

# Dichotomous Estimation of Prostate Volume: A Diagnostic Study of the Accuracy of the Digital Rectal Examination

Michael Z. Su, Daniel Lenaghan, Henry H. Woo

Sydney Adventist Hospital Clinical School, University of Sydney, Sydney, NSW, Australia

**Purpose:** To assess the diagnostic reliability of a dichotomous digital rectal examination (DRE) tool in assessing prostate volume (PV) compared to gold-standard transrectal ultrasound (TRUS) volume measurement.

**Materials and Methods:** Male patients presenting to a single tertiary referral centre urology practice requiring TRUS were prospectively recruited between January 2010 and August 2011. Size was estimated by DRE immediately prior to TRUS measurement. DRE measurements were classed into four groups: < 30 mL, 30 to 49 mL, 50 to 99 mL and > 100 mL. The primary outcomes were sensitivity, specificity, and the positive and negative likelihood ratios for a 30 mL cut-off.

**Results:** Three hundred and three patients were recruited to the study. The median age of the study group was 64.9 years. On TRUS analysis, 244 patients had a PV larger than 30 mL and 139 of them, larger than 50 mL. The median PV was 47 mL with a median International Prostate Symptom Score of 10 and prostatic specific antigen (PSA) of 5.7. When analysed for the ability to identify a gland larger than 30 mL, DRE had a high sensitivity and low negative likelihood ratio. The median PSA level (ng/L) for prostates measured by DRE with a 30 mL cut-off was significantly different with higher median PSA values for volumes larger than 30 mL.

**Conclusions:** DRE is a reliable tool for dichotomous assessment of prostatic volumes above 30 mL and 50 mL. These results illustrate the value of re-examining the role of categorical DRE estimations in benign prostatic hyperplasia patients.

**Key Words:** Diagnostic tests, routine; Digital rectal examination; Lower urinary tract symptoms; Prostatic hyperplasia

## INTRODUCTION

Benign prostatic hyperplasia (BPH) is one of the most common progressive conditions among ageing men. Given that most developed countries are experiencing a trend towards an ageing population, we should expect the incidence of BPH to increase dramatically in the coming

decades, with a major public health impact on healthcare providers. The natural history of BPH requires accurate disease staging to guide decision making between watchful waiting, or medical or surgical therapy. The combination of a burgeoning elderly population and associated BPH disease burden will require an innovative solution to ensure appropriate treatment decisions are made with the

Received: Sep 5, 2013; Revised: Oct 30, 2013; Accepted: Nov 4, 2013

Correspondence to: Henry H. Woo

Sydney Adventist Hospital Clinical School, University of Sydney, Suite 406, SAN Clinic, 185 Fox Valley Road, Wahroonga, Sydney, New South Wales 2076, Australia,  
Tel: +61-2-9473-8765, Fax: +61-2-9473-8969, E-mail: hwoo@urologist.net.au

Copyright © 2013 Korean Society for Sexual Medicine and Andrology  
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

least possible economic impact on the health system.

Prostate volume (PV) has been demonstrated to be an important disease predictor of BPH [1]. A threshold of 30 mL has been determined to be a categorical threshold that defines an increased risk of clinical BPH progression [2]. Determining prostate enlargement is therefore an integral part of patient evaluation. Prostate size can be estimated by transrectal ultrasound (TRUS) or, less accurately, by transabdominal ultrasound. Both of these are expensive resources that are generally not available at initial patient evaluation. A simple alternative, cost-effective way of estimating prostate size would therefore be desirable.

A digital rectal examination (DRE) has been shown to be a relatively poor predictor of exact prostate size [3]. However, the practical evaluation and management of BPH does not require exact PV estimation. Instead, it is important to correctly categorise prostate size above or below a given volume defined as 'enlargement'.

The following study was designed to examine the reliability of DRE as a simple, minimally invasive and cost-effective tool to provide guidance for clinicians on differentiating between set PV thresholds and to compare its performance to the gold standard TRUS measurements for patients presenting with lower urinary tract symptoms (LUTS).

## MATERIALS AND METHODS

### 1. Patient selection

Patients presenting at a tertiary centre practice and undergoing DRE and TRUS were recruited between January 2010 and August 2011 with their informed consent. Inclusion criteria were men who were having prostate size formally measured prior to BPH-related surgery or in conjunction with prostate biopsy. Patients that refused participation or had incomplete data on collection were excluded from the study.

### 2. Data collection

Data were collected on demographics, World Health Organization International Prostate Symptom Score (IPSS) and serum total prostatic specific antigen (PSA) levels. There was no intervention additional to their standard of care for the purposes of this study, other than recording

size estimations on the basis of DRE immediately prior to insertion of the TRUS probe.

### 3. Outcomes

DREs were performed by either an urologist or a senior urology trainee. Each patient was categorically graded to one of four groups: < 30 mL (0), 30 to 49 mL (1), 50 to 99 mL (2), or > 100 mL (3). These results were compared to gold-standard TRUS-measured PV calculated as the product of 0.52 and the transverse, anterior-posterior and sagittal length dimensions. The primary outcome was diagnostic test accuracy, defined as sensitivity, specificity, and the positive and negative likelihood ratios of DRE estimation of PVs at greater or less than the 30-mL cut-off.

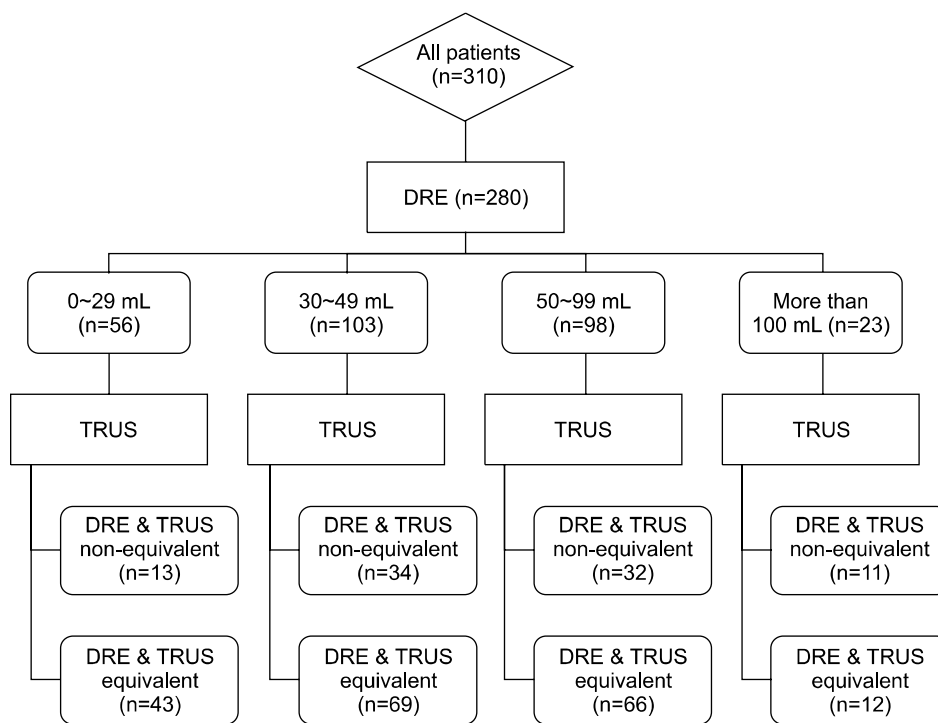
### 4. Statistics

Statistical analysis was performed using IBM SPSS Statistics version 21 (IBM Co., Armonk, NY, USA). The median values of population demographics were reported with interquartile ranges. Diagnostic test accuracy was presented as sensitivity, specificity, and conventional positive and negative likelihood values with 95% confidence intervals (CIs). Kappa statistic scores were calculated to compare the correlation of DRE volume estimates to TRUS volume measurements. Scores from 0.41 to 0.60 were considered to be in moderate agreement, 0.61 to 0.80 in substantial agreement, and 0.81 to 0.99 in almost perfect agreement [4]. The median PSA values were compared between dichotomous volume groups using the Mann-Whitney U test. A p value of less than or equal to 0.05 was considered statistically significant.

**Table 1.** Patient characteristics

Variable	Value
Age (yr)	65 (59~71)
IPSS (points)	10 (5~17)
QL score (points)	2 (1~4)
PSA (ng/L)	5.7 (3.6~8.0)
PV (mL)	47 (34~67)

Values are presented as median (interquartile range). IPSS: International Prostate Symptom Score, QL: quality of life, PSA: prostatic specific antigen, PV: prostate volume.



**Fig. 1.** Study flowchart. DRE: digital rectal examination, TRUS: transrectal ultrasound.

**Table 2.** Diagnostic test results above the PV cut-off

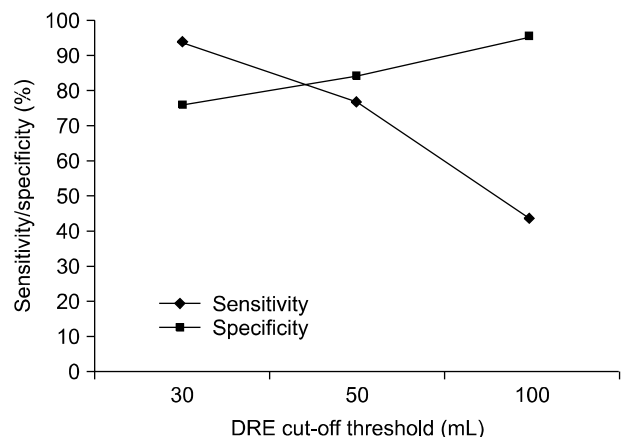
Variable	DRE PV ≥30 mL
Sensitivity	94.3 (90.1 ~ 96.8)
Specificity	78.2 (64.6 ~ 87.8)
Positive likelihood ratio	3.97 (2.51 ~ 6.28)
Negative likelihood ratio	0.08 (0.05 ~ 0.13)

Values are presented as percent (95% confidence interval). PV: prostate volume, DRE: digital rectal examination.

## RESULTS

A total of 280 patients out of 310 eligible patients were included in this analysis. Seven patients were excluded from the analysis because their dataset was incomplete, while 23 were excluded as they had been assessed solely by a trainee. Table 1 summarises the characteristics of the study population. The study protocol with patient numbers at each stage of assessment is shown in Fig. 1 ('DRE & TRUS equivalent' refers to the DRE estimate being equal to the TRUS volume).

Diagnostic test accuracy parameters for a DRE cut-off of 30 mL are shown in Table 2. DRE performed favourably at measured volumes of 30 mL or greater by demonstrating high sensitivity and low negative likelihood ratio values.



**Fig. 2.** Sensitivity and specificity plot against DRE cut-offs. DRE: digital rectal examination.

Fig. 2 illustrates the sensitivity and specificity of three different cut-off thresholds for comparison. A 30-mL threshold combined a more ideal balance of these outcomes compared to 50-mL and 100-mL cut-offs. The Kappa statistic demonstrated substantial agreement between DRE and TRUS at this threshold ( $\kappa=0.72, p<0.01$ ). The trainee estimates varied markedly from consultant DRE estimates ( $\kappa=0.30, p=0.02$ ) and TRUS volumes ( $\kappa=0.21, p=0.03$ ), and thus we subsequently excluded the patients that had been examined solely by a trainee due to the poor correla-

**Table 3.** PSA values by diagnostic tool and prostate volume

Variable	PSA	p value <sup>a</sup>
DRE (mL)		
>0, <30	4.30 (2.30~6.80)	0.02
≥30	5.90 (4.12~8.46)	
TRUS (mL)		
>0, <30	3.70 (2.15~6.25)	<0.01
≥30	6.06 (4.18~8.63)	

Values are presented as median (interquartile range).

PSA: prostatic specific antigen, DRE: digital rectal examination, TRUS: transrectal ultrasound.

<sup>a</sup>Mann-Whitney U test.

tion between trainee and consultant estimates.

The difference between median PSA levels (ng/L) for prostates above and below a 30-mL DRE cut-off was statistically significant (Table 3). In comparison, there was a significant difference between median PSA groups when the same cut-offs were used for TRUS volumes.

## DISCUSSION

DRE is a commonly used examination technique in the assessment of a patient presenting with LUTS. It is recommended by the latest guidelines from the American Urological Association and European Association of Urology. DRE is also important in excluding malignancy and neurological disease. There is limited literature that addresses the accuracy of DRE for measurement of PV. Common criticisms of DRE include high interobserver variability and significant underestimation of PVs [5-7]. Roehrborn et al [3] have investigated DRE as compared to TRUS volumes in several large scale population studies. Despite its limitations, the authors found that both DRE estimated volumes and surface area were significantly correlated with TRUS volumes below 30 mL. This study is the first to suggest DRE with a specific cut-off as a potentially reliable test for high-risk PV, particularly with a cut-off at 30 mL as measured by TRUS.

We evaluated the accuracy of dichotomous DRE measurement as a diagnostic method with a similar study design to those conducted by Roehrborn et al [3,8] and Roehrborn [5]. Our study had a larger population compared to the 121-patient 2001 prospective study, and adapted the primary outcomes of the previous two population registry

analysis studies. Utilising larger estimation categories, our study findings were consistent with the previous literature in evaluating volumes above 30 mL. Outcomes using this cut-off threshold were highly correlated with gold standard measurements ( $\kappa=0.705$   $p=0.00$ ) supporting a dichotomous over a continuous or multiple ordinal grading scale. These results are also supported by a recent smaller Canadian study [9].

A PV cut-off of 50 mL has been advocated by previous population studies. The Krimpen study evaluated 1,688 men aged between 50 and 70 years with benign prostatic disease to compare the test performance of DRE and serum PSA in determining dichotomous PVs against planimetric TRUS [10]. Their findings suggested that 40-mL and 50-mL cut-offs were superior in all test characteristics compared to a 30-mL cut-off. A 30-mL cut-off demonstrated the highest sensitivity (39.8%, 95% CI 36.2 to 43.4) and lowest specificity (81.6%, 95% CI 78.5 to 84.4). These results are in contrast to our findings of a higher sensitivity and low negative likelihood ratio for a 30-mL threshold. These results demonstrate a lower false negative rate and excellent ability to exclude patients below this clinically significant volume to ensure that most patients treated are not from a group that would be unlikely to benefit from therapy. It is not clear whether inclusion of men with known or suspected prostate cancer would contribute to the differences in our study outcome, given that the Krimpen study excluded patients with a history of prostate or bladder cancer.

PSA levels have been advocated as a surrogate marker of prostate size. There are numerous independent factors that influence this as a volume marker including age, malignancy, and 5-alpha-reductase inhibitor (5-ARI) therapy. Numerous studies have assessed the reliability of PSA as a predictor of PV when malignancy is excluded [11-13]. Bohnen et al [14] found that PSA was highly correlated with PVs greater than 30 mL and found even better results at volumes greater than 40 mL or 50 mL. PSA values of 1, 2.5 and 4 ng/mL were advocated as cut-offs for 30 mL, 40 mL and 50 mL, respectively, by Bosch et al [10] in a population with benign prostatic disease. Our mean PSA levels reflect these findings with median values consistent for clinically larger volumes. The high average PSA level was likely due to a large proportion of subjects that presented for assessment of suspected malignancy. PSA appears to

be useful as a measure of PV in benign prostatic disease, but DRE remains an important modality to use in conjunction due to the potential confounding factors of age and malignancy.

Ethnicity also appears to be a contributing factor to PSA and DRE PV estimations. African-Americans have a higher contribution of PSA from benign prostatic tissue compared to Caucasians on a volume-volume analysis [15]. Asian males have also been demonstrated to have different PSA levels relative to PV. In a large 5,716 patient study, Chung et al [16] found that Korean males had marginally lower PSA and smaller PV compared to Caucasians. Our study population was predominantly Caucasian, with a small group of East Asian patients. We were unable to statistically analyse this group of patients due to the small sample size. Nonetheless, this is an important topic that can be explored in future studies.

Our study had several limitations. Despite a large sample size, our patients were evaluated at a single tertiary centre predominantly by a single clinician. This could be improved through future studies involving a range of clinicians from attending urologists to junior trainees. Such a study design would allow us to explore more comprehensively the interobserver reliability between different examiners and patterns among different ethnic groups. The inclusion group predominantly presented for TRUS biopsy with a large proportion with suspected malignancy to ensure adequate sample size. Nonetheless, our study maintains strong external validity due to several factors. Our group had a mean IPSS of 11.86, which is consistent with clinically significant LUTS. The study was designed to investigate outcomes specifically related to volume in comparison with TRUS. Our results support the reliability of DRE-based estimations above a threshold cut-off. Future study designs may further improve the external validity by including multicentre cohorts that exclude patients with lower urinary tract malignancies.

## CONCLUSIONS

DRE remains an important, rapid, cost-effective tool in the examination of a patient presenting with LUTS. It is useful for determining whether the prostate has a categorical volume greater than 30 mL and can be used in this

context to guide diagnosis of patients at risk of disease progression that would be appropriate for 5-ARI therapy. The results of our study illustrate the value of re-examining the role of DRE in estimating PV in patients presenting for assessment of BPH.

## ACKNOWLEDGEMENTS

We thank Karen Byth for assistance in data analysis.

## REFERENCES

1. Kumar V, Marr C, Bhuvangiri A, Irwin P. A prospective study of conservatively managed acute urinary retention: prostate size matters. *BJU Int* 2000;86:816-9.
2. Jacobsen SJ, Jacobson DJ, Girman CJ, Roberts RO, Rhodes T, Guess HA, et al. Natural history of prostatism: risk factors for acute urinary retention. *J Urol* 1997;158:481-7.
3. Roehrborn CG, Girman CJ, Rhodes T, Hanson KA, Collins GN, Sech SM, et al. Correlation between prostate size estimated by digital rectal examination and measured by transrectal ultrasound. *Urology* 1997;49:548-57.
4. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360-3.
5. Roehrborn CG. Accurate determination of prostate size via digital rectal examination and transrectal ultrasound. *Urology* 1998;51:S19-22.
6. Varenhorst E, Berglund K, Löfman O, Pedersen K. Inter-observer variation in assessment of the prostate by digital rectal examination. *Br J Urol* 1993;72:173-6.
7. Kijvikai K. Digital rectal examination, serum prostatic specific antigen or transrectal ultrasonography: the best tool to guide the treatment of men with benign prostatic hyperplasia. *Curr Opin Urol* 2009;19:44-8.
8. Roehrborn CG, Sech S, Montoya J, Rhodes T, Girman CJ. Interexaminer reliability and validity of a three-dimensional model to assess prostate volume by digital rectal examination. *Urology* 2001;57:1087-92.
9. Ahmad S, Manecksha RP, Cullen IM, Flynn RJ, McDermott TE, Grainger R, et al. Estimation of clinically significant prostate volumes by digital rectal examination: a comparative prospective study. *Can J Urol* 2011;18:6025-30.
10. Bosch JL, Bohnen AM, Groeneveld FP. Validity of digital rectal examination and serum prostate specific antigen in the estimation of prostate volume in community-based men aged 50 to 78 years: the Krimpen Study. *Eur Urol* 2004;46:753-9.
11. Collins GN, Lee RJ, McKelvie GB, Rogers AC, Hehir M. Relationship between prostate specific antigen, prostate volume and age in the benign prostate. *Br J Urol* 1993;71:445-50.
12. Pinsky PF, Kramer BS, Crawford ED, Grubb RL, Urban DA,

- Andriole GL, et al. Prostate volume and prostate-specific antigen levels in men enrolled in a large screening trial. *Urology* 2006;68:352-6.
13. Yu HJ, Chiang GJ, Chiu TY, Lai MK. Relationship between serum prostate specific antigen concentration and prostate volume. *J Formos Med Assoc* 1995;94:666-70.
  14. Bohnen AM, Groeneveld FP, Bosch JL. Serum prostate-specific antigen as a predictor of prostate volume in the community: the Krimpen study. *Eur Urol* 2007;51:1645-52.
  15. Fowler JE Jr, Bigler SA, Kilambi NK, Land SA. Relationships between prostate-specific antigen and prostate volume in black and white men with benign prostate biopsies. *Urology* 1999;53:1175-8.
  16. Chung BH, Hong SJ, Cho JS, Seong DH. Relationship between serum prostate-specific antigen and prostate volume in Korean men with benign prostatic hyperplasia: a multi-centre study. *BJU Int* 2006;97:742-6.