Outcomes of Variation in Technique And variation in accuracy of measurement in Penile Length Measurement

Mohamad Habous¹ MD, Gordon Muir² FRCS(urol), Tarek Soliman³ Mohammed Farag⁴ MD, Ben Williamson⁵ MBBS BSc., Saleh Binsaleh⁶ MD, Waleed Elhadek⁷ FECSM, Saad Mahmoud⁸ MD, Hussein Ibrahim⁹ FECSM, Osama Abdelwahab¹⁰ MD, Ziad Abdelrahman¹¹ FECSM, Mohamed Abdelkader¹² FECSM, Lawrence C Jenkins⁷ MD, and John P. Mulhall⁷ MD

1-Elaj medical centres, Urology and Andrology department, KSA
2- King’s college, Urology department, London, UK
3- Benha University, Urology department Benha, Egypt
4- Al-Azhar Faculty of Medicine, Urology Department, Egypt
5- University Hospitals Birmingham, Birmingham, UK
6- division of Urology, department of Surgery, Faculty of Medicine, King Saud University
   Riyadh, Saudi Arabia
7- Sexual and reproductive medicine program, Memorial Sloan Kettering cancer center, NY, USA

Corresponding Author: Dr Mohamad Habous; MD, FEBU, FECSM
consultant urologist ; Elaj medical group

Contact:
email: drmhabos@hotmail.com
mobile: 00966599950131
Fax: 00966126928192
Postal address: 2527 Qurysh street, Albawady district
postal code: 21391, Jeddah
mail pox: 117364
Saudi Arabia; Jeddah

Conflict of interest: No disclosure
Abstract:
Introduction: Accurate data regarding the size of the erect penis is of great importance to several disciplines working with male patients, but little data exists on the best technique to measure penile length. While some previous small studies have suggested good correlation between stretched penile length, others have shown significant variability. Penile girth has been less well studied, and little data exists on the possible errors induced by differing observers and different techniques. Much of the published data report penile length measured from the penopubic skin junction to glans tip (STT) rather than pubic bone to tip (BTT). We wished to assess the accuracy of different techniques of penile measurements with multiple observers. Methods: men who achieved full erection using dynamic penile Doppler ultrasound for the diagnosis of sexual dysfunction or a desire for objective penile measurement. Exclusion criteria were penile scarring, curvature or congenital abnormality. In each case, the penis was measured by one of seven andrology specialists in a private air conditioned (21°C) environment. Each patient had three parameters measured: circumference (girth) of the penile shaft; length from suprapubic skin to distal glans (skin-to-tip); and pubis to distal glans (bone-to-tip). The three measurements were recorded in the stretched flaccid state, and the same three measurements were then repeated in the fully erect state, following induction of full erection with intracavernosal injection. We analyzed the accuracy of each flaccid measurement using the erect measurements as a reference, for the overall patient population and for each observer. Results: 201 adult men (mean age 49.4 years) were included in this study. Assessing the penis in the stretched and flaccid state gave a mean underestimate of the erect measurement of approximately 20% (skin-to-tip length 23.39%; bone-to-tip length 19.86%; and circumference 21.38%). Conclusion: In this large, multicenter, multi-observer study of penis size, flaccid measurements were only moderately accurate in predicting erect size. They were also significantly observer dependent. Measuring penile length from pubic bone to tip of glans is more accurate and reliable, the discrepancy being most notable in overweight patients.
Introduction

Concern and insecurity over penis size is ubiquitous among men in numerous cultures across the globe, though many are unaware of what actually represents ‘normal’ size. Indeed, despite a reasonable body of research being carried out over the previous 50 years, the majority of studies have paid little attention to the methodological challenges in penile length assessment [1-3].

The majority of studies examining penile anthropometric data have relied on measurements in the flaccid state. The standard approach involves stretching of the flaccid penis and recording its length from suprapubic skin to the distal glans. This technique is inherently flawed by its use of stretching, as different investigators will vary the degree to which they stretch the penis. Some attempts have been made to standardize this aspect of measuring the flaccid penis, including development of an engineering model to approximate the optimal tensile force to be applied[4].

While previous studies have developed nomograms in a variety of areas, all published studies have some problems in methodology, as shown in the table 1 below.

Most studies have used the length measured from the pubopenile skin junction to the tip of the glans (skin to tip, STT), while some have used the pubic bone to tip of glans (bone to tip, BTT) [26]. To date, there has not been a study which compared both STT and BTT measurements in the same cohort of patients.

We were concerned by the possible inaccuracy both in measuring penis length between flaccid and erect, and also the potential for observer bias in measuring and advising men.
Patients and Methods:

201 adult male patients who presented to a network of urology clinics across the Middle East were analyzed in this cross-sectional observational study. All men either had erectile dysfunction or wished either to be advised on the size of the penis and whether they were “normal” or not. Some also actively expressed a desire to seek penile augmentation. 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 center) is automatically controlled with central air condition. Data was collected and recorded in centimeters to the nearest five millimeters. Each patient had three parameters measured: circumference (girth) of the penile shaft; length from suprapubic skin to distal glans (skin-to-tip); and pubis to distal glans (bone-to-tip). The three measurements were recorded in the stretched flaccid state, and the same three measurements were then repeated in the fully erect state, following induction of full erection with intracavernosal injection (ICI). ICI is done for most of our Erectile Dysfunction (ED) patients as part of optional investigations for ED. Consecutive patients who underwent ICI and met the inclusion criteria were assessed.

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.

Exclusion criteria were penile scarring, clinical evidence of Peyronie’s disease, previous surgery (excluding circumcision – all patients were circumcised) and congenital curvature or hypospadias. Men who did not get a full erection were not included.

Data were collated and subsequently analyzed using the SPSS software package (IBM, SPSS Statistics 20). For each patient, the measurements in the erect state were considered the “real” values and those in the stretched flaccid state were then compared to the corresponding ‘real’ value, with absolute and percentage error recorded for each. All measurements were then placed into subgroups according to the examining andrologist. Variability in measurements was then assessed.

No extra tests were carried out beyond routine clinical care. Consent was collected from all participants and the study was approved by institutional ethics committee.
Results

The mean age of patients was 49.6 years (range 20-75; SD 12.9).

Overall results are shown in the table 2

A significant discrepancy was noted between erect measurements and flaccid state measurements, as seen in table 3

Assessing the penis in the stretched and flaccid state gave a mean underestimate of the erect measurement of approximately 20% (skin-to-tip length 23.39%; bone-to-tip length 19.86%; and circumference 21.38%).

All penile measurements correlated: STT and BTT lengths (r=0.910, p<0.01). Correlation existed between STT length and age (r=-0.176, p<0.01), BTT length and age (r=-0.100, p<0.01) and STT length and BMI (r=-0.283, p<0.01). The other notable correlation was that between age and BMI – the older the patient, the higher the mean BMI (r=0.102, p<0.01). After controlling for age, correlation remained between STT length and BMI (r=-0.270, p<0.01); this means effectively the penis “looks” shorter in overweight patients. After controlling for BMI, correlation remained between age and STT length (r=-0.155, p<0.01), as well as between age and BTT length (r=-0.094, p<0.01).

To assess variation in accuracy of measurement we analysed variability in stretched and erect state measurements as shown in table 4

Notable variation in the difference between stretched and erect state measurements is immediately apparent, as is the marked interobserver variation. Assuming the erect size to be “accurate,” skin-to-tip length measurements were inaccurate by between 16.44% and 26.82%. Bone-to-tip length measurements were inaccurate by between 14.23% and 23.23%.
Discussion:

Penile length is a concern for some men [27]. Penis size is a symbol of masculinity, and the perception of having a large penis has been linked to higher self-esteem and sexual identity [1, 28, 29]. From a psychological perspective, however, the crucial issue may not be actual penis size but rather a man’s perception of the size of his penis relative to other men’s. However, penile length measurement is fraught with methodological challenges: room temperature, stretched flaccid versus erect length, pubic bone versus penopubic skin junction, proximal point to glans tip or coronal sulcus and also number of observers doing the measuring. We have shown here that the pubic bone to tip of glans measurement (BTT) is more accurate than the use of the penopubic skin junction as the proximal measurement point (STT). We have also shown that BMI is a major factor in limiting the accuracy of the STT measurement; furthermore, our data support the concept that aging reduces the length of the erect penis.

Variation in accuracy of measurement was most marked when comparing skin-to-tip length. We hypothesized 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 standardize the force, using a technique proposed by Chen et al [24], 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 interobserver variation in measuring the skin-to-tip length is that of the somewhat variable proximal end point for measuring at the suprapubic 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.

Nonetheless, a significant degree of variation in accuracy of measurement remained when comparing bone-to-tip length measurements, even though 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 somewhat labile end point of the suprapubic skin).

Although there is no standard technique for measuring penile length, numerous researchers measure penile length along the dorsum of the penis beginning from the penopubic skin junction to the tip of the glans (skin-to-tip, STT) [26]. This measurement has been used to measure the flaccid, stretched flaccid and erect states. Other authorities advocate for the use of pubic bone to tip of glans (bone-to-tip, BTT). A meta-analysis of 17 studies required the use of BTT measurement for inclusion, all used SFL and 3 used both SFL and erect length [3].

Wessells et al measured flaccid and erect penile dimensions in 80 physically normal men before and after pharmacological erection. They found that the mean flaccid length was 8.8 cm., stretched length 12.4 cm. and erect length 12.9 cm; there was no significant difference between both stretched and erect measurements: stretched length most closely correlated with erect length. [2]
Promodu and his colleagues had different results when they studied penile dimensions in 93 healthy men: the mean flaccid length was found to be 8.21 cm, mean stretched length 10.88 cm and circumference 9.14 cm. Mean erect length was found to be 13.01 cm and erect circumference 11.46 cm. These results, like ours, show significant variation in penile dimensions between both stretched and erect state. (14)

Our data supports the Promodu study findings, with larger numbers, and demonstrates the potential for misdiagnosis introduced by observer error. What is interesting is that few “penile augmentation” techniques claim to give lengthening of more than a centimeter or so – such measurements are well within the realm of observer variability for flaccid measurements.

The fact that we have demonstrated that BTT is longer than STT measurements and that STT correlated with BMI is intuitive. While the differences may be intuitive this paper establishes the magnitude of the difference. The greater the BMI, likely the thicker the pre-pubic fat pad will be and thus the shorter the STT penile length measurement. However, we believe that we are the first to demonstrate a clear correlation between patient age and erectile BTT penile length. It thus appears that erect length decreases as men age.

Our study could be improved by using the same group of patients and having each patient assessed by each andrologist in turn, thus negating any artefact from using different patients. Unfortunately most men are reluctant to undergo physical examination of their genitalia and dislike intracavernosal injection induction of erection. Therefore, practical constraints and respect for patient wishes preclude this methodology.

This is a sample of men all of whom are circumcised, and from a culture where objective confirmation of a “normal” penis is often requested prior to marriage (which often happens late in life.) It is possible that different results would be obtained in a more heterogeneous or circumcised cohort although there is no obvious reason to expect this.

Also, we may be criticized 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-observer variability when measuring the erect penis, but due to geographic, logistical and ethical issues this was again not possible. It does however seem logical to assume that a fully erect penis, not requiring stretching or support, should be the easiest to measure. It is also the case that most men are concerned with their erect penile dimensions more than flaccid, although a minority does focus on flaccid size, as do nearly all intervention studies so far reported.

The difference in penile circumference between the stretched flaccid state and the erect state may predominantly be due to the actual morphological changes that occur with engorgement, rather than operator error. Our comparison of circumference 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 variability.
All our men were given intracavernosal injections as part of ED investigations or for counselling about penile size. This is our standard practice in the Middle East for men concerned by penile size or ED unresponsive to oral medication. 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.
Conclusion

This large multicenter, multi-observer study of penile length measurement, demonstrates that flaccid measurements are unreliable and erect measurements are superior. Based on our data, measuring penile length from pubic bone to glans tip is the method of choice. Furthermore, variation in accuracy of measurement (inter-observer variability) is such that a single evaluator is encouraged in such studies.
References


14. Promodu K, Shanmughadas KV, Bhat S, Nair KR.: Penile length and circumference: an Indian study; Int


16. Sengezer, M., S. Ozturk, and M. Deveci, Accurate method for determining functional penile length in

17. Smith AM, Jolley D, Hocking J, Benton K, Gerofi J.: Does penis size influence condom slippage and

18. da Ros, C., Teloken, C., Sogari, P., Barcelos, M., Silva, F., Souto, C.; Caucasian penis: what is the normal


men cause problems in condom use? a prospective measurement of penile dimensions in 111 young and


length is normal in most men seeking penile lengthening procedures. International Journal of Impotence
Research2002; 14: 283–286


Legends:

Table 1: previous studies published in penile size methodology and their errors

Table 2: Overall results

Table 3: stretched and flaccid vs erect measurements for the whole patient cohort

Table 4: interobserver variability in assessing stretched versus erect size
<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Number of Participants</th>
<th>Measurement State</th>
<th>Age Range</th>
<th>Marked Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mehraban [8]</td>
<td>2007</td>
<td>1500</td>
<td>Flaccid</td>
<td>20-40</td>
<td>Narrow age range</td>
</tr>
<tr>
<td>Kamel [10]</td>
<td>2009</td>
<td>1047</td>
<td>Flaccid</td>
<td>-</td>
<td>Flaccid only</td>
</tr>
<tr>
<td>Khan [12]</td>
<td>2011</td>
<td>609</td>
<td>Flaccid</td>
<td>16-90</td>
<td>Flaccid only</td>
</tr>
<tr>
<td>Ajmani [13]</td>
<td>1985</td>
<td>320</td>
<td>Flaccid</td>
<td>17-23</td>
<td>Young population</td>
</tr>
<tr>
<td>Promodu [14]</td>
<td>2007</td>
<td>301</td>
<td>Some Erect</td>
<td>18-60</td>
<td>Mixed results</td>
</tr>
<tr>
<td>Sengezer [16]</td>
<td>2002</td>
<td>200</td>
<td>Erect</td>
<td>20-22</td>
<td>Young population</td>
</tr>
<tr>
<td>Smith [17]</td>
<td>1998</td>
<td>184</td>
<td>Erect</td>
<td>-</td>
<td>Self-reported</td>
</tr>
<tr>
<td>Da Ros [18]</td>
<td>1994</td>
<td>150</td>
<td>Erect</td>
<td>-</td>
<td>Unpublished - abstract only</td>
</tr>
<tr>
<td>Choi [19]</td>
<td>2011</td>
<td>144</td>
<td>Flaccid</td>
<td>21-89</td>
<td>Flaccid only</td>
</tr>
<tr>
<td>Schneider [21]</td>
<td>2001</td>
<td>111</td>
<td>Erect</td>
<td>18-19</td>
<td>Young population</td>
</tr>
<tr>
<td>Shah [22]</td>
<td>2002</td>
<td>104</td>
<td>Flaccid</td>
<td>17-84</td>
<td>Flaccid only</td>
</tr>
<tr>
<td>Wessells [2]</td>
<td>1996</td>
<td>80</td>
<td>Erect</td>
<td>21-82</td>
<td>All patients have “sexual dysfunction”</td>
</tr>
<tr>
<td>Mondaini [23]</td>
<td>2002</td>
<td>67</td>
<td>Flaccid</td>
<td>16-55</td>
<td>Flaccid only</td>
</tr>
<tr>
<td>Chen [24]</td>
<td>2000</td>
<td>55</td>
<td>Erect</td>
<td>21-78</td>
<td>Invalidated stretching-measuring device</td>
</tr>
<tr>
<td></td>
<td>Rect Skin-Tip</td>
<td>Rect Bone-Tip</td>
<td>Rect Girth</td>
<td>Stretched Skin-Tip</td>
<td>Stretched Bone-Tip</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------</td>
<td>---------------</td>
<td>------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>12.36</td>
<td>14.30</td>
<td>11.61</td>
<td>9.72</td>
<td>11.69</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>1.990</td>
<td>2.023</td>
<td>1.465</td>
<td>1.639</td>
<td>1.648</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>5-17</td>
<td>8-19</td>
<td>8-16</td>
<td>6-15</td>
<td>8-17</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>BMI</td>
<td>Rect Skin-Tip</td>
<td>Rect Bone-Tip</td>
<td>Rect Girth</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>---------------</td>
<td>---------------</td>
<td>------------</td>
</tr>
<tr>
<td>Mean</td>
<td>9.64</td>
<td>8.77</td>
<td>2.51</td>
<td>4.43</td>
<td>1.74</td>
</tr>
<tr>
<td>Range</td>
<td>0-75</td>
<td>6-51</td>
<td>7-17</td>
<td>7-19</td>
<td>7-16</td>
</tr>
<tr>
<td>Andrologist Identifier</td>
<td>Number of Patients</td>
<td>Skin-Tip Difference as %</td>
<td>Range</td>
<td>SD</td>
<td>Bone-Tip Difference as %</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>-------</td>
<td>----</td>
<td>--------------------------</td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>3.29</td>
<td>25.56</td>
<td>1-6</td>
<td>1.06</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>2.30</td>
<td>19.31</td>
<td>0-5</td>
<td>1.18</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>3.47</td>
<td>26.82</td>
<td>0-6</td>
<td>1.63</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>3.27</td>
<td>25.73</td>
<td>1-5</td>
<td>1.09</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>2.12</td>
<td>16.44</td>
<td>1-4</td>
<td>0.76</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>2.38</td>
<td>19.93</td>
<td>-4-6</td>
<td>2.69</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>2.50</td>
<td>21.74</td>
<td>-4-6.5</td>
<td>2.30</td>
</tr>
<tr>
<td>Total</td>
<td>201</td>
<td>2.93</td>
<td>23.39</td>
<td>-4-6.5</td>
<td>1.63</td>
</tr>
</tbody>
</table>