

# Inter and Intra rater reliability of measurement of testicular volume

Akbarian-Tefaghi L<sup>1</sup>, De Silva S<sup>1</sup>, Langley N<sup>2</sup>, Stanton A<sup>2</sup>, Wright NP<sup>3</sup>, Elder CJ<sup>1,3</sup>

University of Sheffield<sup>1</sup>, Sheffield Hallam University<sup>2</sup> Sheffield Children's Hospital<sup>3</sup>, Sheffield, UK

## Introduction and objectives

Measurement of testicular volume (TV) by Prader orchidometer is a practical and inexpensive method of pubertal staging, with testicular size correlating to pubertal stage (figure 1).

Despite its popularity and widespread use there is a paucity of evidence as to its reliability in different hands <sup>1,2,3</sup>.

Studies have highlighted that standardised training on optimising visual inspection, isolation and measurement of the testes could improve accuracy <sup>3,4</sup>.

A pilot study revealed high inter and intra-practitioner variability.

Figure 1: Prader orchidometer.



## Research Questions

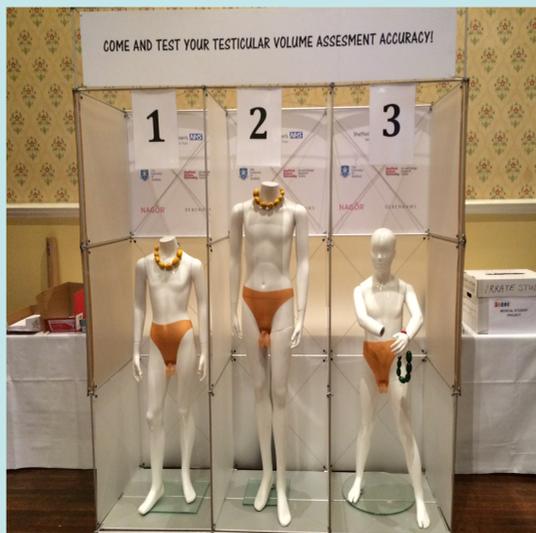
1. Is there significant inter-rater variability in estimating testicular volume?
2. Is there significant intra-rater variability in estimating testicular volume?
3. Do factors such as gender, experience or training influence the reliability of testicular volume measurement?
4. What is the acceptability and feasibility of our model as a training tool?

## Manufacturing testicles

Creating realistic synthetic testicles was a considerable challenge! Collaborations were forged with engineers at Sheffield Hallam University who trialed a number of different materials and techniques. Molded silicon testes encased in multiple layered latex scrotum most closely mimicked reality. Further improvements were made following a pilot study and the prosthetic testicles made by the UK company Nagor using the SHU molds. Scrotums were displayed on donated shop manikins using latex pants (figure 2).

## Methods

Figure 2: Three mannequins used in the BSPED study with synthetic testicles attached to shop manikins using latex pants.



- The study was conducted over the 3-day BSPED meeting in November 2015. Three child-sized mannequins displayed latex scrotums containing prosthetic testicles of 3ml, 4ml, 5ml, 10ml and 20ml.
- Demographic data, paediatric endocrinology experience, TV examination training and information on examination technique were collected anonymously and TV estimations performed.
- Delegates were asked to repeat their measurements again during the meeting and were blinded to the size of the testicles. Scrotum order was changed daily to minimise recall bias.

## Results

Figure 3: Percentage of participants at different levels of clinical experience

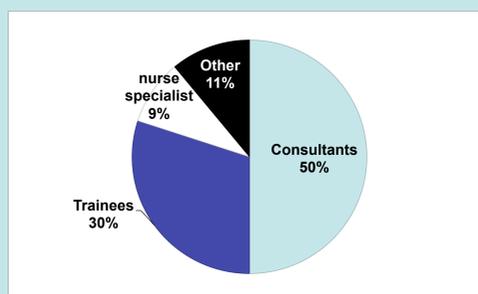
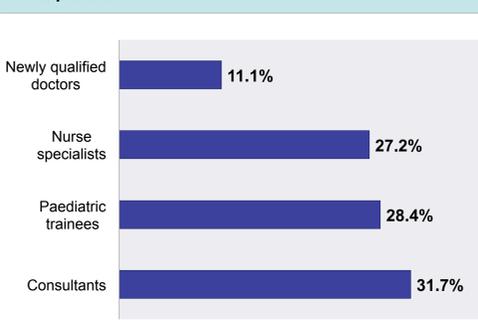


Figure 4: Accuracy of TV measurements at different levels of experience



- 208 delegates participated (158F, 50M), with varying levels of experience (figure 3).
- 60% had less than 5 years experience in clinical endocrinology, 19% had practiced for over 10 years.
- 25.5% had received formal training in TV estimation.
- Examination technique varied. The majority preferring the patient recumbent (54%) and using both look and feel for estimation (59%).
- TV was measured accurately on **30%** of occasions. Underestimations were made on **40%** of measurements and overestimation on **30%**.
- Males and females were similarly inaccurate, males estimating correctly on 31.4% of occasions and females on 29.7%.
- Experience improved accuracy (figure 4).
- Inaccuracies were highest at the smallest volumes: **57%** overestimating 3ml testis (largest coefficient of variation, 57.8%), compared with **17%** for the 20ml testis (smallest coefficient of variation, 31.9%).

## Discussion

### Learning points:

This study showed significant variability between participants in measuring testicular volume, with only 30% of measurements being accurate. Seniority improves accuracy of measurements. Smaller volume testicles of 3ml and 4ml were measured less accurately which has implications for clinical decision making.

### The future:

Delegate feedback supported formal training and use of these prosthetic models as an educational tool. A further study looking at the effect of training on TV estimation accuracy is being undertaken.

## References

1. Karaman MI, et al. (2005). "Measurement of pediatric testicular volume with Prader orchidometer: comparison of different hands." *Pediatr Surg Int* 21(7):517-520
2. Carlsen E, et al. (2000). "Inter-observer variation in the results of the clinical andrological examination including estimation of testicular size." *Int J Androl* 23(4):248-253
3. Tatsunami S, et al. (2006). "Inter/intra investigator variation in orchidometric measurements of testicular volume by ten investigators from five institutions." *Asian J Androl* 8(3):373-378
4. Nahata L, et al. (2015). "The Male Genital Examination: Overcoming Barriers to Prevent Missed Diagnoses." *Clinical paediatrics* 54(13): 1237-1239

