ADHERENCE TO GUIDELINES STRONGLY IMPROVES REPRODUCIBILITY OF BRACHIAL ARTERY FLOW-MEDIATED DILATION

Arno Greyling [a, b], Anke C.C.M van Mil [a, c, d], Peter L. Zock [b, c], Daniel J. Green [e], Lorenzo Ghiadoni [f], Dick H. Thijssen [a, g *]

on behalf of the **TIFN International Working Group on Flow Mediated Dilation**

[A)] Department of Physiology, Radboud University Medical Centre, Nijmegen, The Netherlands

[B)] Unilever R&D Vlaardingen, Vlaardingen, The Netherlands

[C)] TI Food and Nutrition, Wageningen, The Netherlands

[D)] Maastricht University Medical Centre, Maastricht, The Netherlands

[E)] School of Sports Science, Exercise and Health, The University of Western Australia, Crawley, Australia

[F)] University of Pisa, Pisa, Italy

[G)] Research institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United Kingdom

Short title: Reproducibility of flow mediated dilation

Word count text: 3504

Word count abstract: 244

Figures: 2

Tables: 4

***Author for correspondence:**

Prof. Dick H.J. Thijssen, Department of Physiology, Radboud University Medical Center, Nijmegen, the Netherlands. Tel: +31243614222, Email: Dick.Thijssen@radboudumc.nl

ABSTRACT

Background: Brachial artery FMD is widely used as a non-invasive measure of endothelial function. Adherence to expert guidelines is believed to be of vital importance to obtain reproducible measurements. We conducted a systematic review of studies reporting on the reproducibility of the FMD in order to determine the relation between adherence to current expert guidelines for FMD measurement and its reproducibility.

Methods: Medline-database was searched through July 2015 and 458 records were screened for FMD reproducibility studies reporting the mean difference and variance of repeated FMD measurements. An adherence score was assigned to each of the included studies based on reported adherence to published guidelines on the assessment of brachial artery FMD. A Typical Error Estimate (TEE) of the FMD was calculated for each included study. The relation between the FMD TEE and the adherence score was investigated by means of Pearson correlation coefficients and multiple linear regression analysis.

Results: Twenty-seven studies involving 48 study groups and 1,537 subjects were included in the analyses. The adherence score ranged from 2.4 to 9.2 (out of a maximum of 10) and was strongly and inversely correlated with FMD TEE (adjusted R^2=0.36, P<0.01). Use of automated edge-detection software, continuous diameter measurement, true peak diameter for %FMD calculation, a stereostatic probe holder, and higher age emerged as factors associated with a lower FMD TEE.

Conclusions: These data demonstrate that adherence to current expert consensus guidelines and applying contemporary techniques for measuring brachial artery FMD decreases its measurement error.

Keywords: cardiovascular disease; atherosclerosis; endothelial function; reproducibility; methodology

INTRODUCTION

The endothelium is a key regulator of vascular homeostasis and endothelial dysfunction is an early manifestation of atherosclerosis [1]. Currently, the most widely used technique to study endothelial function *in vivo* is the flow-mediated dilation (FMD) of the brachial artery. This is a non-invasive, ultrasound-based method which correlates with endothelial function of the coronary arteries [2, 3] and independently predicts cardiovascular disease (CVD) [4, 5]. The technique is attractive as a surrogate end-point, especially since changes in FMD can be detected across a relatively short timeframe [6]. Despite its popularity, minor changes in the methodological approach may critically impact variability and decrease reproducibility of the FMD response [7-9].

Previous expert consensus guidelines have made important contributions to standardize the technical approach and to set minimum standard requirements for FMD measurements [10, 11]. However, not all studies on FMD apply these recommendations, or only in part. The impact of adherence to these guidelines on the reproducibility of FMD measurements is currently unclear, but may importantly contribute to the measurement error of the FMD technique Furthermore, little is known about the relative importance of the individual aspects of the expert-consensus guidelines to contribute to the reproducibility of the FMD. Better quantitative data on this matter can help reduce variation within and between studies, which will increase the statistical power of studies on FMD to detect changes and, subsequently, decrease chances for type II errors.

In light of these considerations, we hypothesized that adherence to expert consensus guidelines is related to better reproducibility of FMD measurements [10, 11]. Therefore, we performed a systematic search for published studies that reported data on reproducibility of FMD measurements, and investigated the relation between (full or partial) adherence to current expert

consensus guidelines and reproducibility of the FMD. Secondly, we explored which subject- and methodology-related factors were related to FMD reproducibility.

METHODS

Search Strategy

The MEDLINE bibliographic database was searched (January 2000 through July 2015) for studies that assessed the reproducibility of the FMD using the following search terms: *"flow mediated dilation", "flow mediated dilatation", "flow mediated vasodilation", "flow mediated vasodilatation", "endothelial function", "endothelial dysfunction", "FMD", "FMV", "brachial artery", "reproducibility", "reliability", "repeatability", "coefficient of variation", "CV", and "variance"*. The search was limited to studies in human adults published in the English language. Additionally, we supplemented the search by hand-searching references of included studies and relevant reviews and meta-analyses on this topic.

Selection of Studies

Included studies were identified by means of a two-step selection process. During the first step, two reviewers (ACCMvM, AG) independently screened titles, abstracts and keywords of publications to identify potentially eligible studies. Studies were included if the mean difference and variance of repeated FMD measurements of the brachial artery were reported. During step 2 of the selection, both reviewers examined the full text of these publications to gauge eligibility based on two additional inclusion criteria: FMD was determined through noninvasive ultrasound imaging, and a reactive hyperaemia protocol (with an ischemia duration of 4 to 5 minutes) was used to elicit the shear stress stimulus required for FMD. Thus, studies that adopted (ischemic) hand-grip exercise, passive movement and/or skin warming protocols to elicit (brachial) artery

dilation were not included in our analysis. In cases of discrepancy between the reviewers, eligibility was discussed along with a third reviewer (DHJT) until consensus was reached.

Data extraction

Study and subject characteristics: A standardized data collection sheet was used to extract general publication details (author, year of publication, country) and specific study- and subject characteristics: number of subjects, mean age (in years); CVD risk status of the study population (defined as presence of diagnosed CVD, hypertension or diabetes); baseline brachial artery diameter (in mm); % brachial artery FMD and its associated variance for each repeated measurement; and the mean absolute difference between repeated FMD measurements and its associated variance.

Adherence to guidelines: We extracted information from the methods sections of the individual papers to assess the adherence to current expert-consensus guidelines. Based on recent guidelines [11], we scored each individual study on the reporting of 19 different factors which were divided over 4 categories. The categories were related to: 1. Subject preparation (10 items), 2. Image acquisition (4 items), 3. Data analysis (3 items), and 4. Laboratory (2 items). Before performing the systematic literature search, values were assigned to each factor proportional to its perceived importance for valid assessment of the FMD. This was done through expert consensus discussion within the Working Group (AG, LG and DHJT) (see online data supplement). The *"Adherence Score"* that could be assigned to a study ranged from 0 to 10 points, depending on how many of the 19 different factors that were reported. In addition, we counted the number of previous studies on FMD published by the principal author of each study included in the systematic review. This

number served as a measure of the perceived experience in FMD measurements for each centre at the time of publication of the reproducibility data included here.

Statistical analysis

Reported measures of FMD reproducibility varied between studies. Many studies presented the coefficient of variation (CV) of repeated measurements, although this measure was calculated in a number of different ways, precluding direct comparisons. Measures of reproducibility included the technical error of the measurement (TEM), Pearson- and intraclass correlation coefficients (ICC), and limits of agreement. In order to make valid comparisons between studies, we defined as primary outcome measure the typical error of estimate (TEE) of FMD ,which is calculated as standard deviation of the paired differences/$\sqrt{2}$ [12].

Data are presented as mean ± SD or median (range) as appropriate for continuous variables and as frequencies for categorical variables. FMD TEE data were highly skewed (Shapiro-Wilk test, $P<0.0001$) and were log transformed prior to the analyses. Relations between log-FMD TEE and continuous variables were determined by Pearson correlation coefficients analysis. For categorical variables, the statistical significance of differences in FMD TEE between different levels were assessed by the Mann-Whitney *U* test. Significant correlates were entered in a multivariate linear regression analyses with backward elimination to identify independent predictors of FMD TEE. All analyses were conducted using JMP version 11.0 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Our systematic search identified 446 potentially relevant publications and an additional 12 were obtained through review of references of included studies, relevant reviews and meta-analyses. Twenty-seven studies [13-39] with 48 relevant study groups met our inclusion criteria and were included in our analysis (Figure 1). Characteristics of the included study groups are presented in Table 1. The 48 study groups comprised a total of 1,537 subjects (mean sample size 32; range, 8-135) with a mean age of 41.5 years (range, 22-79 years). Eleven study groups included subjects with increased CVD risk, i.e. presence of diagnosed CVD, hypertension or diabetes. The other remaining 37 study groups consisted of healthy subjects. The time between repeated FMD measurements ranged from 25 minutes to 9 months. Mean baseline brachial diameter was 3.9 mm (range, 3.5 to 4.7 mm) and mean baseline FMD (i.e. this first of the two repeated measurements) was 7.1% (range, 1.8 to 19.9%). The FMD TEE ranged from 0.33 to 4.83% across study groups, with a mean value of 1.4%. The level of experience for each centre at the time of publication of the reproducibility study in question varied widely (number of previously published studies on FMD ranging from 0 to 71, median of 3).

Methodology-related factors *versus* variation in FMD

There was considerable variation in the methodological factors between studies. Adherence scores ranged from 2.4 to 9.2, with a mean of 5.3 (out of a maximum of 10). The adherence score was inversely correlated with log FMD TEE (adjusted R^2=0.36, P<0.01, Figure 2).

To explore the impact of the different aspects of the adherence score on the FMD TEE, we compared the FMD TEE between adherence (Yes *vs* No) to various methodological variables. Statistically significant differences in FMD TEE were found for use of the true peak diameter to

calculate %FMD, continuous brachial artery diameter measurement over the cardiac cycle, use of automated edge detection software and smoking cessation prior to measurements (Table 2).

Subject-related factors *versus* variation in FMD

For the remaining methodology related factors and subject characteristics, there were weak, but statistically significant correlations of log FMD TEE with age (adjusted R^2= -0.18, P<0.01) and with baseline FMD (adjusted R^2 0.11, P=0.013). In addition, FMD TEE was significantly smaller in the subgroup of studies that applied a stereostatic probe holder, and in studies performed by groups with more experience according to number of earlier publications on FMD. The %FMD TEE of studies above and below the median duration between repeated measurements (7 days) was not significantly different (Table 3), and there was no correlation between %FMD TEE and the time between repeated measurements (adjusted R^2= -0.02, P<0.75).

We constructed a stepwise multivariate regression model with log FMD TEE as the dependent variable and all factors that significantly influenced FMD TEE based on the individual analyses (adherence score, age, baseline FMD, probe holder and previous experience). The stepwise multivariate regression model predicted 51% of the variability in log FMD TEE. Adherence score (β= -0.16), age (β = -0.01) and probe holder (β = -0.19) remained as statistically significant (P<0.05) predictors in the model (Table 4).

DISCUSSION

Measurement of the FMD of the brachial artery has obtained in the recent years a well-established predictive capacity for future CVD events. Despite this and its relatively

straightforward and non-invasive approach, the clinical use of the FMD is hampered by its sensitivity to variations in methodology.

Our systematic analysis of previous studies that explored the FMD reproducibility provides us with a number of novel observations. First, we found considerable variation in the methodology applied to measure FMD and consequently, differences between studies in the adherence to current expert consensus guidelines. Secondly, these data show a robust inverse association between adherence to the guidelines and FMD reproducibility, with higher adherence to guidelines being related to smaller variation in FMD. Thirdly, we identified methodological factors that were associated with smaller variation in FMD. Specifically, the use of automated edge detection software, continuous measurement of brachial artery diameter over the cardiac cycle, calculating %FMD by means of the true peak diameter and use of a stereostatic probe holder were related to a better reproducibility. Taken together, our study provides strong scientific data that highlight the importance of rigorous application of standardized contemporary methodology to reduce measurement error of the FMD and, consequently, improve its use in (pre)clinical studies.

To our knowledge, no previous study has explored the (relative) importance of adherence to expert consensus guidelines for measures of vascular health, including frequently used techniques like intima-media thickness, pulse wave velocity, and finger photoplethysmography. Taking all studies on the reproducibility of the FMD together, involving 1,537 subjects, we found a TEE of 1.4% based on an average FMD of 7.1%. This indicates an overall good-to-acceptable reproducibility of the FMD. However, significant variation was observed between studies, with adherence to the expert consensus guidelines representing an important determinant of this

variation. Our data suggests that roughly 36% of the variation in FMD reproducibility can be explained through adherence to the guidelines alone. The presence of a linear relation between adherence to the guidelines and variation of the FMD suggests that measurement error would be further reduced with stricter adherence to the guidelines. Our data also indicate that even with full adherence to current expert consensus guidelines, some level of measurement error remains present. Nonetheless, a significant amount of variation in the FMD can be prevented by strong adherence to guidelines.

Our analysis provides further insight into methodological factors that determine within-person error of the FMD measurement. For example, we found that taking the true peak artery diameter (rather than a fixed time point), continuous diameter measurement and automated edge-detection contribute to minimizing measurement error. The importance of these methodological factors have already been acknowledged in previous work. For example, Black *et al.* found that the peak diameter following cuff release differs between young and older subjects [7]. Consequently, calculating the FMD% at an arbitrary time point (e.g. 60 seconds) may lead to misleading conclusions compared to a more sophistical approach in which diameter of the brachial artery is recorded continuously, allowing for the detection of the true peak dilation. Furthermore, previous work demonstrated that the adoption of edge-detection software to perform (observer-independent) analysis leads to smaller variation compared to the application of manual calipers (a technique highly sensitive for measurement bias) [35, 40, 41]. Whilst these studies highlight the importance of considering these factors for valid use of FMD, the present study highlights the importance of considering these factors to lower variation. Therefore, our study provides an additional rationale to perform continuous assessment of the diameter and the adoption of edge-detection software when performing valid and reproducible assessments of the FMD.

Another important observation in our study was that previous experience of a laboratory with the FMD resulted in a smaller variation in FMD. A potential explanation for this finding is that experienced laboratories are more likely to demonstrate better adherence to the expert consensus guidelines. Indeed, when all factors were included in the final regression analysis (including adherence to the guidelines), previous experience of a laboratory with FMD did not emerge as an independent predictor of FMD reproducibility. Another factor that contributed to a smaller variation of the FMD was the use of a probe holder. The use of such devices is largely dependent on the personal preference of the laboratory and the effect on measurement reproducibility is a complex topic, since highly skilled operators with years of experience are able to conduct FMD measurements with exceptional reproducibility, regardless of the use of a probe holder [18]. One may speculate that sonographers' learning curves will likely differ depending on whether a probe holder is used or not and also depending on the design and construction of the probe holder itself. Therefore, despite the significant inverse association in our analysis, it remains difficult to ascertain whether use of a probe holder leads to a smaller variation in FMD *per se*. Further studies are needed to confirm the importance of using a probe holder to reduce variability of the FMD.

Of the subject-related factors (age, diameter and baseline FMD), only age contributed independently to the variation in FMD. Notably, higher age of subjects was associated with a smaller variation in FMD. Older age is typically associated with a lower FMD [42, 43], which may contribute to a smaller (biological) variation and/or less ability to change in response to hemodynamic stimuli, consequently leading to a smaller measurement error. However, previous work suggests the presence of larger variability for measurements of vascular health in clinical

groups. For example, Craiem *et al.* found that subjects with CVD, despite comparable baseline FMD% values, demonstrate a larger coefficient of variation compared to healthy controls [21]. At least, our data suggest that the reproducibility of the FMD may differ between (clinical) groups.

Interestingly, the time duration between repeated measurements did not significantly affect FMD reproducibility in our analyses. This might seem counterintuitive as poorer reproducibility is expected as the time duration between repeated measurements increases. Indeed, a recent study specifically designed to determine FMD reproducibility over short (48 hours), medium (3 months) and long (9 months) time frames did find poorer reproducibility at 9 months between repeated measurements [14]. Reproducibility was comparable for the shorter time periods however, which is in agreement with a recent Italian multicenter study which found no differences in FMD reproducibility up to 30 days between measurements [15]. It should be noted that there was a large heterogeneity in time between measurements in the included study groups, with the majority ranging between one and 15 days (n = 32) and some up to 30 (n=11), 90 (n = 4) and 270 days (n = 1). Excluding these last 16 studies from the analyses did not appreciably change our findings our findings however (data not shown).

Limitations. An obvious limitation of our systematic review is that the degree of adherence to expert consensus guidelines was assessed from information as provided in the papers. If a methodological description omitted one or more of the 19 different scoring factors, no points were assigned for those factors. As a consequence some studies with sparse methodological descriptions received lower scores. Inconsequent reporting of methodological details might therefore have confounded our outcomes. It should also be acknowledged that our estimation of the experience of a laboratory with FMD measurements does not necessarily reflect the

283 experience of an individual sonographer. However, a laboratory more experienced in performing
284 FMD measurements will generally require a level of skill and training for their sonographers that
285 will meet at least the standard of their previous work. This highlights the importance of the level
286 of experience in performing studies with FMD as an outcome variable. Another limitation is that
287 our analysis on the relative importance of individual subject- and/or methodology-related factors
288 could only be based on a between-study comparison of factors contributing to the reproducibility
289 of the FMD. Various other factors may have influenced this analysis. Therefore, future studies
290 are necessary to further explore the importance of (some of) the methodology-related factors,
291 including the effects factors which we could not examine with the current dataset such as the
292 observer/analyst, the time of cuff occlusion and changes in baseline brachial artery diameter.
293
294 In conclusion, this systematic review shows that adherence to current expert consensus guidelines
295 significantly reduces measurement error when assessing brachial artery FMD in humans.
296 Moreover, when adopting the guidelines, we found that the use of contemporary techniques (i.e.
297 continuous diameter recording, edge-detection and wall-tracking software and possibly also the
298 use of a probe holder) is crucial to improve reproducibility of the FMD measurement.
299 Considering these factors will importantly decrease measurement error of the FMD and,
300 consequently, decrease chances for type II errors in studies that rely on FMD as their primary
301 outcome parameter. In other words, ignoring current expert-consensus guidelines causes
302 significant variability of the FMD and, consequently, may lead to spurious conclusions. This
303 study delivers important insight that should be taken into account when developing future updates
304 to expert-consensus guidelines.
305

SOURCES OF FUNDING:

Ms Anke van Mil is financially supported by a Top Institute for Food and Nutrition-grant.

Dr. Dick Thijssen is financially supported by the Netherlands Heart Foundation (E Dekker-stipend, 2009T064). Professor Green receives Fellowship and grant funding from the National Heart Foundation of Australia (APP1045204).

DISCLOSURES:

AG and PLZ are employed by Unilever R&D Vlaardingen B.V. No conflicts of interest, financial or otherwise, are declared by the remaining authors.

316 **REFERENCES:**

317

318 [1] Deanfield, JE, Halcox, JP and Rabelink, TJ, Endothelial function and dysfunction: testing
319 and clinical relevance, Circulation, 2007;115:1285-1295.

320 [2] Celermajer, DS, Sorensen, KE, Gooch, VM, et al., Non-invasive detection of endothelial
321 dysfunction in children and adults at risk of atherosclerosis, Lancet, 1992;340:1111-1115.

322 [3] Takase, B, Uehata, A, Akima, T, et al., Endothelium-dependent flow-mediated
323 vasodilation in coronary and brachial arteries in suspected coronary artery disease, Am J Cardiol,
324 1998;82:1535-1539, A1537-1538.

325 [4] Ras, RT, Streppel, MT, Draijer, R, et al., Flow-mediated dilation and cardiovascular risk
326 prediction: a systematic review with meta-analysis, Int J Cardiol, 2013;168:344-351.

327 [5] Inaba, Y, Chen, JA and Bergmann, SR, Prediction of future cardiovascular outcomes by
328 flow-mediated vasodilatation of brachial artery: a meta-analysis, Int J Cardiovasc Imaging,
329 2010;26:631-640.

330 [6] Bianchini, E, Giannarelli, C, Bruno, RM, et al., Functional and structural alterations of
331 large arteries: methodological issues, Curr Pharm Des, 2013;19:2390-2400.

332 [7] Black, MA, Cable, NT, Thijssen, DH, et al., Importance of measuring the time course of
333 flow-mediated dilatation in humans, Hypertension, 2008;51:203-210.

334 [8] Doshi, SN, Naka, KK, Payne, N, et al., Flow-mediated dilatation following wrist and
335 upper arm occlusion in humans: the contribution of nitric oxide, Clin Sci (Lond), 2001;101:629-
336 635.

337 [9] Mullen, MJ, Kharbanda, RK, Cross, J, et al., Heterogenous nature of flow-mediated
338 dilatation in human conduit arteries in vivo: relevance to endothelial dysfunction in
339 hypercholesterolemia, Circ Res, 2001;88:145-151.

[10] Corretti, MC, Anderson, TJ, Benjamin, EJ, et al., Guidelines for the ultrasound assessment of endothelial-dependent flow-mediated vasodilation of the brachial artery: a report of the International Brachial Artery Reactivity Task Force, J Am Coll Cardiol, 2002;39:257-265.

[11] Thijssen, DH, Black, MA, Pyke, KE, et al., Assessment of flow-mediated dilation in humans: a methodological and physiological guideline, Am J Physiol Heart Circ Physiol, 2011;300:H2-12.

[12] Hopkins, WG, Measures of reliability in sports medicine and science, Sports Medicine, 2000;30:1-15.

[13] Kanahara, M, Harada, H, Katoh, A, et al., New methodological approach to improve reproducibility of brachial artery flow-mediated dilatation, Echocardiography, 2014;31:197-202.

[14] Charakida, M, de Groot, E, Loukogeorgakis, SP, et al., Variability and reproducibility of flow-mediated dilatation in a multicentre clinical trial, Eur Heart J, 2013;34:3501-3507.

[15] Ghiadoni, L, Faita, F, Salvetti, M, et al., Assessment of flow-mediated dilation reproducibility: a nationwide multicenter study, J Hypertens, 2012;30:1399-1405.

[16] Onkelinx, S, Cornelissen, V, Goetschalckx, K, et al., Reproducibility of different methods to measure the endothelial function, Vasc Med, 2012;17:79-84.

[17] Lima, JC, Martins, WP, Nastri, CO, et al., Pulsatility index change of brachial artery shows better reproducibility than flow-mediated vasodilation, Ultrasound Med Biol, 2010;36:2036-2041.

[18] Thijssen, DH, Dawson, EA, Tinken, TM, et al., Retrograde flow and shear rate acutely impair endothelial function in humans, Hypertension, 2009;53:986-992.

[19] Donald, AE, Halcox, JP, Charakida, M, et al., Methodological approaches to optimize reproducibility and power in clinical studies of flow-mediated dilation, J Am Coll Cardiol, 2008;51:1959-1964.

[20] Simova, I, Nossikoff, A and Denchev, S, Interobserver and intraobserver variability of flow-mediated vasodilatation of the brachial artery, Echocardiography, 2008;25:77-83.

[21] Craiem, D, Chironi, G, Gariepy, J, et al., New monitoring software for larger clinical application of brachial artery flow-mediated vasodilatation measurements, J Hypertens, 2007;25:133-140.

[22] Harris, RA, Padilla, J, Hanlon, KP, et al., Reproducibility of the flow-mediated dilation response to acute exercise in overweight men, Ultrasound Med Biol, 2007;33:1579-1585.

[23] Meirelles Cde, M, Leite, SP, Montenegro, CA, et al., Reliability of brachial artery flow-mediated dilatation measurement using ultrasound, Arq Bras Cardiol, 2007;89:160-167, 176-183.

[24] Donald, AE, Charakida, M, Cole, TJ, et al., Non-invasive assessment of endothelial function: which technique?, J Am Coll Cardiol, 2006;48:1846-1850.

[25] Harris, RA, Padilla, J, Rink, LD, et al., Variability of flow-mediated dilation measurements with repetitive reactive hyperemia, Vasc Med, 2006;11:1-6.

[26] Leeson, CP, Robinson, M, Francis, JM, et al., Cardiovascular magnetic resonance imaging for non-invasive assessment of vascular function: validation against ultrasound, J Cardiovasc Magn Reson, 2006;8:381-387.

[27] Elsen, BM, Scholten, HJ, Schilder, JCM, et al., Reproducibility of B-mode ultrasound brachial Flow Mediated Dilation measurements, Atherosclerosis Supplements, 2005;6:126-126.

[28] Sejda, T, Pit'ha, J, Svandova, E, et al., Limitations of non-invasive endothelial function assessment by brachial artery flow-mediated dilatation, Clin Physiol Funct Imaging, 2005;25:58-61.

[29] Stoner, L, Sabatier, M, Edge, K, et al., Relationship between blood velocity and conduit artery diameter and the effects of smoking on vascular responsiveness, J Appl Physiol (1985), 2004;96:2139-2145.

388 [30] West, SG, Wagner, P, Schoemer, SL, et al., Biological correlates of day-to-day variation
389 in flow-mediated dilation in individuals with Type 2 diabetes: a study of test-retest reliability,
390 Diabetologia, 2004;47:1625-1631.
391 [31] Sidhu, JS, Newey, VR, Nassiri, DK, et al., A rapid and reproducible on line automated
392 technique to determine endothelial function, Heart, 2002;88:289-292.
393 [32] Beux, F, Carmassi, S, Salvetti, MV, et al., Automatic evaluation of arterial diameter
394 variation from vascular echographic images, Ultrasound Med Biol, 2001;27:1621-1629.
395 [33] de Roos, NM, Bots, ML, Siebelink, E, et al., Flow-mediated vasodilation is not impaired
396 when HDL-cholesterol is lowered by substituting carbohydrates for monounsaturated fat, Br J
397 Nutr, 2001;86:181-188.
398 [34] Herrington, DM, Fan, L, Drum, M, et al., Brachial flow-mediated vasodilator responses in
399 population-based research: methods, reproducibility and effects of age, gender and baseline
400 diameter, J Cardiovasc Risk, 2001;8:319-328.
401 [35] Woodman, RJ, Playford, DA, Watts, GF, et al., Improved analysis of brachial artery
402 ultrasound using a novel edge-detection software system, J Appl Physiol, 2001;91:929-937.
403 [36] Lind, L, Hall, J, Larsson, A, et al., Evaluation of endothelium-dependent vasodilation in
404 the human peripheral circulation, Clin Physiol, 2000;20:440-448.
405 [37] Preik, M, Lauer, T, Heiss, C, et al., Automated ultrasonic measurement of human arteries
406 for the determination of endothelial function, Ultraschall Med, 2000;21:195-198.
407 [38] Liang, YL, Teede, H, Kotsopoulos, D, et al., Non-invasive measurements of arterial
408 structure and function: repeatability, interrelationships and trial sample size, Clin Sci (Lond),
409 1998;95:669-679.

[39] Hardie, KL, Kinlay, S, Hardy, DB, et al., Reproducibility of brachial ultrasonography and flow-mediated dilatation (FMD) for assessing endothelial function, Aust N Z J Med, 1997;27:649-652.

[40] Gemignani, V, Faita, F, Ghiadoni, L, et al., A system for real-time measurement of the brachial artery diameter in B-mode ultrasound images, IEEE Trans Med Imaging, 2007;26:393-404.

[41] Sonka, M, Liang, W and Lauer, RM, Automated analysis of brachial ultrasound image sequences: early detection of cardiovascular disease via surrogates of endothelial function, IEEE Trans Med Imaging, 2002;21:1271-1279.

[42] Black, MA, Cable, NT, Thijssen, DH, et al., Impact of age, sex, and exercise on brachial artery flow-mediated dilatation, Am J Physiol Heart Circ Physiol, 2009;297:H1109-1116.

[43] Thijssen, DH, Carter, SE and Green, DJ, Arterial structure and function in vascular ageing: are you as old as your arteries?, J Physiol, 2015.

424 **FIGURE LEGENDS:**

425 **Figure 1.** Flow diagram of the study selection procedure.

426

427 **Figure 2.** Linear correlation between the Typical Error of the Flow Mediated Dilation Estimate

428 (FMD TEE) and adherence to expert guidelines (Adherence Score) in 27 studies (involving 48

429 study groups) of FMD reproducibility.

430

431

TABLES:

Table 1. General characteristics of the FMD reproducibility studies included in the systematic review.

Source	Health status	Number of subjects	Mean age (years)	Mean baseline FMD (%)	Mean baseline diameter (mm)	Time between measurements (days)	TEE
Kanahara 2014 [13]	Healthy	32	40	7.90	3.83	14	1.28
Charakida 2013 [14]		67	61	4.10	4.55	2	0.94
Charakida 2013	CVD, Diabetes	67	61	4.10	4.60	90	1.04
Charakida 2013		67	61	4.10	4.65	270	1.47
Ghiadoni 2012 [15]	Healthy	135	32	6.52	3.53	1 hour	0.83
Ghiadoni 2012		135	32	6.52	3.55	30	1.15
Onkelinx 2012 [16]	CVD	18	68	6.80	3.92	0.5 hour	0.94
Onkelinx 2012		18	68	7.13	3.91	2	0.88
Lima 2010 [17]	Healthy	31	25	13.17	3.57	2	2.91
Thijssen 2009 [18]	Healthy	10	24	6.83	4.28	0.5 hour	0.89
Donald 2008 (true peak diameter) [19]	Healthy	32	43	8.10	3.70	6 hours	0.79

Source	Health status	Number of subjects	Mean age (years)	Mean baseline FMD (%)	Mean baseline diameter (mm)	Time between measurements (days)	TEE
Donald 2008 (true peak diameter)		34	43	7.50	3.70	7	0.79
Donald 2008 (true peak diameter)		37	43	8.10	3.75	30	0.53
Donald 2008 (true peak diameter)		35	43	7.80	3.80	90	0.74
Donald 2008 (60 sec)		32	43	7.30	3.70	6 hours	1.08
Donald 2008 (60 sec)		34	43	6.70	3.70	7	0.95
Donald 2008 (60 sec)		37	43	7.50	3.75	30	0.63
Donald 2008 (60 sec)		35	43	7.10	3.80	90	0.87
Simova 2008 [20]	CVD, Hypertension	40	62	6.05	3.84	0.25 hour	0.85
Craiem 2007 [21]	Healthy	10	32	7.60	3.95	1 hour	0.80
Craiem 2007		10	32	8.10	3.89	7	0.91

Source	Health status	Number of subjects	Mean age (years)	Mean baseline FMD (%)	Mean baseline diameter (mm)	Time between measurements (days)	TEE
Craiem 2007	CVD	26	44	6.98	3.97	1 hour	1.34
Craiem 2007		26	44	5.66	4.15	30	0.96
Harris 2007 [22]	Healthy	9	57	7.80	4.11	2	1.32
Meirelles 2007 [23]	Healthy	10	33	19.90	3.50	1.5 hours	2.70
Meirelles 2007		13	33	16.50	3.55	3	2.50
Donald 2006 [24]	Healthy	16	28	7.30	3.55	1	1.63
Harris 2006 [25]	Healthy	16	23	9.88	3.74	2 hours	0.71
Leeson 2006 [26]	Healthy	17	32	4.74	4.05	20	1.22
Elsen 2005 [27]	Healthy	15	23	4.61	4.04	1	0.63
Sejda 2005 [28]	Healthy	18	28	5.95	4.04	7	3.89
Sejda 2005		18	28	4.23	4.15	7	1.63
Stoner 2004 [29]	Healthy	9	23	10.20	3.90	2	3.26
West 2004 [30]	Diabetes	18	55	5.57	4.01	7	0.81
West 2004		18	55	5.57	4.01	14	1.07

Source	Health status	Number of subjects	Mean age (years)	Mean baseline FMD (%)	Mean baseline diameter (mm)	Time between measurements (days)	TEE
Sidhu 2002 [31]	Healthy	12	36	5.38	3.94	20	0.37
Sidhu 2002	CVD	12	62	1.80	4.29	20	0.33
Beux 2001 [32]	Healthy	38	44	6.62	4.41	1 hour	1.97
Beux 2001		38	44	4.32	4.41	1 hour	1.22
De Roos 2001 [33]	Healthy	34	27	4.13	3.90	25	2.01
Herrington 2001 [34]	Healthy	127	79	2.63	4.53	7	0.79
Herrington 2001		30	45	7.87	4.35	7	1.46
Woodman 2001 [35]	Healthy	24	55	6.60	4.06	7	0.71
Lind 2000 [36]	Healthy	10	22	7.40	3.55	2 hours	2.19
Lind 2000		10	22	7.40	3.55	21	2.82
Preik 2000 [37]	Healthy	8	28	10.60	3.62	20	1.06
Liang 1998 [38]	Healthy	30	44	10.80	3.84	18	2.01
Hardie 1997 [39]	Healthy	19	36	3.00	3.78	90	4.83

434

435 **Table 2.** Relationship of individual components of the adherence score with FMD TEE

Adherence Score Characteristic	Median (IQR) %FMD TEE						
Subject preparation	**n**		**No**	**n**		**Yes**	**p**
Fasting state (>6h)	21	1.08	(0.83-2.10)	27	0.96	(0.80-1.47)	0.38
No smoking/tobacco consumption prior to measurement (>6h)	22	0.89	(0.73-1.22)	26	1.30	(0.89-2.55)	**<0.01**
No habitual exercise prior to measurement (>48h)	31	1.22	(0.89-1.97)	17	0.87	(0.72-1.36)	0.07
No food/beverages that contain alcohol and/or caffeine for >12 h	31	1.06	(0.79-1.97)	17	1.04	(0.82-1.40)	0.6
No polyphenol-rich food/beverages (cocoa, tea, fruit juices) for >18 h	45	1.04	(0.79-1.63)	3	1.15	(0.83-2.91)	0.6
No vitamins for at least 72h	44	1.05	(0.80-1.63)	4	0.99	(0.68-2.47)	0.8
Vasoactive medications withheld/noted on the morning of the study	26	1.01	(0.79-1.98)	22	1.06	(0.84-1.51)	0.8
Supine position; ≥15 min rest in a quiet, temperature controlled room	30	1.01	(0.80-1.72)	18	1.10	(0.82-1.60)	1.0
Repeated measurements standardised to timing of the menstrual cycle	36	1.01	(0.79-1.89)	12	1.06	(0.84-1.59)	0.7
Repeated measurements done in fixed time windows (same time of day)	7	1.22	(0.94-1.47)	41	0.96	(0.79-1.80)	0.5
Image acquisition							
Diameter measurements recorded continuously over the cardiac cycle	35	1.22	(0.85-2.01)	13	0.88	(0.75-0.95)	**<0.01**
Diameter measurements obtained during end diastole only	15	0.89	(0.79-1.15)	33	1.22	(0.83-2.01)	0.06

Adherence Score Characteristic		**Median (IQR) %FMD**			**TEE**		
Simultaneous acquisition of pulse-wave Doppler velocity signal for quantification of shear stimulus	20	1.40	(0.86-2.15)	28	0.94	(0.79-1.21)	0.05
Image analysis							
Analysis using automated edge detection and wall tracking software	13	2.19	(1.47-2.87)	35	0.91	(0.79-1.22)	**<0.01**
FMD calculation point (true peak diameter)	17	1.63	(0.94-2.76)	31	0.91	(0.79-1.28)	**<0.01**
Lab data							
Use of experienced sonographers reported	20	1.09	(0.73-2.38)	28	1.05	(0.82-1.47)	0.7
Same sonographers paired to same subjects for repeated measurements	8	1.10	(0.86-1.44)	40	1.01	(0.79-1.89)	0.8

436

Table 3. Relationship of subject- and methodology-related characteristics with FMD TEE

Continuous variables	**Adjusted Pearson R^2**	**P-values**
Age (years)	-0.18	**<0.01**
Baseline FMD (%)‡	0.11	**0.01**
Baseline diameter (mm)	-0.02	0.15
Number of subjects (n)	-0.001	0.33

Categorical variables	**Median (IQR) %FMD TTE**					
	n	**No**		**n**	**Yes**	**P-values**
CVD risk	37	1.15	(0.79-2.01)	11	0.94 (0.85-1.07)	0.31
Distal occlusion cuff placement	5	2.01	(0.91-2.6)	43	1.04 (0.79-1.47)	0.17
Stereostatic probe holder	18	1.82	(1.02-2.85)	30	0.92 (0.73-1.22)	**<0.01**
Experienced centre*	23	1.32	(0.88-2.5)	25	0.91 (0.80-1.19)	**0.01**
Time between repeated measurements above median†	18	0.94	(0.81-1.72)	30	1.06 (0.71-1.61)	0.77

‡ Baseline FMD refers to the first of the two repeated measurements

*Centre experience was defined as the number of previous studies on FMD published by the principle author of each included study. The effect of centre experience was examined by comparing the %FMD TEE of studies below (No) and above (yes) the median number of previously published FMD studies.

443 †The effect of the time duration between studies was examined by comparing the %FMD TEE of

444 studies below (no) and above (yes) the median duration of 7 days.

445

446 **Table 4**. Relation of the adherence score, subject- and methodological factors with the
447 reproducibility of the FMD measurement

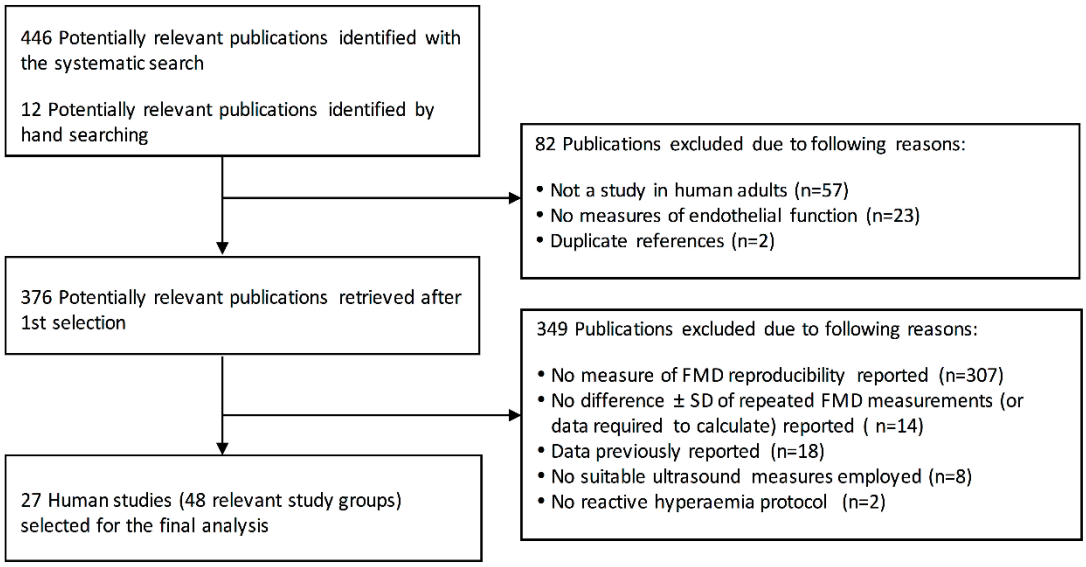
Stepwise Regression Analysis (model Adj R^2=0.51)			
Variable	**β**	**95% CI**	**P-value**
Adherence Score (unit)	-0.16	-0.24; -0.07	**<0.01**
Age (year)	-0.01	-0.02; -0.001	**0.03**
Stereostatic probe holder (yes)	-0.19	-0.06; -0.33	**<0.01**

448 The regression coefficient β represents the increase in the log FMD TEE per unit increase in each
449 factor. Baseline FMD and Centre experience did not remain in the model

450

451 **FIGURES:**

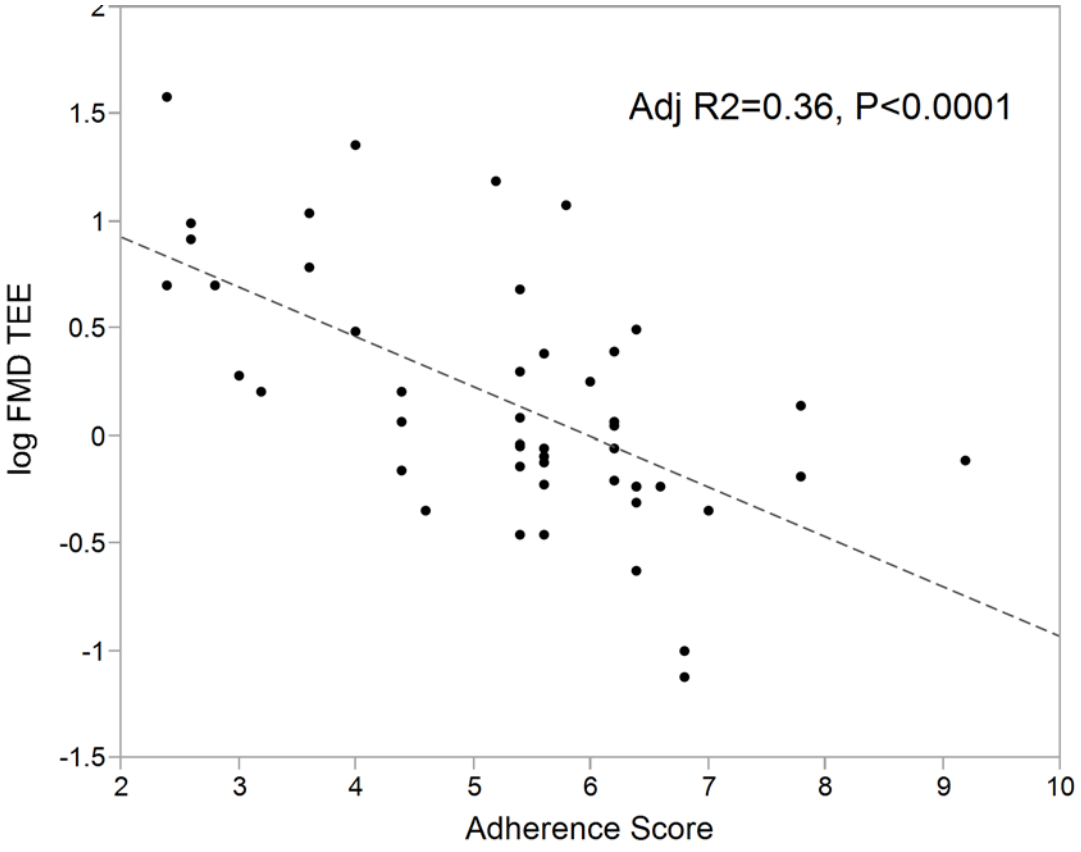
452 **Figure 1:**



453

454

455 **Figure 2:**



456