



DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU



HARVARD LIBRARY
Office for Scholarly Communication

Electrophysiological Studies in Thyroid Associated Orbitopathy: A Systematic Review

The Harvard community has made this
article openly available. [Please share](#) how
this access benefits you. Your story matters

Citation	lao, Tiara W. U., Shi Song Rong, An Ni Ling, Mårten E. Brelén, Alvin Lerrmann Young, and Kelvin K. L. Chong. 2017. "Electrophysiological Studies in Thyroid Associated Orbitopathy: A Systematic Review." <i>Scientific Reports</i> 7 (1): 12108. doi:10.1038/s41598-017-11998-0. http://dx.doi.org/10.1038/s41598-017-11998-0 .
Published Version	doi:10.1038/s41598-017-11998-0
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:34492266
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

SCIENTIFIC REPORTS



OPEN

Electrophysiological Studies in Thyroid Associated Orbitopathy: A Systematic Review

Tiara W. U. Iao¹, Shi Song Rong^{1,3}, An Ni Ling¹, Mårten E. Brelén^{1,2}, Alvin Lerrmann Young^{1,2} & Kelvin K. L. Chong^{1,2}

Received: 12 August 2016

Accepted: 31 August 2017

Published online: 21 September 2017

Dysthyroid optic neuropathy (DON) is the commonest cause of blindness in thyroid associated orbitopathy (TAO). While diagnosis remains clinical, objective tests for eyes with early or equivocal findings are lacking. Various electrophysiological studies (EPS) have been reported, yet the types and parameters useful for DON remain inconclusive. We performed a systematic literature search in MEDLINE, EMBASE and the Cochrane databases via the OVID platform up to August 20, 2017. 437 records were identified for screening and 16 original studies (1327 eyes, 787 patients) were eligible for review. Pattern visual evoked potential (pVEP) was the most frequently studied EPS. Eyes of TAO patients with DON showed delayed P100 latencies, decreased P100 amplitudes or delayed N75 latencies during pVEP, compared to those without or healthy controls. Due to study heterogeneity, no quantitative analysis was possible. This review highlights the most common type (pVEP) and useful parameters (P100 latency and amplitude) of EPS, and supports further research on them using standardized testing conditions.

Dysthyroid optic neuropathy (DON) is the commonest blinding complication affecting 4–8% of patients with thyroid associated orbitopathy (TAO), with an estimated annual incidence of 0.6–1.3 cases per 100,000 population^{1,2}. While exact mechanisms of DON remain elusive, apical compression by enlarged extraocular muscles and/or fat (crowding)^{3,4}, ischemia due to increased retrobulbar pressure, mechanical stretch due to proptosis and perineural inflammation have been proposed⁵. Empirical treatments including surgical apical decompression, systemic steroids and orbital radiotherapy are often effective to restore vision. It is thus imperative to confirm diagnosis early to avoid irreversible visual loss and unnecessary treatments in alternative causes⁵. Ancillary tests, for example optical coherence tomography⁶, orbital imaging⁷ and electrophysiological studies (EPS), including visual evoked potential (VEP) and electroretinogram (ERG) were attempts to objectively assess the presence, predict the development and correlate with the severity of DON^{8–23}. However, methodologies and results were heterogeneous across studies. In this systematic review, we studied published reports on EPS in DON.

Results

Characteristics of included studies. Our search yielded 768 reports from databases. After removing 331 duplicated records, we studied 437 publications. Among them 415 studies were found to be irrelevant according to our eligibility criteria (see Methods below). For the remaining 22 studies, 8 reports were excluded: 1 report on a duplicated study population²⁴, 1 case report²⁵, 1 review article²⁶, and 5 studies with irrelevant or insufficient results^{1,27–30}. 2 additional studies were identified from manual search of references^{8,9}. 16 studies were finally included for the systematic review (Fig. 1)^{8–23}. No clinical trial was identified.

The pooled sample included 787 patients (1,327 eyes). The age of patients with DON ranged from 14 to 77 years old¹³. VEP was used in 14 studies^{8–13,15–22}. 3 studies tested pERG^{12,14,23}. No study was found on flash or multifocal ERG (Table 1).

Phenotypic definition of subjects. Study populations were phenotypically defined as patients with DON, TAO, or healthy subjects. Clinical features of DON included optic disc swelling, relative afferent pupillary defect,

¹Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China.

²Department of Ophthalmology and Visual Sciences, The Prince of Wales Hospital, Hong Kong, China. ³Present address: Harvard Medical School, and Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA. Correspondence and requests for materials should be addressed to K.K.L.C. (email: chongkamlung@cuhk.edu.hk)

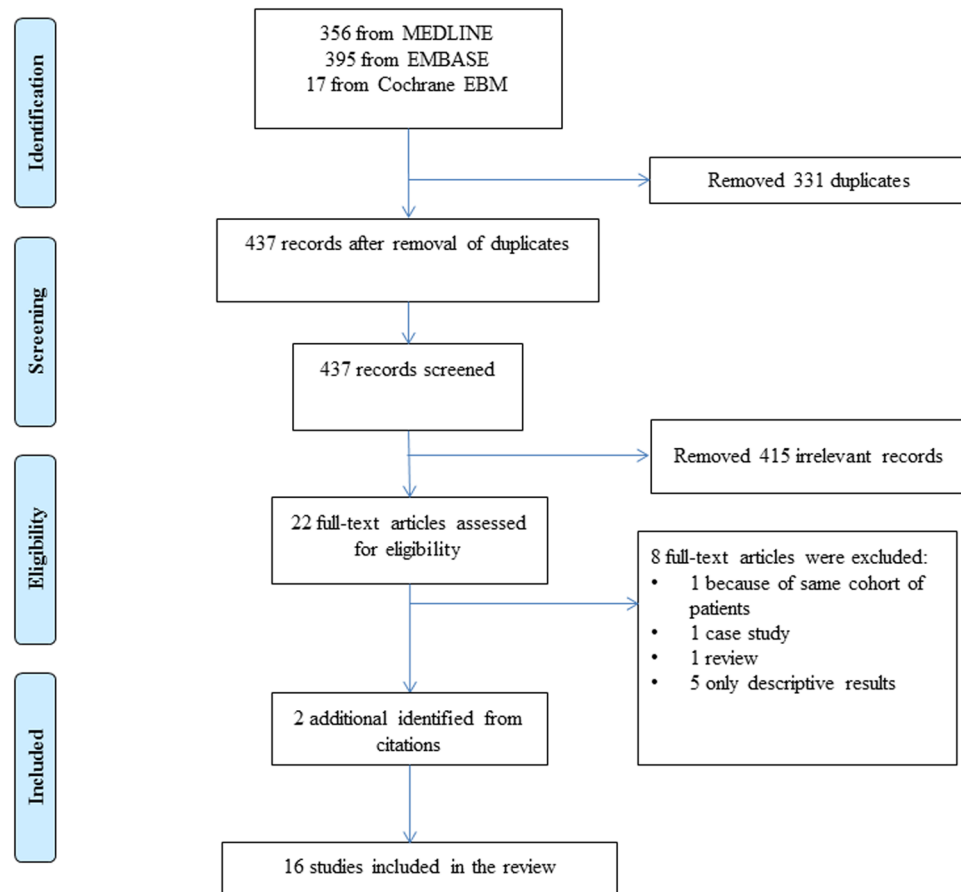


Figure 1. Flowchart of the literature search and study selection process.

decreased visual acuity, impaired color vision, and visual field defect³¹. DON was considered “definite” if there was optic disc swelling or 2 of the four other clinical features above without alternative explanation in a patient with TAO³². Subclinical or “equivocal” DON was proposed by some as the presence of optic nerve dysfunction in TAO patients without the full-blown clinical features of DON¹⁵, often identified by abnormal electrophysiological changes^{8,12,13,16,18,22}.

Flash VEP (fVEP) in TAO & DON. Only 2 earlier studies reported use of fVEP in TAO and DON patients (Tables 2 and 3)^{9,11}. Alteration in P2 amplitude was reported in clinically evident DON¹¹. Tsaloumas *et al.* found significantly smaller P2 amplitude in DON eyes which improved either after orbital decompressions (6.83 ± 0.92 vs. 13.12 ± 1.65 μV ; $P < 0.05$) or 2 weeks of high-dose systemic steroids (7.00 ± 1.10 vs. 9.61 ± 1.43 μV ; $P < 0.05$ ¹¹). However, treatment-related improvement was not shown in Setala’s study after decompression or radiotherapy⁹.

Pattern VEP (pVEP) in TAO & DON. *Comparison of pVEP results in DON, TAO, and normal controls.* P100 latency, P100 amplitude, and N75 latency were compared between DON and normal controls in 3 studies^{10,11,17}. An increase in P100 latency of patients with DON was reported by Shawkat *et al.* (115.2 ± 5.7 vs. 103.2 ± 4.3 ms, $P = 0.0005$)¹⁰, Tsaloumas *et al.* (129.2 ± 7.1 vs. 108.2 ± 1.2 ms, $P < 0.005$)¹¹, and Ambrosio *et al.* ($P < 0.0001$)¹⁷. A decrease in P100 amplitude was found in eyes with DON compared to control by Tsaloumas *et al.* (3.67 ± 0.81 vs. 8.97 ± 0.59 μV , $P < 0.001$)¹¹ and Ambrosio *et al.* ($P < 0.0001$)¹⁷.

Comparisons between TAO eyes with or without DON were reported in 3 studies^{10,11,15}. Significant increases in P100 latency in eyes with DON were shown by Shawkat *et al.* (115.2 ± 5.7 vs. 110.3 ± 5.1 ms, $P = 0.043$)¹⁰, Tsaloumas *et al.* (129.2 ± 7.13 vs. 111 ± 1.86 ms, $P < 0.005$)¹¹, and Rutecka-Debniak *et al.* (124.4 ± 15.4 vs. 114.9 ± 11.2 ms, $P = 0.05$)¹⁵. Significant decreases in P100 amplitude in DON patients were reported by Shawkat *et al.* (11.9 ± 6.4 vs. 21.2 ± 9.7 μV , $P = 0.018$)¹⁰ and Tsaloumas *et al.* (3.67 ± 0.81 vs. 8.55 ± 0.73 μV , $P < 0.001$)¹¹. Moreover, the mean N75 latency of eyes with DON was also increased (90.0 ± 17.9 vs. 80.3 ± 14.7 ms, $P = 0.01$)¹⁵.

Five studies reported significant increases in P100 and N75 latencies comparing eyes from TAO patients without DON to healthy eyes (Table 2)^{8,12,13,16,18}. Wijngaarde *et al.* first reported significant increase in P100 latency of TAO to healthy eyes ($P < 0.01$)⁸. Spadea *et al.* (126.7 ± 10.7 vs. 118.5 ± 5.7 ms, $P < 0.05$)¹², Salvi *et al.* (105.6 ± 0.5 vs. 102.0 ± 0.5 ms, $P < 0.001$)¹³, Acaroglu *et al.* (122.0 ± 14.4 vs. 105.9 ± 7.7 ms, $P = 0.0004$)¹⁶, and Pawlowski *et al.* (106.2 ± 4.4 vs. 102.4 ± 2.7 ms, $P < 0.01$) also found increased P100 latencies in eyes from TAO subjects without clinical evidence of DON when compared with controls¹⁸. In addition, Pawlowski *et al.* found an increase in N75 latency (79.0 ± 3.7 vs. 73.9 ± 2.8 ms, $P < 0.001$)¹⁹, while Spadea *et al.* showed a decrease in P100 amplitude

	Author (year)	Country/region	Study Design	EPS tested	ISCEV	Sample size	TAO Age range (mean)	DON Age range (mean)	Subgroups	Outcomes
1	Wijngaarde <i>et al.</i> ⁸	The Netherlands	Prospective case series	pVEP	✗	53	n.a.	—	TAO/Control	Correlation between P100 (latency) and VA
2	Setala <i>et al.</i> ⁹	Finland	Prospective case series	fVEP	✗	31	28–66	—	TAO	The differences in N60 & P120 (amplitude & latency) before and after TAO treatment
3	Shawkat <i>et al.</i> ¹⁰	England	Prospective case-control study	pVEP	✗	20	37–62 (47.3)	37–62 (47.3)	DON/TAO/Control	The differences in P100 (amplitude & latency) among DON, TAO and control
4	Tsaloumas <i>et al.</i> ¹¹	UK	Retrospective case series	fVER, pVEP	✗	43	23–68 (45.1)	26–73 (49.1)	DON/TAO/Control	The differences in P2 (amplitude & latency) among DON, TAO and control; The differences in P2 (amplitude & latency) before and after DON treatment
5	Spadea <i>et al.</i> ¹²	Italy	Prospective case series	pVEP, pERG	✗	49	(57.2)	—	TAO/Control	The differences in P100 (amplitude & latency) between TAO and control
6	Salvi <i>et al.</i> ¹³	Italy	Retrospective case series	pVEP	✗	117	14–77 (45.3)	—	TAO/Control	The differences in P100 (amplitude & latency) between TAO and control
7	Genovesi-Ebert <i>et al.</i> ¹⁴	Italy	Prospective case series	pERG	✗	44	(51.9)	—	TAO/Control	The significant difference in amplitude between TAO and control
8	Rutecka-Debniak <i>et al.</i> ¹⁵	Poland	Prospective case series	pVEP	✗	110	18–74	18–74	DON/TAO	The differences in N75 & P100 (latencies) between DON and TAO; The differences in N75 & P100 (latencies) before and after DON & TAO treatment
9	Acaroglu <i>et al.</i> ¹⁶	Turkey	Prospective case series	pVEP	✗	31	20–65 (41.7)	—	TAO/Control	The difference in P100 (latency) between TAO and control; Correlation between P100 (latency) and CAS
10	Ambrosio <i>et al.</i> ¹⁷	Italy	Prospective case-control study	pVEP	✗	63	(36.3)	(42.5)	DON/TAO/Control	The differences in P100 (amplitude & latency) between DON and control
11	Pawlowski <i>et al.</i> ¹⁸	Poland	Prospective case series	pVEP	✓ 2004 ⁴⁹	27	35.6 ± 11.3	—	TAO/Control	The differences in N75 (latency) & P100 (amplitude & latency) between TAO and control; Correlation between N75 & P100 (latencies) and IOP & degree of proptosis
12	Liao <i>et al.</i> ¹⁹	China Taiwan	Retrospective case series	pVEP	n.a.	22	—	30–76 (58.4)	DON	The differences in P100 (latency) before and after DON treatment
13	Wei <i>et al.</i> ²⁰	China Taiwan	Prospective case series	pVEP	n.a.	76	22–79 (46.7)	—	TAO	Correlation between P100 (latency) and VA, degree of proptosis, color test, visual field test, OCT and extraocular muscles measurements
14	Lipski <i>et al.</i> ²¹	Germany	Retrospective case series	pVEP	✓ 2004 ⁴⁹	15	—	43–76 (55)	DON	The differences in P100 (amplitude & latency) before and after DON treatment
15	Perez-Rico <i>et al.</i> ²²	Spain	Prospective case series	mfVEP	—	65	47.5 ± 11.5	—	TAO/Control	The difference in latency between TAO and control
16	Pawlowski <i>et al.</i> ²³	Poland	Prospective case-control study	pERG	✓ 2012 ⁴¹	21	24–55 (36)	—	TAO/Control	The difference in P50 amplitude between TAO and control

Table 1. Characteristics of included studies in the systematic review. DON = dysthyroid optic neuropathy; EPS = electrophysiological studies; fVEP = flash visual evoked potential; ISCEV = International Society for the Clinical Electrophysiology of Vision Standard; mfVEP = multifocal visual evoked potential; n.a. = not available; OCT = optical coherence tomography; pERG = pattern electroretinography; pVEP = pattern visual evoked potential; TAO = thyroid associated orbitopathy; VA = visual acuity.

No.	Author (year)	Age range (mean)			Sample size (eyes)			VEP outcome	Mean \pm SD		
		DON	TAO	Control	DON	TAO	Control		DON	TAO	Control
fVEP											
4	Tsaloumas <i>et al.</i> ¹¹	26–73 (49.1)	23–68 (45.1)	22–68 (46.1)	8 (13)	15 (30)	20 (40)	P2 amplitude (μ V)	6.83 \pm 0.92 ^{†‡}	12.40 \pm 1.05	11.72 \pm 1.16
								P2 latency (ms)	112.0 \pm 4.46	110.1 \pm 2.65	109.6 \pm 2.08
pVEP											
1	Wijngaarde <i>et al.</i> ⁸		n.a.	n.a.		33 (66)	20 (40)	P100 amplitude (μ V)	n.a.	Data n.a.	n.a.
								P100 latency (ms)	n.a.	Data n.a. [‡]	n.a.
3	Shawkat <i>et al.</i> ¹⁰	37–62 (47.3)	37–62 (47.3)	37–62 (47.3)	10 (10)	10 (10)	10 (10)	P100 amplitude (μ V)	11.9 \pm 6.4 [*]	21.2 \pm 9.7	Data n.a.
								P100 latency (ms)	115.2 \pm 5.7 [§]	110.3 \pm 5.1	103.2 \pm 4.3
4	Tsaloumas <i>et al.</i> ¹¹	26–73 (49.1)	23–68 (45.1)	22–68 (46.1)	8 (13)	15 (30)	20 (40)	P100 amplitude (μ V)	3.67 \pm 0.81 ^{†§}	8.55 \pm 0.73	8.97 \pm 0.59
								P100 latency (ms)	129.2 \pm 7.13 [‡]	111 \pm 1.86	108.2 \pm 1.19
5	Spadea <i>et al.</i> ¹²		(57.2)	41–60		9 (18)	40 (40)	P100 amplitude (μ V)	n.a.	3.47 \pm 3.81 [‡]	9.78 \pm 4.26
								P100 latency (ms)	n.a.	126.7 \pm 10.7 [‡]	118.5 \pm 5.7
6	Salvi <i>et al.</i> ¹³		14–77 (45.3)	14–73 (41.8)		88 (172)	29 (56)	P100 amplitude (μ V)	n.a.	10.2 \pm 0.3	11.3 \pm 0.6
								P100 latency (ms)	n.a.	105.6 \pm 0.5 [§]	102.0 \pm 0.5
8	Rutecka-Debniak <i>et al.</i> ¹⁵	18–74	18–74	18–74	12 (21)	13 (26)		N75 latency (ms)	90.0 \pm 17.9 [*]	80.3 \pm 14.7	n.a.
								P100 latency (ms)	124.4 \pm 15.4 [*]	114.9 \pm 11.2	n.a.
9	Acaroglu <i>et al.</i> ¹⁶		20–65 (41.7)	23–65 (42.3)		16 (32)	15 (30)	P100 latency (ms)	n.a.	122.0 \pm 14.4 [§]	105.9 \pm 7.7
10	Ambrosio <i>et al.</i> ¹⁷	(42.5)		(44.3)	14 (28)		20 (40)	P100 amplitude (μ V)	Data n.a. [§]	n.a.	n.a.
								P100 latency (ms)	Data n.a. [§]	n.a.	n.a.
11	Pawlowski <i>et al.</i> ¹⁸		(35.6)	(28.6)		15 (30)	12 (24)	N75 latency (ms)	n.a.	79.0 \pm 3.7 [§]	73.9 \pm 2.8
								P100 amplitude (μ V)	n.a.	7.3 \pm 3.5	6.5 \pm 2.5
								P100 latency (ms)	n.a.	106.2 \pm 4.4 [‡]	102.4 \pm 2.7
13	Wei <i>et al.</i> ²⁰		22–79 (46.7)			76 (151)		P100 latency (ms)	n.a.	103.7 \pm 10.0	n.a.
mfVEP											
15	Perez-Rico <i>et al.</i> ²²		(47.5)	(48.1)		34 (65)	31 (62)	mfVEP latency (ms)	n.a.	6.57 \pm 1.90 [‡]	2.12 \pm 1.72

Table 2. Summary outcomes of observational case series and case-control studies on the use of VEP in DON/TAO. DON = dysthyroid optic neuropathy; fVEP = flash visual evoked potential; mfVEP = multifocal visual evoked potential; ms = millisecond; n.a. = not available; pERG = pattern electroretinography; pVEP = pattern visual evoked potential; TAO = thyroid associated orbitopathy; VEP = visual evoked potential; μ V = microvolts. * $P < 0.05$ compared to TAO without DON, [†] $P < 0.001$ compared to TAO without DON, [‡] $P < 0.05$ compared to Control, [§] $P < 0.001$ compared to Control.

(3.47 \pm 3.81 vs. 9.78 \pm 4.26 μ V, $P < 0.05$) in TAO patients comparing to normal subjects¹². However, the differences between eyes from TAO patients and normal controls in N75 and P100 latencies were insignificant in other studies (Shawkat *et al.*¹⁰ and Tsaloumas *et al.*¹¹). While these TAO patients did not show clinical evidence of DON, abnormal pVEP in particular prolonged P100 latencies may present electrophysiological evidence of early or subclinical optic nerve dysfunction in TAO patients.

Correlation of pVEP latencies with clinical parameters. Four studies investigated correlation between pVEP latencies and clinical parameters (Table 4)^{8,16,18,20}. Wijngaarde *et al.* reported a mild but significant correlation ($r = 0.27$, P value not available) of P100 latency with Snellen visual acuity⁸, while Wei *et al.* reported a similar degree of correlation without statistical significance ($r = 0.278$, $P > 0.05$) using LogMAR visual acuity²⁰. In the latter study, correlation of P100 latency was moderate and statistically significant with total cross-sectional areas of all extraocular rectus muscles (EOM-A) ($r = 0.496$, $P < 0.01$); moderate but insignificant with ratio between the total cross-sectional area of all extraocular rectus muscles and the orbital area ($r = 0.482$, $P > 0.05$), mild and insignificant with total error of 100-hue color sensation ($r = 0.363$, $P > 0.05$) and with mean deviation of retinal sensitivity (MD) in perimetry ($r = -0.342$, $P > 0.05$). On the other hand, the correlation between peripapillary nerve fiber layer thickness and degree of exophthalmos with P100 latency was insignificant²⁰. Acaroglu *et al.* reported a mild but significant correlation between the disease activity (clinical activity score) and P100 latency ($r = 0.364$, $P = 0.04$)¹⁶.

The correlation between degree of exophthalmos and pVEP varied among studies. Pawlowski *et al.* reported a moderate and significant correlation between degree of proptosis and N75 latency ($r = 0.51$, $P < 0.01$) but not with p100 latency¹⁸. On the other hand, Wijngaarde *et al.* described a mild correlation coefficient between degree of proptosis and P100 latency (r and P value not available)⁸, while Wei *et al.* reported poor and insignificant correlation ($r = -0.126$, P value not available)²⁰.

pVEP after treatments. Four studies reported the pVEP results before and after treatments including high-dose steroids, orbital radiotherapy and/or decompression (Table 3)^{11,15,19,21}. While treatment strategies varied, increase in p100 amplitude and/or decrease in p100 latency post-treatment were generally observed. More improvements were observed in eyes with DON than those without. Three studies reported more than 10% decrease in P100 latency after treatment of DON. Tsaloumas *et al.* reported a significant decrease (from 129.2 \pm 7.13 to 114.0 \pm 4.47 ms, $P < 0.01$)¹¹, and so did Rutecka-Debniak *et al.* (from 126.0 \pm 15.9 to 108.0 \pm 5.3 ms, $P = 0.01$)¹⁵ and Liao *et al.* (from 134.8 \pm 22.1 to 107.3 \pm 4.0 ms, $P < 0.001$)¹⁹. Rutecka-Debniak *et al.* also

No.	Author (year)	Definition of cases	Age group	Sample size (eyes)	Treatment	VEP outcome	Reported values (mean \pm SD)	
							Pre-treatment	Post-treatment
fVEP								
2	Setala <i>et al.</i> ⁹	TAO	49–66 (55.8)	7 (13)	Decompression	N60 amplitude (μ V)	15.8 \pm 6.1	13.8 \pm 6.9
						N60 latency (ms)	83.1 \pm 21.6	81.9 \pm 16.6
						P120 amplitude (μ V)	8.0 \pm 4.0	7.7 \pm 4.2
						P120 latency (ms)	130.0 \pm 21.3	129.4 \pm 20.5
		TAO	50–64 (55.3)	3 (6)	Irradiation	N60 amplitude (μ V)	16.8 \pm 5.7	13.8 \pm 6.0
						N60 latency (ms)	80.9 \pm 7.0	87.5 \pm 7.4
4	Tsaloumas <i>et al.</i> ¹¹	DON	26–73 (49.1)	6	Decompression	P2 amplitude (μ V)	6.83 \pm 0.92	13.12 \pm 1.65*
						P2 latency (ms)	112.0 \pm 4.46	106.7 \pm 3.34
		DON	26–73 (49.1)	10	2 weeks High-dose steroids	P2 amplitude (μ V)	7.00 \pm 1.10	9.61 \pm 1.43*
						P2 latency (ms)	118.4 \pm 5.79	108.3 \pm 5.47
pVEP								
4	Tsaloumas <i>et al.</i> ¹¹	DON	26–73 (49.1)	6	Decompression	P100 amplitude (μ V)	3.67 \pm 0.81	6.50 \pm 0.67*
						P100 latency (ms)	129.2 \pm 7.13	114.0 \pm 4.47*
		DON	26–73 (49.1)	10	2 weeks High-dose steroids	P100 amplitude (μ V)	5.30 \pm 0.89	8.06 \pm 0.80*
						P100 latency (ms)	116.1 \pm 4.71	111.4 \pm 4.89
8	Rutecka-Debniak <i>et al.</i> ¹⁵	DON	18–74	12 (21)	Unspecified	N75 latency (ms)	93.3 \pm 18.7	78.8 \pm 7.7*
						P100 latency (ms)	126.0 \pm 15.9	108.0 \pm 5.3*
		TAO	18–74	13 (18)	Unspecified	N75 latency (ms)	81.7 \pm 16.6	74.6 \pm 7.9
						P100 latency (ms)	114.8 \pm 12.6	107.3 \pm 13.2*
12	Liao <i>et al.</i> ¹⁹	DON	30–76 (58.4)	22 (38)	Decompression	P100 latency (ms)	134.8 \pm 22.1	107.3 \pm 4.0 [†]
14	Lipski <i>et al.</i> ²¹	DON	43–76 (55)	15 (30)	Decompression	P100 amplitude (μ V)	4.45 \pm 2.3	8.8 \pm 6.32*
						P100 latency (ms)	130.2 \pm 11.22	127.8 \pm 12.07

Table 3. Summary outcomes of longitudinal case series comparing VEP changes before and after treatment for DON/TAO. DON = dysthyroid optic neuropathy; fVEP = flash visual evoked potential; ms = millisecond; No. = number; pVEP = pattern visual evoked potential; TAO = thyroid associated orbitopathy; VEP = visual evoked potential; μ V = microvolts. * $P < 0.05$ compared to pre-treatment, [†] $P < 0.001$ compared to pre-treatment.

reported a significant decrease in N75 latency in eyes with DON after treatment (from 93.3 \pm 18.7 to 78.8 \pm 7.7 ms, $P = 0.01$)¹⁵. Significant increase in P100 amplitude over 50% was reported by Tsaloumas *et al.* after decompression (from 3.67 \pm 0.81 to 6.50 \pm 0.67 μ V, $P < 0.01$) and high-dose steroids treatment (from 5.30 \pm 0.89 to 8.06 \pm 0.80 μ V, $P < 0.01$)¹¹. Lipski *et al.* also reported significant increase in P100 amplitude after bony orbital decompression (from 4.45 \pm 2.3 to 8.8 \pm 6.32 μ V, $P < 0.05$)²¹.

In TAO eyes with no clinical evidence of DON but prolonged P100 latency, Rutecka-Debniak *et al.* reported a significant decrease after treatment (from 114.8 \pm 12.6 to 107.3 \pm 13.2 ms, $P = 0.05$)¹⁵. There was no post-treatment change in TAO eyes with normal pre-treatment VEP.

Multifocal VEP (mfVEP) in TAO. In 2012, Perez-Rico *et al.* first reported the use of mfVEP in TAO patients without DON²². There was a significant increase in mean latency in TAO group compared to age-matched control (2.12 \pm 1.72 vs. 6.57 \pm 1.90 ms, $P < 0.05$) and 23 eyes (35.4%) had abnormal mfVEP amplitude and/or latency. By interocular comparison, 12.3% of TAO eyes showed decreased amplitude and 13.8% of them showed increased latency. Visual acuity was significantly related to mfVEP amplitude changes (mean difference = -0.104 , $P = 0.018$), while intraocular pressure measured at upgaze was significantly related to mfVEP latency changes (mean difference = 2.595, $P = 0.028$). No statistically significant relationship was observed between mfVEP parameters and standard automated perimetry results or nerve fiber layer thickness measured on optical coherence tomography²².

Electroretinography (ERG) in TAO. Comparing TAO eyes with controls, Spadea *et al.* found significant decreases in amplitudes for both P50 (1.17 \pm 0.58 vs. 1.74 \pm 0.50 μ V, $P < 0.05$) and N95 (1.71 \pm 1.10 vs. 2.37 \pm 0.59 μ V, $P < 0.05$)¹². No significant difference was found in latency¹². Genovesi-Ebert *et al.* reported significantly

No.	Author (year)	Definition of subjects	Age range (mean)	Sample size (eyes)	VEP latency	Clinical measurement	Correlation	
							P value	r
1	Wijngaarde <i>et al.</i> ⁸	TAO	n/a	66	P100	VA	Significant	0.270
9	Acaroglu <i>et al.</i> ¹⁶	TAO	20–65 (41.7)	32	P100	CAS	0.0406	0.364
11	Pawlowski <i>et al.</i> ¹⁸	TAO	(35.6)	30	N75	Exophthalmos	<0.01	0.510
					P100	IOP	Insignificant	—
					P100	Exophthalmos	Insignificant	—
13	Wei <i>et al.</i> ²⁰	TAO	22–79 (46.7)	151	P100	logMAR	<0.1	0.278
					P100	Exophthalmos	Insignificant	−0.126
					P100	total error	<0.1	0.363
					P100	MD	<0.1	−0.342
					P100	ON	Insignificant	−0.055
					P100	M/O ratio	<0.1	0.482
					P100	EOM-A	<0.01	0.496

Table 4. Correlations between pVEP latencies and clinical measurements of DON/TAO. CAS = clinical activity score; DON = dysthyroid optic neuropathy; EOM-A = cross-sectional area of all extraocular rectus muscles; IOP = intraocular pressure; logMAR = logarithm of the minimal angle of resolution; MD = mean deviation of retinal sensitivity; M/O ratio = ratio between the cross-sectional area of all extraocular rectus muscles and the orbital area; No. = number; ON = peripapillary nerve fiber thickness; pVEP = pattern visual evoked potential; r = correlation coefficient; TAO = thyroid associated orbitopathy; total error = total error of 100-hue color sensation; VA = visual acuity; VEP = visual evoked potential.

smaller ($P < 0.0001$) pERG amplitude in TAO eyes without providing numerical results¹⁴. They also described a negative correlation of pERG amplitude with optic nerve diameter measured by ultrasonography. Pawlowski *et al.* reported significant decrease in P50 amplitude in TAO eyes (2.04 ± 0.99 vs. $2.69 \pm 0.88 \mu\text{V}$, $P < 0.05$) but not in N95 amplitude or latencies²³. 3 studies reported drop in P50 amplitude^{12,14,23}, with statistical significance shown by Spadea *et al.* and Pawlowski *et al.*^{12,23}.

Assessment of the quality of study and grading of clinical recommendation. The 12 studies on VEPs were assessed according to the NOS (Newcastle-Ottawa Scale) quality assessment of case-control studies³³ (Table 5). The study with best quality was carried out by Tsaloumas *et al.* in 1994¹¹. Clinical recommendation of EPS in detecting and monitoring visual dysfunction in TAO was rated according to the American Academy of Ophthalmology on preparing Preferred Practice Pattern (PPP) guidelines (Table 6)³⁴. pVEP was given level A importance in application and level II in strength of evidence.

Discussion

Clinical features of DON may include impaired visual acuity and color vision, visual field, afferent and relative affect pupillary defect (APD/RAPD), optic disc hyperemia or swelling^{5,31,35}. In practice, these features rarely co-exist while ocular co-morbidities often confound with clinical assessment³⁵. The European Group on Graves' Orbitopathy (EUGOGO) was the first to propose that the presence of optic disc swelling alone or any other two of the above abnormalities without an alternative explanation suggested the presence of DON in any TAO patient^{35,36}. Among the 94 eyes recruited, impaired visual acuity ($< 20/40$), color vision, visual field defects, relative afferent pupillary defect and optic disc swelling were present in only 73%, 77%, 71%, 45%, and 56% of eyes subsequently diagnosed to have "definite DON". On the other hand, these abnormalities were also found in 32%, 7%, 13%, 0% and 5% of eyes subsequently diagnosed to have "no DON". These results implied that none of the individual findings of optic nerve dysfunction was found to be sensitive or specific enough to diagnose or exclude DON. Proptosis or increased clinical activity scores ($\geq 3/7$) were absent in more than one-third of eyes with "definite" DON³⁵. Despite its serious visual consequences, no widespread consensus on the diagnostic criteria of DON is available to date. The challenge in diagnosing DON at its early stage or in patients with ocular comorbidities remains.

Electrophysiological studies (EPS), including visual evoked potential (VEP) and electroretinogram (ERG) were adopted to provide objective evaluation and correlation with the presence and/or severity of DON. VEP refers to the electrophysiological signals extracted from visual cortex during visual stimulation over the retina³⁷. Any disturbance along the visual pathway or visual cortex results in VEP abnormalities (decrease in amplitude or increase in latency). It was first reported in 1972 by Halliday *et al.* to assess optic neuritis³⁸. Subsequently it was used in patients with DON in 1980 by Wijngaarde *et al.*⁸. Three types of VEP have been used: flash VEP (fVEP), pattern VEP (pVEP), and multifocal VEP (mfVEP) (Table 7). fVEP uses a diffuse flash stimulating the entire retina for a mass response. Therefore, localized abnormal response may be averaged out and left undetected. pVEP uses checkerboard pattern reversal stimulation covering the central 15° visual field. The major components of pVEP are a large positive wave at peak latency of about 100 milliseconds (P100) and a negative wave peaking at 70 milliseconds (N70). Any delay in P100 latency or decrease in amplitude measured from N70 to P100 suggests the presence of optic neuropathy³⁷. Since the first report on pVEP in assessing visual function in TAO patients by Wijngaarde *et al.* in ref.^{8,9} other studies were published comparing the use of pVEP in TAO patients with or

Author	Newcastle-Ottawa Quality Assessment Scale (NOS) for Case-control Study ³³									
	Selection				Comparability		Exposure			Total Stars
(Year of Publication)	Case Definition	Representativeness of Cases	Selection of Controls	Definition of Controls	Comparability of cases and controls (a)	Comparability of cases and controls (b)	Ascertainment of exposure	Same method of ascertainment	Non-Response rate	
Wijngaarde <i>et al.</i> ⁸	—	—	—	—	—	—	—	*	n.a.	1
Shawkat <i>et al.</i> ¹⁰	—	—	—	—	*	*	—	*	n.a.	3
Tsaloumas <i>et al.</i> ¹¹	*	*	*	*	*	*	—	*	n.a.	7
Spadea <i>et al.</i> ¹²	*	—	—	*	*	—	—	*	n.a.	4
Salvi <i>et al.</i> ¹³	—	*	—	*	*	—	—	*	n.a.	4
Genovesi-Ebert <i>et al.</i> ¹⁴	—	—	—	—	*	—	—	*	n.a.	2
Rutecka-Debniak <i>et al.</i> ¹⁵	—	—	—	*	—	—	—	*	n.a.	2
Acaroglu <i>et al.</i> ¹⁶	*	—	—	*	*	*	—	*	n.a.	5
Ambrosio <i>et al.</i> ¹⁷	*	—	—	*	*	—	—	*	n.a.	4
Pawlowski <i>et al.</i> ¹⁸	*	—	*	*	*	*	—	*	n.a.	6
Perez-Rico <i>et al.</i> ²²	*	*	—	*	*	*	—	*	n.a.	6
Pawlowski <i>et al.</i> ²³	*	—	*	*	*	*	—	*	n.a.	6

Table 5. Quality Assessment for Included Case-control Studies. *A star is awarded when the study meets the quality standard of an item. Details of the requirements of each item can be found in NOS for Case-control Study checklist³³. n.a.: not available. Note: A study may be awarded a maximum of one star for each item within the Selection and Exposure categories. A maximum of two stars may be given for Comparability. A score of ≥ 7 stars is indicative of a high-quality study³³.

Clinical care	Recommendation	Evidence rating
Detecting and monitoring visual dysfunction in TAO	The use of fVEP	[B:II]
Detecting and monitoring visual dysfunction in TAO	The use of pVEP	[A:II]
Detecting and monitoring visual dysfunction in TAO	The use of mfVEP	[B:II]
Detecting and monitoring visual dysfunction in TAO	The use of pERG	[C:II]

Table 6. Clinical recommendation of VEP or ERG in detecting visual dysfunction in TAO. A = most important application; B = moderately important application; C = relevant but not critical application; II = well-designed cohort or case-control analytic studies, preferably from more than one center, or multiple-time series with or without the intervention.

without DON (Table 2). mfVEP records signals from multiple stimuli given simultaneously across 20° to 25° of the central visual field enabling assessment of small local defects³⁹.

ERG records the electrical response of the retina upon light stimulation by various types of corneal electrodes. ERG is widely used in retinal disorders but rarely in TAO⁴⁰. Pattern electroretinogram (pERG) uses reversing black and white checkerboard stimulus to collect signals from inner retina and indirectly measure retinal ganglion cell function. Commonly used parameters of pERG include a prominent positive wave at approximately 50 millisecond (P50) and a larger negative wave at about 95 millisecond (N95)⁴¹. pERG was used for evaluating early ganglion cell dysfunction in glaucoma patients since 1980s^{42,43}. pERG alteration was also reported in animal models of optic nerve transection during retrograde degeneration of retinal ganglion cells^{44,45}. In clinical practice, combined interpretation of pVEP and pERG helps to differentiate retinal (abnormal pVEP and pERG) from optic nerve disorders (abnormal pVEP and normal pERG)⁴⁶.

Here we report the first systematic review on the use of EPS in DON. pVEP has been the most widely reported EPS in DON. Case-control studies reported significant differences of pVEP parameters among eyes with DON, TAO only and from controls^{8,10–13,15–18,22}. Prolonged P100 latency was found comparing either eyes with DON to eyes without from TAO patients or eyes from TAO patients to control. P100 latency correlated with visual acuity, clinical activity score, color vision, visual field, and orbital imaging parameters^{8,20}. Significant improvement in pVEPs were found in patients after successful treatment of DON^{11,15,19,21}.

We acknowledge insufficient evidence to support the use of pVEP as part of the diagnostic criteria of DON due to its limited availability and inherent variability. To improve generalizability for meta-analysis, future studies should adopt testing protocols by the International Society for the Clinical Electrophysiology of Vision (ISCEV) standards^{37,41,47–49}, include age and/or gender-specific reference ranges, post-treatment follow-up results and all clinical parameters recommended by the EUGOGO^{5,31,35,37}. Longitudinal follow-up of pVEP on TAO patients with equivocal or early clinical features of DON may shed insight on the natural history, treatment response and clinical implication on the evolving entity of “subclinical” DON.

In conclusion, pVEP was the most studied EPS in DON. Latency and amplitude of P100 were shown to be promising for the diagnosis and monitoring of DON. Future studies on pVEP using standardized settings will be required to fully evaluate its diagnostic accuracy and clinical utility in the management of DON.

EPS test	Key features	No. of studies	Reported parameters	Reference
Visual evoked potential				
fVEP	Diffuse flash stimulus, full-field, one response, examine whole visual pathway	2	amplitude & latency of P2, N60, P120	9,11
pVEP	Checkerboard pattern reversal stimulus, central $\geq 15^\circ$ field, one response, examine whole visual pathway	12	amplitude & latency of N75, P100	8,10–13,15–21
mfVEP	16 checks times 60 sectors stimulus, central 20 to 25° field, 60 topographic responses, examine whole visual pathway in 60 sectors	1	amplitude & latency	22
Electroretinography				
fERG	Diffuse flash stimulus, full-field, one response, examine retinal cells	0	n/a	n/a
pERG	Checkerboard pattern reversal stimulus, central 15° field, one response, examine retinal cells	3	amplitude & latency of N35-P50, P50-N95	12,14,23
mfERG	103 scaled hexagons stimulus, central 25° field, 103 topographic responses, examine retinal cells in 103 sectors	0	n/a	n/a

Table 7. Features of included studies. EPS = electrophysiological studies; fVEP = flash visual evoked potential; pVEP = pattern visual evoked potential; mfVEP = multifocal visual evoked potential; pERG = pattern electroretinography; mfERG = multifocal electroretinography; No. = number; n/a. = not applicable.

No.	Search terms
1	((thyroid associated or thyroid-associated) and (orbitopathy or orbitopathies or ophthalmopathy or ophthalmopathies)).mp.
2	(endocrine and (orbitopathy or orbitopathies or ophthalmopathy or ophthalmopathies or exophthalmos)).mp.
3	ophthalmic Graves disease.mp.
4	(thyroid and (orbitopathy or orbitopathies or ophthalmopathy or ophthalmopathies)).mp.
5	(Graves adj1 (orbitopathy or orbitopathies or ophthalmopathy or ophthalmopathies)).mp.
6	(dysthyroid and (orbitopathy or orbitopathies or ophthalmopathy or ophthalmopathies)).mp.
7	1 or 2 or 3 or 4 or 5 or 6
8	(optic adj1 (neuropathy or neuropathies)).mp.
9	(optic adj1 nerve adj1 (disease or disorder)).mp.
10	8 or 9
11	7 and 10
12	(dysthyroid adj1 (optic adj1 (neuropathy or neuropathies))).mp.
13	11 or 12

Table 8. Search strategies used in MEDLINE and EMBASE.

Methods

Literature search. Literature search was performed in MEDLINE, EMBASE, and the Cochrane databases via Ovid platform. We formulated sensitive search strategies using the Boolean logic and search terms with controlled vocabularies (Medical Subject Heading terms): (“thyroid associated” OR “endocrine” OR “dysthyroid” OR “Graves”) AND (“orbitopathy[ies]” OR “ophthalmopathy[ies]”) OR (“ophthalmic Graves’ disease”) in combination with “optic neuropathy(ies)” (Table 8). The search was supplemented by manual screening of the reference lists of the relevant articles and reviews. Language filter was not applied in the search. We identified records published from January 1st, 1977 to August 20th, 2017.

Eligibility criteria. Studies were included in the systematic review according to the following criteria: (1) studies that used electrophysiological tests (e.g. VEP or ERG) to evaluate optic nerve dysfunction in patients with TAO or DON; and (2) studies can be observational case series, case-control study, cohort study, interventional case series, and clinical trials. Animal studies, case reports, reviews, abstracts, conference proceedings, and editorials were excluded.

Assessment of the quality of study and level of evidence. NOS (Newcastle-Ottawa Scale)³³ was adopted to evaluate the quality of the case-control studies. The clinical recommendation of VEP or ERG in detecting and monitoring visual dysfunction in TAO were rated from 2 aspects, “importance to the care process” and “the strength of evidence in the available literature”, according to the American Academy of Ophthalmology on preparing Preferred Practice Pattern (PPP) guidelines³⁴. “Importance to the care process” represents the value of this application to improve the quality of the patient’s care in a meaningful way. Level A indicates the most important; level B indicates moderately important and level C indicates relevant but not critical application. “Strength of evidence” was rated in 3 levels. Level I includes evidence obtained from at least one properly conducted, well-designed, randomized, controlled trial. It also includes meta-analysis

of randomized controlled trials. Level II includes well-designed controlled trials without randomization, well-designed cohort or case-control analytic studies, preferably from more than one center, or multiple-time series with or without the intervention. Level III includes evidence obtained from descriptive studies or case reports.

References

1. Neigel, J. M. *et al.* Dysthyroid optic neuropathy. The crowded orbital apex syndrome. *Ophthalmology* **95**, 1515–1521 (1988).
2. Bartley, G. B. The epidemiologic characteristics and clinical course of ophthalmopathy associated with autoimmune thyroid disease in Olmsted County, Minnesota. *Trans Am Ophthalmol Soc* **92**, 477–588 (1994).
3. Birchall, D., Goodall, K. L., Noble, J. L. & Jackson, A. Graves ophthalmopathy: intracranial fat prolapse on CT images as an indicator of optic nerve compression. *Radiology* **200**, 123–127 (1996).
4. Feldon, S. E., Muramatsu, S. & Weiner, J. M. Clinical classification of Graves' ophthalmopathy. Identification of risk factors for optic neuropathy. *Archives of ophthalmology* **102**, 1469–1472 (1984).
5. Wiersinga, W. M., Kahaly, G. *Graves' Orbitopathy: A Multidisciplinary Approach - Questions and Answers*. (ed. Wiersinga WM, Kahaly G. 2nd ed. Basel: Karger (2010).
6. Park, K. A., Kim, Y. D., In Woo, K., Kee, C. & Han, J. C. Optical coherence tomography measurements in compressive optic neuropathy associated with dysthyroid orbitopathy. *Graefes Archive for Clinical & Experimental Ophthalmology* **254**, 1617–1624 (2016).
7. Weis, E. *et al.* Quantitative computed tomographic predictors of compressive optic neuropathy in patients with thyroid orbitopathy: a volumetric analysis. *Ophthalmology* **119**, 2174–2178, <https://doi.org/10.1016/j.optha.2012.04.021> (2012).
8. Wijngaarde, R. & van Lith, G. H. Pattern EPS in endocrine orbitopathy. *Doc Ophthalmol* **48**, 327–332 (1980).
9. Setälä, K., Raitta, C., Välimäki, M., Katevuo, V. & Lamberg, B. A. The value of visual evoked potentials in optic neuropathy of Graves' disease. *J Endocrinol Invest* **15**, 821–826 (1992).
10. Shawkat, F. *et al.* The role of macular- and paramacular-derived visual evoked potentials in the detection of dysthyroid optic neuropathy. *Clinical Vision Sciences* **8**, 235–244 (1993).
11. Tsaloumas, M. D., Good, P. A., Burdon, M. A. & Misson, G. P. Flash and pattern visual evoked potentials in the diagnosis and monitoring of dysthyroid optic neuropathy. *Eye* **8**, 638–645 (1994).
12. Spadea, L., Bianco, G., Dragani, T. & Balestrazzi, E. Early detection of P-VEP and PERG changes in ophthalmic Graves' disease. *Graefes Archive for Clinical & Experimental Ophthalmology* **235**, 501–505 (1997).
13. Salvi, M. *et al.* The study of visual evoked potentials in patients with thyroid-associated ophthalmopathy identifies asymptomatic optic nerve involvement. *The Journal of clinical endocrinology and metabolism* **82**, 1027–1030 (1997).
14. Genovesi-Ebert, F. *et al.* Standardized echography, pattern electroretinography and visual-evoked potential and automated perimetry in the early diagnosis of Graves' neuropathy. *Ophthalmologica* **212**(Suppl 1), 101–103 (1998).
15. Rutecka-Debniak, A., Lubinski, W. & Krzystolik, Z. Visual evoked potentials in diagnosis and monitoring of optic neuropathy in the course of thyroid ophthalmopathy. *Klinika Oczna* **101**, 361–365 (1999).
16. Acaroglu, G., Simsek, T., Ozalp, S. & Mutluay, A. Subclinical optic neuropathy in Graves' orbitopathy. *Japanese journal of ophthalmology* **47**, 459–462 (2003).
17. Ambrosio, G., Ferrara, G., Vitale, R. & De Marco, R. Visual evoked potentials in patients with Graves' ophthalmopathy complicated by ocular hypertension and suspect glaucoma or dysthyroid optic neuropathy. *Documenta Ophthalmologica* **106**, 99–104 (2003).
18. Pawlowski, P., Mysliwiec, J., Mrugacz, M., Bakunowicz-Lazarczyk, A. & Gorska, M. Pattern visual evoked potentials in the early diagnosis of optic neuropathy in the course of Graves' ophthalmopathy. *Endokrynologia Polska* **57**, 122–126 (2006).
19. Liao, S. L., Chang, T. C. & Lin, L. L. K. Transcaruncular orbital decompression: an alternate procedure for Graves ophthalmopathy with compressive optic neuropathy. *American journal of ophthalmology* **141**, 810–818 (2006).
20. Wei, Y. H., Chi, M. C. & Liao, S. L. Predictability of visual function and nerve fiber layer thickness by cross-sectional areas of extraocular muscles in graves ophthalmopathy. *American journal of ophthalmology* **151**, 901–906.e901 (2011).
21. Lipski, A. *et al.* Course of pattern-reversed visual evoked cortical potentials in 30 eyes after bony orbital decompression in dysthyroid optic neuropathy. *British Journal of Ophthalmology* **95**, 222–226 (2011).
22. Perez-Rico, C., Rodriguez-Gonzalez, N., Arevalo-Serrano, J. & Blanco, R. Evaluation of multifocal visual evoked potentials in patients with Graves' orbitopathy and subclinical optic nerve involvement. *Documenta ophthalmologica. Advances in ophthalmology* **125**, 11–19 (2012).
23. Pawlowski, P., Janusz, M., Alina, B. L. & Maria, G. Pattern electroretinogram (PERG) in the early diagnosis of optic nerve dysfunction in the course of Graves' orbitopathy. *Klinika Oczna* **115**, 9–12 (2013).
24. Setälä, K., Raitta, C., Välimäki, M., Katevuo, K. & Lamberg, B. A. Visual evoked potentials in Basedow's ophthalmopathy before and after orbital decompression or retrobulbar irradiation. *Exp Clin Endocrinol* **97**, 344–347, <https://doi.org/10.1055/s-0029-1211090> (1991).
25. Batch, J. A. & Lepre, F. Early diagnosis of Graves' optic neuropathy using visual evoked responses. *Postgraduate medical journal* **66**, 664–666 (1990).
26. Boulos, P. R. & Hardy, I. Thyroid-associated orbitopathy: a clinicopathologic and therapeutic review. *Curr Opin Ophthalmol* **15**, 389–400 (2004).
27. Leone, C. R. Jr. & Bajandas, F. J. Inferior orbital decompression for dysthyroid optic neuropathy. *Ophthalmology* **88**, 525–532 (1981).
28. Heinemann, B., Bischoff, P., Jörgensen, J. & Weder, B. Diagnosis of optic neuropathy in endocrine orbitopathy. *Klin Monbl Augenheilkd* **198**, 460–461 (1991).
29. Wu, Z., Yan, J., Yang, H. & Mao, Y. The outcome of orbital decompression for thyroid associated ophthalmopathy. *Zhonghua Yan Ke Za Zhi* **38**, 399–401 (2002).
30. Clauser, L. C., Tieghi, R., Galie, M., Franco, F. & Carinci, F. Surgical decompression in endocrine orbitopathy. *Visual evoked potential evaluation and effect on the optic nerve. J Craniomaxillofac Surg* **40**, 621–625 (2012).
31. European Group on Graves' Orbitopathy (EUGOGO) *et al.* Clinical assessment of patients with Graves' orbitopathy: the European Group on Graves' Orbitopathy recommendations to generalists, specialists and clinical researchers. *Eur J Endocrinol* **155**, 387–389 (2006).
32. The European Group on Graves' Orbitopathy (EUGOGO), Clinical Evaluation EUGOGO Atlas, (2009) http://www.serm.md/ghiduri/tiroida/Clinical_Evaluation_GO_EUGOGO_Atlas.pdf (Date of access: 20/08/2017).
33. Wells, G. A. *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses, http://www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf (Date of access: 20/08/2017)(2011).
34. Feder, R. S. *et al.* Comprehensive Adult Medical Eye Evaluation Preferred Practice Pattern[®] Guidelines, [http://www.aaojournal.org/article/S0161-6420\(15\)01269-5/pdf](http://www.aaojournal.org/article/S0161-6420(15)01269-5/pdf) (Date of access: 20/08/2017) (2015).
35. McKeag, D. *et al.* Clinical features of dysthyroid optic neuropathy: a European Group on Graves' Orbitopathy (EUGOGO) survey. *The British journal of ophthalmology* **91**, 455–458 (2007).
36. Dayan, C. M. & Dayan, M. R. Dysthyroid optic neuropathy: a clinical diagnosis or a definable entity? *The British journal of ophthalmology* **91**, 409–410 (2007).

37. Odom, J. V. *et al.* ISCEV standard for clinical visual evoked potentials: (2016 update). *Doc Ophthalmol* **133**, 1–9 (2016).
38. Halliday, A. M., McDonald, W. I. & Mushin, J. Delayed Visual Evoked Response in Optic Neuritis. *The Lancet* **299**, 982–985 (1972).
39. Hood, D. C., Odel, J. G. & Winn, B. J. The multifocal visual evoked potential. *J Neuroophthalmol* **23**, 279–289 (2003).
40. Arden, G. B., Barrada, A. & Kelsey, J. H. New clinical test of retinal function based upon the standing potential of the eye. *Br J Ophthalmol* **46**, 449–467 (1962).
41. Bach, M. *et al.* ISCEV standard for clinical pattern electroretinography (PERG): 2012 update. *Doc Ophthalmol* **126**, 1–7 (2013).
42. Bobak, P. *et al.* Pattern electroretinograms and visual-evoked potentials in glaucoma and multiple sclerosis. *Am J Ophthalmol* **96**, 72–83 (1983).
43. Wanger, P. & Persson, H. E. Pattern-reversal electroretinograms in unilateral glaucoma. *Invest Ophthalmol Vis Sci* **24**, 749–753 (1983).
44. Mafei, L. & Fiorentini, A. Electroretinographic responses to alternating gratings before and after section of the optic nerve. *Science* **211**, 953–5 (1981).
45. Tobimatsu, S., Celesia, G. G., Cone, S. & Gujrati, M. Electroretinograms to checkerboard pattern reversal in cats: physiological characteristics and effect of retrograde degeneration of ganglion cells. *Electroencephalogr Clin Neurophysiol* **73**, 341–352 (1989).
46. Ryan, S., Arden, G. B. & Weinstein, G. W. Electrophysiological discrimination between retinal and optic nerve disorders. *Metab Pediatr Syst Ophthalmol* **12**, 69–73 (1989).
47. Hood, D. C. *et al.* ISCEV standard for clinical multifocal electroretinography (mfERG) (2011 edition). *Doc Ophthalmol* **124**, 1–13 (2012).
48. McCulloch, D. L. *et al.* ISCEV Standard for full-field clinical electroretinography (2015 update). *Doc Ophthalmol* **130**, 1–12 (2015).
49. Odom, J. V. *et al.* Visual evoked potentials standard. *Doc Ophthalmol* **108**, 115–123 (2004).

Author Contributions

T.W.U.I. did the literature search, quality assessment, data collection, and data interpretation. She wrote the main manuscript text and prepared the tables and figures. S.S.R. did the literature search and data interpretation. He critically revised the manuscript. A.N.L. did the quality assessment and data collection. M.E.B. critically revised the manuscript. A.L.Y. critically revised the manuscript. K.K.L.C. conceived the study design, supervised the data collection and data interpretation and critically revised the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017