Baseline Detrending for the Photopic Negative Response

Jessica Tang1,2, Flora Hui1,2, Michael Coote1,2, Jonathan G. Crowston1,2, and Xavier Hadoux1,2

1 Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia
2 Ophthalmology, Department of Surgery, The University of Melbourne, Melbourne, Australia

Correspondence: Jessica Tang, Glaucoma Research Unit, Centre for Eye Research Australia, Level 7, Peter Howson Wing, 32 Gisborne St, East Melbourne, VIC 3002, Australia. e-mail: j.tang8@student.unimelb.edu.au

Purpose: The photopic negative response (PhNR) of the light-adapted electroretinogram (ERG) holds promise as an objective marker of retinal ganglion cell function. We compared baseline detrending methods to improve PhNR repeatability without compromising its diagnostic ability in glaucoma.

Methods: Photopic ERGs were recorded in 20 glaucoma and 18 age-matched control participants. A total of 50 brief, red-flashes (1.6 cd.s/m²) on a blue background (10 photopic cd/m²) were delivered using the RETeval device. Detrending methods compared were: (1) increasing the high-pass filter from 1 to 10 Hz and (2) estimating and removing the trend with an increasing polynomial (order from 1–10) applied to the prestimulus interval, prestimulus and postsignal interval, or the whole ERG signal. Coefficient of repeatability (COR%), unpaired Student’s t-test, and area under the receiver operating characteristic curve (AUC) were used to compare the detrending methods.

Results: Most detrending methods improved PhNR test–retest repeatability compared to the International Society for Clinical Electrophysiology of Vision (ISCEV) recommended 0.3 to 300 Hz band-pass filter (COR% ≤ 200%). In particular, detrending with a polynomial (order 3) applied to the whole signal performed the best (COR% ≤ 44%) while achieving similar diagnostic ability as ISCEV band-pass (AUC 0.74 vs. 0.75, respectively). However, over-correcting with higher orders of processing can cause waveform distortion and reduce diagnostic ability.

Conclusions: Baseline detrending can improve the PhNR repeatability without compromising its clinical use in glaucoma. Further studies exploring more complex processing methods are encouraged.

Translational Relevance: Baseline detrending can significantly improve the quality of the PhNR.

Introduction

The photopic negative response (PhNR) is a component of the full-field electroretinogram (ERG) and holds promise as an objective, noninvasive functional marker in glaucoma and other optic neuropathies. However, like all bioelectrical signals, the ERG, and, therefore, the PhNR, can be contaminated by baseline drift, which is a low-frequency noise arising from sources, such as the acquisition equipment and electrode impedance. This negatively impacts on the accurate estimation of the PhNR, especially when measured with respect to the baseline.

To our knowledge, there is no comparative analysis of baseline trend corrections for the ERG. The International Society for Clinical Electrophysiology of Vision (ISCEV) recommends a high-pass filter cutoff of 0.3 Hz for full-field ERG recordings, and the recently released extended protocol for the PhNR suggests this cutoff could be lower to minimize distortion and signal attenuation. However, while a lower limit to the high-pass filter may reduce some low frequency noise it would be unable to completely remove baseline trend containing higher frequency components. Further, linear detrending has been applied in some studies and ERG recording...
systems, but this approach also would be inadequate in dealing with nonlinear drifts.

In this study, we sought to compare methods of baseline correction to improve the PhNR repeatability while maintaining its diagnostic ability in glaucoma. Two main approaches were applied: (1) increasing the high-pass filter cutoff and (2) estimating the trend using a polynomial of increasing order and subtracting it from each trace. These methods were compared to applying the ISCEV band-pass filter (0.3–300 Hz) only. We showed that appropriate baseline correction can be achieved readily and should be considered in all future studies investigating the role of the PhNR in glaucoma and other inner retinal disorders.

**Methods**

The research followed the tenets of the Declaration of Helsinki and was approved by the Human Research Ethics Committee at the Royal Victorian Eye and Ear Hospital, Melbourne, Australia (13/1121H). Participants were recruited from two sites: public outpatient clinics at the Royal Victorian Eye and Ear Hospital or a private ophthalmology clinic (Melbourne Eye Specialists, Australia). Informed consent was obtained from all participants before examination. Exclusion criteria for all participants were visual acuity worse than 6/12, diabetes, ophthalmic surgery in the last 6 months, and other eye conditions (except visually insignificant cataracts).

**Glaucoma Patients**

A total of 20 glaucoma patients (mean age ± SD, 69 ± 8 years) were recruited. All patients had glaucomatous disc appearance assessed by glaucoma experts and reproducible visual field (VF) defects defined as $P < 0.05$ for pattern standard deviation (PSD) or an abnormal glaucoma hemifield test (GHT; 24-2, Humphrey Field Analyzer; Carl Zeiss Meditec, Inc., Dublin, CA). All participants had intraocular pressure ≤ 21 mm Hg with or without treatment. Baseline retinal nerve fiber layer thickness (RNFL) measurement was obtained with spectral-domain optical coherence tomography (SD-OCT; Spectralis, Heidelberg Engineering, Dossenheim, Germany; or RS-3000 Advance, NIDEK, Aichi, Japan). Where both eyes were eligible for the study, the eye with the better visual field mean deviation (MD) was chosen for inclusion. Average MD was $-5.7 ± 4.4$ dB and RNFL thickness was $74 ± 14$ μm. All patients returned for repeat ERG testing within 1 month.

**Control Participants**

All 18 healthy controls included in the study (age, 72 ± 6 years) had normal slit-lamp inspection, tonometry, funduscopy, VF, and SD-OCT imaging. Where both eyes were eligible for the study, one was chosen at random. A total of 14 participants returned for repeat ERG testing within 1 month.

**Full-Field ERG Recording**

Pupils were dilated using tropicamide 0.5% (Mydriacyl; Alcon Laboratories, NSW, Australia). A DTL-like electrode (22/1 dtex; Shieldex Trading, Palmyra, NY) was placed along the lower lid margin with reference and ground gold-cup skin electrodes (Grass Technologies, Astro-Med Inc., West Warwick, RI) placed on the temple and forehead, respectively. Participants were adapted to ambient light in the clinic for at least 10 minutes, followed by preadaptation to the blue background light (photopic 10 cd/m²) for 1 minute before testing. Monocular, full-field stimulation was produced using the RETeval (LKC Technologies, Inc., Gaithersburg, MD). The stimulus consisted of brief (<4 ms), red-flashes (1.6 cd.s/m²) on a steady blue background (photopic 10 cd/m²); 50 flashes were delivered at a frequency of 2 Hz. The recording was repeated if there were excessive artefacts or noise. Photopic luminance was calibrated using an International Light photometer (model ILT-1700; International Light Technologies, Newburyport, MA). Signals were acquired at a sampling frequency of 2 kHz and extracted for offline processing using Matlab (R2016b; Mathworks, Natick, MA).

**Baseline Detrending Methods**

**High-Pass Filter (HP)**

An HP filter attenuates signals with frequencies lower than the specified cutoff. The PhNR was shown to contain low temporal frequencies near approximately 11 Hz with smaller contributions from frequencies in the 2 to 8 Hz range. Therefore, a comparison of increasing the high-pass cutoff from 1 to 10 Hz (in 1 Hz increments) was evaluated. The ISCEV low-pass cutoff of 300 Hz was retained.

**Trend Estimation Using Polynomial Fitting**

Before further processing, the ISCEV band-pass (0.3–300 Hz) was applied. Baseline trend then was
estimated from each trace using a polynomial of increasing order from 1 to 10 (in 1° steps); that is, order 1 is a linear regression, order 2 is quadratic, order 3 is cubic, and so forth. The estimated trend was removed from the trace. The polynomial was fitted to specific parts of the ERG waveform including (Fig. 1):

1. Prestimulus interval (PS): sampling points from −100 ms before the flash up to the time of flash (0 ms);
2. Prestimulus and postsignal interval (PP): prestimulus interval (−100 to 0 ms) and sampling points after the PhNR to the end of recording (200 to 375 ms); and
3. Whole signal (WS): sampling points from prestimulus signal to the end of recording (−100 to 375 ms).

**Signal Analysis**

After baseline correction, outlier traces were removed using a multivariate analysis based on robust principal component analysis (rPCA). In brief, we removed outlier traces, which were defined in the two-dimensional rPCA space as traces whose Mahalanobis distance was >2.4. This distance corresponds to the square root of the 5% upper tail of the \( \chi^2 \) distribution with 2\( ^\circ \) of freedom. The remaining traces were averaged, and the following parameters were extracted from the average trace (Fig. 1): a-wave, defined as the trough in the time window 0 to 30 ms and measured from the detrended baseline; b-wave, defined as the maximum after the a-wave and measured from a-wave trough to the maximum; and the PhNR, defined as the trough in the time window 60 to 90 ms after the b-wave, measured from the detrended baseline. As the PhNR trough can be broad, we averaged 11 consecutive sample points centered at the trough (i.e., 5.5 ms on either side of the minimum) to obtain the PhNR amplitude. Note, the a-wave and PhNR amplitudes were measured with respect to the “zero” baseline, even if there was drift in the waveform (see Fig. 2A).

**Statistical Analysis**

The 95% coefficient of repeatability was calculated, which is where we expect 95% of intersession measurement differences to lie. Unpaired \( t \)-tests were performed between glaucoma and control groups for each processing method. The area under the receiving operator characteristic curve (AUC) was derived to assess the performance of the PhNR to discriminate between glaucoma and control participants.

**Results**

Figure 2 shows a representative of each detrending method with the individual ERG traces (gray) and the average ERG waveform (red) after ISCEV band-pass filtering without additional baseline correction (Fig. 2A), HP filtering (Fig. 2B), WS polynomial (Fig. 2C), PP polynomial (Fig. 2D), and PS polynomial (Fig. 2E). When only ISCEV band-pass filtering was used, a clear downward trend on the average ERG waveform was observed (Fig. 2A). Except for prestimulus polynomial at higher orders, each method qualitatively reduced the general baseline trend. However, there were some caveats to the processing methods. Increasing the HP filter and WS polynomial above degree 4 led to significant distortion around the prestimulus baseline and PhNR (Figs. 2B, 2C). Increasing the order of PP polynomial progressively increased signal flaring of individual traces around the PhNR (Fig. 2D). When a PS polynomial above degree 2 distorted the ERG waveform completely (Fig. 2E), and thus, analysis was capped for this method.

Figure 3 illustrates selected traces from Figure 2A, demonstrating that while the average waveform shows a linear trend, it is apparent that some individual traces are nonlinear (blue and red bold
The estimated trends for the selected polynomial methods are represented by the dashed lines in Figures 3A–E, with the detrended waveform for each method shown in Figures 3F–J. Qualitatively, detrending by applying a linear regression to the whole signal (Fig. 3A), to the prestimulus and postsignal intervals (Fig. 3C), or the prestimulus interval (Fig. 3E) did not adequately estimate the trend. On the other hand, PP and WS polynomial order 3 appeared to provide a better fit (Figs. 3B, 3D).

We examined the effect of baseline detrending on the a-wave, b-wave, and PhNR amplitudes (Fig. 4). The a-wave and PhNR amplitudes obtained from the ISCEV band-pass were larger compared to the other methods (also see the Table). However, this was likely related to poor baseline correction especially when there was negative drift (Fig. 2A). There was a trend for increasing a-wave amplitude with increasing HP filtering (Fig. 4A) and WS polynomial (Fig. 4B) while the amplitudes remained similar for PP polynomial (Fig. 4C) and PS polynomial (Fig. 4D). There were no amplitude differences for the b-wave across all methods (Figs. 4E–H). Meanwhile, increasing the HP filtering and WS polynomial gradually reduced the differences in PhNR amplitude between the control and glaucoma groups (Figs. 4I, 4J). The effect on the PhNR was more variable with increasing PP polynomial processing (Fig. 4K), while the separation between groups appeared the largest for PS polynomial degrees 1 and 2 (Fig. 4L).

Figure 5 illustrates the effect of increasing the order of processing on the COR% for the a-wave, all methods improved the COR% to approximately ±50% compared to ISCEV band-pass filtering without additional baseline correction (COR% ± 145%, Figs. 5A–D). The processing methods had no effect on b-wave repeatability (COR% ± 34 to 37%, Figs. 5E–H). The PhNR COR% reduced from ±200% with ISCEV band-pass only to ±38% and ±31% with the highest order of processing for HP filter (Fig. 5I) and WS polynomial (Fig. 5J), respectively. For PP polynomial, PhNR COR% reduced initially to ±93%, but increased again above degree 6 (Fig. 5K). This is most likely due to the signal flaring observed around the PhNR (Fig. 2D). For PS polynomial, repeatability improved for degree 1 (COR% ± 78%, Fig. 5L), but increased again for degree 2.

The Table is a selected comparison of ISCEV band-pass only with the lowest order and the best performing order for each detrending method (except for PS polynomial) based on COR% and AUC. All methods shown had similar AUC ranging from 0.69 to 0.78, with overlapping confidence intervals. However, the ISCEV band-pass only had significantly larger COR% for the a-wave and PhNR compared to other methods, indicating significant measurement.
variability. Compared to their lowest order, increasing to HP4 and WS3 reduced the PhNR variability by approximately half. On the other hand, there was no difference with increasing the order for PP polynomial. Overall, WS3 afforded the best repeatability with the lowest PhNR COR% ± 44%.

**Discussion**

To our knowledge, this is the first study to compare methods for baseline correction for the ERG with an emphasis on PhNR in glaucoma. While we focused on glaucoma, we believed our approach to baseline correction is applicable to other inner retinal disorders, particularly those with longitudinal measurements. Based on our comparison, we would recommend whole signal polynomial degree 3 for baseline detrending as it resulted in the best combination of PhNR repeatability (COR ± 44%) and diagnostic ability (AUC 0.74). This is a significant improvement compared to our earlier study, where we reported COR ± 148% for the baseline PhNR using the linear drift removal on an Espion system (Diagnosys LLC, Lowell, MA), although fewer traces also were recorded. PS polynomial, even at the lowest order, performed poorly as the trend was extrapolated from limited data points. On the other hand, trends were interpolated with PP and WS polynomials, which provided better approximations. However, a linear regression in both cases (i.e., PP1 and WS1) did not improve repeatability as it would inadequately estimate nonlinear drifts (Fig. 3).

Few studies on the PhNR describe their approach to baseline correction, although it is possible that linear drift removal available in certain recording systems may have been used without specification. Some have acknowledged baseline trend without applying further correction. Others have used manual approaches, such as resetting the amplifier to zero before each stimulus presentation, to minimize drift or rejecting signals where the baseline did not return after the PhNR trough. While these measures may correct some trend, they are time-consuming and subjective. Preiser et al. estimated the trend from the prestimulus data only (time window 45–0 ms) and subtracted this from the whole trace. Rangaswamy et al. performed a similar correction over a shorter period (17 ms before the stimulus). However, we showed that linear regression applied either to the prestimulus or to the prestimulus and post-signal baseline resulted in less repeatable measurements compared to whole signal polynomial degree 3 (Fig. 5).

Baseline trend correction requires a compromise between trend removal, which improves signal repeatability, and waveform distortion which reduces discriminatory ability. Using discrete wavelet analysis, Kundra et al. showed that the PhNR was characterized by components located near the 11 Hz band (approximate range, 8–16 Hz), with minor contribution from lower frequencies (approximate range, 2–8 Hz). Thus, setting a high pass above 2 Hz attenuated the PhNR by removing these lower frequencies (Fig. 4I). However, we compared up to 10 Hz since this was below the major component described by Kundra et al., and also to determine
the point at which the reduction in amplitude would no longer separate the groups (above 4 Hz). Similarly, significant waveform distortion occurred above degree 3 for WS polynomial. In HP and WS, while increasing the order of processing improved COR% (Fig. 5I, 5J), there also was a corresponding decrease in PhNR amplitude (Figs. 4I, 4J) due to attenuation and distortion, respectively. Further, this reduction in PhNR also would have the effect of improving COR% by simply limiting the range of PhNR values.

Therefore, care must be taken as overprocessing ERG signals would reduce its clinical utility. The other detrending methods (PP and PS polynomial) retained the separation between groups at the cost of poor repeatability (Figs. 4, 5) and, thus, limiting its use for longitudinal monitoring.

An important consideration with detrending using whole signal polynomial is that there still may be some overestimation of the trend. For this reason, we chose the lowest polynomial order possible to achieve

---

**Table.** Selected Processing Methods to Compare PhNR COR%, Amplitude and AUC

<table>
<thead>
<tr>
<th>Processing Method</th>
<th>a-wave COR% (95% CI)</th>
<th>b-wave COR% (95% CI)</th>
<th>PhNR COR% (95% CI)</th>
<th>Mean PhNR Amplitude, µV (SD) Control</th>
<th>Mean PhNR Amplitude, µV (SD) Glaucoma</th>
<th>AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISCEV band-pass only</td>
<td>145 (101–190)</td>
<td>37 (26–49)</td>
<td>200 (135–265)</td>
<td>16.8 (10.2)</td>
<td>8.8 (10.4)</td>
<td>0.75 (0.59–0.90)</td>
</tr>
<tr>
<td>HP1</td>
<td>80 (55–104)</td>
<td>35 (24–46)</td>
<td>112 (76–48)</td>
<td>12.5 (6.0)</td>
<td>8.0 (4.8)</td>
<td>0.72 (0.56–0.89)</td>
</tr>
<tr>
<td>HP4</td>
<td>40 (28–52)</td>
<td>36 (25–47)</td>
<td>51 (35–68)</td>
<td>11.6 (3.0)</td>
<td>9.2 (3.4)</td>
<td>0.69 (0.52–0.86)</td>
</tr>
<tr>
<td>WS1</td>
<td>46 (32–60)</td>
<td>36 (25–47)</td>
<td>82 (55–107)</td>
<td>14.1 (6.0)</td>
<td>9.4 (4.8)</td>
<td>0.73 (0.57–0.89)</td>
</tr>
<tr>
<td>WS3</td>
<td>40 (28–52)</td>
<td>35 (24–46)</td>
<td>44 (30–59)</td>
<td>12.0 (3.2)</td>
<td>9.1 (3.4)</td>
<td>0.74 (0.58–0.90)</td>
</tr>
<tr>
<td>PP1</td>
<td>58 (40–76)</td>
<td>36 (25–47)</td>
<td>102 (70–133)</td>
<td>15.7 (8.7)</td>
<td>9.3 (6.4)</td>
<td>0.72 (0.56–0.88)</td>
</tr>
<tr>
<td>PP3</td>
<td>49 (34–64)</td>
<td>35 (24–46)</td>
<td>94 (64–125)</td>
<td>14.7 (6.3)</td>
<td>8.9 (5.1)</td>
<td>0.78 (0.64–0.93)</td>
</tr>
<tr>
<td>PS1</td>
<td>44 (30–57)</td>
<td>36 (25–46)</td>
<td>78 (53–104)</td>
<td>15.9 (7.2)</td>
<td>10.1 (5.0)</td>
<td>0.74 (0.58–0.90)</td>
</tr>
</tbody>
</table>

HP1, high-pass filter 1 Hz; HP4, high-pass filter 4 Hz; WS1, whole signal polynomial degree 1; WS3: whole signal polynomial degree 3, PP1: prestimulus and postsignal polynomial degree 1; PP3, prestimulus and postsignal polynomial degree 3; PS1, prestimulus polynomial degree 1.
a fair COR% without excessive distortion of the waveform. Therefore, while the COR of 44% for the baseline PhNR measure is significantly better than previously reported,\textsuperscript{12,14} it still can be improved upon. Further, the length of the ERG recording can influence the trend estimate. That is, the shorter the sweep length, the more influence the ERG signal will have on the estimate leading to greater waveform distortion. Therefore, we recommend using the longest sweep length that is achievable to reduce this effect.

For easier implementation in the clinic, we used a portable, hand-held ERG device in this study. Our signals may have contained more low frequency artefacts than those recorded using conventional table-top ERG systems as the device was not grounded. However, adequate ERGs were recorded by ensuring proper skin preparation and minimizing external electrical interferences by switching off other electrical devices in the room. In addition, measures were repeated when the trace shown on the RETeval device demonstrated excessive noise or distortion.

In summary, we highlighted the importance of proper baseline correction in clinical ERG recordings. Of the methods tested, correction using whole signal polynomial degree 3 provided the optimal balance in reducing baseline drift and improving PhNR repeatability while still retaining the clinical differences between control and glaucoma participants. This should be incorporated into signal processing protocols to improve PhNR recording and enhance its clinical use in diagnostics and longitudinal monitoring in glaucoma and other inner retinal disorders.

Acknowledgments

Supported by the Australian Government Research Training (RTP) Scholarship and the Jean Miller Foundation. The Centre for Eye Research Australia receives Operational Infrastructure Support from the Victorian Government.

Disclosure: J. Tang, None; F. Hui, None; M. Coote, None; J.G. Crowston, None; X. Hadoux, None

References


25. Moss HE, Park JC, McAnany JJ. The photopic negative response in idiopathic intracranial hy-
