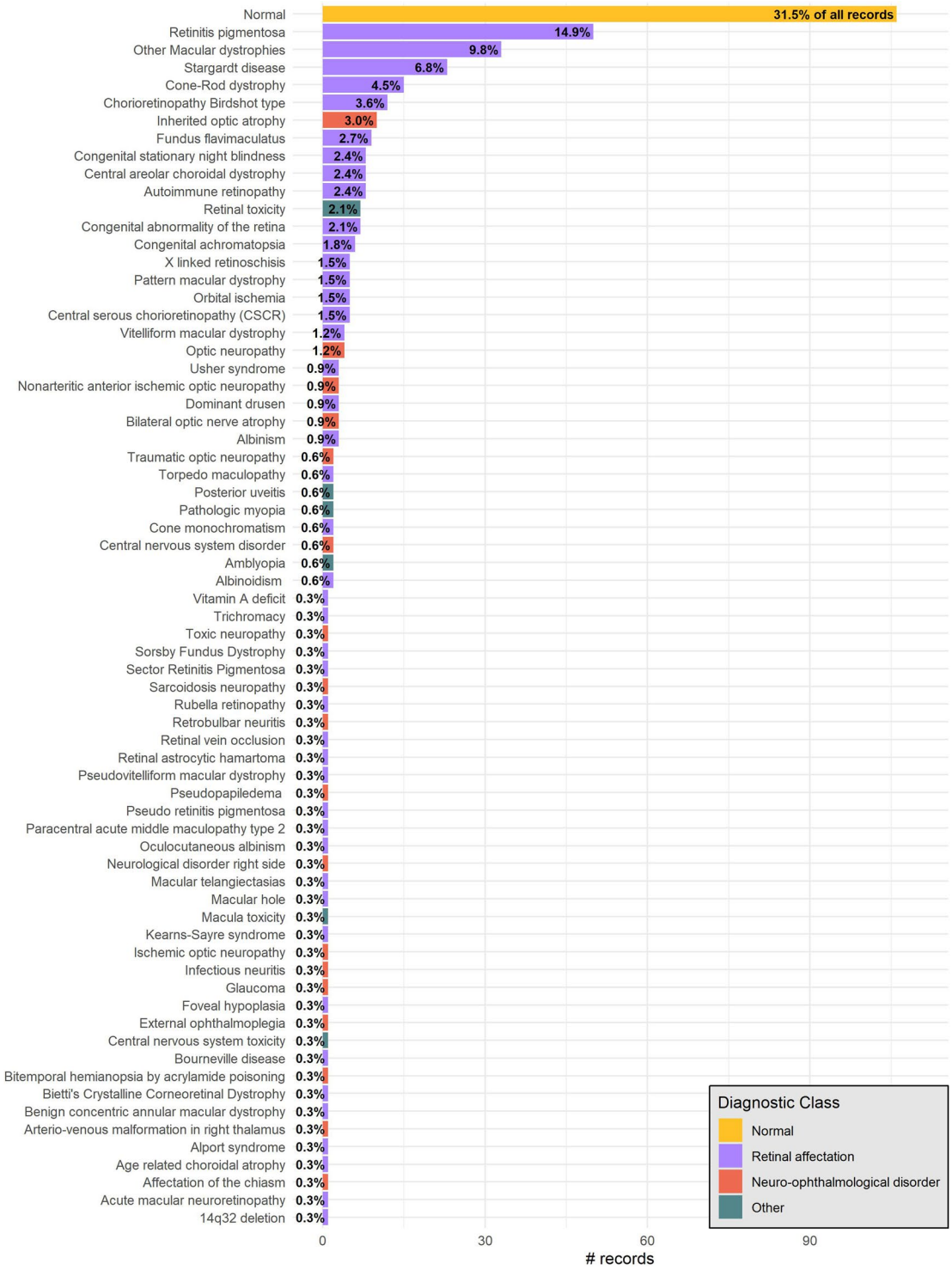

Fig. 2 Distribution of PERG-IOBA diagnostics. The color scheme differentiates between diagnostic classes, including Normal, Retinal Affectation, Neuro-Ophthalmic Disorder, and Other Class.

noise. In this context, averaging involves combining multiple repetitions of the same stimulus presentation. As noise is random, the signal, which remains consistent across repetitions, becomes clearer from the background noise, contributing to enhancing the signal-to-noise ratio. The presence of noise in signals is often characterized by rapid and random changes in amplitude from point to point within the signal. In contrast, the signal amplitudes typically exhibit a smoother, more gradual change. Therefore, the use of smoothing techniques can be useful for evaluating the presence of high-frequency variations or noise in a signal. In terms of the frequency components of a signal, a smoothing operation serves as a low-pass filter, reducing high-frequency components while preserving low-frequency components. This results in a naturally smoother signal, characterized by a slower step response to

	Age			Visual acuity RE			Visual acuity LE		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
N	336	160	176	317	152	165	317	152	165
Mean	37.07	37.33	36.84	0.34	0.30	0.37	0.32	0.30	0.34
Standard Deviation	18.28	17.89	18.68	0.55	0.47	0.61	0.50	0.47	0.53
Minimum	4.00	4.00	5.00	−0.10	−0.10	−0.10	−0.10	−0.10	−0.10
25th Percentile	21.00	22.00	20.00	0.00	−0.02	0.00	0.00	0.00	0.00
Median	38.00	38.00	38.00	0.16	0.16	0.16	0.14	0.11	0.14
75th Percentile	51.00	50.25	51.00	0.50	0.51	0.46	0.48	0.42	0.50
Maximum	86.00	81.00	86.00	3.00	3.00	3.00	3.00	3.00	3.00

Table 3. Description of demographic and clinical data.

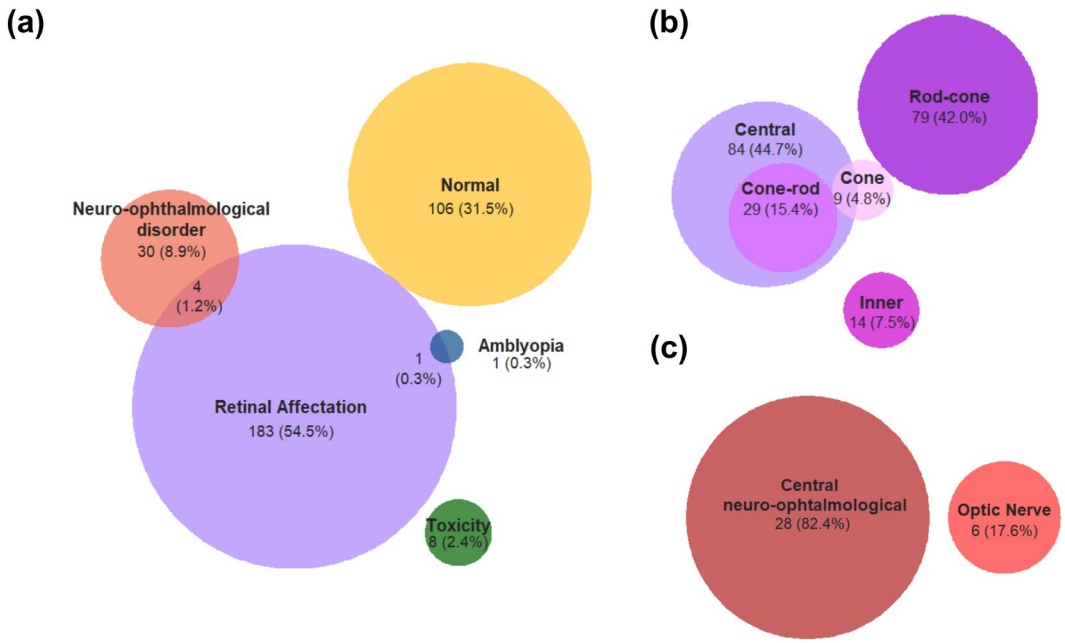

Fig. 3 Distribution of PERG-IOBA records within diagnostic categories and subcategories: (**a**) Diagnostic classes, (**b**) Subclasses of retinal impairments, and (**c**) Subclasses of neuro-ophthalmic disorders.

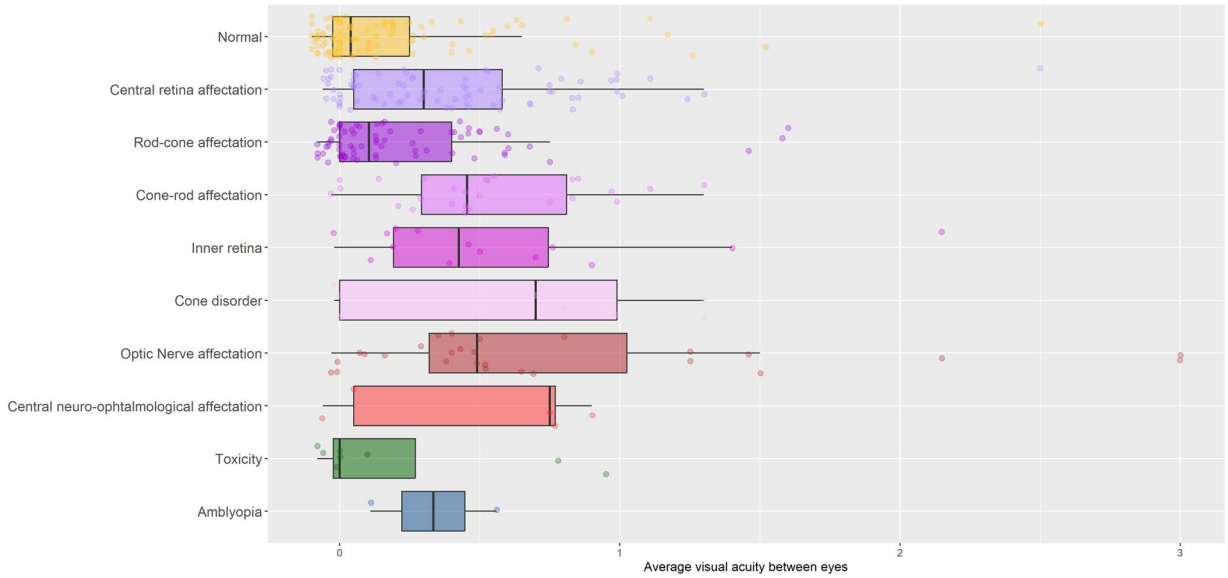

Fig. 4 Distribution of the average visual acuity between eyes within diagnostic subcategories.

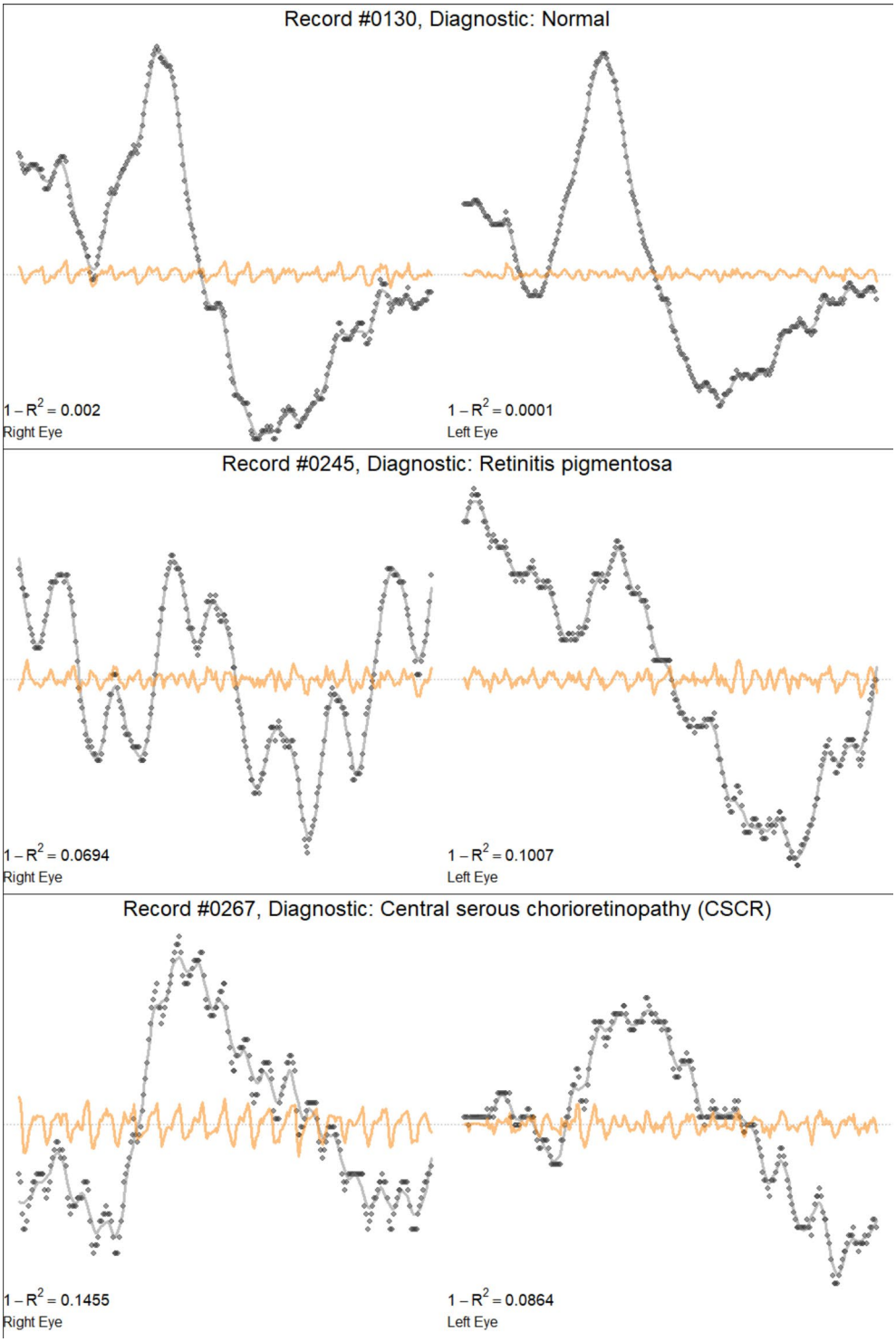

Fig. 5 Decomposition of PERG data into signal and noise components from three illustrative recordings exhibiting different levels of background noise. Points: observed data, black line: smoothed signal, orange line: noise component.

signal changes. In this section we use smoothing to assess the presence of noise in the signals of the PERG-IOBA dataset. Note that this smoothing process has not been applied to the actual signals available in the dataset.

We assume that the signals are distorted by an additive noise. Therefore, the observed signal, denoted by $X(t_i)$ for time points $t_1 < t_2 < \ldots < t_n$, is a sum of both the signal and noise components. Locally weighted regression (loess) is used to estimate the filtered signal from the observed data, with a bandwidth of 10% and a quadratic polynomial fitted via least squares. This technique performs estimations on a point-wise fashion, using the neighboring known values for each specific data point. Figure 5 illustrates the decomposition into signal

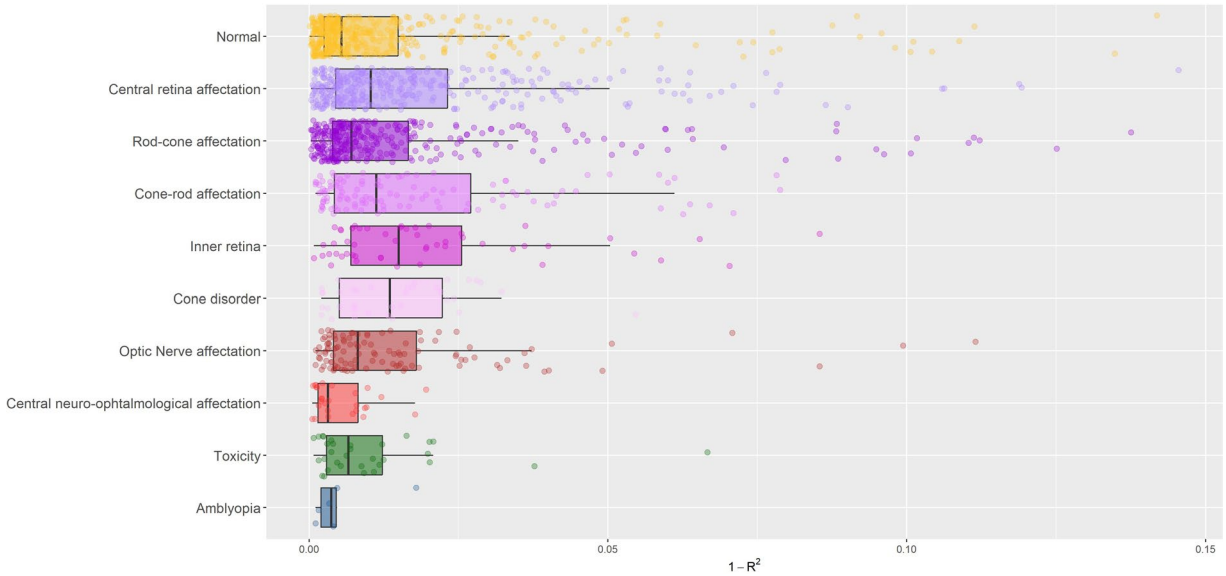

Fig. 6 Distribution of the $1 - R^2$ across diagnostic subcategories.

and noise components for three representative examples from the PERG-IOBA dataset. Each record exhibits a different level of background noise, demonstrating a consistently high signal-to-noise ratio in each case.

To assess the predictive performance of the smoothed signal, the coefficient of determination (R^2) can be used as a measure of goodness of fit. It is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(X(t_i) - \hat{X}(t_i))^2}{\sum_{i=1}^{n}(X(t_i) - \overline{X})^2}$$

where $\hat{X}(t_i)$ represents the smoothed signal value at time-point t_i, $i = 1, \ldots, n$ and $\overline{X}$ is the average of the observed signal. The R^2 value ranges between 0 and 1 and can be interpreted as the proportion of variability that is explained by the smoothed signal. Therefore, $1 - R^2$ can be considered as a measure of the residual error or the proportion of variability left unexplained by the predicted signal. Figure 6 shows the distribution of $1 - R^2$ across the different diagnostic subcategories considered within the PERG-IOBA dataset. The values of $1 - R^2$ appear to be consistently low, with the 75th percentile of $1 - R^2$ below 0.03 for each subcategory. All diagnostic subcategories exhibit values lower than 0.15.

As expected, a certain dependence on the diagnosis is observed. When a patient has difficulty seeing the stimuli clearly, the retinal ganglion cells may produce a weaker and less distinguishable signal. These challenges in visual perception can stem from factors like poor vision, refractive errors, other visual impairments, or even symptoms of inattention[20]. The weakened signal becomes more susceptible to interference from various sources, such as electrical interference, physiological artifacts, or even stray light in the environment.

Usage Notes

The PERG-IOBA dataset was created for the primary objective of developing and assessment of automated diagnostic algorithms relying on PERG signals. In a field where repositories of ocular electrophysiological signals are limited, this extensive dataset stands as a valuable resource, holding the potential to drive substantial progress in the realm of ophthalmology research. Its accessibility opens up fresh avenues for the exploration of a wide range of eye-related conditions and diseases. This, in turn, facilitates the advancement of diagnostic techniques, treatment approaches, and a more profound comprehension of ocular electrophysiology.

To download and explore this dataset, users can visit the following url: https://physionet.org/content/perg-ioba-dataset/1.0.0/. Anyone is permitted to use, share, and build upon the data for commercial, research, or other purposes, provided that appropriate attribution is given to the original data owner.

Code availability

No custom code was generated for this work.

Received: 29 January 2024; Accepted: 4 September 2024;
Published online: 18 September 2024

References

1. Kremers, J., McKeefry, D. J., Murray, I. J. & Parry, N. R. Developments in non-invasive visual electrophysiology. *Vision Research* **174**, 50–56, https://doi.org/10.1016/j.visres.2020.05.003 (2020).
2. Mahroo, O. A. Visual electrophysiology and "the potential of the potentials". *Eye* **37**, 2399–2408, https://doi.org/10.1038/s41433-023-02491-2 (2023).

3. Yu, M., Creel, D. J. & Iannaccone, A. *Handbook of Clinical Electrophysiology of Vision*, https://doi.org/10.1007/978-3-030-30417-1 (Springer, 2019).
4. Parisi, V. *et al*. Morphological and functional retinal impairment in alzheimer's disease patients. *Clinical Neurophysiology* **112**, 1860–1867, https://doi.org/10.1016/S1388-2457(01)00620-4 (2001).
5. Nightingale, S., Mitchell, K. & Howe, J. Visual evoked cortical potentials and pattern electroretinograms in parkinson's disease and control subjects. *Journal of Neurology, Neurosurgery & Psychiatry* **49**, 1280–1287, https://doi.org/10.1136/jnnp.49.11.1280 (1986).
6. Barton, J. L., Garber, J. Y., Klistorner, A. & Barnett, M. H. The electrophysiological assessment of visual function in multiple sclerosis. *Clinical Neurophysiology Practice* **4**, 90–96, https://doi.org/10.1016/j.cnp.2019.03.002 (2019).
7. Robson, A. G., Li, S., Neveu, M. M., Yin, Z. Q. & Holder, G. E. The electrophysiological characteristics and monitoring of ethambutol toxicity. *Investigative Ophthalmology & Visual Science* **55**, 6213–6213 (2014).
8. Silverstein, S. M., Demmin, D. L., Schallek, J. B. & Fradkin, S. I. Measures of retinal structure and function as biomarkers in neurology and psychiatry. *Biomarkers in Neuropsychiatry* **2**, 100018, https://doi.org/10.1016/j.bionps.2020.100018 (2020).
9. Holder, G. E. Pattern electroretinography (PERG) and an integrated approach to visual pathway diagnosis. *Progress in Retinal and Eye Research* **20**, 531–561, https://doi.org/10.1016/S1350-9462(00)00030-6 (2001).
10. Bach, M. *et al*. ISCEV standard for clinical pattern electroretinography (PERG): 2012 update. *Documenta Ophthalmologica* **126**, 1–7, https://doi.org/10.1007/s10633-012-9353-y (2013).
11. Holder, G. E. *et al*. ISCEV standard for clinical pattern electroretinography-2007 update. *Documenta Ophthalmologica* **114**, 111–116 (2007).
12. Bach, M. *et al*. Standard for pattern electroretinography. *Documenta Ophthalmologica* **101**, 11–18 (2000).
13. Holder, G. Electrophysiological assessment of optic nerve disease. *Eye* **18**, 1133–1143, https://doi.org/10.1038/sj.eye.6701573 (2004).
14. Thompson, D. A. *et al*. ISCEV standard for clinical pattern electroretinography (2024 update). *Documenta Ophthalmologica* **148**, 75–85 (2024).
15. Canedo, C., Fernández, I., Coco, R. M., Cuadrado, R. & Rueda, C. Novel modeling proposals for the analysis of pattern electroretinogram signals. In *Statistical Methods at the Forefront of Biomedical Advances*, 255–273, https://doi.org/10.1007/978-3-031-32729-2_11 (Springer, 2023).
16. Fernández, I., Cuadrado Asensio, R., Larriba, Y., Rueda, C. & Coco Martín, R. M. A comprehensive dataset of pattern electroretinograms for ocular electrophysiology research: The PERG-IOBA dataset (version 1.0.0), https://doi.org/10.13026/d24m-w054 (2024).
17. Goldberger, A. L. *et al*. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **101**, e215–e220, https://doi.org/10.1161/01.CIR.101.23.e215 (2000).
18. Robson, J. Guidelines for calibration of stimulus and recording parameters used in clinical electrophysiology of vision. *Documenta Ophthalmologica* **107**, 185–193 (2003).
19. McCulloch, D. L. *et al*. ISCEV guidelines for calibration and verification of stimuli and recording instruments (2023 update). *Documenta Ophthalmologica* 1–12, https://doi.org/10.1007/s10633-023-09932-z (2023).
20. Bubl, E. *et al*. Elevated background noise in adult attention deficit hyperactivity disorder is associated with inattention. *PLoS One* **10**, e0118271, https://doi.org/10.1371/journal.pone.0118271 (2015).

Acknowledgements

This work was supported by a biomedical research grant from the Eugenio Rodriguez Pascual Foundation, awarded in the 2021.

Author contributions

Data acquisition: R.C.-A. and R.C.-M.; Creation and maintenance of the dataset: I.F. and Y.L.; Verified PERG signal quality: R.C.-A. and R.C-M.; Diagnosis and classification assignment: R.C.-A. and R.C.-M.; Supervision of the project: I.F., C.R. and R.C.-M.; Manuscript preparation: I.F.; Critical comments and revision of manuscript: all authors.

Competing interests

The authors declare no competing interests. The funder had no role in the design of the study; in the collection of the data; in the writing of the manuscript, or in the decision to publish the results.

Additional information

Correspondence and requests for materials should be addressed to I.F. or Y.L.

Reprints and permissions information is available at www.nature.com/reprints.

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