CLINICAL SCIENCE

Evaluation of the validity and reliability of A-scan ultrasound biometry with a single use disposable cover

K Cass, C M Thompson, C Tromans, I C J Wood

Background: The UK Medical Devices Agency has suggested that ophthalmic practitioners should, where practicable and not compromising clinical outcome, restrict corneal contact devices to single patient use to minimise a remote theoretical risk of transmission of new variant Creutzfeldt-Jakob disease (vCJD). This study reports on a modified technique of ultrasound A-scan biometry that complies with the MDA recommendations.

Methods: The right eyes of 37 consecutive hospital patients had a series of biometry readings taken with a Humphrey 820 A-scan instrument with a plane wave transducer used conventionally and with the addition of a disposable latex cover.

Results: Intrasequential repeatability of axial length measurements was similar for conventional readings—mean difference 0.027 mm, 95% confidence intervals (CI) ± 0.44 mm and those taken with a disposable cover (0.028 mm, CI ± 0.38). Intersessional repeatability was equivalent with (0.002 mm, CI ± 0.51) and without a cover (0.03 mm, CI ± 0.51). Readings with a cover were not significantly different from those without (paired t test; p >0.05), but tended to be greater (mean difference 0.083 mm, CI ± 0.60).

Conclusions: These findings suggest that corneal contact biometry with a disposable cover is a viable and theoretically safer alternative to the conventional technique.
method exactly doubles each recorded measure using method
necessarily a measure of agreement. This method was preferred to
throughout the recorded range. This method was preferred to
fixation which prevented full data being collected (n = 3).
The biometer probe required two applications of viscous
coupling solution, one before and a second after the placement
of the disposable cover, to obtain an adequate signal. Care was
also taken to ensure a reservoir of fluid was not held beneath
the cover by shaking off excess coupling fluid before
placement of the cap and squeezing out any remaining reser-
voir once the cap was in place. This was achieved by pressing
the protected biometer probe against the sterile cap mount.
There was no noticeable attenuation of the signal. It was not
necessary to increase the signal gain from the 60% default
value.

In automatic mode the Humphrey 820 model only accepts
measurements that meet its programmed criteria for consist-
ency and amplitude. These are presented as a frozen display
and the operator is warned by an auditory signal. In the cur-
cent study all “frozen” readings were accepted by depressing
the instrument foot pedal once. Five such readings were
accrued and stored. On obtaining five readings the instrument
would automatically compute an average after discarding the
lowest two stored axial length measurements. If an error sig-
nal was obtained (signified by the appearance of “The axial
length criteria not met” message) the investigator would
review the stored measurements and traces and discard any
doubtful traces and extreme readings. Replacement readings
were then taken. At the conclusion of the reassessment
appointment a final measurement was taken both without
(study reading 5) and with the disposable cover (reading 6) in
order to obtain an intersessional comparison of readings.

Statistical analysis
Our data are represented in the now widely accepted Bland
and Altman style. The difference between the two measures
plotted versus the mean represents the degree of agreement,
95% of the differences between the two measurements fall
within the confidence limits which are calculated from the
mean difference ± 1.96 x standard deviation of differences.
This allows any bias between the two methods to be viewed
throughout the recorded range. This method was preferred to
correlation coefficients as the measure of correlation is not
necessarily a measure of agreement. If, for example, a second
method exactly doubles each recorded measure using method
one there will be perfect correlation but no agreement. Finally,
any bias was assessed statistically as the mean of differences
compared to zero (t test).

Disposable cover parameters
On completion of the clinical data collection, two separate
investigations into the mean thickness and within batch vari-
ability of the disposable covers were carried out on covers from
the same batch as used in the main study. The first of these
applied an optical technique involving travelling microscopy
in order to provide data on the predicted maximum thickness
with the cover is in its dry non-stretched state. Initially, two
observers recorded their interpretation of 10 separate readings
using a travelling microscope’s vernier measuring scale blind
to their colleague’s results. Subsequently, a disposable cover
had a vertical cut made in its end and was then mounted so as
to enable the unstretched thickness to be measured using the
travelling microscope. The two observers then each took five
repeated measures of the thickness. The procedure was
repeated for a further nine covers, making a total of 10 covers.
The mean thickness and variability of the covers, intraobserver
variability, and intraobserver agreement on vernier measures
were then analysed using a series of paired t tests.

Since readings from covers in an unstretched state presum-
ably represent a greater thickness than that actually obtained
during the data collection process two more sophisticated
measurement systems—laser interferometry and coordinate
measurement—were used to gain information on their thick-
ness when stretched in a manner similar to the conditions
under which they were used. For these methods it was neces-
sary to manufacture a dummy transducer from silver steel
that was then clamped into position for use with the Hewlett
Packard model HP5529A laser interferometer and also the
LKG90-C coordinate measuring machine (Fig 2). Both
techniques are capable of resolving thickness to 1/10th µm
(0.0001 mm).

In the interferometric technique the position of the end of
the dummy transducer in relation to the laser was measured,
with the assistance of “dynamic calibrator” computer soft-
ware, and the measurement gauge set to zero before
placement of the disposable cover. After wetting the trans-
ducer and placing a disposable cover over it a further applica-
tion of coupling solution was used on the end of the cover, as
had been performed during the data collection. The position of
the end of the covered transducer in relation to the laser was
measured 10 times, the change in the distance, as recorded by
the gauge (that is, half the change in path length), being equal
to the thickness of the disposable cover. After carefully remov-
ing the cover the measurement gauge of the instrument was
rechecked to ensure that the position of the dummy probe had
not changed during positioning or removal of the cover. The

Figure 2 Close up view of the dummy transducer with disposable
cover in place mounted on the coordinate measurement test rig
during a measurement.

Figure 1 (A) “Oculofilm” disposable cover in unstretched form
shown adjacent to the Humphrey ultrasound biomier model 820
transducer probe. (B) The oculofilm disposable cover shown in situ,
stretched over the biomier probe.
procedure was repeated for nine further disposable covers. An equivalent method was used for the coordinate measurement technique with the constraint that the outer surface of the cover could not be wetted. A further four samples taken from different batches were subjected to the same travelling microscopy and coordinate measurement techniques in order to provide data on interbatch variability.

RESULTS

Reliability and validity

Analysis of the Bland and Altman\* plot of the repeatability data (Fig 3) illustrates the even distribution of the data points for repeated measures throughout the measured range regardless of whether or not a protective cover was used. When comparisons are made of the means with and without the cover the mean difference and 95\% confidence intervals do, however, increase to 0.1 mm and CI ± 0.6 mm respectively for the intrasessional validity comparisons (Fig 4 and Table 1). The bias seen is for slightly longer readings (on average 0.085 mm) to be obtained with the disposable cover in place, therefore suggesting that the thickness of the disposable cover does perhaps have some influence on the measurement of axial length.

In order to assess the variation in any repeated measure, paired \( t \) tests were carried out for intrasessional repeatability of the biometry measures with and without the disposable cover using the initial paired readings. \( T \) tests on intersessional measures compared the first biometry measure from session one with and without the disposable cover with the respective measure in the final session. To validate the biometry measures with the disposable cover \( t \) tests were then performed comparing the first measure without the disposable cover and the first measure with the disposable cover. None of these \( t \) tests found the distribution of the obtained readings to be significantly different at the \( p = 0.05 \) significance level (Table 1). This lack of significance suggests very little variation in the repeated biometry measures regardless of whether they were taken either with or without the disposable cover. The mean difference of repeated measures remained in the order of ≅ 0.03 mm for both the intrasessional and intersessional conditions.

### Table 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean difference (mm)</th>
<th>95% confidence intervals (mm)</th>
<th>( t ) test result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrasessional repeatability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without cover</td>
<td>0.027</td>
<td>± 0.44</td>
<td>( p=0.20 ) NS</td>
</tr>
<tr>
<td>(reading 1 – reading 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with cover</td>
<td>0.028</td>
<td>± 0.38</td>
<td>( p=0.59 ) NS</td>
</tr>
<tr>
<td>(reading 3 – reading 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrasessional validity</td>
<td>0.07</td>
<td>± 0.59</td>
<td>( p=0.82 ) NS</td>
</tr>
<tr>
<td>with cover</td>
<td>0.10</td>
<td>± 0.60</td>
<td>( p=0.37 ) NS</td>
</tr>
<tr>
<td>(reading 6 – reading 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersessional repeatability</td>
<td>0.03</td>
<td>± 0.51</td>
<td>( p=0.20 ) NS</td>
</tr>
<tr>
<td>without cover</td>
<td>0.002</td>
<td>± 0.51</td>
<td>( p=0.72 ) NS</td>
</tr>
<tr>
<td>(reading 1 – reading 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( NS \) = not significant.
Study reading 1 = initial uncovered; 2 = repeat uncovered; 3 = initial covered; 4 = repeat covered; 5 = final uncovered; 6 = final covered.

---

**Figure 3** Intrasessional repeatability of readings taken without (crosses) and with (open circles) the disposable cover during session 1. Difference data are plotted versus mean data. Reading “A” indicates the first and “B” the second of a pair. The overall mean differences were virtually identical for each condition (centre dotted line), the outer lines represent 95\% confidence limits without (solid line) and with the cover (broken line). Table 1 summarises numerical data relating to this figure.

**Figure 4** Intrasessional validity of readings taken during session 1. The data are computed from the first of each pair of readings. The centre line represents the mean difference and the outer lines the 95\% confidence limits. A slight preponderance of positive data (ordinate) indicates a tendency for longer readings being obtained with the disposable cover in place. Table 1 summarises numerical data relating to this figure.
Tonopen cover thickness and observer agreement

The mean thickness of 10 unstretched Tonopen covers from the main study sample was found to be 0.069 mm, CI ± 0.020 mm (Table 2). The overall range for this sample extended from 0.055 mm to 0.083 mm. The mean of the differences between the two observers was 0.003 mm. A two tailed t test comparing the readings of the two observers found no statistical difference (p = 0.14). When assessing the agreement of two observers repeatedly recording the same vernier reading, on four of 10 occasions the observers were in exact agreement, with each other, while three readings differed by 0.01 mm and three by 0.02 mm—that is, the mean difference was 0.009 mm CI ± 0.0018 mm. The mean unstretched thickness ranged between 0.069 and 0.088 mm for the remaining four batches (Table 2).

When the disposable covers were placed on the dummy transducer in a scenario most likely to replicate the actual conditions under which data were collected the mean thickness of the 10 disposable covers in our study sample was 0.012 mm (12 µm) with CI ± 0.0029 mm (2.9 µm). The range extended from 0.009 to 0.019 mm.

The results of the coordinate measuring technique were encouragingly close at 0.013 mm (13 µm) but with a greatly reduced confidence intervals 0.0012 mm (1.2 µm). The latter therefore became our first choice method for the remaining four batches. Details of these batches, both unstretched and stretched, and the results of the interferometry technique used on the main study sample are summarised in Table 2.

**Table 2** Summary of average thickness readings (and 95% confidence intervals), in unstretched and stretched state, for samples of 10 disposable oculofilm covers derived from five separate batches

<table>
<thead>
<tr>
<th>Batch study sample number</th>
<th>Unstretched readings (mm)</th>
<th>Stretched readings (mm)</th>
<th>Interferometry (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Travelling microscopy (CI)</td>
<td>Coordinate (CI)</td>
<td></td>
</tr>
<tr>
<td>1 (used in main study)</td>
<td>0.069 (0.020)</td>
<td>0.013 (0.0012)</td>
<td>0.012 (0.0029)</td>
</tr>
<tr>
<td>2</td>
<td>0.079 (0.020)</td>
<td>0.010 (0.0008)</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>0.080 (0.025)</td>
<td>0.008 (0.0008)</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>0.088 (0.024)</td>
<td>0.009 (0.0018)</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>0.069 (0.018)</td>
<td>0.010 (0.0024)</td>
<td>–</td>
</tr>
</tbody>
</table>

DISCUSSION

Humphrey Instruments claims an inherent instrument accuracy of ± 0.034 mm for the biometer model 820 in their instrument specification literature and a patient measurement accuracy of ± 0.10 mm. Our largest mean difference for repeatability and reproducibility was in the order of 0.03 mm and as such compares favourably with the optimum accuracy. Two other groups have reported repeatability for the Humphrey 820 biometer. The two experimenters in Rudnicka’s study produced absolute mean differences of 0.42 mm and 0.008 and 95% confidence intervals in the order of ± 0.25 mm and ± 0.20 mm, respectively, when considering the data of both eyes. Zadnik and coworkers did not report directly on axial length but their separate data produced 95% confidence intervals of ± 0.29 mm for anterior chamber depth, ± 0.20 mm for lens thickness, and ± 0.37 mm for vitreous chamber depth.

The earlier studies differ in their methodology—one using a slit lamp mounted probe, the other a handheld approach similar to that used in our study. This may explain the closer resemblance of results from the two studies using the handheld probe. The pragmatic nature of this investigation and its sample size mean that there are no very short or very long eyes represented in the study population. The available data do not, however, suggest that the axial length of individual eyes is of relevance to the findings.

Both the previous studies differ from the present study by using subjects free of pathology, which we presume would give less variation in the reflected wave. In view of the current MDA advice our study population were, of necessity, restricted to patients undergoing biometry as part of their standard clinical investigations. Forty three per cent of the eyes measured for the study had a best corrected visual acuity of 6/24 or worse. Forty eight per cent of the patients had a best corrected acuity in the left, fixing eye, of 6/24 or worse with an overall range of acuities from 6/5 to hand movements, which may have had a detrimental effect on the fixational stability of the eye. It was considered, however, that the use of a distance spotlight as the fixation target throughout the study would help to reduce the impact of any such instability. Since the standard clinical protocol during ultrasound biometry would not normally include correction of ametropia, the levels of vision experienced in the fixing eye of some “ophthalmologically normal” observers during the procedure may not differ from those of our observers. With the exception of the three patients who were excluded from the study on account of very poor fixation, no difficulty was noted in obtaining readings from our “typical” clinical population, either with or without the disposable cover in situ. Considering all of the study population, readings were obtained with equal ease whether or not a cover was in use during measurement.

Whereas the average unstretched thickness of the disposable covers was equivalent to 90% of the mean discrepancy between readings taken with and without the cover, the average stretched thickness of the covers in the simulated interferometric test rig was equal to only 12% of this difference. The extent to which the oculofilm covers thin (to an average of 10 µm) compared with their original dry state (average 77 µm) when wet and stretched over the dummy transducer was a source of initial surprise. This change may be due to a reduction in the friction between the latex cover and dummy probe on wetting the inner surface. It is also apparent from Table 2 that the thickness of the stretched covers cannot readily be estimated from the unstretched thickness as when ranked in order of thickness unstretched the covers do not correspond with the ranking when stretched. One would assume the cover surface profile would be a discontinuous surface of peaks and troughs; hence, as the coordinate measurement takes a reading the final recorded thickness will be highly dependent on the region being measured be it a peak or a trough. Perhaps the cover could have been removed and replaced between each measurement to overcome this problem, though it is unclear as to whether this would have ultimately strained the cover, weakening the polymer and giving rise to an unnaturally thin reading or perhaps even damaging the cover, preventing collection of the full set of measurements.

The possibility of the 0.085 mm increase in axial length being related to the properties of ultrasound transmission through latex was considered. The speed of sound in latex rubber is in the order of 1550 m/s (± 10 m/s) compared with
the 820 biometer’s assumed speed of 1532 m/s for vitreous and aqueous. One would not therefore anticipate an appreciable increase in the measured axial length with the cover in place over and above the actual cover thickness, in fact when using a typical cover of 10 µm this would result in an increase of 0.11 µm to the cover thickness.

The clinical implications of our finding of a marginal increase in measured axial length with the disposable cover is not the only method of avoiding direct corneal contact when undertaking ultrasound biometry. A further possible technique, which is of particular value in the examination of unanaesthetised infants, is to incorporate into the on-screen software of the biometer used in this study a shift towards hyperopia in the order of 0.16 dioptres (CI ± 1.27D) occurs (Fig 5). This is, of course, smaller than the power increments for a typical IOL, usually 0.50D. Omission of the three readings falling outside the confidence intervals results in a mean difference of only +0.03D. The differences in refractive error found based on the two axial length readings are in fact independent of the magnitude of the patients’ keratometry readings, providing the keratometry readings remain constant for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia for each of the two calculations. The tendency for a bias in the measurement generated by this method, a shift towards hyperopia.