

Observer Variability in Head Circumference Measurement Using Routine Versus Non-Stretchable Tapes in Children

Thakkar P¹, Yagnik K², Parmar NT³, Das RR⁴, Thakkar UP⁵

Abstract

Introduction: Observer variability is known in taking head circumference. Stretchable tapes do not provide accurate results after being used repeatedly. The aim of this study was to compare the intra-observer and inter-observer variability in measuring head circumference (HC) or occipito-frontal circumference (OFC) by using routine stretchable and non-stretchable tapes. **Material and Methods:** In this cross-sectional study, all stable children in the age group of 1 month to 5 years attending the pediatric department of a tertiary care teaching hospital over six months period were included. The main outcome measures were: intra-observer & inter-observer variability in HC measurements by routinely used stretchable and non-stretchable tapes, and accuracy of readings of old stretchable tapes. **Results:** A total of 301 children were included. There was good correlation in the measurements taken by two observers (inter-observer), and three different readings of the same observer (intra-observer), correlation coefficient being > 0.99. With stretchable tape, 26% subjects had both intra-observer and inter-observer variability of ≥ 0.5 cm, which was 9.9% and 15.9% with the use of non-stretchable tape, respectively. Non-stretchable tape compared to stretchable tape had significantly less intra-observer ($p < 0.001$) and inter-observer variability ($p=0.038$). **Conclusions:** Observer variability was significantly less with the use of non-stretchable tapes compared to routinely used stretchable tapes. Old stretchable tapes used repeatedly over certain periods were not able to provide accurate readings.

Key words: Anthropometry, occipito-frontal circumference, cross-sectional study.

Introduction

Head circumference (HC) or occipito-frontal circumference (OFC) monitoring of children is very crucial, and is a basic tool for early detection of neuro-developmental disorders^{1,2,3}. An abnormal rate of head growth suggest a pathological disorder, and would require further evaluation⁴. Hence, there is need of a valid, precise, accurate, and reliable tool for OFC measurement in children.

Currently used measuring tapes for taking anthropometric

¹Dr. Pareshkumar Thakkar, MBBS. MD, Associate Professor, Department of Paediatrics, Medical College Baroda and SSG Hospital, Vadodara. ²Dr. Kinjal Yagnik, MBBS. DCH, Resident, Department of Paediatrics, Medical College Baroda and SSG Hospital, Vadodara, ³Dr. Niyati T Parmar, MD, Tutor, Department of Community Medicine, Medical College Baroda and SSG Hospital, Vadodara, ⁴Dr. Rashmi Ranjan Das, MD, Assistant Professor, Department of Paediatrics, All India Institute of Medical Sciences, Bhubaneswar, India, ⁵Dr. Ukti P Thakkar, BPT, Physiotherapist, VIP Physiotherapy Clinic, Vadodara, Gujarat.

Address for correspondence

Dr. Pareshkumar A Thakkar
Associate Professor
Department of Paediatrics
Sir Sayajirao General Hospital & Medical College, Vadodara, Gujarat, India.
E-mail: drpareshthakkar123@gmail.com
Phone no: 91-9426032150

Acknowledgements: None

Funding: Nil

Conflict of Interest: None

Permission from IRB: Yes

How to cite

Thakkar P, Yagnik K, Parmar NT, Das RR, Thakkar UP. Observer Variability in Head Circumference Measurement Using Routine Versus Non-Stretchable Tapes in Children. J Nepal Paediatr Soc 2017;37(3):238-243.

doi: <http://dx.doi.org/10.3126/jnps.v37i3.19556>

This work is licensed under a Creative Commons Attribution 3.0 License.



measurements do not give accurate readings after being used for few months because of their stretchability¹. In addition, there are no recent guidelines on when to replace the tape. Besides these, inter-observer (different observer) and intra-observer (same observer) variations are very common because of different forces (stretch) applied while taking OFC measurements^{1,5,6}. Observer variations in anthropometric measurements of > 0.5 cm are common. This might result in missing an at-risk child or might be a false alarm and subject a child to unnecessary work-up and impose significant stress among parents. This emphasizes the need of a tool which gives accurate OFC measurement with no or minimal variability, and is able to provide reliable measurement even after repetitive use.

Though it is known since long that significant observer variability exists in anthropometric measurements, and stretchable tapes are not reliable after being used over a period of time, there are no published studies on how to minimize this variability. Non-stretchable exclusive tapes for OFC measurement have recently become available⁷. The aim of this study was to find whether significant observer variability exists and whether use of non-stretchable tape can minimize the observer variability or not. We also intended to study whether the stretchable and non-stretchable tapes are able to provide accurate readings after being used repeatedly over a period of time.

Material and Methods

This cross-sectional study was carried out over a period of six months (November 2013 to April 2014) at a tertiary care teaching hospital of Western India. All stable children in the age group of one month to five years whose parents were willing to participate and ready to give consent, were included in this study. All the children meeting the inclusion criteria or sample selection criteria were enrolled by the observers (post-graduate students) on a particular day, the limit being up to five children per day (this figure was chosen to prevent observer fatigue on the given day). The study was approved by Scientific Review Committee and Institutional Ethics Committee on Human Research. Participants were enrolled in the study after obtaining consent from the parents.

Outcome variables and their measurement: The outcome variables were: (a) the intra-observer and inter-observer variability in OFC measurement by routinely used stretchable and non-stretchable tapes, and (b) accuracy of readings of old stretchable OFC measurement tapes.

Initially, we assessed the accuracy of various tapes by comparing them against the standard steel

rule (scale). We took two new stretchable, two new non-stretchable and ten old stretchable OFC tapes. The stretchable tape (Butterfly Brand Tape, China) has the following specifications: material (soft fiber glass plus plastic cover), measuring range/length (0 – 150 cm), width (1.5 cm), graduation (measuring rod: 1 mm/0.1 cm), and weight (60 g). The non-stretchable tape (Ibis Medical, India) has the following specifications: material (flexible PVC), measuring range (0 – 55 cm), graduation (measuring rod: 1 mm/0.1 cm), and measure (W x H x D: 25 x 1 x 663 mm). The old tapes selected were the tapes which had been used repeatedly for at least 3 months by the resident doctors in the paediatric wards, clinics and outpatient departments. Measurement of OFC was carried out by the new stretchable and non-stretchable tapes. The non-stretchable exclusive OFC tapes were also periodically (monthly) assessed for their reliability and accuracy during the study period.

Two observers were required for testing of the inter-observer variations. At the start of the study, both the observers (post-graduate students) were trained for taking OFC measurement accurately by the guide (faculty). They were taught about the important landmarks for measurement of OFC. Both the observers used the same stretchable and non-stretchable exclusive tapes for OFC measurements.

The stretchable tapes were discarded after their use for a period of one month, and thereafter a new stretchable tape was used. Non-stretchable exclusive OFC tapes were assessed every month for their reliability and accuracy against the standard steel rule (scale) for OFC measurement. It was decided, if repeated use of non-stretchable exclusive OFC tape was not able to provide accurate reading, then it would be discarded and a new tape would be used. Necessary measures and precautions were taken to avoid possible bias which could have arose while taking OFC measurements.

There was a possibility of bias if observer-2 had access to the readings taken by observer-1. To avoid this, observer-2 was blinded to the readings taken by observer-1. Observer-1 checked the inclusion criteria, enrolled the patient, entered the primary details in Proforma, and took first reading of OFC. Observer-1 provided observer-2 just the list of participants in a separate sheet. Observer-2 did not have access to the readings taken by observer-1. Observer-2 took independent measurement while observer-1 was not present, and entered in a separate sheet. Observer-2 used the same measuring tapes which were used by observer-1. The readings were taken on the same day, but the timings were different.

For studying intra-observer variability, the observer-1 took three readings of OFC of same patients at different times of the same day. This also possibly could lead to bias in two ways. First, if the observer-1 has access to his/ her previous readings, he/she might give closer readings. To decrease/eliminate this bias, observer-1 entered the first reading in the Proforma, and subsequent two readings (entered in separate sheets) were handed over to the guide (faculty) for entry into the Proforma. Second, was the possibility of recall bias. To decrease/eliminate this, observer-1 enrolled nearly ten patients on a given day. The subsequent readings were taken after significant lapse of time of at least 4-6 hrs. Both the observers took all the readings by using both stretchable and non-stretchable exclusive OFC tapes. All the readings were entered into the Proforma by guide (faculty).

Three different readings taken by the same observer were analyzed for intra-observer variability. For analyzing inter-observer variability, the first reading out of three readings taken by observer-1 was compared with one reading taken by observer-2. Then we compared the difference in observer variability with the use of stretchable and non-stretchable exclusive OFC tapes.

Sample size was calculated keeping the allowable error alpha 5% and power of study 80%. The required sample size was $130(N=4pq/L^2)$. We have enrolled total 301 patient which gives power of 96% to our study. Data were entered into Microsoft Excel Sheet (version 10.0). The data was checked for accuracy by two persons involved in the study. MedCalc for windows software (version 12.5.0) was used for statistical analysis. The data was analysed quantitatively using paired t-test, unpaired t-test, analysis of variance (ANOVA), and post-hoc tests. The data was analysed for inter-observer and intra-observer variability. The differences between the measurements, their correlation, and agreement were analysed using the correlation coefficient (Pearson) and intra-class correlation coefficient. Inter-observer and intra-observer measurements were compared by ANOVA, with the F value \geq F critical indicating significant difference^{8,9}.

Results

Three-hundred and one children aged one month to five years were enrolled in this study: 42.5% were in the age group of 1 to 6 months (128/301). The male (55.1%): female (44.9%) sex ratio was 1.23:1.

Observations on accuracy of various measuring tapes

At the beginning of the study, we assessed

accuracy of all available stretchable tapes which were used repeatedly over different period of time. We looked at the difference in the reading by comparing various tapes against standard steel (scale) rule at 35cm and 50 cm level. When we compared accuracy of new stretchable tape and new non-stretchable tape, we found no difference in readings of both the tapes at 35 cm and 50 cm against the steel (scale) rule. We took ten old stretchable tapes which had been used for a period ranging from 3 to 8 months. The tapes which had been used for a period 3-4 months showed a difference of 0.1 to 0.4 cm at the level of 35 cm and 50 cm. Whereas, the tapes which had been used for 5 to 8 months showed a difference of 0.5 to 1 cm (Table 1 and Figure 1).

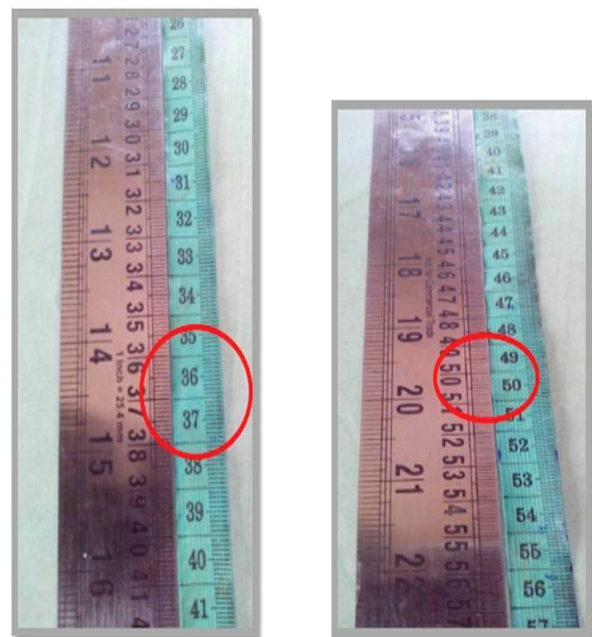


Fig 1: Old stretchable tapes showing difference of 1 cm when compared to standard steel rule at 35 and 50 cm.

It is obvious that use of such old tapes would lead to significant difference in the observations, even under ideal conditions. Such a big difference or inaccuracy in the reading is clinically not acceptable. To avoid this, in the present study, we used new stretchable tapes, and changed them every month. The non-stretchable exclusive OFC tapes were assessed for their accuracy at monthly interval, during the entire study period (Table 2). The non-stretchable exclusive OFC tapes gave consistently accurate readings even after being used repeatedly over a period of six months.

Intra-observer variability

The first observer took head circumference measurements at three different times with both stretchable and non-stretchable exclusive OFC tapes. Observer was masked to own previous readings. ANOVA and post-hoc tests were applied to see whether

any significant difference exist in the measurements obtained by the same observer at different times.

Stretchable tape: The measurements obtained by the same observer using stretchable tape at three different times were consistent without any significant difference ($p = 0.999$, F value = 0.001). The post-hoc analysis compared each one of the three readings with the rest of the two readings, but didn't show significant difference.

Non-stretchable tape: There was no significant intra-observer variability with the use of non-stretchable tape ($p = 0.998$, F value = 0.002). The post-hoc analysis compared each one of the three readings with the rest of the two readings, but didn't show significant difference.

Good correlation was observed for intra-observer variability for both stretchable and non-stretchable tapes (Table 2).

Comparison of stretchable vs. non-stretchable tape: When intra-observer variability was checked for stretchable and non-stretchable tapes, the difference observed was highly significant ($p < 0.001$) (Table 3). This suggests that intra-observer variability is significantly less with the use of non-stretchable exclusive OFC tape. This means, multiple readings obtained by the same observer are more consistent with the use of non-stretchable compared to stretchable tape.

Observed variability in clinical context: The intra-observer variability in the measurement of ≥ 0.5 cm using stretchable tapes was seen in 79 (26%) cases and using non-stretchable tapes was seen in 30 (9.9%) cases, the difference being significant ($p < 0.001$). The difference of

≥ 0.5 cm for any case is not acceptable clinically (Table 4).

Inter-observer variability

OFC measurements were taken by two different observers independently using both stretchable and non-stretchable exclusive OFC measuring tapes. Independent t-test was used to analyse the variability in the measurements obtained by two different observers. No significant difference was observed with the use of stretchable ($p = 0.86$) and non-stretchable ($p = 0.81$) tapes. It is worth mentioning that, the stretchable tapes used were new tapes, and both the observers were trained for measurement of OFC.

Good correlation was observed for inter-observer variability for both stretchable and non-stretchable tapes (Table 2).

Comparison of stretchable vs. non-stretchable tape: The mean difference observed in the measurement of OFC between two observers using stretchable tape was 0.27 cm, and using non-stretchable tape was 0.23 cm. The difference observed was statistically significant ($p = 0.038$) (Table 3). This suggests, with the use of non-stretchable tape, observation remains more consistent between two observers, as compared to stretchable tape.

Observed variability in clinical context: The inter-observer variability in the measurement of ≥ 0.5 cm using stretchable tapes was seen in 79 (26%) cases and using non-stretchable tapes was seen in 48 (15.9%) cases, the difference being significant ($p = 0.003$). The difference of ≥ 0.5 cm for any case is not acceptable clinically (Table 4).

Table 1: Accuracy of used stretchable tapes

Sr. No.	Difference (in cm) when measured against standard steel rule at 35 cm and 50 cm		Approximate duration of use (in months)
	35 cm	50 cm	
1.	0.1	0.1	3
2.	0.2	0.2	3
3.	0.2	0.2	4
4.	0.4	0.5	4
5.	0.5	0.5	5
6.	0.5	0.6	6
7.	0.7	0.7	6
8.	0.8	0.8	6
9.	0.8	0.9	8
10.	1.0	1.0	8

Table 2: Pearson correlation for the observer variability

Variability	Type of tape	Pearson's correlation coefficient r	Limits of agreement
Intra-observer	Stretchable	0.997	0.9973 to 0.9983
	Non stretchable	0.9988	0.9985 to 0.9990
Inter-observer	Stretchable	0.9966	0.9957 to 0.9973
	Non stretchable	0.9973	0.9966 to 0.9979

Table 3: Comparison of observer variability – stretchable versus non-stretchable

	Stretchable tape in cm, mean (\pm SD)	Non-stretchable tape in cm, mean (\pm SD)	p- value
Intra-observer variability	0.33 (0.21)	0.27 (0.26)	<0.001
Inter-observer variability	0.25 (0.18)	0.23 (0.25)	<0.038

Table 4: Proportion of observer variability ≥ 0.5 cm - comparisons

Variability of ≥ 0.5 cm	Stretchable tape	Non-stretchable tape	p-value
Intra-observer Variability	79(26%)	30(9.9%)	<0.001
Inter-observer variability	79(26%)	48(15.9%)	0.003

Discussion

In the present study, we found good correlation in the measurements taken by two observers, and three different measurements of the same observer, correlation co-efficient being > 0.99 . However, with the stretchable tape, 26% subjects had intra-observer and inter-observer variability of ≥ 0.5 cm, which was 9.9% and 15.9% with the use of non-stretchable tape, respectively. Intra-observer variability and inter-observer variability were significantly less with the use of non-stretchable tape compared to stretchable tape.

Similar finding have been observed by a previous study using that compared the stretchability between the old and new tape¹. The authors observed the intra-observer variability to be statistically insignificant, but in 5% measurements it was clinically significant (differed by ≥ 0.65 cm). Similar was their finding with inter-observer variability, which was statistically insignificant, but in 9% measurements it was clinically more significant (differed by ≥ 1 cm)¹.

In another study using blank tape measures for anthropometric measurements (including OFC) in term infants, the intra-examiner mean absolute difference in OFC measurements was lesser than that of inter-examiner measurements (intra-examiner: no difference, and inter-examiner: 0.01 cm difference)⁶. This means repeated measurements by same individual was more accurate than measurements by different individuals. However, we could not find any study comparing observer difference using non-stretchable exclusive OFC tape and any study comparing stretchable versus non-stretchable tapes.

In a similar study done previously, the authors made a statement that; their study was not designed to address the issue of timing of replacement of the tapes¹. In the present study, stretchable tapes which have been used over a period of 3-4 months showed differences of 0.1 to 0.4 cm, and tapes used over 5 to 8 months showed the differences of 0.5 to 1 cm. Non-stretchable tapes gave consistently accurate readings even after being used repeatedly over a period of 6 months. In routine clinical practice, stretchable measuring tapes are widely used, and only one reading is taken. In a previous study, the authors applied normal range of force required in clinical practice (calculated using Newton meters) to the new as well old stretchable tapes. They found that, in the clinical force range, the old tapes stretched significantly compared to new tapes¹.

It is easy to understand that stretchable tapes after repeated use are not reliable as they will not give accurate readings. This element of inaccuracy due to stretchability added to the element of observer variability would make the measurements highly unreliable. It is obvious that use of such old tapes will result in significant difference in the observation even under an ideal condition. Such a large difference or inaccuracy is clinically not acceptable. In the present study, we indirectly found the duration till which stretchable tapes give reliable readings. However, there is no specific guideline regarding after what duration/period, a stretchable tape should be discarded, and a new tape should be used. Based on the findings of the present study, we should either use non-stretchable tapes or, if we are using stretchable tapes, they should be new, periodically assessed for their reliability, and replaced

frequently. In the literature, we did not find studies using non-stretchable exclusive OFC tape.

The strengths of the present study are following: adequate sample size, use of non-stretchable OFC tapes and comparison with stretchable tapes, assessment of both intra- and inter-observer variability in same setting with robust methodology, and data from a tertiary care setting of a developing country (can be generalised to such settings). However, the study is not without limitations. First, there was no true value of measurements which could have served as gold standard for comparing the inter-observer and intra-observer variations. Second, we did not try to quantify the observer variability from the "true value". Third, did not aim to study after what duration a stretchable tape do not give accurate reading or the time when it should be discarded.

Future research should be directed to address the issue after what duration stretchable tapes should be

discarded, and how to minimize the observer variability. Taking average from the multiple readings would give value near to the "true value". Future research should also try to address the question regarding "minimum" number of readings taken to find an average value that is near to the "true value", and the reading is clinically acceptable.

Conclusions

Variability in head circumference (OFC) measurement is significantly less with the use of non-stretchable exclusive tapes compared to routinely used stretchable tapes. The intra-observer and inter-observer variability though statistically insignificant, was significant in the clinical context. Stretchable tapes used repeatedly over certain periods (5 to 8 months) were not able to provide accurate readings, while non-stretchable tapes could provide accurate readings at the end of study period of six months.

References

1. Bartram JL, Rigby AS, Baxter PS. The "Lasso-o" tape : Stretchability and Observer Variability in Head Circumference Measurement. *Arch Dis Child* 2005;90:820-21. DOI: 10.1136/adc.2004.063743.
2. Gokhale R, Kirschner BS. Transition of care between paediatric and adult gastroenterology. Assessment of growth and nutrition. *Best Pract Res Clin Gastroenterol* 2003;17:153-62.
3. Ariza AJ, Greenberg RS, LeBailly SA, Binns HJ; Pediatric Practice Research Group.. Parent perspectives on messages to be delivered after nutritional assessment in pediatric primary care practice. *Ann Fam Med* 2005;3:S37-9. DOI: 10.1370/afm.354.
4. Block RW, Krebs NF; American Academy of Pediatrics Committee on Child Abuse and Neglect.; American Academy of Pediatrics Committee on Nutrition. Failure to thrive as a manifestation of child neglect. *Pediatrics* 2005;116:1234-7. DOI: 10.1542/peds.2005-2032.
5. Bhushan V, Paneth N. The reliability of neonatal head circumference measurement. *J Clin Epidemiol* 1991;44:1027-35.
6. Johnson TS, Engstrom JL, Gelhar DK. Intra- and interexaminer reliability of anthropometric measurements of term infants. *J Pediatr Gastroenterol Nutr* 1997;24:497-505.
7. Neonatal Head circumference tape/Occipitofrontal Circumference measuring Tape. Available from: www.info@ibismedical.net (accessed 26th May 2017).
8. Bland JM, Altman DG. Agreed statistics: measurement method comparison. *Anesthesiology* 2012; 116: 182-5. DOI:10.1097/ALN.0b013e31823d7784
9. Moczko JA. Methods for assessing reliability and agreement of measurement. *Przegl Lek* 2013;70:875-79.