Eur J Vasc Endovasc Surg (2011) 41, 704–710

Observer Agreement of Lower Limb Venous Reflux Assessed by Duplex Ultrasound Scanning using Manual and Pneumatic Cuff Compression in Patients with Chronic Venous Disease and Controls

R. Broholm [a,*], S. Kreiner [b], N. Bækgaard [a], L. Panduro Jensen [a], H. Sillesen [a]

[a] *Vascular Clinic, Department of Vascular Surgery, Rigshospitalet and Gentofte, University of Copenhagen, Copenhagen, Denmark*
[b] *Department of Biostatistics, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark*

Submitted 28 September 2010; accepted 4 January 2011
Available online 17 February 2011

KEYWORDS
Venous reflux assessment;
Observer agreement;
Duplex ultrasound;
Manual compression;
Pneumatic cuff compression;
Chronic venous disease

Abstract *Objectives:* The study aimed to evaluate observer agreement between two experienced ultrasound operators examining deep venous reflux assessed by duplex ultrasound (DU) using either manual or pneumatic cuff compression. In addition, the two methods were compared with each other with regard to immediate "eyeballing" and direct measurements of reflux time from Doppler flow curves.
Design: This was a case control study.
Material and Methods: Cases were found among patients admitted to our department with deep venous thrombosis of the iliac, femoral or popliteal veins during the period 1999–2006. Controls were departmental staff. DU was used to assess valve function in the common femoral, femoral and popliteal veins in the standing position using manual and pneumatic cuff compression. The investigators were blinded to the other's observations. Observer agreement was assessed using the Rasch model for binary items.
Results: Twenty patients and 20 controls participated in the study and were analysed by the Rasch model. Quantitative measurement was found to be more reliable than "eyeballing", and cuff compression was more reproducible in identifying reflux than manual compression. We found that assessment by manual measurement by one investigator functioned at a lower level of expertise than for the other investigator.
Conclusions: Cuff measurements were more accurate in diagnosing deep venous reflux than manual measurements, and measurement was more accurate than "eyeballing". The fact that

* Corresponding author. Vascular Clinic, Department of Vascular Surgery, Rigshospitalet and Gentofte Hospital, Niels Andersens Vej 65, 2900 Hellerup, Denmark. Tel.: +45 28 49 48 70; fax: +45 39 77 76 07.
E-mail address: rbroholm@dadlnet.dk (R. Broholm).

doi:10.1016/j.ejvs.2011.01.014

assessment by manual compression by one investigator functioned at a lower level of expertise suggests that cuff measurement might be the optimal assessment method, especially in the difficult cases.

Over the past three to four decades, duplex ultrasound has developed into an invaluable diagnostic tool in vascular departments for assessment of the arterial and venous circulations. It has become the principal tool in examining patients with chronic venous disease, and it is useful for investigation of venous reflux, both in the superficial and deep venous system.[1,2] Several studies have led to the definition of reflux as a retrograde flow in the reverse direction to physiological flow lasting more than 0.5 s,[3,4] and others have proposed a cut-off value of 1.0 s.[5] Venous reflux can be elicited either after release of manual or pneumatic cuff compression with the patient standing or by having the patient perform a Valsalva manoeuvre in the supine position. Pneumatic cuff compression has been used to assess quantitative measurements of venous reflux.[3,6,7] Despite the fact that duplex investigation is operator dependent, only a few studies have evaluated the reproducibility of the duplex measurements,[4,8,9] and inter-observer variability in reflux assessment is not always perfect. A recent study compared a new pneumatic cuff inflation/deflation device with manual compression in the diagnosis of venous reflux.[10]

The aim of this study was to evaluate the observer agreement between two experienced ultrasound operators examining deep venous reflux assessed by duplex ultrasound, using either manual or pneumatic cuff compression in the standing position. In addition, the two methods were compared with each other with regard to immediate "eyeballing" and direct measurements of reflux time from ultrasound Doppler flow curves.

Materials and methods

Between April 2008 and July 2009, deep venous valve function was studied in 40 individuals: 20 patients with a previous history of deep venous thrombosis (DVT) treated with warfarin and 20 healthy persons with no prior history of venous disease. Patients were selected among a group of patients admitted to our department of vascular surgery with DVT of the iliac, femoral (FV) or popliteal (PV) veins in the years 1999–2006.

We examined the common femoral vein (CFV), the FV a few centimetres below the junction of the profunda femoris vein and the femoral vein and the PV. Both lower limbs were assessed. All volunteers were categorised according to the clinical classification (C) for reporting standards in venous disease, CEAP (Clinical, etiologic, anatomic and pathophysiologic data).[11,12] Further characteristics, such as age and gender, were also recorded (Table 1).

The observer agreement was obtained by comparing duplex measurements performed by two experienced ultrasound operators, blinded for each others' results, in the 20 healthy persons and in the 20 individuals with a previous history of DVT. All persons were examined with both manual calf compression and with pneumatic calf cuff compression by the two examiners (RB and NB) (Fig. 1).

The study was approved by the national ethics committee (journal number H-D-2007-0120), and all participants provided written informed consent for their inclusion in the study.

Duplex ultrasound

A colour duplex ultrasound scanner (Phillips iU22, Bothell, Washington, USA), with a 9–3 MHz linear array transducer, was used in the examination of venous reflux in the CFV, in the FV and in the PV in all cases.

All participants were scanned in a standing position, holding on to a support and with their weight on the leg contralateral to that being examined. The first examination was performed with distal manual compression of the calf followed by a sudden release to elicit retrograde flow.

In the second examination, a standardised pneumatic cuff (Venopulse, STR Teknik, Norway) with a cuff size of 17 cm, was placed around the calf and then rapidly inflated followed by deflation. This procedure was managed with a foot switch (Fig. 2). Again, the volunteers were examined in the standing position with the weight on the contralateral leg. The cuff was inflated to 150 mmHg and inflation was automatically held for 3 s, followed by rapid deflation. A pressure of 150 mmHg was applied because this is sufficient to overcome the hydrostatic and gravitational pressures and to produce upward blood flow.

Clinical examination of the participants was first performed by one investigator (RB), both with manual and cuff compression; and after data had been saved on the hard drive, they were deleted from the ultrasound system before the second investigator entered the room. Then, the second investigator (NB) performed his examination blinded to the outcome of the first investigator (RB).

All data were recorded digitally onto the hard drive of the ultrasound system and later transferred onto digital video discs (DVDs). In addition, black-and-white pictures of frozen Doppler spectra recorded during compression and after release were printed. At every examination, the investigator made an immediate estimate ("eyeballing") of whether reflux was present or not and recorded his findings on the case report form (CRF). In addition, every Doppler curve was analysed more thoroughly after the clinical examination, measuring the exact reflux time (Fig. 1). This was later compared to the "eyeballing" result. The direct measurements were made by one investigator (RB), using the software Sante DICOM Editor version 3.0.4.

In addition to the presence or absence of reflux judged by immediate "eyeballing", the reflux time (measured directly from the stored Doppler spectra) and the diameter of the vein were recorded in each venous segment.

Reflux was considered to be present if the duration of retrograde flow was above 0.5 s in the FV and in the PV, as proposed by van Bemmelen,[3] and a reflux time above 1.0 s in the CFV was considered pathological.[5]

Table 1 Characteristics of participants.

	Patients (20) (40 limbs)	Controls (20) (40 limbs)
Age (median)	32 years (range: 24—63)	32 years (range: 21—55)
Female:Male ratio	17:3	18:2
Previous DVT (no. limbs)	20 (23)	0
CEAP	19 C3, 1 C5	19 C0, 1 C3
Calf circumference (cm, median)	40.5 (range: 30.5—46.5)	36.5 (range: 32—47)
Vein diameter, mean (cm)	CFV 1.36	CFV 1.47
	FV 1.06	FV 1.34
	PV 0.88	PV 0.91

Statistical analysis

Power calculations were performed to determine the number of participants, based on the assumptions that two experienced investigators would agree in the diagnosis of venous reflux in 90% of cases, and that reflux would be detected in 20—30% of cases. Type I error (alpha) is 0.05, and type II error (beta) is 0.10. If we assumed that reflux would be present in 30% of cases, 15 participants would be required. However, if venous reflux was present in only 10% of cases, 32 participants would be required. Therefore, a total of 40 volunteers was included in the study.

We chose not to use Cohen's kappa coefficient (κ) as a measure of observer agreement, as the value of κ is restricted to the specific cases and controls that have been examined. Instead, observer agreement between the two experienced ultrasound operators was analysed by Rasch's model for binary items.[13—16] Westergaard et al. provide two applications of Rasch models for observer agreement.[13] The model describes each observer's evaluation of a specific case of reflux assessment, with the probability given by a logistic regression model. Here, θ_i can be regarded as a random patient effect containing information both on the presence of reflux and on the difficulty of the case. Reflux is

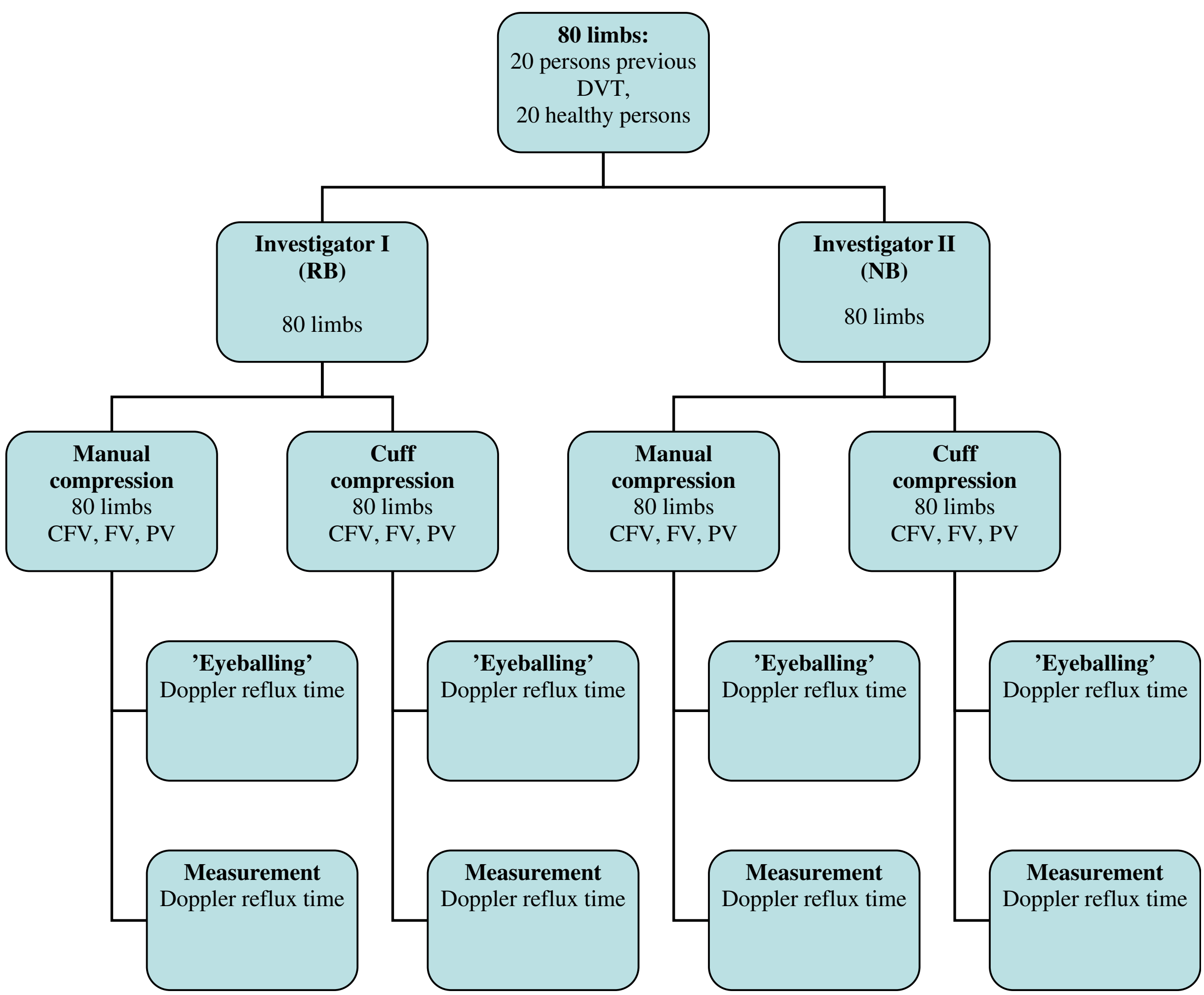

Figure 1 Diagram illustrating the examination pathway.

Figure 2 The Venopulse Model (STR Teknik, Norway). The pneumatic cuff is placed around the calf with the patient in standing position. Inflation and deflation are controlled by a foot switch.

present if θ_i is positive and not present if θ_i is negative. The closer θ_i is to zero, the more difficult the case. Variation among the observers is represented by the observer parameter α_j, representing observer bias. Bias leans towards reflux if $\alpha_j > 0$ and towards no reflux if $\alpha_j < 0$. There is no observer bias if $\alpha_j = 0$ for all observers.

The model is given by

$$p_{ij} = \frac{\exp(\theta_i + \alpha_j)}{1 + \exp(\theta_i + \alpha_j)}$$

The Rasch model assumes that the cases have the same effect on observers. However, to distinguish between expert and novice, the degree of observer expertise (β_j) can be added to the model

$$p_{ij} = \frac{\exp(\beta_j(\theta_i + \alpha_j))}{1 + \exp(\beta_j(\theta_i + \alpha_j))}$$

If $\beta_j = 0$, the level of expertise is at its lowest, as rating here is equivalent to the flipping of a coin.

The initial analysis of observations consists of a test of the Rasch model followed by a test that all bias parameters are equal to zero. Conditional likelihood ratio tests were used primarily to test the fit of the Rasch model and then to test that observations were unbiased.[17]

Results

Among the 20 patients with a history of DVT, three patients had bilateral thrombosis. Patients included 17 women and 3 men with a median age of 32 years (range, 24–63 years), and 19 was categorised as CEAP C3 and one patient as C5. Volunteers were department staff and included 18 women and 2 men, with a median age of 32 years (range, 21–55 years). One control was categorised as C3, and all other controls had no signs or symptoms of chronic venous disease, and were therefore categorised as C0. Further details can be seen in Table 1.

In every volunteer, three venous segments were examined in both lower limbs (CFV, FV and PV), with two different methods (manual and pneumatic cuff); and with both exact measuring and "eyeballing", several parameters were obtained for each of the two investigators; 'manual measure', 'manual eyeballing', 'cuff measure' and 'cuff eyeballing', corresponding to 48 parameters for each patient and a total of 1920 data on whether reflux was present or not. The vein segment was categorised as normal or with reflux according to the cut-off values mentioned above. Values of reflux time above 2.5 s were classified as indefinitely long and were not recorded more accurately.

The conditional likelihood ratio tests provided marginal evidence against the Rasch model ($p = 0.023$), corresponding to good fit of the model. The hypothesis that bias parameters were equal to zero was rejected ($p = 0.006$), but the hypothesis that the investigators were not biased relative to each other was accepted ($p = 0.85$). The bias that was found was therefore related to the different methods. Table 2 shows the estimates of the bias parameters together with odds ratio coefficients describing observer bias relative to cuff measurement.

Two methods may be used for evaluating the different assessments of reflux. The so-called outfit statistic is one of those which are often used in health research, and the more powerful fit statistic is Loevinger's *H* coefficient defined in non-parametric item response theory.[18] Both of these points at manual measurement by investigator RB as an assessment where the level of expertise is less than the level of expertise for the other assessments (Outfit = 2.05, $p = 0.002$; $H = 0.81$, $p = 0.002$).

Results are presented in Fig. 3. With the cuff method used as the reference and compared with this, the figure illustrates that it is more reproducible to identify reflux by cuff measurement (black line) than by manual measurement (blue line), although this did not reach statistical significance ($\chi^2 = 3.3$, df = 1 and $p = 0.069$). It was more reproducible to identify reflux by cuff measurement than by cuff eyeballing (black vs. purple line) ($\chi^2 = 4.1$, df = 1 and $p = 0.042$). Reflux assessment using manual measurement was more reproducible than by manual eyeballing (blue vs. green line) ($\chi^2 = 6.5$, df = 1, $p = 0.011$). Furthermore, cuff eyeballing was more reproducible than manual eyeballing (purple vs. green line) ($\chi^2 = 5.5$, df = 1, $p = 0.019$). Finally, the figure illustrates that assessment by manual measurement by one investigator (RB) may function at a lower level of expertise than for the other investigator (red line with less steep slope).

In addition, we found that when two cuff measurements agreed on "reflux", the manual measurement did not agree

Table 2 Estimates of bias parameters and odds ratios describing observer bias relative to cuff measurement. The bias parameters can be interpreted in terms of odds ratios. Measured by the odds ratio, the chance that reflux is indicated by Manual Measuring is for instance only 100*exp (0.112 − 0.847) = 48% of the chance that reflux is indicated by Cuff Measuring. Fit of the Rasch Model: $p = 0.023$.

Compression	Assessment	Bias α_j	Odds ratio
Manual	Measurement	0.112	0.48
Manual	Eyeballing	−0.983	0.16
Cuff	Measurement	0.847	Reference
Cuff	Eyeballing	0.024	0.44

in 15% of cases, and when two cuff measurements agreed on "no reflux", the manual measurement would not agree in 2.4% of these cases.

Discussion

Our results suggested that it was more accurate to identify reflux by measuring than by "eyeballing" the outcome of duplex ultrasonography. This could be explained by the fact that "eyeballing" classified some veins as having no reflux when measurement revealed that reflux was present, according to the definitions we used. The reason for this finding is not surprising, as reflux time judged by immediate "eyeballing" on screen is more likely to be underestimated with a spectral Doppler curve of normal appearance, even though reflux time is above the given definitions. The study by van Bemmelen et al. from 1989 analysed the duration of reflux in the deep veins of healthy volunteers and found that the median duration of reflux in the PV was 0.19 s, and

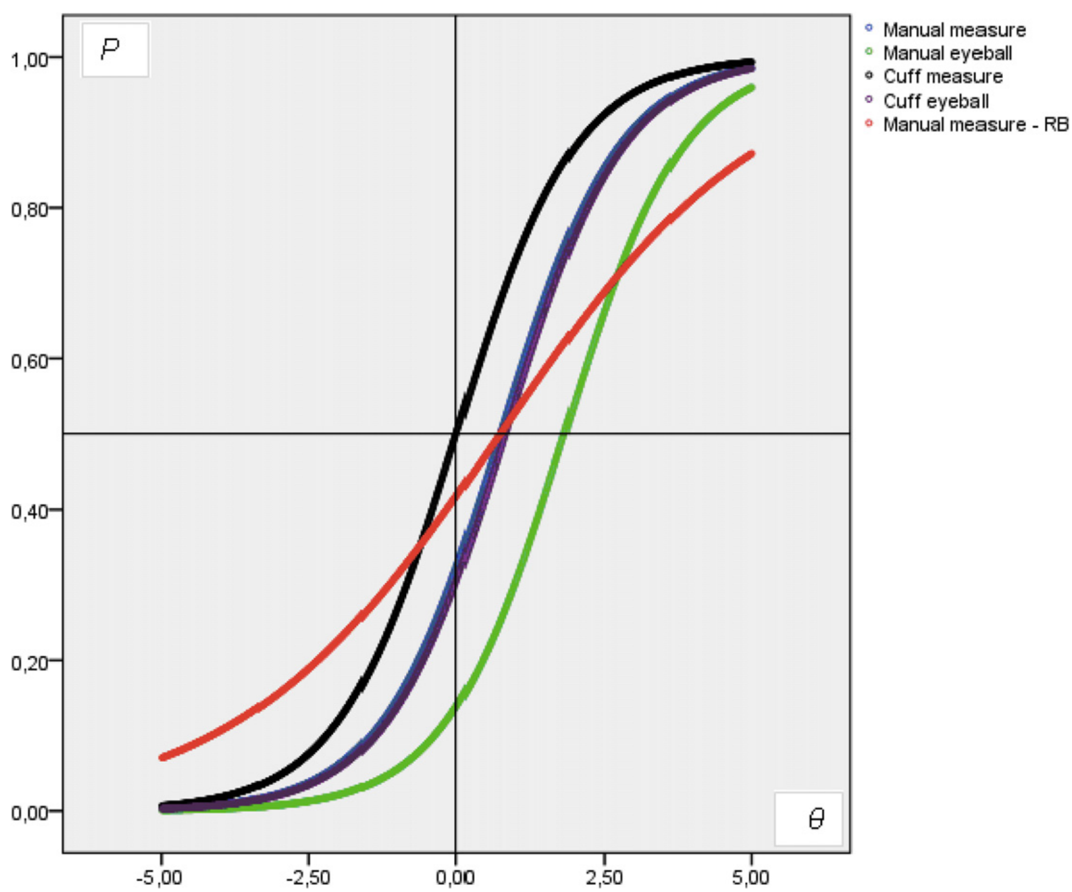

Figure 3 Item characteristic curves of the Rasch Model. Reflux is present if θ_i is positive and not present if θ_i is negative. The closer θ_i is to zero, the more difficult the case. At one end of the curve are the easy cases with reflux and at the other end of the curve are the easy cases without reflux. In the middle area (θ around zero) the curves differ, corresponding to the difficult cases where it has been difficult to identify whether reflux is present or not and where the two investigators disagree. Cuff measurement (black line) is more reproducible in identifying reflux than manual measurement (blue line) and cuff eyeballing (purple line). Manual measurement (blue line) is more reproducible than manual eyeballing (green line) in assessment of reflux. The red line illustrates analysis of manual compression for investigator RB. The curve has a less steep slope (degree of observer expertise, β_j) meaning that manual compression functions at a lower level of expertise ("novice") for investigator RB.

that 95% of the values in the PV was less than 0.66 s.[3] Furthermore, our results showed that it was more reproducible to identify reflux by cuff compression than by manual compression, and that manual measurements functioned at a lower level of expertise for one investigator (RB). The reason for this finding may have a good explanation because of the fact that the female investigator, RB, did not have the same amount of physical strength in her hands compared with the male investigator, NB. Therefore, it may have been more difficult for investigator RB to compress the veins sufficiently with her hands when investigating for venous reflux. Another factor that may have affected the results is the calf circumference. It may be assumed that a large calf circumference makes it more difficult to achieve consistency, and may also explain why assessment by manual compression by investigator RB functioned at a lower level of expertise. This is one factor that is believed to have caused lack of agreement in the peak reflux velocity measured with manual and cuff compression in the study by Yamaki and colleagues.[19]

In the literature, manual and cuff compression have been compared with varying results. Yamaki and colleagues found a similar duration of reflux whether elicited by manual compression or by the cuff deflation method.[19] Araki et al. studied reflux in the PV with both manual and pneumatic cuff compression in the standing and supine position. Overall, they reported no significant difference between the pneumatic cuff and manual compression either in normal limbs or in limbs with chronic venous insufficiency.[20] However, their findings demonstrated that manual compression did require diligent effort to achieve consistency (36% sensitivity with manual compression at the thigh). It was therefore concluded that the pneumatic cuff probably provided better test standardisation than manual compression, but with greater technical experience, manual compression was considered as an alternative.[20] Similarly, van Bemmelen et al. concluded that the duration of popliteal reflux with manual compression in the supine position was so variable that it could not be evaluated.[3] In the study by Markel and colleagues, the cuff method was found to be superior to a Valsalva manoeuvre and limb compression, as 20–25% of cases with reflux in the CFV or PV detected by the cuff method remained undetected by using the supine methods.[21] This corresponds well with our results where cuff compression is more accurate in identifying reflux compared with the manual method, as 15% of cases with reflux identified by cuff measurements was not identified with manual measurements. However, all our examinations have been performed in the standing and not in the supine position.

Kakkos and co-workers recently showed that pneumatic cuff and manual compression were equally effective in inducing and diagnosing reflux. Good agreement between the two methods was found with a Cohen's kappa of 0.68.[10] However, it must be noted that all reflux examinations were performed with the patient lying in a reverse Trendelenburg position. In case of no reflux, the test was repeated in the standing position. The authors found that manual compression required the assistance of an additional experienced person, who could squeeze the leg adequately to achieve substantial forward flow. Because of the large volume of patients being examined in the vascular laboratories every day, the pneumatic compression method was expected to reduce the total costs.[10]

The limitations of our study include difference between investigators in the physical ability to perform sufficient compression of the calf muscle as mentioned earlier. In addition, not all veins were re-canalised completely, and residual thrombus and wall thickening may have influenced the reflux measurements. However, this was not examined any further. There is no "gold standard" in reflux assessment, but, based on the assumption that it is more precise to investigate venous valve function using a method that can be standardised, the cuff method was used as the reference in our study.

In conclusion, we found that reflux assessed by cuff compression was more reproducible than when assessed by manual compression. Immediate "eyeballing" did not reveal reflux in as many cases as exact measurements of the Doppler flow curve did. The fact that assessment by manual compression by one investigator functioned at a lower level of expertise suggests that cuff measurement might be the optimal assessment method, especially in difficult cases.

Conflict of interest

The authors have no conflict of interest.

Funding

None.

Acknowledgements

The authors thank Ulla Walseth and John Leth, vascular technicians at the Department of Vascular Surgery, Gentofte Hospital, for assisting with ultrasound examinations.

References

1 Cavezzi A, Labropoulos N, Partsch H, Ricci S, Caggiati A, Myers K, et al. Duplex ultrasound investigation of the veins in chronic venous disease of the lower limbs—UIP consensus document. Part II. Anatomy. *Eur J Vasc Endovasc Surg* 2006;**31**:288–99.
2 Coleridge-Smith P, Labropoulos N, Partsch H, Myers K, Nicolaides A, Cavezzi A. Duplex ultrasound investigation of the veins in chronic venous disease of the lower limbs—UIP consensus document. Part I. Basic principles. *Eur J Vasc Endovasc Surg* 2006;**31**:83–92.
3 van Bemmelen PS, Bedford G, Beach K, Strandness DE. Quantitative segmental evaluation of venous valvular reflux with duplex ultrasound scanning. *J Vasc Surg* 1989;**10**:425–31.
4 Sarin S, Sommerville K, Farrah J, Scurr JH, Coleridge Smith PD. Duplex ultrasonography for assessment of venous valvular function of the lower limb. *Br J Surg* 1994;**81**:1591–5.
5 Labropoulos N, Tiongson J, Pryor L, Tassiopoulos AK, Kang SS, Ashraf MM, et al. Definition of venous reflux in lower-extremity veins. *J Vasc Surg* 2003;**38**:793–8.
6 van Bemmelen PS, Mattos MA, Hodgson KJ, Barkmeier LD, Ramsey DE, Faught WE, et al. Does air plethysmography correlate with duplex scanning in patients with chronic venous insufficiency? *J Vasc Surg* 1993;**18**:796–807.
7 van Bemmelen PS, Beach K, Bedford G, Strandness Jr DE. The mechanism of venous valve closure. Its relationship to the velocity of reverse flow. *Arch Surg* 1990;**125**:617–9.

8 Haenen JH, van LH, Janssen MC, Wollersheim H, van't hof MA, van Asten WN, et al. Venous duplex scanning of the leg: range, variability and reproducibility. *Clin Sci (Lond)* 1999;**96**:271—7.
9 Masuda EM, Kistner RL, Eklof B. Prospective study of duplex scanning for venous reflux: comparison of Valsalva and pneumatic cuff techniques in the reverse Trendelenburg and standing positions. *J Vasc Surg* 1994;**20**:711—20.
10 Kakkos SK, Lin JC, Sparks J, Telly M, McPharlin M, Reddy DJ. Prospective comparison of the pneumatic cuff and manual compression methods in diagnosing lower extremity venous reflux. *Vasc Endovascular Surg* 2009;**43**:480—4.
11 Porter JM, Moneta GL. Reporting standards in venous disease: an update. International Consensus Committee on Chronic Venous Disease. *J Vasc Surg* 1995;**21**:635—45.
12 Eklof B, Rutherford RB, Bergan JJ, Carpentier PH, Gloviczki P, Kistner RL, et al. Revision of the CEAP classification for chronic venous disorders: consensus statement. *J Vasc Surg* 2004;**40**: 1248—52.
13 Westergaard HB, Langhoff-Roos J, Lingman G, Marsal K, Kreiner S. A critical appraisal of the use of umbilical artery Doppler ultrasound in high-risk pregnancies: use of meta-analyses in evidence-based obstetrics. *Ultrasound Obstet Gynecol* 2001;**17**:466—76.
14 Agresti A. Modelling patterns of agreement and disagreement. *Stat Methods Med Res* 1992;**1**:201—18.
15 Gjorup T, Kelbaek H, Nielsen D, Kreiner S, Godtfredsen J. Interpretation of the electrocardiogram in suspected myocardial infarction: a randomized controlled study of the effect of a training programme to reduce interobserver variation. *J Intern Med* 1992;**231**:407—12.
16 Petersen JH, Larsen K, Kreiner S. *Assessing and quantifying inter-rater variation for dichotomous ratings using a Rasch Model*. Department of Biostatistics, University of Copenhagen; 2007. Report No.: 07/6.
17 Andersen EB. A goodness of fit test for the Rasch model. *Psychometrika* 1973;**38**:123—40.
18 Christensen KB, Kreiner S. Monte Carlo tests of the Rasch model based on scalability coefficients. *Br J Math and Stat Psychol* 2010;**63**:101—11.
19 Yamaki T, Nozaki M, Sakurai H, Takeuchi M, Soejima K, Kono T. Comparison of manual compression release with distal pneumatic cuff maneuver in the ultrasonic evaluation of superficial venous insufficiency. *Eur J Vasc Endovasc Surg* 2006;**32**:462—7.
20 Araki CT, Back TL, Padberg Jr FT, Thompson PN, Duran WN, Hobson RW. Refinements in the ultrasonic detection of popliteal vein reflux. *J Vasc Surg* 1993;**18**:742—8.
21 Markel A, Meissner MH, Manzo RA, Bergelin RO, Strandness Jr DE. A comparison of the cuff deflation method with Valsalva's maneuver and limb compression in detecting venous valvular reflux. *Arch Surg* 1994;**129**:701—5.