



Reliability of Schirmer and Phenol Red Thread test value on dry eye patients



Authors:

Urvashni Nirghin¹ 
Tharuna Govender¹ 
Yeshalya Hirachund¹ 
Rorisang Mosololi¹ 
Hlengiwe Nxumalo¹ 
Sumaya Patel¹ 
Philani Zwane¹ 
Khathutshelo P. Mashige¹ 

Affiliations:

¹Department of Optometry,
Faculty of Health Science,
University of KwaZulu-Natal,
Durban, South Africa

Corresponding author:

Urvashni Nirghin,
nirghinu@ukzn.ac.za

Dates:

Received: 15 May 2024
Accepted: 25 Nov. 2024
Published: 29 Jan. 2025

How to cite this article:

Nirghin U, Govender T,
Hirachund Y, et al. Reliability
of Schirmer and Phenol Red
Thread test value on dry eye
patients. Afr Vision Eye
Health. 2025;84(1), a946.
[https://doi.org/10.4102/
aveh.v84i1.946](https://doi.org/10.4102/aveh.v84i1.946)

Copyright:

© 2025. The Author(s).
Licensee: AOSIS. This work
is licensed under the
Creative Commons
Attribution License.

Read online:



Scan this QR
code with your
smart phone or
mobile device
to read online.

Background: The repeatability of the Schirmer test (ST) and Phenol Red Thread test (PRT) has been investigated thoroughly on patients with normal tear function; however, it is limited to dry eye patients.

Aim: This study aimed to compare the reliability of the ST and PRT on young adults with dry eyes.

Setting: University of KwaZulu-Natal Eye Clinic.

Methods: Forty-eight young adults (mean age 20.65 ± 1.71 years) participated in the study. Preliminary tests excluded dry eye-related abnormalities. Novesin (0.4%) was instilled into the right eye of each participant, followed by the administration of the ST and PRT, 5 min apart, then re-administered after a 10-min interval. The retest was conducted 1 week later with the PRT being performed first to minimise bias.

Results: The intra-session correlation of the ST and the PRT was poor ($ICC < 0.5$) and the difference was statistically significant ($P < 0.05$). The inter-session correlation for the PRT was poor ($ICC = 0.432$) while the ST was moderate ($ICC = 0.567$) with no statistical difference observed for all parameters ($P < 0.05$). The Bland–Altman plots for intra- and inter-session indicate that 96% of observations fell within the 95% limits of agreement.

Conclusion: The ST and PRT demonstrated good intra-session and inter-session reliability, with ST showing better inter-session agreement than the PRT.

Contribution: The ST and the PRT provide consistent and reproducible results on dry eye patients when conducted on the same individuals at different time points.

Keywords: dry eye; Schirmer test; Phenol Red Thread test; reliability; intra-session; intersession.

Introduction

Dry eye disease (DED), or Keratoconjunctivitis sicca (KCS), is a chronic condition caused by disruption in the ocular tear film layer associated with poor quality of tears (Evaporative Dry Eye [EDE]) and or insufficient quantity of tears (Aqueous Deficient Dry Eye [ADDE]).¹ Dry eye disease may lead to inflammation and damage to the outer structures of the eye, resulting in symptoms that include discomfort and blurred vision.¹ The prevalence of DED varies globally among different populations and has been reported to range from 5% to 50%.² The aetiology of the condition is linked to infections, allergies, systemic diseases, environmental factors, age and genetic associations, among others.¹ Considering its multifactorial nature, the diagnosis of DED involves a battery of tests including but not limited to, symptom questionnaires, staining of the ocular surface, tear osmolarity and objective assessment of tear volume or production.^{3,4}

While a range of tests are available for dry eye assessment, the use of appropriate clinical tests is essential for the accurate diagnosis and management of dry eyes. The reliability of these tests needs to be assessed before being used for clinical and research purposes. Reliability or precision refers to the close alignment between repeated measurements and is expressed as repeatability and reproducibility.^{5,6} Various factors contribute to discrepancies in repeated measurements including time intervals between measurements and different instruments used.⁶ Repeatability is a measure of consistent results of an instrument conducted on the same participant by the same examiner under the same testing conditions. In contrast, reproducibility is the consistency of measurements despite a varied set-up such as a different time interval between measurements.⁶

The Schirmer test (ST) and Phenol Red Thread test (PRT) are two common diagnostic tests used to determine ADDE. The ST utilises a filter paper which has some drawbacks including extended testing period (5 min) and patient discomfort.⁷ The PRT employs a thread impregnated with phenol red, is quick to perform in less than 15s with less discomfort and can be easily administered in children.^{8,9} Reflex tearing can occur during both tests; therefore, anaesthesia helps mitigate the discomfort, limits reflex tearing and provides more objective and reliable results measuring true basal tear production.^{4,10}

Nichols et al. conducted a study on inter-session repeatability of the PRT and ST, without anaesthetic, on dry eye patients and reported poor repeatability of the ST.¹¹ The authors suggested that to increase the reliability of the test, reflex tearing should be controlled with the use of an aesthetic. The authors, however, did not investigate intra-session repeatability. Saleh et al.¹² reported poor agreement between ST and PRT values on young adults and elderly Caucasian patients who presented with symptoms of dry eye; however, no inter-session reproducibility was conducted. Senchyna et al.¹³ reported better repeatability in PRT values compared to ST, and concluded that the PRT was reliable in diagnosing dry eye. Vashisht et al.¹⁴ conducted a comparative study on 40–80-year-old Indian patients and found that there was a strong agreement between ST and PRT values. The study, however, did not report on the inter-session reproducibility of the two tests.¹⁴ The above-mentioned studies suggest that the reliability of the PRT and ST test has yielded conflicting results and intra-session and inter-session analyses are not always reported. Furthermore, the repeatability and reproducibility of these tests on dry eye patients are somewhat limited, while the use of topical anaesthetic, to limit the effect of reflex tearing for more objective results, has not been extensively explored in reliability studies. The current study therefore aims to assess and compare the intra-session repeatability and inter-session reproducibility of the PRT and ST test, with the use of anaesthetic, on dry eye patients.

Ethical considerations

Ethical approval to conduct the study was obtained from the Biomedical Research and Ethics Committee (BE144/19). The study followed the principles of the Declaration of Helsinki, with informed consent obtained from participants after an explanation of the study's purpose and potential implications.

Methods

This was a quantitative, observational and comparative study conducted among young adults with ADDE. The participants were selected using a nonprobability purposive sampling methodology. A sample-size estimate for this study was $n = 48$ based on the within-subject standard deviation (S_w) to achieve a 20% margin of error at a 95% confidence level using the precision experiment formula,

$$1.96 \frac{S_w}{\sqrt{2n(n' - 1)}} = 0.2S_w, \text{ for repeated measurements } (n').^6$$

The preliminary assessment included: (1) case history to identify any confounding variables, (2) completion of a range of questionnaires (Ocular Surface Disease Index [OSDI], Standard Patient Evaluation of Eye Dryness [SPEED] and McMonnies) to identify symptomatic patients, as relying on a single questionnaire may result in potentially missed patients,² (3) slit-lamp examination to rule out ocular pathology,⁴ (4) and Tear Break Up Time (TBUT) of $< 10 \text{ mm}^4$ to exclude EDE patients, and finally, (5) ST evaluation, without anaesthetic, to assess for the presence of ADDE with a reading of $\leq 10 \text{ mm}$ on ST.¹⁴ Upon adherence to the inclusion criteria, healthy males and females, between the ages of 18 and 30 years, identified with the presence of ADDE and symptomatic, were enlisted in this study. Those with ocular infection or inflammation, prior surgery, pregnant or lactating females, and those taking medication such as antihistamines, diuretics, oral contraceptives, eye cosmetics and contact lens wearers were excluded from the study.

Data collection procedure

Following the preliminary screening, the recruited participants proceeded to the PRT and ST to assess the reliability of these tests. The ST was performed first followed by the PRT, and the timing of the tests was monitored with a digital stopwatch. Novesin (0.4%) was considered the anaesthetic of choice for this study because of its lower concentration; hence, reduced drug absorption rate and risk of epithelial damage or allergic responses, particularly in DE patients.^{15,16} Furthermore, its duration of action of about 15 min allowed sufficient time for the completion of the testing procedures,^{15,16} Novesin (0.4%) was instilled into the right eye of each participant. A cotton-wisp test was performed to ensure the cornea was desensitised, followed by the placement of the Schirmer strip in the inferior fornix of that eye. The eye remained closed throughout the procedure to minimise eye movement for more consistent results.^{4,15} Following a 5-min timing, the strip was removed, and the wet portion below the folded edge was measured immediately. Five minutes later, the cotton-wisp test was performed again ensuring the cornea was still anaesthetised and the PRT was conducted on the same eye. The timing between the ST and PRT was 5 min to avoid potential interference with results, as considered in previous studies.^{10,12,17} The thread was positioned on the lateral part of the lower eyelid, with the eye remaining closed during the entire process. Upon contact with the tears, the PRT altered its colour from yellow to red, and after 15 s, the thread was removed, and the red portion below the bend was measured concluding session one. For session two, 10 min later, the procedure was repeated with the PRT conducted first followed by the ST. During and after the testing procedures, patients were advised to report any discomfort they experienced, but no discomfort was reported.

A repeat evaluation was administered a week later, with the PRT conducted first, followed by ST as a measure to eliminate any bias in results. Test and retest were conducted in the same room and under a similar uniform environment. Furthermore, to eliminate intra-examiner variability, a single

examiner performed all tests on both occasions. In addition, a predetermined protocol was followed to minimise systematic differences.

Data management and analysis

The Statistical Package for Social Sciences version 26.0 (SPSS, Inc., Chicago, IL, United States) was used to conduct statistical analysis. Normal distribution was tested using the Shapiro–Wilk test and descriptive statistics was used to express means \pm standard deviations (s.d.), while intra-session and inter-session reproducibility were assessed using the Bland–Altman analysis, One-sample *t*-test (significant Level at $P < 0.001$) and intra-class correlation coefficient (ICC) at a 95% confidence interval (CI). The ICC levels were interpreted as poor (< 0.5), moderate (0.50 to 0.75), good (0.75 to 0.90) and excellent (> 0.90).¹⁸ The paired-sample *t*-test established the comparative reliability of the PRT and ST with $P < 0.001$ considered as statistically significant.

Results

A total of 48 subjects including 40% males ($n = 19$) and 60% females ($n = 29$) participated in the study. Their ages ranged from 18 to 25 years (mean = 20.65 ± 1.71 years) with 23% being Indian ($n = 11$) and 77% African ($n = 37$).

Table 1 shows the intra-class correlation and one-sample *t*-test results for intra-session repeatability of PRT and ST. These tests were conducted using two measurements (1 and 2) taken 10 min apart on the first day (Day 1) and repeated 1 week later (Day 2). The intra-session means \pm s.d. for PRT values and ST values for Day 1 were $2.21 \text{ mm} \pm 3.04 \text{ mm}$ and $1.81 \text{ mm} \pm 2.66 \text{ mm}$, while for Day 2 were $1.58 \text{ mm} \pm 2.83 \text{ mm}$ and $1.10 \text{ mm} \pm 2.44 \text{ mm}$, respectively. The ICCs for PRT and ST measurements were < 0.5 suggesting poor intra-session association on both days with the one-sample *t*-test revealing statistically significant differences ($P < 0.001$) for all parameters except for ST on Day 2 ($P = 0.003$). An intra-session repeatability comparison of PRT and ST for Days 1 and 2 displays a poor correlation (ICC < 0.50) that is statistically different ($P < 0.001$) (Table 1).

The Bland-Altman plots were used to evaluate the agreement between the first and second readings of PRTs and STs for Days 1 and 2. The results are displayed in Figure 1a–d. The three horizontal lines depict the mean (solid line) and the upper and lower Limits of Agreements (LoA) (dashed lines) which are set at mean ± 1.96 s.d., representing a 95% CI for PRT (Figure 1a and b), 96% (46/48) and 100%

(48/48) of the measurements fall within the 95% LoA for both days. For ST (Figure 1c and d), on Days 1 and 2, 100% (48/48) and 93.75% (45/48) of the measurements are within the 95% LoA, respectively. The mean differences for PRT on Days 1 and 2 are 2.21 mm and 1.58 mm while ST is 1.81 mm and 1.10 mm.

Table 2 shows the inter-session reproducibility and comparison of the PRT and ST for the two sessions (Days 1 and 2). The difference in means for PRT and ST is 0.326 and 0.243, respectively with PRT having poor inter-session association (ICC = 0.432), while ST being moderate (ICC = 0.567). No statistical difference was observed for both PRT and ST ($P > 0.001$) measurements. Comparatively (Table 2), PRT and ST readings reflect a statistically significant difference ($P < 0.05$) with a poor association (ICC = 0.053) between the two tests.

The Bland–Altman plots (Figure 2) were used to evaluate the inter-session reproducibility of the PRT and ST conducted on the same participants but on different days. The three horizontal lines depict the mean (solid line) and the upper and lower LoA (dashed lines), positioned at the mean ± 1.96 standard deviations, indicating a 95% CI. The plots display that 96% (46/48) of the PRT (Figure 2a) and 97.9% (47/48) of the ST (Figure 2b) values lie within the 95% LoA.

Discussion

Dry eye is a prevalent condition affecting millions globally. Among numerous dry eye diagnostic tests available, PRT and ST are commonly used to measure tear production and volume; however, their reliability appears inconsistent. Furthermore, the repeatability and reproducibility of these tests on dry eye patients, even with topical anaesthetics, are somewhat constrained, with this study providing evidence of good reliability of PRT and ST.

The greater participation of females in this study may be attributed to dry eye being more prevalent in females than males.⁹ The contributing factors include a higher susceptibility of females to systemic health issues, which may increase their risk of developing dry eye.⁹ With controlled variables in place, a possible explanation for the presence of dry eyes could be hormonal.¹⁹ Young adults (mean age of 20.65 ± 1.71 years) with ADDE were included in the study despite the risk of dry eye being lower in the younger population.² Considering restricted repeatability studies on young adults, and the prevalence of dry eye in younger adults and school children on the rise,² possibly related to the increased use of

TABLE 1: Intra-session repeatability and comparison of the Phenol Red Thread and Schirmer tests.

Day	Test	Intra-class correlation coefficient			Mean	One-sample <i>t</i> -test		Intra-class correlation coefficient			Paired-sample <i>t</i> -test	
		ICC	Lower bound	Upper bound		s.d.	<i>P</i>	ICC	Lower bound	Upper bound	<i>P</i>	
1	PRT	0.146	0.078	0.377	2.21	3.04	0.000	-0.003	-0.162	0.193	0.000	
	ST	0.340	0.065	0.567	1.81	2.66	0.000	-	-	-	-	
2	PRT	0.134	-0.097	-0.370	1.58	2.83	0.000	0.056	-0.109	0.254	0.000	
	ST	0.330	0.065	0.056	1.10	2.44	0.003	-	-	-	-	

PRT, Phenol Red Thread test; ST, Schirmer test; s.d., standard deviation.

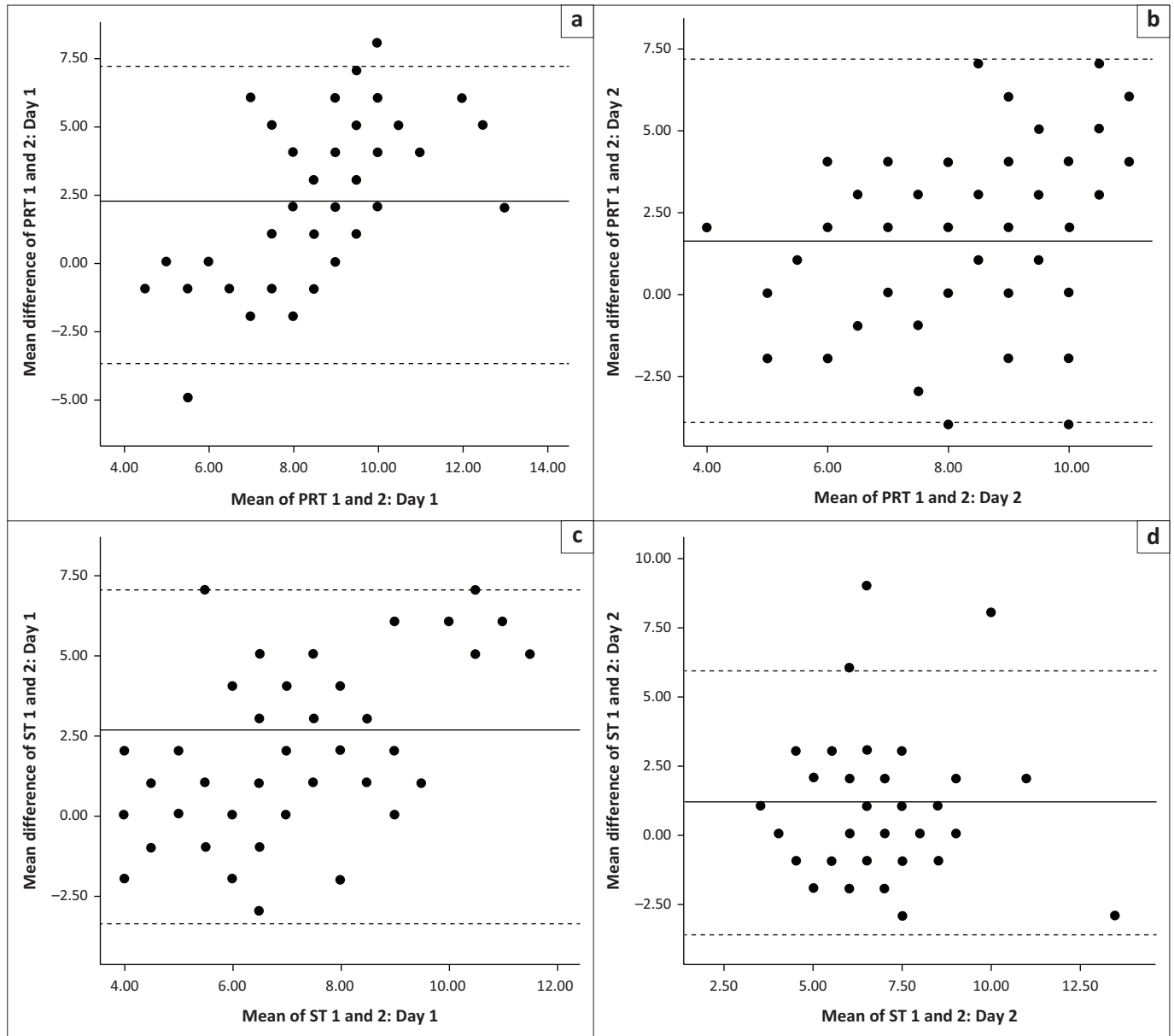


FIGURE 1: Bland–Altman plots showing intra-session repeatability of Phenol Red Thread test (a and b) and Schirmer test (c and d) on Days 1 and 2.

TABLE 2: Inter-session reproducibility and comparison of the Phenol Red Thread test and Schirmer test.

Test	Intra-class correlation coefficient (ICC)			One-sample <i>t</i> -test			Intra-class correlation coefficient (ICC)			Paired-sample <i>t</i> -test
	ICC value	Lower bound	Upper bound	Mean	s.d.	<i>P</i>	ICC value	Lower bound	Upper bound	<i>P</i>
PRT	0.432	0.177	0.635	0.3264	1.56	0.153	0.053	-0.097	0.239	0.000
ST	0.567	0.342	0.731	0.2431	1.43	0.246	-	-	-	-

PRT, Phenol Red Thread test; ST, Schirmer test; s.d., standard deviation.

digital devices, more studies are encouraged in this younger cohort. The majority of participants were African, followed by Indian, with no other racial groups represented. A previous study, in a similar setting, found a greater occurrence of dry eye among the African population compared to other race groups.²⁰ This supports the representation of more Africans with ADDE in this study.

Intra-session repeatability

The intra-session correlation of both the PRT and ST is poor (ICC < 0.5) with the difference between measurements for

each test being statistically significant ($P < 0.001$). Furthermore, the intra-session comparison of PRTs and STs tests reveals the same. Previous studies reported similar findings^{10,12} indicating a lack of association between repeated measurements per test per session and between the two tests. Statistical differences suggest that the observed differences in results are unlikely to have occurred by chance. While Masmali et al.¹⁰ associated this difference with the participants in their study being non-dry eye patients, together with Saleh et al.,¹² the authors reported that each test may be measuring different components of the tear film hence the difference.

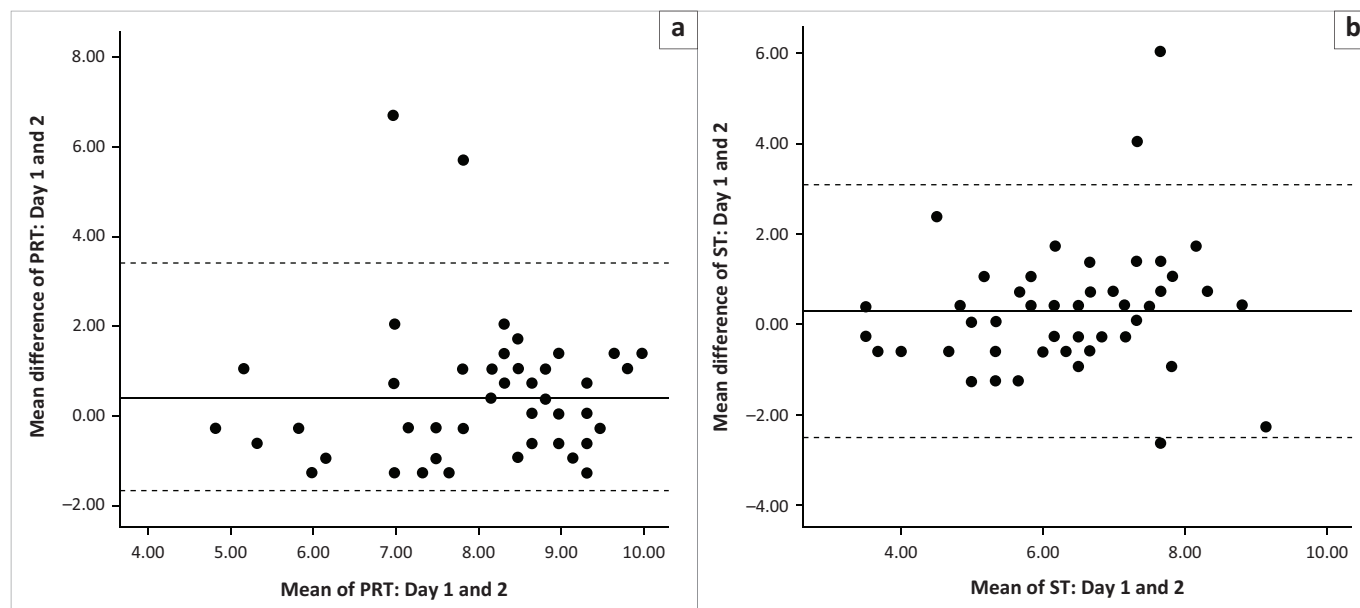


FIGURE 2: Bland–Altman plots showing inter-session reproducibility of Phenol Red Thread test (a) and Schirmer test (b) for Days 1 and 2.

The use of Bland–Altman plots is a common and preferred method to assess the agreement and repeatability of tests, providing a graphical representation that is less sensitive to outliers as opposed to ICC and *T*-tests.²¹ The Bland–Altman plots used to evaluate the intra-session agreement of the PRT and ST suggest that while there may be a poor association for each test measurement within each session (Day 1 or 2), there is good agreement in terms of the differences between measurements for each test except for ST on Day 2. Despite the use of anaesthetic, the ST may still induce some reflex tearing impacting its reliability.¹⁰ Furthermore, the 95% confidence limit for sessions one and two for PRT and ST provides a range within which the true values for the intra-session repeatability of the tests are likely to lie. Vashisht et al.¹⁴ also found a strong intra-session agreement between PRT and ST tests on dry eye patients; however, no anaesthetic was used in the procedure. The difference in means and range of the LoA for both the PRT and ST shows a greater variation in intra-session measurements with first readings being greater than the second. The current study, unlike Vashisht et al.,¹⁴ did not grade the severity of the dry eye, and together with the use of anaesthetic could have had a contributory effect on the poor intra-session correlation. Masmali et al.¹⁰ stated that despite the use of anaesthetic, the awareness of the test in the eye may itself be a stimulus to some reflex tearing. Even though there is good intra-session agreement for PRT and ST, the differences observed within sessions should be considered by averaging repeated measurements and considering a combination of tests in improving diagnosis, especially in patients with ADDE.

Inter-session reproducibility

The means \pm s.d. for PRTs and STs for inter-session reproducibility is lower compared to intra-session with ST being even lower with moderate association (ICC = 0.567) between the two sessions while PRT displays a poor

association (ICC = 0.432). This suggests a lesser variation is observed for inter-session than intra-session evaluations and for ST over PRT for ADDE patients; however, considering the differences observed are not statistically significant, the reproducibility of both the test measurements, 1 week apart, are similar. This is further supported by the Bland–Altman plots, which show that most measurements fall within the 95% LoA for both the PRT and ST when anaesthetic is used. This indicates that both tests are reliable for assessing ADDE over two sessions spaced 1 week apart, with the ST test exhibiting slightly better reproducibility. Similarly, Nichols⁷ found that ST was more reliable than the PRT test particularly when ST measurements were lower. A study by Kecskemet et al.,²² on the other hand, found that despite the good reproducibility of both tests, the PRT was more reproducible compared to ST as the level of protein in the tears was less likely to contaminate PRT results than ST.

Finally, the paired-sample *t*-test for inter-session comparison of PRT and ST displays a poor correlation between tests with measurements being statistically different ($P < 0.001$). As reflected in intra-session comparison, the discrepancy between tests may be attributed to each test assessing distinct components of the tears, and the consciousness of the test in the eye inducing some reflex tearing despite the use of anaesthetic.^{10,14}

As this study showed good intra-session repeatability and inter-session reproducibility of both PRT and ST on ADDE patients, the poor correlation especially for intra-session measurements for ST and PRT could have been influenced by the severity of the ADDE which was not controlled for in this study. The instillation of topical anaesthetic on DE patients increases the risk of epithelial damage; however, the use of one drop of Novesin (0.4%) at a time, and patients being advised not to rub their eyes during and after the testing procedure, as well as the application of ocular lubricant

ensured no epithelial complication. Furthermore, post-staining revealed no epithelial compromise.

Conclusion

The study shows that both PRT and ST are reliable measures for ADDE patients. These tests demonstrate consistency in repeatability and reproducibility when conducted on the same individuals at different time points. However, it's important to note that the final diagnosis of dry eye should not be based solely on one test.

Acknowledgements

The authors would like to thank all participants who took part in the study.

This article is based on the author's thesis entitled, 'A comparative study of the repeatability and reproducibility of Phenol Red Thread test and Schirmer test values on dry eye patients', towards the degree of Bachelor of Optometry in the Department of Optometry, University of KwaZulu-Natal, South Africa, December 2019, with supervisors Dr. U. Nirghin and Prof. K.P. Mashige. It is available here: <https://library.ukzn.ac.za/>.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors' contributions

U.N., T.G., Y.H., R.M., H.N., S.P., P.Z. and K.P.M. were responsible for the conception or design of the work, the drafting of the article as well as the data collection. U.N. and K.P.M. contributed to the critical revision of the article as well as the final approval of the study.

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability

The data that support the findings of this study are available from the corresponding author, U.N., upon reasonable request.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research.

The article does not necessarily reflect the official policy or position of any affiliated institution, funder, agency or that of the publisher. The authors are responsible for this article's results, findings and content.

References

- Craig JP, Nichols KK, Nichols JJ, et al. TFOS DEWS II. Definition and classification report. *Ocul Surf*. 2017;30:276e283. <https://doi.org/10.1016/j.jtos.2017.05.008>
- Stapleton F, Alves M, Bunya VY, et al. TFOS DEWS II. Epidemiology report. *Ocul Surf*. 2017;15:334–365. <https://doi.org/10.1016/j.jtos.2017.05.003>
- Sheppard J, Lee BS, Periman LM. Dry eye disease: Identification and therapeutic strategies for primary care clinicians and clinical specialists. *Ann Med*. 2023;55(1):241–252. <https://doi.org/10.1080/07853890.2022.2157477>
- Wolffsohn JS, Arita R, Chalmers R, et al. TFOS DEWS II diagnostic methodology report. *Ocul Surf*. 2017;15(3):539–574. <https://doi.org/10.1016/j.jtos.2017.05.001>
- Mashige KP. Repeatability and reproducibility of axial length, anterior chamber depth, and crystalline lens thickness measurements using the Nidek US-500 Echoscanner. *Afr Vis Eye Health*. 2015;74(1):1–6. <https://doi.org/10.4102/aveh.v74i1.16>
- McAlinden C, Khadka J, Pesudovs K. Precision (repeatability and reproducibility) studies and sample-size calculation. *J Cataract Refract Surg*. 2015;41(12):2598–2604. <https://doi.org/10.1016/j.jcrs.2015.06.029>
- Nichols KK, Mitchell GL, Zadnik K. The repeatability of clinical measurements of dry eye. *Cornea*. 2004;23(3):272–285. <https://doi.org/10.1097/00003226-200404000-00010>
- Hamano H, Hori M, Mitsunaga S, Kojima S, Maeshima J. Tear test (preliminary report). *J Jpn Contact Lens Soc*. 1982;24:103–107.
- Hao Y, Jin T, Zhu L, et al. Validation of the phenol red thread test in a Chinese population. *BMC Ophthalmol*. 2023;23(1):498. <https://doi.org/10.1186/s12886-023-03250-3>
- Masmali A, Alqahtani TA, Alharbi A, El-Hiti GA. Comparative study of the reliability of phenol red thread test versus Schirmer test in normal adults in Saudi Arabia. *Eye Contact Lens*. 2014;40(3):127–131. <https://doi.org/10.1097/ICL.0000000000000025>
- Nichols KK, Nichols JJ, Mitchell GL. The lack of association between signs and symptoms in patients with dry eye disease. *Cornea*. 2004;23(8):762–770. <https://doi.org/10.1097/01.icc.0000133997.07144.9e>
- Saleh TA, McDermott B, Bates AK, Ewings P. Phenol red thread test vs Schirmer's test: A comparative study. *Eye*. 2006;20(8):913–915. <https://doi.org/10.1038/sj.eye.6702052>
- Senchyna M, Wax MB. Quantitative assessment of tear production: A review of methods and utility in dry eye drug discovery. *J Ocul Biol Dis Inform*. 2008;1(1):1–6. <https://doi.org/10.1007/s12177-008-9006-2>
- Vashisht S, Singh S. Evaluation of phenol red thread test versus Schirmer test in dry eyes: A comparative study. *Inter J Appl Basic Med Res*. 2011;1(1):40–42. <https://doi.org/10.4103/2229-516X.81979>
- Li N, Deng XG, He MF. Comparison of the Schirmer I test with and without topical anesthesia for diagnosing dry eye. *Int J Ophthalmol*. 2012;5(4):478–481.
- Booyesen DJ, Booyesen JL. *The Southern African guide to topical ophthalmic drugs*. Chennai: Notion Press; 2022.
- Ghislandi GM, Lima GC. Comparative study between phenol red thread test and the Schirmer's test in the diagnosis of dry eyes syndrome. *Rev Bras Oftalmol*. 2016;75:438–442. <https://doi.org/10.5935/0034-7280.20160088>
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiroprac Med*. 2016;15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Versura P, Fresina M, Campos EC. Ocular surface changes over the menstrual cycle in women with and without dry eye. *Gynecol Endocrinol*. 2007;23(7):385–390. <https://doi.org/10.1080/09513590701350390>
- Castelyn B, Majola S, Motilal R, Naidu MT, Ndebele SA, Vally TA. Prevalence of dry eye amongst black and Indian university students aged 18–30 years. *Afr Vis Eye Health*. 2015;74(1):6. <https://doi.org/10.4102/aveh.v74i1.14>
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Keckskemet G, Toth-Molnar E, Janaky T, Szabo Z. An extensive study of phenol red thread as a novel non-invasive tear sampling technique for proteomics studies: Comparison with two commonly used methods. *Int J Mol Sci*. 2022;23:8647. <https://doi.org/10.3390/ijms23158647>