



King's Research Portal

DOI:

[10.1111/jsm.13005](https://doi.org/10.1111/jsm.13005)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Habous, M., Muir, G., Tealab, A., Williamson, B., Elkhoully, M., Elhadek, W., Mahmoud, S., Laban, O., Binsaleh, S., Abdelwahab, O., Mulhall, J. P., & Veale, D. (2015). Analysis of the Interobserver Variability in Penile Length Assessment. *JOURNAL OF SEXUAL MEDICINE*, 12(10), 2031–2035. <https://doi.org/10.1111/jsm.13005>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Habous, M, Tealab, A, Williamson, B, Bin Saleh, S, El Dawy, S, Mahmoud, S, Abdelwahab, O, Nassar, M, Mulhall, JP, Veale, D & Muir, G.

Analysis of the inter-observer variability in penile length assessment.

Journal of Sexual Medicine, 12(10): 2031-5. [http://dx.doi.org/ 10.1111/jsm.13005](http://dx.doi.org/10.1111/jsm.13005)

Abstract

Introduction: Many men seek penile augmentation treatments – a standard tool for their counselling is to inform them of what is “normal.” Although some studies suggest good correlation between stretched and erect penile length, those that have measured stretched and erect length *simultaneously* have shown significant variability. Aim: To assess the accuracy of differing penile measurements with multiple observers. Methods: We recruited 201 adult men (mean age 49.6 years) who achieved full erection using intra-cavernosal injection. Main Outcome Measures: Penile measurements were taken by one of seven andrology specialists in a private, temperature controlled (21°C, 72°F) environment. Stretched flaccid and erect length and circumference were measured. We analysed the accuracy of each flaccid measurement using the erect measurements as a reference, for the overall patient population and for each observer. Results: The mean underestimate of length from stretched flaccid to erect was 2.64 cm (21.4%) and girth 2.27cm (19.5%). Inter-observer variability ranged from a mean underestimate of 16-27% (length) and 15-27% (girth). Conclusions: In this large, multicentre, multi-observer study of penis size, flaccid measurements under-estimated erect size. It also seems likely that there is significant inter-observer variability. We believe erect penis measurements should be used for men being considered for treatment of small penis anxiety.

Keywords: Methodology, penile size measurement, inter-observer errors, underestimate

Introduction

Urologists and psychotherapists often encounter patients who complain about the length of their penis, but these patients are usually well within the typical range for penis size¹. Most men seeking advice on penile size, or requesting penile augmentation, have normal penile size². Assessing penis length is critical when reassuring men about the “norm,” and also when counselling such men. Men concerned about their penile size may fall into two groups³⁻⁶ – those with small penis anxiety (SPA), and those with more severe symptoms of body dysmorphic disorder (BDD) who may benefit from formal psychiatric assessment and treatment (including cognitive behaviour therapy and selective serotonergic reuptake inhibitor medication)^{7,8}. Both these groups, and in particular those men with BDD have particular vulnerabilities possibly making them prone to exploitation with unethical and unproven treatments.

Cosmetic phalloplasty is regarded, at best, as experimental for men with small penis anxiety without any adequate outcome measures or evidence of safety⁹. There are no simple or risk free strategies for penile augmentation, and great potential for harm¹⁰. In general, penile enhancement surgery can cause, at the very most, an apparent 1-2-cm increase in flaccid penile length and up to a 2.5-cm augmentation of penile girth. Unwanted outcomes and complications, namely penile deformity, paradoxical penile shortening, disagreeable scarring, granuloma formation, migration of injected material, and sexual dysfunction were reported frequently in these studies^{11,12}. Being able to help these men contextualise their concerns is vital in their management. Knowing what is “normal” and being able to help a man place himself on the scale of size is of key importance¹³.

The first description of standard penile length for age was used by Schonfeld and Beebe¹⁴. The majority of studies examining penile anthropometric data have relied on measurements in the flaccid state. Veale et al¹³ recently conducted a systematic review and

constructed nomograms for flaccid and erect penis length and circumference. In this meta-analysis, the stretched flaccid length appeared to be an excellent estimate of erect penile length (a ratio of 0.98). However, only three studies¹⁵⁻¹⁷ actually measured the stretched and erect penile length simultaneously¹⁵⁻¹⁷. Of these only the original study of Wessells, Lue¹⁷ showed a good correlation between erect length and stretched length. Both Promodu, Shanmughadas¹⁵ and Sengezer, Öztürk¹⁶ significantly under-estimated erect length by 15-29% respectively.

Some attempts have been made to standardise this aspect of measuring the flaccid penis, including development of an engineering model to approximate the optimal tensile force to be applied¹⁸. While it has been accepted that self-measurement may be inaccurate both due to variable technique and also potential exaggeration, no studies have addressed the issue of third party inter-observer variability.

We were concerned by the potential inaccuracy both in measuring penis length between flaccid stretched and erect, and also the potential for observer bias in measuring penile size. Our hypothesis was that there may be significant error in predicting erect size by the stretched or flaccid measurements and that there may be significant measurement variation.

Methods

Study Population: Men presenting to a group of andrology clinics between January 2011 and June 2012 were evaluated by history and physical examination including penile length measurement. All men were having intra-cavernosal injection (ICI) to induce erection as part of their routine care. ICI is done for most (but not all) of our Erectile Dysfunction (ED) patients as part of optional investigations for ED. Organic causes for ED are common and most of our patients request to ICI, since most of them have failed oral medication prior to presentation.

Consecutive patients who underwent ICI and met the inclusion criteria were assessed. Excluded from analysis were men with previous penile surgery (excluding circumcision), Peyronie's disease, congenital curvature or hypospadias, and patients who did not achieve a fully rigid erection with diagnostic ICI. Comorbidities and patient demographics were recorded.

Consent was collected from all participants and the study was approved by institutional ethics committee.

Main Outcome Measures:

Penile Length Measurement: Each patient was assessed by one of seven experienced andrology specialists in a controlled environment: private, air-conditioned consulting rooms at a constant temperature (21°C, 70°F). The temperature in the examination rooms (and all the medical centre) is automatically controlled with central air condition (AC). Data was collected and recorded in centimetres. Each patient had two parameters measured in the flaccid state and the same two parameters in the erect state: (i) pubic bone to tip of glans (bone-to-tip: BTT) and (ii) circumference (girth) at the mid-penile shaft. Alprostadil 10mcg was the injection agent utilized. For stretched flaccid length, the penis was extended to maximum capacity at a 90-degree angle to the body with the patient in the upright standing position. For erect length measurement, a fully rigid erection was required for inclusion. Length was measured with a rigid plastic ruler, while girth was assessed using a disposable paper tape at the base of the penis.

Statistical Analysis: Data were collated and subsequently analysed using SPSS software package (IBM, SPSS Statistics 20). For each patient, the measurements in the erect state were considered the "true" values and those in the stretched flaccid state were then compared to the corresponding 'true' value, with absolute and percentage error recorded for each measurer.

Results

Study Population: 201 patients with a mean age of 49.6 (range 20-75; SD 13.0) years were included. BMI - mean 28.77, range 16-51, SD 5.202. The comorbidity breakdown was: hypertension 12%, dyslipidemia 24%, diabetes 13%, obesity 24%, coronary artery disease 7%, and cigarette smoking 12%.

Examiners: The mean age and duration in clinical practice of the andrologists was 41 and 14 years respectively (range 34-53; 8-26) years. Six of seven practitioners were formally fellowship trained and all were fellows of the European Society of Sexual Medicine (ESSM). The mean number of patients examined by each andrologist was 29 (range 12-59; SD 14).

Penile Length Data: The mean dimensions are listed in Table 1. The mean flaccid dimensions were BTT 11.7 ± 1.6 (8-17) cms and circumference 9.4 ± 1.5 (5-14) cms. The mean erect dimensions were: BTT 14.3 ± 2.0 (8-19) cms and circumference 11.6 ± 1.5 (8-16) cms. The mean inter-observer discrepancy between stretched flaccid and erect BTT length was 2.9 ± 0.6 (range 2.0-3.5). The mean inter-observer discrepancy between stretched and erect circumference was 2.5 ± 0.5 (range 1.7-3.1).

Comorbidity: The numbers with hypertension were 12% (24 patients); dyslipidaemia 24% (48 patients); diabetes 13% (26 patients); obesity 24% (48 patients); coronary artery disease 7% (14 patients); and cigarette smoking 12% (24 patients).

Discussion

Our study suggests that not only is stretched length inaccurate in predicting erect length, but that even experienced andrologists are likely to produce significantly different measurements. Put in context, a 20% under-estimate of length can easily introduce an error of 3cm in assessing patient's true penile length – much more than the even the most enthusiastic proponents of penile augmentation would ever claim (9). The inter-observer variation was most marked when comparing skin-to-tip length. We hypothesise that this measurement

generated the widest variation due to two principal factors, the first being the problem of how much axial traction the andrologist should apply. Whilst it is possible to standardise the force, using a technique proposed by Chen, Gefen¹⁸, this complicated methodology was neglected in our study in the interests of patient comfort. This concern for patient comfort is no doubt one of the main problems associated with measuring the stretched flaccid penis.

The second problem we associate with the wide inter-observer variation in measuring the skin-to-tip length is that of the somewhat variable proximal end point for measuring at the supra-pubic skin. Different andrologists are likely to employ slightly different techniques, such as whether the patient should remain in full expiration for the duration of the measurement, or indeed the posture the patient should adopt. It can also be tricky to fully stretch the penis while simultaneously compressing the infrapubic fat pad with the measuring device.

Nonetheless, a significant degree of inter-observer variation remained when comparing bone-to-tip stretched length measurements, despite the fact that this parameter in theory negates perhaps the most marked source of error inherent in the skin-to-tip measurement noted above (i.e. that there is a definitive proximal end point in the bony pubis, as opposed to the variable end point of the supra-pubic skin). We have previously shown that bone to tip measurement seems more accurate than skin to tip in the erect penis, due to correction for BMI and thus the amount of infra-pubic fat (data on file, manuscript submitted).

Our data supports the Promodu, Shanmughadas¹⁵ and Sengezer, Öztürk¹⁶ findings, with a larger sample of both participants and andrologists. We have demonstrated the potential for misdiagnosis and misinformation of patients introduced by observer error.

Limitations

Our study could be improved by using the same group of patients and having each participant measured by two or more examiners in order to calculate the intra-class correlation coefficient. Many men are however reluctant to undergo repeated physical examination of their genitalia and dislike intra-cavernosal injection induction of erection, apart from the ethics of repeated injection. A training programme might improve inter-observer reliability for the measurement of stretched length, although our preference is for the use of erect measurements.

We may be criticised for assuming that measurement of the erect penis is the reference value. Ideally a number of specialists would have measured the same erect penises on a number of occasions to determine inter-rater reliability when measuring the erect penis, but due to geographical and ethical issues this was not possible, and probably never will be. It does however seem logical to assume that a fully erect penis, not requiring stretching or support, should be the easiest to measure. Men may be concerned with both erect and flaccid size, although a minority focus on flaccid size alone⁴, as do nearly all physical intervention studies so far reported.

The difference in penile girth between the stretched flaccid state and the erect state is likely to be due to the morphological changes that occur with engorgement, rather than operator error. Our comparison of girth in this study was more for completeness and should be considered as a separate piece of information generated, rather than a measure to reflect on inter-observer variation. Our data in this respect confirm that men can be reassured that flaccid girth does not correlate with erect girth. Previous studies have differed when assessing the correlation between erect and stretched penis length. This study could suggest that inter-observer variability may be a factor in such discrepancies.

All our men were given intra-cavernosal injections as part of routine andrology care. This is also our practice in the Middle East for men concerned by penile size when

counselling them. In other cultures video sex stimulation may well be a less invasive substitute, but there remain countries where the use of erotic images, even for medical reasons, is unacceptable. The study group was composed largely of men of Middle Eastern origin, but there is no reason to expect a different result with other ethnic groups even if the mean size measurements were larger or small.

Lastly measurement of penile size is of course only one aspect of assessment of men with small penis anxiety. Equally as important is male genital image satisfaction^{19, 20}; an understanding of the beliefs and attitudes about penile size³; the frequency of avoidance and safety seeking behaviours to prevent the risk of shame or humiliation and the need to screen for body dysmorphic disorder²¹.

Conclusions

This large multicentre, multi-observer study of penis size, demonstrates that flaccid penile measurements are both unreliable and observer dependent. For the first time we have shown a further disadvantage of flaccid state measurement - marked and clinically significant inter-observer variation. Thus a “normal” penis to one doctor may be assessed as “small” by another specialist, resulting in potential physical and psychological harm. We therefore recommend that when counselling men about penile augmentation and reassuring men with regards to penis size, erect measurements should be made, recorded, and compared with population based nomograms.

Acknowledgements

DV acknowledges salary support from the National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and the Institute of Psychiatry, Psychology and Neuroscience, King's College London.

References

- [1] Lever J, Frederick DA, Peplau LA. Does size matter? Men's and women's views on penis size across the lifespan. *Psychology of Men and Masculinity*. 2006;**7**: 129-43.
- [2] Ghanem H, Shamloul R, Khodeir F, El Shafie H, Kaddah A, Ismail I. Structured management and counseling for patients with a complaint of a small penis. *Journal of Sexual Medicine*. 2007;**4**: 1322-27.
- [3] Veale D, Eshkevari E, Read J, et al. Beliefs about penis size: validation of a scale for men with shame about the size of their penis. *Journal of Sexual Medicine*. 2014;**11**: 84-92.
- [4] Veale D, Miles S, Read J, et al. Phenomenology of men with body dysmorphic disorder concerning penis size compared to men anxious about their penis size and to controls: a cohort study. *Body Image*. 2015;**13**: 53-61.
- [5] Veale D, Miles S, Read J, et al. Sexual functioning and behaviour of men with body dysmorphic disorder concerning penis size and men anxious about penis size with controls: a cohort study. *Sexual Medicine*. 2015.
- [6] Veale D, Miles S, Read J, et al. Environmental and Physical Risk Factors for men to develop Body Dysmorphic Disorder concerning penis size compared to men anxious about their penis size and men with no concerns: a cohort study. *Journal of Obsessive-Compulsive and Related Disorders*. 2015.

- [7] Veale D, Anson M, Miles S, Pieta M, Costa A, Ellison N. Efficacy of cognitive behaviour therapy versus anxiety management for body dysmorphic disorder: a randomised controlled trial. *Psychotherapy and Psychosomatics*. 2014;**83**: 341-53.
- [8] Phillips KA, Albertini RS, Rasmussen SA. A randomized placebo-controlled trial of fluoxetine in body dysmorphic disorder. *Archives of General Psychiatry*. 2002;**59**: 381-88.
- [9] Ghanem H, Glina S, Assalian P, Buvat J. Position paper: management of men complaining of a small penis despite an actually normal size. *Journal of Sexual Medicine*. 2013;**10**: 294-303.
- [10] Vardi Y, Harshai Y, Gil T, Gruenwald I. A Critical Analysis of Penile Enhancement Procedures for Patients with Normal Penile Size: Surgical Techniques, Success, and Complications. *Eur Urol*. 2008;**54**: 1042-50.
- [11] Chevallier D, Haertig A, Faix A, Droupy S. Cosmetic surgery of the male genitalia. *Progres en urologie : journal de l'Association francaise d'urologie et de la Societe francaise d'urologie*. 2013;**23**: 685-95.
- [12] Wessells H, Lue TF, McAninch JW. Complications of Penile Lengthening and Augmentation Seen at 1 Referral Center. *The Journal of Urology*. 1996;**155**: 1617-20.
- [13] Veale D, Miles S, Bramley S, Hodson J. Am I normal? A systematic review and construction of nomograms for flaccid and erect penis length and circumference in up to 15,521 men. *BJU International* 2015;**115**.
- [14] Schonfeld WA, Beebe GW. Normal growth and variation in the male genitalia from birth to maturity. *The Journal of Urology*. 1942;**48**: 759-79.
- [15] Promodu K, Shanmughadas KV, Bhat S, Nair KR. Penile length and circumference: An Indian study. *Int J Impot Res*. 2007;**19**: 558-63.
- [16] Sengezer M, Öztürk S, Deveci M. Accurate method for determining functional penile length in Turkish young men. *Annals of Plastic Surgery*. 2002;**48**: 381-85.

- [17] Wessells H, Lue TF, McAninch JW. Penile length in the flaccid and erect states: Guidelines for penile augmentation. *J Urol*. 1996;**156**: 995-97.
- [18] Chen J, Gefen A, Greenstein A, Matzkin H, Elad D. Predicting penile size during erection. *Int J Impot Res*. 2000;**12**: 328-33.
- [19] Davis SN, Paterson LQ, Binik YM. Male genital image: Measurement and implications for medical conditions and surgical practice. *Sexologies*. 2012;**21**: 43-47.
- [20] Herbenick D, Schick V, Reece M, Sanders SA, Fortenberry JD. The Development and Validation of the Male Genital Self-Image Scale: Results from a Nationally Representative Probability Sample of Men in the United States. *The Journal of Sexual Medicine*. 2013;**10**: 1516-25.
- [21] Veale D, Miles S, Read J, et al. Penile dysmorphic disorder: development of a screening scale. *Archives of Sexual Behavior*. 2015;**44**: 1573-2800.

Table 1:

	Age	BMI	Flaccid Length	Flaccid Circumference	Erect Length	Erect Circumference	Flaccid-Erect Length Discrepancy	Flaccid-Erect Circumference Discrepancy
Mean	49.6	28.8	11.7	9.4	14.3	11.6	2.9	2.5
SD	13.0	5.2	1.6	1.5	2.0	1.5	0.6	0.5
Range	20-75	16-51	8-17	5-14	8-19	8-16	2.0-3.5	1.7-3.1

Table 1 *Studies that conducted simultaneous measures of stretched and erect length included in the nomogram*

Study	Population	Country	N	Age <i>Mean ± SD</i> (Years)	Flaccid length <i>Mean ± SD</i> (cm)	Stretched length <i>Mean ± SD</i> (cm)	Erect length <i>Mean ± SD</i> (cm)
Promodu, 2007	Sexual dysfunction clinic	India (Kerala)	301	31.58 ± 6.38 (Range = 18-60)	8.21 ± 1.44 (Range = 4.5-13)	10.88 ± 1.42 (Range = 6.5-16)	12.93 ± 1.63 (Range = 10.5-17)
Sengezer, 2002	Volunteers	Turkey	200	21.2 (Range = 20-22)	6.80 ± 0.08 (Range= 4-9)	8.98 ± 0.09 (Range = 6.5-12.5)	12.73 ± 0.11 (Range= 9.5-17)
Wessells, 1996	Urology patients	USA	80	54 ± 14.37 (Range = 21-82)	8.85 ± 2.38 (Range = 5-15.5)	12.45 ± 2.71 (Range = 7.5-19)	12.89 ± 2.91 (Range = 7.5-19)
Totals for all 17 studies (Veale et al, in press)	-	-	1552	Range = 17-91	9.16 ± 1.57	13.24 ± 1.89	13.12 ± 1.66