

TITLE

Maximum likelihood perimetric progression analysis: Using raw (trial-by-trial) response data to estimate progression more robustly

RUNNING TITLE

Maximum likelihood perimetric progression analysis

LIST OF AUTHORS

Pete R. Jones¹

¹*Division of Optometry and Visual Sciences, School of Health Science, City, University of London, London, England, EC1V 0HB.*

CORRESPONDENCE

*Pete R Jones, Division of Optometry and Visual Sciences, School of Health Sciences, City, University of London, Northampton Square EC1V 0HB, London, UK; peter.jones@city.ac.uk

SUPPLEMENTAL MATERIAL

1 item of supplemental code: *MLprogressionAnalysis_SuplCode.m*

MEETING PRESENTATION(S)

< none >

FINANCIAL SUPPORT

< none >

CONFLICTS OF INTEREST

No conflicting relationship exists for any author.

DISCLOSURES

P.R. Jones, None

ABSTRACT

1 **Purpose:** To describe and demonstrate a more efficient (Maximum Likelihood) method for
2 quantifying visual field progression.

3 **Design:** Monte Carlo simulation.

4 **Methods:** Trial-by-trial response data were simulated using a stochastic psychometric model (a
5 “simulated observer”). Simulated Differential Light Sensitivity (DLS) decreased between tests to
6 mimic long-term visual field progression. Progression slopes were fitted, either by fitting a
7 regression slope to independent DLS estimates from each test (conventional method), or by fitting
8 all the raw data combined in a single model (proposed maximum likelihood method).

9 **Results:** The proposed ML method seldom performed worse than a conventional, regression-based
10 approach, and often performed better. For an idealized observer with a lapse (false negative) rate of
11 0 and a guess (false positive) rate of 0, both methods were equally precise. However, as lapse rate
12 increased, the ML method exhibited less random measurement error. For small numbers of trials
13 this increase in precision translated to a negative progression slope being detected with 95%
14 confidence at least one year/assessment sooner. The only time the ML method was observed to
15 perform worse was when very few trials ($N = 4$) were combined with very high lapse rates ($\lambda = 0.3$):
16 an unlikely but not inconceivable scenario.

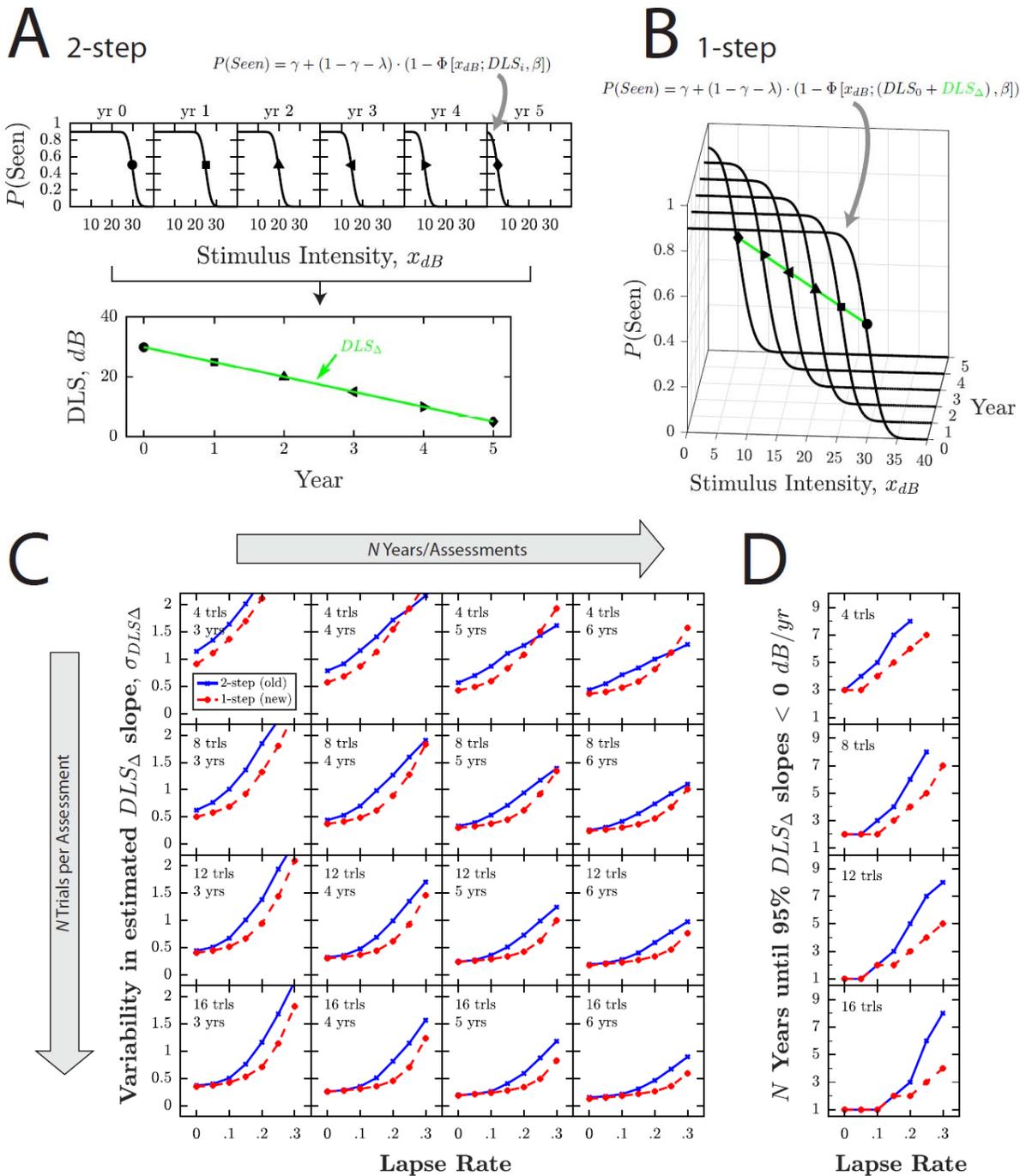
17 **Conclusions:** Combining raw, trial-by-trial response data in a single ML model can provide a more
18 robust estimate of visual field progression than conventional methods (e.g., linear regression), at no
19 additional cost to the patient or clinician (i.e., no additional trials).

20 **KEY WORDS:** *Visual Fields; Perimetry; Progression; Glaucoma; Psychophysics; Maximum Likelihood*

21 **Maximum likelihood perimetric progression analysis: Using raw (trial-by-trial) response data to**
22 **estimate progression more robustly**

23 The ability to accurately monitor perimetric rate of change ('progression') is key for the effective
24 management of glaucoma. Typically, rate of progression (e.g., change in Differential Light Sensitivity,
25 DLS, at a specific retinal location, or in some overall summary metric, such as Mean Deviation, MD)
26 is estimated by fitting a curve through a series of point-estimates (scalar 'best guess' values), with
27 each value representing a single assessment/hospital-visit. This approach is suboptimal, however,
28 since there is additional information contained in the raw, trial-by-trial, test data; information that is
29 effectively discarded when a point-estimate is generated at the end of each assessment.

30 In principle, this additional information could also be factored in when estimating rate of
31 progression. Thus, rather than using the raw data from each test to compute independent point-
32 estimates, $[DLS_0, DLS_1, \dots, DLS_n]$, and then fitting a line through this time-series (a 'two step'
33 solution, illustrated in **Fig 1A**), all of the raw data from every test could instead be used together to
34 fit a single, maximum likelihood (ML) model, consisting of a single psychometric function whose
35 threshold is determined by two parameters: an initial value, DLS_0 , and a rate of progression value,
36 DLS_{Δ} (a 'one step' solution, illustrated in **Fig 1B**). MATLAB code instantiating both of these methods is
37 given as **Supplemental Code**, and this same code was used to generate the simulations reported
38 below.



39

40 **Figure 1.** Methods and Results. **(A)** Schematic illustrations of conventional progression analysis: a two-step approach, in
 41 which DLS thresholds are first estimated independently for each assessment (e.g., using a maximum likelihood
 42 algorithm or adaptive staircase), and then a curve is fitted to this series of point-estimates (e.g., a least-squares linear
 43 regression slope). **(B)** Our proposed maximum likelihood method: a one-step approach in which the raw, trial-by-trial
 44 data from every test are used together to fit a single psychometric function, the threshold/location parameter of which,
 45 DLS_{dB} , is determined by a starting value, DLS_0 , and rate of progression value, DLS_{Δ} . **(C)** Simulation Results: variability in
 46 estimated rate of change, $\sigma_{DLS_{\Delta}}$, for the conventional method (blue crosses) and proposed ML method (red circles), as a
 47 function of lapse rate. Each datapoint represents 10,000 simulations (2.8M simulations in total), with a given N Trials per
 48 assessment, 4, 8, 12, or 16 trials, N Assessments, 3, 4, ..., or 12 years, and Lapse Rate, 0.05, 0.10, ..., 0.30. **(D)**
 49 Further analysis. Same underlying data as (C), but here showing the number of years until 95% of simulations (i.e., 9,500
 50 simulations) produced a negative slope estimate: $DLS_{\Delta} < 0$. This can be thought of as the number of years until the
 51 negative progression slope was detected with 95% confidence.

52 To assess whether this 'one step' (ML) solution has the potential to meaningfully reduce
53 measurement error, we ran Monte Carlo simulations using a simulated observer, whose probability
54 of responding correctly to a stimulus of magnitude x_{dB} was determined, in the typical fashion, by a
55 psychometric function of the form:

$$P(\textit{Seen}) = \gamma + (1 - \gamma - \lambda) \cdot (1 - \Phi [x_{dB}; (DLS_0 + DLS_{\Delta}), \beta]) \quad (\text{Eq. 1})$$

56 where Φ is a Gaussian cumulative density function (CDF); DLS_0 is the initial threshold sensitivity at
57 first assessment, fixed at 30 dB; DLS_{Δ} is the rate of progression, fixed at -1 dB/year (i.e., DLS = 25 dB
58 after 5 years; Note that the sum of DLS_0 and DLS_{Δ} together constituted the mean/location
59 parameter of the CDF); γ is the guess rate (0 to 1), fixed at 0; β is the psychometric slope parameter
60 (the standard deviation of the CDF), fixed at 2 dB (Note, not be confused with the rate of
61 progression slope parameter: DLS_{Δ}); and λ is the lapse rate (0 to 1), which was systematically
62 manipulated across simulations as follows: $\{0, 0.05, \dots, 0.3\}$.

63 Each simulated observer was used to generate $\{3, 4, \dots, 12\}$ 'years' of DLS estimates (1
64 assessment per year), with each assessment consisting of a single DLS estimate (a '1 point grid')
65 made using either $\{8, 16, 32, \text{ or } 64\}$ trials. This resulted in 280 unique conditions ($7 \textit{ Lapse Rates} \times$
66 $10 \textit{ N Years/Assessments} \times 4 \textit{ N Trials Per Assessment}$). Each of these 280 conditions was
67 independently simulated 10,000 times, yielding 2.8M simulated 'patients' in total (each of whom
68 performed 3–12 DLS assessments). Within each assessment, stimulus selection was determined
69 by a conventional QUEST+ routine¹ (in this instance equivalent to ZEST²), using a bimodal prior taken
70 from Vingrys & Pianta (1999)³. The final threshold estimate (i.e., DLS , in dB) was computed as the
71 mean of the probability mass function. See **Supplemental Code** for full technical details.

72 For each simulated patient, rate of progression was then computed, *post hoc*, in two ways: First,
73 using a conventional two-step linear regression procedure (**Fig 1A**). Second, by refitting all of the
74 raw, trial-by-trial data from every assessment using the proposed ML procedure (**Fig 1B**). This ML
75 procedure differed from the conventional QUEST+/ZEST procedure used within each individual
76 assessment only inasmuch as the stimulus domain had two dimensions (*Stimulus Intensity* in dB,
77 and *N Years*) rather than one (*Stimulus Intensity*) and the model free-parameters domain had two
78 dimensions (DLS_0 and DLS_{Δ}) rather than one (DLS).

79 For each set of 10,000 rate of progression estimates, the standard deviation, $\sigma_{DLS_{\Delta}}$, was computed
80 as an index of random measurement error. Ideally, $\sigma_{DLS_{\Delta}}$ would be 0 (same estimated rate of
81 progression on every simulation), with higher values indicating lower precision.

82 The results are shown in **Figure 1C,D**. In brief, the proposed ML method seldom performed worse
83 than a conventional, regression-based approach, and often performed better. For an idealized
84 observer with a lapse (false negative) rate of 0 and a guess (false positive) rate of 0, both methods
85 were equally precise (**Fig 1C, leftmost data point in each panel**). However, as lapse rate increased,
86 the ML method exhibited less random measurement error. For small numbers of trials this increase
87 in precision translated to a negative progression slope being detected with 95% confidence at least
88 one year/assessment sooner (**Fig 1D**). Unsurprisingly, this difference between the two methods was
89 attenuated when large, clinically unrealistic, numbers of trials were used to estimate thresholds,
90 since in such circumstances measurement error for both methods converged towards zero (**Fig**
91 **1C,D, bottommost panels**). Though some advantage for the ML method was still observed for very
92 high lapse rates. The only time the ML method was observed to perform more poorly than the
93 conventional method was when very few trials ($N = 4$) were combined with very high lapse rates (λ
94 $= 0.3$): an unlikely but not inconceivable scenario.

95 Other than the data shown in **Figure 1**, no other conditions (i.e., observer parameters or
96 psychometric algorithm settings) were simulated: The goal was not to systematically assess the
97 exact level of benefit under all possible scenarios, but to assess whether any meaningful benefit is
98 possible under a single, broadly realistic scenario. Interested readers can, however, modify the
99 **Supplemental Code** to simulate other scenarios.

100 These simulations, though not comprehensive, illustrate that combining raw, trial-by-trial response
101 data in a single ML model can provide a more robust estimate of visual field progression than a
102 conventional ‘two-step’ (e.g., linear regression) approach, at no additional cost to the patient or
103 clinician (i.e., without requiring any additional stimulus presentations, or varying current test
104 durations or methods in any way).

105 REFERENCES

- 106 1. Watson, A. B. QUEST+: A general multidimensional Bayesian adaptive psychometric
107 method. *J. Vis.* 2017;17:10
- 108 2. Turpin, A., McKendrick, A. M., Johnson, C. A. & Vingrys, A. J. Properties of perimetric
109 threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by
110 computer simulation. *Invest. Ophthalmol. Vis. Sci.* 2003;44:4787–4795
- 111 3. Vingrys, A. J. & Pianta, M. J. A new look at threshold estimation algorithms for automated
112 static perimetry. *Optom. Vis. Sci.* 1999;76:588–595