

# Inter and intra-Rater Reliability of Accuracy of Testicular volume Evaluation (IRRATE): a simulation study

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

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

## Introduction

The measurement of testicular volume (TV) by Prader orchidometer is a practical and inexpensive method of pubertal staging, with testicular size correlating to pubertal stage (fig. 1). Despite its popularity and wide spread use over the last 50 years 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>



Fig. 1. Prader orchidometer with pubertal stages indicated

## 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. Does training improve reliability of testicular volume measurement?
4. Do factors such as gender or experience influence the reliability of testicular volume measurement?
5. What is the acceptability and feasibility of our model as a training tool?

## Making testicles....

Manufacturing synthetic testicles proved to be our greatest challenge. After water-filled gloves, synthetic testicular implants displayed in thick flesh-coloured tights and cadaveric porcine testicles didn't make for realistic, practical simulation options we forged collaborative links with engineers from Sheffield Hallam University (SHU). An innovative solution of casting silicon moulds, of known volume, and encasing them in an oiled latex scrotum for realistic movement were designed and manufactured (fig. 3). The scrotums were displayed on shop dummies using latex pants (fig. 2).

## Methods

A pilot study was conducted at an endocrine training day on puberty, which included a workshop on TV assessment. Participants were asked to measure four pairs of testicles (fig. 2) using an orchidometer. Participants were blinded to the TVs. The assessment was repeated after attending the training session.



Fig. 2. Pilot study mannequins with scrotums



Fig. 3. Scrotums manufactured by Sheffield Hallam Art and Design Research Centre.

## Results

- 16 participants: 10 junior doctors, 4 consultant paediatric endocrinologists, 2 specialist endocrine nurses.
- 12 female participants, 4 male.
- 13 repeated TV measurements following training.
- 46% of participants overestimated TV, 7% underestimated and 46% were accurate. (fig. 4).

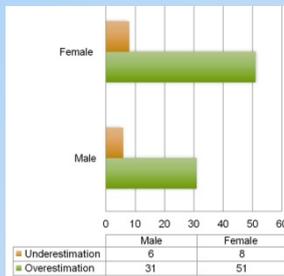


Fig. 4. Graph showing correlation between gender and accuracy

- Males were significantly more accurate than females ( $p=0.04$ ).
- Larger volume testicles were more accurately measured: 20ml volume had the smallest variability (CV 12%) and 3ml volume the largest (CV 50%).
- Variation was higher at all volumes in the left testicle.
- Experience and training did not significantly impact on accuracy.

## Discussion

### Limitations of the study:

- Manufactured prototypes
  - Scrotum oil leaking
  - Testicles too hard
- Small sample size

### Feedback:

Participants felt they required more training as the pilot study revealed lack of confidence in the reliability of their assessments. They reported narrowing their estimated volumes down to two beads.

Some participants felt that prototypes could be made more realistic by including more detailed external genitalia, spermatic cord and textured skin.

### The future:

- Feedback led to a second prototype manufactured in collaboration with SHU engineers and NAGOR, a UK company who produces prosthetic testes. These specifically engineered prototypes will be employed at BSPED Sheffield 2015 in a larger study.
- A future study is proposed to investigate the effects of training on reliability.
- Commercialisation for use as simulation tools for training.

## Conclusions

- **We found considerable variation in the estimation of testicular volume between subjects and at smaller volumes.**
- **Males are significantly more accurate testicular staging than females.**

## 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). "Interintra investigator variation in orchidometric measurements of testicular volume by ten investigators from five institutions." *Asian J Androl* 8(3):373-379
4. Nahata L, et al. (2015). "The Male Genital Examination: Overcoming Barriers to Prevent Missed Diagnoses." *Clinical paediatrics* 54(13): 1237-1239