

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Ophthalmology* 2012;119(8):1552-7 following peer review. The definitive publisher-authenticated version is available online at www.sciencedirect.com

Systematic review of the agreement of tonometers with Goldmann applanation tonometry

Cook JA¹, Botello AP¹, Elders A¹, Alia A¹, Azuara-Blanco A¹, Fraser C¹, McCormack K¹, Burr JM¹, For the Surveillance of Ocular Hypertension Study Group*

¹*Health Services Research Unit, University Of Aberdeen, Aberdeen, UK.*

Address for correspondence/reprints Dr Jonathan Cook, Health Services Research Unit, University of Aberdeen, Health Sciences Building, Foresterhill, Aberdeen, AB25 2ZD.

Telephone: +44 (0)1224 438166

Fax: +44 (0) 1224 438165

E-mail: j.a.cook@abdn.ac.uk

Surveillance of Ocular Hypertension group Jennifer Burr, Paola Pinzon, Yemisi Takwoingi, Rodolfo Hernández, Mario Vazquez-Montes, Andrew Elders, Ryo Asaoka, Katie Banister, Josine van der Schoo, Cynthia Fraser, Anthony King, Hans Lemij, Roshini Sanders, Stephen Vernon, Anja Tuulonen, Aachal Kotecha, Paul Glasziou, David Garway-Heath, David Crabb, Luke Vale, Augusto Azuara-Blanco, Rafael Perera, Mandy Ryan, Jonathan Deeks, Jonathan Cook

Financial Support This review was part of the Surveillance for Ocular Hypertension study funded by the UK NIHR Health Technology Assessment programme (project number 07/46/02). JC held a Medical Research Council UK fellowship (G0601938). AA-B was a grant holder on Astra Zeneca funded study trial of a new medication for Glaucoma. Health Services Research Unit receives core funding from the Chief Scientist Office of the Scottish Government Health Directorates. Views and opinions expressed are those of the authors and

do not necessarily reflect those of the Chief Scientist Office, NIHR HTA programme or the Department of Health. None of the funders had a role in the design or conduct of this research.

Conflicts of Interest No conflicting relationship exists for any author.

Abstract

Objective:

To assess the agreement of tonometers available for clinical practice with Goldmann applanation tonometer (GAT), the most commonly accepted reference device.

Design:

A systematic review and meta-analysis of directly comparative studies assessing the agreement of one or more tonometers with the reference tonometer (GAT).

Participants:

A total of 11,582 participants (15,525 eyes) were included.

Methods:

Summary 95% limits of agreement were produced for each comparison.

Main Outcome measures:

Agreement, recordability and reliability.

Results:

A total of 102 studies, including 130 paired comparisons were included representing eight tonometers (Dynamic Contour Tonometer, Non-contact Tonometer [NCT], Ocular Response Analyser, Ocuton S, Handheld Applanation Tonometer, Rebound Tonometer, Transpalpebral and Tonopen). The agreement (95% limits) appeared to vary across tonometers; 0.2 mmHg (-3.8 to 4.3 mmHg) for NCT to 2.7 mmHg (-4.1 to 9.6 mmHg) for Ocuton S. The estimated proportion within 2mmHg of GAT ranged from 33% (Ocuton S) to 66% and 59% (NCT and Handheld applanation tonometers respectively). Substantial inter- and intra-observer variability was observed for all tonometers.

Conclusions:

Systematic review of agreement of tonometers with GAT

NCT or Handheld applanation tonometers appear to achieve a measurement closest to GAT.

However, there was substantial variability in measurements both within and between studies.

Manuscript

Background

Raised intraocular pressure (IOP) is the most important risk factor for developing glaucoma and is the only one which is treatable. The risk of developing glaucoma and worsening of existing disease rises with increasing IOP.¹⁻⁴ In the United Kingdom (UK), there is considerable debate about the role and optimal organisation of a monitoring service for those patients with ocular hypertension, and whether other health professionals (e.g., nurses, optometrists) might be safely involved in measuring IOP. To be used in such a setting a tonometer needs to be accurate, precise and easy to use.

Goldmann applanation tonometer (GAT), a contact tonometer, is currently the most widely used tonometer by ophthalmologists. A very thick or thin cornea can lead to measurement error in tonometry, including GAT.^{5,6} New tonometers are available which account for the biomechanical properties and thickness of the cornea. In addition, non-invasive self measurement devices are available and may be highly appropriate and relevant as monitoring devices.

The aim of this systematic review was to compare the agreement of the tonometers used in clinical practice with GAT as the reference tonometer.

Methods

Directly comparative studies, i.e., those which assessed the agreement of one or more tonometers, compared with the reference standard tonometer (GAT) in the same group of people (paired data) were included. Clinic (e.g., case-control and cohort design) and population (e.g., cross-sectional) studies were eligible provided they incorporated paired data for GAT and at least one other tonometer which could be used in clinical practice. The following tonometers were not included as they were either not commercially available or were judged not suitable for monitoring ocular hypertension in routine clinical practice: Applanation resonance tonometer;⁷ Ocular blood flow instrument;⁸ Schoitz;⁹ Smartlens;¹⁰

pneumatonometer;¹¹ and manometry.¹² Studies published in a non-English language and conference abstracts were excluded. All patients aged 16 years and over, including those with a diagnosis of ocular hypertension or glaucoma or representative of the general population, were eligible for inclusion. Where the age range was not reported confirmation from the authors was sought. If there was still uncertainty, a formula was applied ($\text{Mean}-3\text{SD}\geq 16$; or $\text{median}-3*\text{[IQR/1.35]}\geq 16$), where SD and IQR are the standard deviation and inter-quartile range respectively) to assess inclusion and prevent exclusion purely on the failure to report the age. Participants with corneal abnormalities were excluded (corneal pathology including: keratoconus; bullous keratoplasty; or post corneal grafts). Measurements performed by any type of examiner (e.g., optometrists; ophthalmologists; nurses; technicians; or patients) were considered. The outcomes of interest were agreement (mean difference and limits of agreement) between a tonometer and the reference standard, the reliability (inter- and intra-observer variation) associated with measurements, and the proportion of participants with a recorded IOP measurement hereafter termed 'recordability'.

Sensitive electronic searches using both thesaurus controlled and text terms were conducted to identify reports of published and ongoing studies on the reliability and agreement of tonometers. The following bibliographic databases were searched from 1987 until February 2010: Medline; Medline-In Process; Embase; Science Citation Index; Biosis; and the Cochrane Central Register of Controlled Trials. In addition, the websites of key journals were screened for relevant in-press publications. Additional searches were undertaken in current research registers including: Clinical Trials; Current Controlled Trials; and WHO International Clinical Trials Registry Platform. An Internet search using Copernic Agent was also undertaken and included key professional organisations and manufacturers of tonometers. The reference list of included studies was also checked. Full details of the search strategies used are available from the authors as is the protocol.

Two reviewers independently screened the titles and abstracts (if available) of all reports identified by the electronic searches. Full text copies of all studies deemed to be potentially relevant were obtained and independently assessed for inclusion by two reviewers. Two

reviewers independently extracted data on study design, participant characteristics, type of tonometer used and outcome data. When outcome data were provided either per eye (right/left) right eye data only was used. If different versions of the same technology were reported, data on the most recent version were included. We conducted a 20% check of all extracted data. Quality of included studies was assessed using a modified checklist adapted from QUADAS tool for diagnostic studies¹³ and a checklist for agreement studies.¹⁴ Discrepancies were resolved by discussion or arbitration.

The primary outcome agreement was assessed by calculating summary 95% limits of agreement (LoA).¹⁵ The 95% LoA interval was calculated for each candidate tonometer from pooled estimates of the mean difference (systematic difference) between a tonometer and a reference standard and of the corresponding variability of agreement (random error). Pooled estimates of mean difference and random error were calculated using the DerSimonian and Laird random effects method.¹⁶ Imputation of within participant correlation coefficients to allow calculation of the standard deviation of differences (SDdiff) was employed, if required, using mean correlation of estimates from other studies of the same tonometer. The 95% limits of agreement (LoA) and the proportion of measurements within 2mmHg of GAT was estimated from the pooled difference and standard deviation assuming a normal distribution. Sensitivity analyses included a fixed-effect analysis and/or imputation of correlations using the minimum correlation coefficient reported from the studies assessing the same tonometer. An approximate 95% prediction interval was calculated for the mean difference and the SDdiff parameters using the estimated tau (standard deviation of the study level distribution) from the random effects analysis to quantify the impact of between study heterogeneity. It provides a range of plausible values for a future study, based upon the current studies (pooled parameter estimate \pm 1.96*tau).

Further sensitivity analyses looked at the impact of excluding studies which used suboptimal methods according to our quality assessment tool (i.e., where at least one of the requirements is clearly not met) and excluding studies with data clustered within persons. An additional

analysis was conducted to correct for repeated measurements by using reported estimates of within-participant variation.¹⁷

Heterogeneity between the study estimates in the meta-analyses was explored by visual inspection of forest plots, calculation of tau and I^2 statistics. Possible reasons for heterogeneity were explored through pre-specified subgroup analyses of central corneal thickness, previous corneal refractory surgery, type of examiner and IOP level with corresponding subgroup meta-analyses being conducted. Due to the observed level of heterogeneity, a further subgroup analysis investigated the impact of manufacturers in studies where multiple manufacturers produced the same type of tonometer. Formal comparison between subgroups was not conducted due to the large level of heterogeneity in the main analyses. Recordability data were tabulated with no quantitative analysis conducted. Reliability data was collected for GAT and the tonometers where reported. No formal synthesis of data was carried out as a variety of measures (e.g., intracluster correlation coefficients and repeatability coefficients) were used.

Data were validated and prepared using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and SPSS version 18 (IBM Corporation, NY, USA). Meta-analyses were carried out using the *metan* command in Stata version 11 (Stata Corporation, College Station, TX, USA).

Results

A total of 642 titles and abstracts were identified from the search of which 143 were selected for full text assessment. An additional 46 potentially eligible studies were identified from the reference lists of included studies. In total, 189 studies were full text assessed. 102 studies (130 paired comparisons) involving 11,582 participants (15,525 eyes) were included in the review. Studies included a variety of individuals; both patient and non-diseased cases, some with treatment and untreated cases of ocular hypertension and glaucoma. A table of characteristics of included studies is available from the authors on request. Eighty seven reports were excluded at the full text assessment stage as they failed to meet one or more of

the inclusion criteria in terms of study participants, study design, candidate tonometers and reference standard as presented in Figure 1.

Included studies compared the reference standard tonometer– GAT (Haag Streit, Koeniz, Switzerland) with eight different types of tonometer: Dynamic contour tonometer - DCT (PASCAL®, SMT Swiss Microtechnology, Switzerland); Rebound tonometer - RT (iCare®, Helsinki, Finland); TonoPen® (Mentor O & O Inc., Santa Barbara, CA, USA [incorporating Bio-Rad]; Medtronic Solan, Jacksonville, FL, USA (incorporating Xomed); Intermedics Intraocular Inc., Pasadena, CA, USA); Ocuton S (EPSa Elektronik & Präzisionsbau, Saalfeld, Germany); Handheld applanation tonometer - HAT (Kowa HA-2, Kowa, Japan; Perkins, Haag Streit, Koeniz, Switzerland); Non-contact tonometer - NCT (Canon USA Inc., Lake Success, NY, USA; Keeler Ltd., Windsor, UK; Nidek Co. Ltd., Gamagori, Japan; Reichert ophthalmic instruments, Buffalo, NY, USA; Topcon Corporation, Tokyo, Japan); Ocular response analyser – ORA (Reichert Inc, Depew, NY, USA); Transpalpebral tonometer which includes the Pressure phosphene tonometer (Proview eye pressure monitor®, Bausch & Lomb Inc., Rochester, NY, USA) and the TGDc-01® (Ryazan State Instrument-making Enterprise, Ryazan, Russia) also known as Diaton® tonometer (BiCOM Inc., Long Beach, NY, USA). Quality assessment results are summarised in Figure 2. Apart from participant selection and accounting for all participations, it was often uncertain whether individual quality criteria were met. Rarely was the non-compliance with a criteria item explicitly reported; for example it was clear in only one study that tonometers were not calibrated whereas for most studies this was not stated.

Ninety-nine studies (125 paired comparisons) were included in the meta-analyses of agreement; three did not report sufficient data. Comparison across tonometers was difficult given the indirect nature of the analysis. A summary of the main analyses for all candidate tonometers is provided in Tables 1 and 2. The proportion (%) of results within 2mmHg of GAT, based upon the main analysis mean difference and random error, is also presented. Based upon the meta-analyses, the expected difference varied across tonometers with NCT having the smallest expected difference (0.2 mmHg) in contrast to Ocuton S which had the

largest difference (2.7 mmHg). There was substantial uncertainty for most of the tonometers. In terms of the estimated random error, results also varied with HAT and NCT having the lowest random error (2.1 mmHg) and Ocuton S the largest random error (3.5 mmHg). For all tonometers, the 95% LoA stretched from at least 4 mmHg less to 4 mmHg higher with Ocuton S and Transpalpebral having the largest intervals. For most tonometers approximately 50% of measurements were estimated to be within 2mmHg though it was lowest for Ocuton S (33%). NCT and HAT were slightly higher than the others (66% and 59% respectively).

Substantial heterogeneity was observed in estimates between studies. The 95% prediction interval for the mean difference and random error are shown in Table 2. All I^2 values were greater than 80% (figures not shown). The 95% prediction intervals illustrate the impact of the heterogeneity between individual studies on the expected mean difference: -4.0 to 9.4 mmHg for Ocuton S whereas for NCT the range of values was only -1.4 to 1.9 mmHg. For all tonometers bar NCT, a mean difference of greater than 2mmHg compared with GAT fell within the 95% prediction interval. Similarly the random error 95% prediction intervals illustrate the substantial difference in the level of variability between studies. The sensitivity analyses did not have substantial impact upon the results nor did the subgroup analyses provide informative results (results for both not shown).

27 studies provided data on recordability. For one RT study of only 36 participants, recordability was worryingly low at 50%. For NCT, Ocuton S and Transpalpebral a value in the range of 70-90% was observed in a single study which could be considered problematic if representative of a monitoring scenario. (Table 3) Reliability data were reported for all except the HAT tonometer, although a variety of metric were reported. Inter- and intra-observer reliability data were available for only five of the eight tonometers (37 studies). Generally relatively large levels of variability were observed for inter- and intra-reliability with GAT appearing to have lower levels of variability than most if not all of the other tonometers.

Discussion

We identified a large body of evidence comparing tonometers to GAT. However, poor reporting limited the assessment of the quality of the included studies and the synthesis of the evidence.

The results of this study suggest that, when compared with GAT, NCT was the tonometer with the least amount of variability in IOP. Approximately two thirds of measurements with NCT were estimated to be within 2 mmHg of the GAT measurement. Second lowest variability was observed with HAT, with 59% of measurements within 2 mmHg which was not surprising because it is also an applanation tonometer. HAT has the same advantages and limitations as GAT, the only substantial difference being that HAT is a portable instrument. Other tonometers had about half or more of the measurement differences greater than 2 mmHg. Ocuton S appeared to have the lowest agreement with GAT with only a third of measurements within 2 mmHg.

Recordability was reported for all tonometers except HAT. Disappointingly, only 27 (26%) studies explicitly stated the number of participants for which a measurement was attempted as opposed to the number for which a measurement was successfully taken. In general, reported recordability was moderate to very high with most studies reporting values of 90% and above. Reliability data were available for all tonometers except HAT. There was a clear suggestion of sizeable inter- and intra-observer variability for all seven tonometers where data were available. It is worth noting that GAT reliability, while often smaller than the corresponding study's candidate tonometer value, was also usually sizeable. This would explain the scale of heterogeneity observed in the agreement meta-analyses to some extent, although the use of repeated measurement for both GAT and the candidate tonometer should have lessened the impact.

Although GAT has a number of limitations for measuring IOP, it is likely to remain the standard in secondary care (i.e., hospital setting) for some time. For this reason, determining which tonometers are close to GAT is useful. Unfortunately, variability between tonometers was substantial. Reliability data showed that variability for repeat measurement (including

GAT) was also non-negligible. Consistent use of the same tonometer during clinical follow-up is arguably almost as important as the choice of tonometer.

To be included in this review, a tonometer had to be judged that it was suitable for monitoring ocular hypertension in routine clinical practice and could potentially replace GAT. As such our findings are only directly relevant to the eight tonometer types included in this review; additional tonometers exist which may be considered as relevant by others. We chose to include studies which directly compared a candidate tonometer against the GAT which we used as a reference standard. In principle this should have provided some consistency across comparisons though the results perhaps suggest that this standard, though widely accepted, is somewhat variable in implementation. Implicitly, any contrast between studies is an indirect comparison and suffers from the limitation of such approaches that observed difference may reflect at least to some degree the difference amongst the studies (e.g., population) which contribute to each comparison. An important finding of the review was the large scale heterogeneity between the results of individual studies which assess the same basic comparison. The meta-analysis quantified the degree to which findings differed between studies and showed the inconsistencies and variations in the published literature. There were a number of limitations in the reporting of individual studies which limited the extent to which we could accurately represent the evidence. A number of studies included more than one eye per participant which resulted in clustering of data within person.¹⁸ In addition, studies used varying numbers of observations both for the candidate tonometer and the reference standard. There are a number of other factors, such as central corneal thickness and the underlying IOP level, which are known to influence IOP measurements and potentially agreement between tonometers which we were unable to formally investigate these due to limitations with the data reported in the published literature.

There is a need to standardise the reporting of comparative studies of tonometers.¹⁹ The necessary statistics for meta-analysis are often not presented. The reporting is inconsistent and in particular basic information is not always presented. Our quality assessment highlighted a lack of reporting of key study characteristics and issues such as the clustering of

eyes with participants and the number of observations used is regularly ignored. Furthermore an in-depth exploration of factors which could influence the pressure measurements is needed for the reference standard and candidate tonometers. This could be addressed by a large primary study but also has the potential to be explored in an individual patient data meta-analysis.¹⁵ Given the level of heterogeneity, it may be the case that a systematic review of limit of agreement studies requires very focussed study inclusion criteria akin to those recently proposed for diagnostic test accuracy.²⁰ Finally, more in-depth evaluation of the role of GAT as the default tonometer in clinical practice seems warranted.

There is a variety of tonometers to evaluate intraocular pressure, and GAT is the current reference standard. NCT or HAT tonometers appear to typically achieve the closest measurement to GAT.

Acknowledgement The authors would like to thank Aachal Kotecha for helpful comments on a draft document which this paper was based upon.

References

1. Grodum K, Heijl A, Bengtsson B. A comparison of glaucoma patients identified through mass screening and in routine clinical practice. *Acta Ophthalmol Scand* 2002;80:627-31.
2. Heijl A, Leske MC, Bengtsson B, et al, Early Manifest Glaucoma Trial Group. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* 2002;120:1268-79.
3. Kass MA, Gordon MO. Intraocular pressure and visual field progression in open-angle glaucoma. *Am J Ophthalmol* 2000;130:490-1.
4. Leske MC, Heijl A, Hyman L, Bengtsson B, Early Manifest Glaucoma Trial Group. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology* 1999;106:2144-53.
5. Duch S, Serra A, Castanera J, et al. Tonometry after laser in situ keratomileusis treatment. *J Glaucoma* 2001;10:261-5.
6. Ehlers N, Hansen FK, Aasved H. Biometric correlations of corneal thickness. *Acta Ophthalmol (Copenh)* 1975;53:652-9.
7. Hallberg P, Eklund A, Bäcklund T, Lindén C. Clinical evaluation of applanation resonance tonometry: a comparison with Goldmann applanation tonometry. *J Glaucoma* 2007;16:88-93.
8. Abbasoglu OE, Bowman RW, Cavanagh HD, McCulley JP. Reliability of intraocular pressure measurements after myopic excimer photorefractive keratectomy. *Ophthalmology* 1998;105:2193-6.
9. Wingert TA, Bassi CJ, McAlister WH, Galanis JC. Clinical evaluation of five portable tonometers. *J Am Optom Assoc* 1995;66:670-4.
10. Troost R, Vogel A, Beck S, et al. Clinical comparison of two intraocular pressure measurement methods: SmartLens dynamic observing tonography versus Goldmann. *Graefes Arch Clin Exp Ophthalmol* 2001;239:889-92.
11. Bayraktar S, Bayraktar Z. Central corneal thickness and intraocular pressure relationship in eyes with and without previous LASIK: comparison of Goldmann applanation tonometer with pneumatonometer. *Eur J Ophthalmol* 2005;15:81-8.
12. Feltgen N, Leifert D, Funk J. Correlation between central corneal thickness, applanation tonometry, and direct intracameral IOP readings. *Br J Ophthalmol* 2001;85:85-7.
13. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* [serial online] 2003;3:25. Available at: <http://www.biomedcentral.com/1471-2288/3/25>. Accessed January 16, 2012.

14. Craig JV, Lancaster GA, Williamson PR, Smyth RL. Temperature measured at the axilla compared with rectum in children and young people: systematic review. *BMJ* 2000;320:1174-8.
15. Williamson PR, Lancaster GA, Craig JV, Smyth RL. Meta-analysis of method comparison studies. *Stat Med* 2002;21:2013-25.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
17. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-60.
18. Murdoch IE, Morris SS, Cousens SN. People and eyes: statistical approaches in ophthalmology. *Br J Ophthalmol* 1998;82:971-3.
19. Kottner J, Audige L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96-106.
20. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008;149:889-97.

Table 1 Pooled estimates and summary 95% limits of agreement (mmHg unless otherwise stated)

Comparator	No. Studies	MD	95% CI		RE	95% CI		95% LoA		% within 2.0 mmHg
DCT	32	1.8	1.4	2.2	2.4	2.1	2.6	-2.9	6.5	48%
HAT	4	-1.2	-2.8	0.4	2.1	1.3	2.8	-5.2	2.8	59%
NCT	26	0.2	-0.1	0.6	2.1	1.8	2.3	-3.8	4.3	66%
Ocuton S	3	2.7	-1.2	6.6	3.5	2.4	4.6	-4.1	9.6	33%
ORA	12	1.5	0.9	2.2	2.8	2.5	3.1	-3.9	7.0	46%
RT	14	0.9	0.4	1.4	2.6	2.1	3.2	-4.3	6.1	52%
TonoPen	14	-0.2	-1.0	0.5	3.1	2.5	3.7	-6.2	5.8	48%
Transpalpebral	20	-0.5	-1.3	0.3	3.3	2.8	3.7	-6.9	5.9	46%

CI = confidence interval;
DCT = dynamic contour tonometer;
HAT = handheld applanation tonometer;
LoA = limits of agreement;
MD = mean Comparator minus mean GAT value;
mmHg = millimetres of mercury;
NCT = non contact tonometer;
ORA = ocular response analyser;
RE = the random error (estimated standard deviation of the differences);
RT = rebound tonometer

Table 2 Pooled estimates with 95% prediction intervals (mmHg unless otherwise stated)

Comparator	No. Studies	Mean Diff	95% Pred Int		Ran Err	95% Pred Int	
DCT	32	1.8	-0.4	4.0	2.4	1.1	3.6
HAT	4	-1.2	-4.4	2.0	2.1	0.6	3.6
NCT	26	0.2	-1.4	1.9	2.1	0.8	3.3
Ocuton S	3	2.7	-4.0	9.4	3.5	1.7	5.3
ORA	12	1.5	-0.6	3.7	2.8	1.6	4.0
RT	14	0.9	-0.9	2.7	2.6	0.6	4.7
TonoPen	14	-0.2	-3.0	2.5	3.1	0.9	5.3
Transpalpebral	20	-0.5	-3.8	2.8	3.3	1.2	5.4

DCT = dynamic contour tonometer;
HAT = handheld applanation tonometer;
Mean Diff = mean Comparator minus mean GAT value;
mmHg = millimetres of mercury;
NCT = non contact tonometer;
ORA = ocular response analyser;
Ran Err = the random error (estimated standard deviation of the differences);
RT = rebound tonometer;
Pred Int = the prediction interval incorporating estimated between study heterogeneity

Table 3 Recordability and study size

Tonometer	No. Studies	Recordability (%) - median (range)	Study size - median (range)
DCT	6	99 (93-100)	148 (63-211)
HAT	0	N/A	N/A
NCT	4	98 (76-100)	81 (45-100)
Ocuton S	2	88 (82-94)	77 (68-85)
ORA	2	98 (98)	57 (50-63)
RT	4	100 (50-100)	145 (36-150)
TonoPen	3	100 (90-100)	146 (103-208)
Transpalpebral	9	95 (76-97)	101 (62-213)
DCT = dynamic contour tonometer; HAT = handheld applanation tonometer; N/A = non applicable; NCT = non contact tonometer; ORA = ocular response analyser; RT=rebound tonometer			

Figure 1 Flow diagram of the selection process

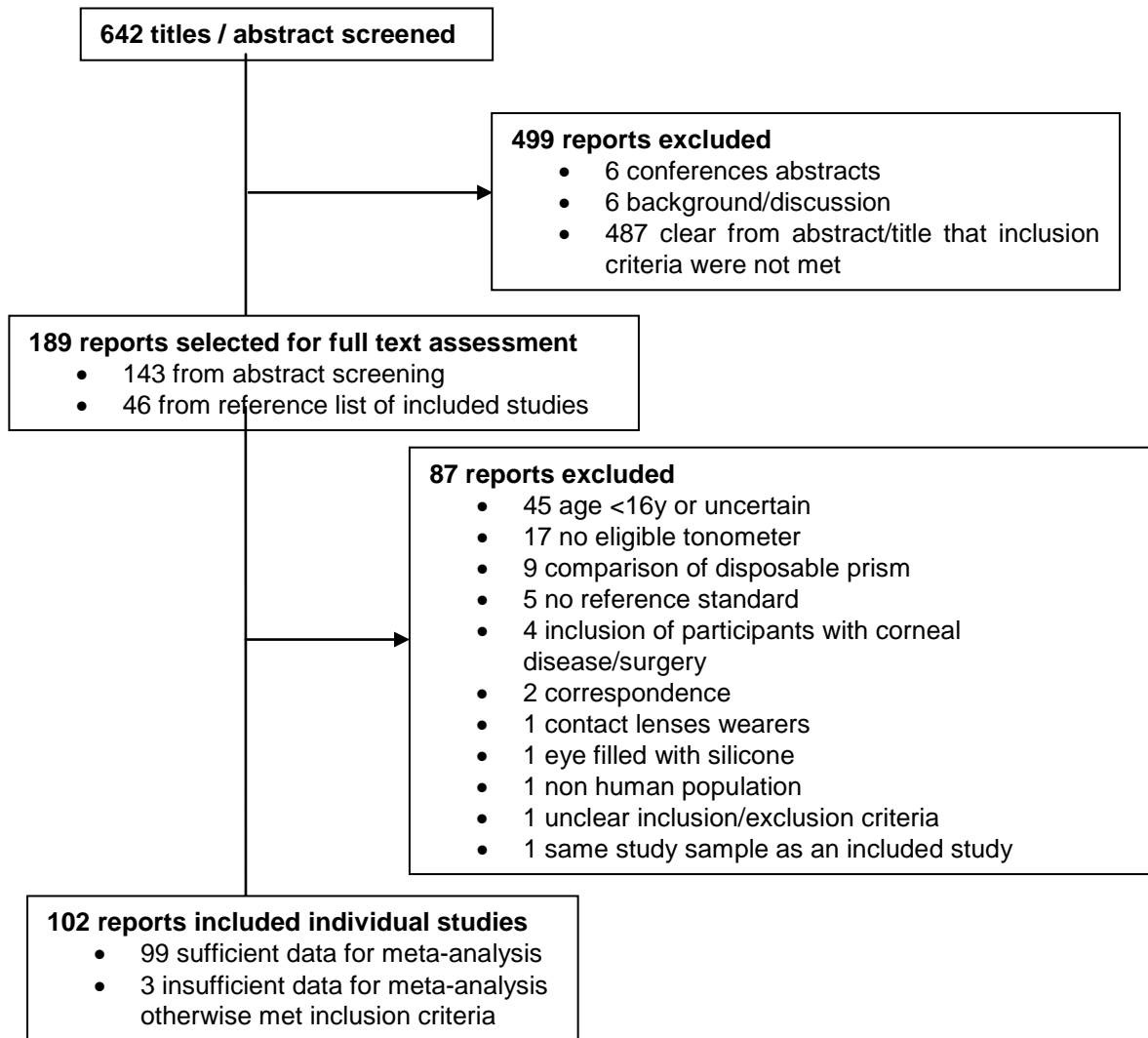


Figure 2 – Quality Assessment of included studies

