Journal of Human Hypertension (2005) 19, 77–82

www.nature.com/jhh

ORIGINAL ARTICLE

Comparing office-based and ambulatory blood pressure monitoring in clinical trials

WM Vollmer[1], LJ Appel[2], LP Svetkey[3], TJ Moore[4], TM Vogt[5], PR Conlin[6], M Proschan[7] and D Harsha[8] for the DASH Collaborative Research Group

[1]*Kaiser Permanente Center for Health Research, Portland, OR, USA;* [2]*Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, MD, USA;* [3]*Duke Hypertension Center and Sarah W. Stedman Center for Nutritional Studies, Duke University Medical Center, Durham, NC, USA;* [4]*Boston University, Boston, MA, USA;* [5]*Kaiser Permanente Center for Health Research, Honolulu, HI, USA;* [6]*Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA;* [7]*Division of Epidemiology and Clinical Applications, NHLBI, NIH, Rockville, MD, USA;* [8]*Pennington Biomedical Research Center, Baton Rouge, LA, USA*

Ambulatory blood pressure monitoring (ABPM) is commonly used in clinical trials. Yet, its ability to detect blood pressure (BP) change in comparison to multiple office-based measurements has received limited attention. We recorded ambulatory and five daily pairs of random zero (RZ) BPs pre- and post-intervention on 321 adult participants in the multicentre Dietary Approaches to Stop Hypertension trial. Treatment effect estimates measured by ambulatory monitoring were similar to those measured by RZ and did not differ significantly for waking *vs* 24-h ambulatory measurements. For systolic BP, the standard deviations of change in mean 24-h ambulatory BP (8.0 mmHg among hypertensives and 6.0 mmHg among nonhypertensives) were comparable to or lower than the corresponding standard deviations of change in RZ-BP based on five daily readings (8.9 and 5.9 mmHg). The standard deviations of change for mean waking ambulatory BP (8.7 and 6.7 mmHg) were comparable to those obtained using three to four daily RZ readings. Results for diastolic BP were qualitatively similar. Ambulatory monitoring was more efficient (ie, a smaller sample size could detect a given BP change) than three to four sets of daily RZ readings and required fewer clinic visits. The average of 33 ambulatory BP readings during the waking hours had an efficiency comparable to that from the mean of four daily pairs of RZ-BPs. Participants readily accepted the ABPM devices, and their use requires less staff training. ABPM provides a useful alternative to RZ-BP measurements in clinical trials.

Journal of Human Hypertension (2005) **19,** 77–82.
doi:10.1038/sj.jhh.1001772
Published online 9 September 2004

Keywords: blood pressure; blood pressure monitoring; ambulatory; clinical trials; methodology

Introduction

Ambulatory blood pressure monitoring (ABPM) is gaining in popularity as an alternative to traditional methods for measuring blood pressure (BP) in clinical trials.[1–4] Its purported benefits in this setting include enhanced precision (allowing for reduced sample size and/or increased study power), elimination of observer bias, and identification and exclusion of individuals with 'white coat' hypertension, also termed 'isolated office' hypertension.[5–9] In particular, several studies suggest that ABPM, compared with traditional BP measurement techniques, should reduce the variability of estimates of BP change in clinical trials.[10–17] Mean 24-h BPs also correlate more closely with measures of hypertensive target organ damage than do office-based BPs.[18] In one large trial, 24-h ABPM was superior to office readings in predicting the regression of left ventricular hypertrophy in hypertensive patients following treatment to reduce BP.[19]

Any advantage in enhanced precision ABPM may have over standard techniques may be offset by reduced effect size estimates. In three small studies of nonpharmacologic therapy, the effect size (BP change) measured by ABPM tended to be smaller than that detected by standard measurements.[20–22] In clinical trials, the ratio of BP change to its standard deviation is key in determining sample size requirements and computing statistical power.

As part of a multicentre trial of the effect of eating patterns on BP, we recorded both ambulatory and random-zero (RZ) BP measurements pre- and post-intervention on 321 adult participants. A previous

Correspondence: Dr WM Vollmer, Center for Health Research, 3800 N. Interstate Ave, Portland, OR 97227-1110, USA.
E-mail: william.vollmer@kpchr.org
Received 18 February 2004; revised 8 April 2004; accepted 15 June 2004; published online 9 September 2004

analysis of the effect of the study diets on ABPM BPs showed statistically significant treatment effects that did not differ significantly in magnitude from those measured using RZ-BP measurements.[23] We now extend those results by comparing results from various ABPM summary measurements with results from RZ-BPs recorded over 1–5 days. The findings are intended to guide researchers in designing future clinical trials in which BP is the primary outcome.

Methods

Dietary Approaches to Stop Hypertension (DASH) was a randomized outpatient feeding trial designed to compare the effects of three dietary patterns on BP among 459 adults who had an average diastolic blood pressure (DBP) of 80–95 mmHg and an average systolic blood pressure (SBP) <160 mmHg and were not taking antihypertensive medications.[24,25] The three dietary patterns included a 'control' diet typical of average US consumption; a 'fruits-and-vegetables' (F&V) diet rich in fruits and vegetables but otherwise similar to the control diet; and a 'combination' diet (the DASH diet) rich in fruits, vegetables, and low-fat dairy foods, and reduced in saturated and total fat.[26] Participants ate the control diet during a 3-week run-in feeding period and their randomly assigned intervention diet for an additional 8 weeks. The study was approved by the institutional review boards of each participating institution and by an NHLBI-appointed protocol review committee, and all participants provided written informed consent.

BP, assessed by trained staff using Hawksley RZ sphygmomanometers, was measured three times during screening, twice during each of the second and third weeks of run-in feeding, and five times during the final 2 weeks of intervention feeding. All measurements were taken with the subject in a seated position and using the right arm (if available).

ABPM measurements were initiated beginning with the second of five feeding cohorts. One measurement was recorded at the end of run-in feeding and a second at the end of intervention feeding. All were made using a Space Labs 90207 device (Redmond, Washington) programmed to take measurements every 30 min and to repeat a reading if the SBP, DBP, or heart rate fell outside predefined acceptable ranges. ABPM placement procedures included three manual RZ readings followed by three readings by the ABPM device while the participant remained seated. Only the first 24 h readings were retained for analysis. If fewer than 14 acceptable readings were obtained, the subject was asked to repeat the monitoring. Among participants with acceptable ABPM measurements, more than 90% of the possible waking and sleeping readings were obtained.[23] Staff were trained and certified on ABPM measurements on a regular basis throughout the study. Participants completed a questionnaire to assess how much wearing the monitor interfered with their daily activities.

Of the 362 participants potentially available for this analysis, 11 were excluded because they did not complete intervention feeding, six more because they did not have technically acceptable ABPM readings at both baseline and end-of-study, and 24 more because they did not have complete RZ-BP data. The final analysis includes the 321 participants with no missing data.

As the ABPM measurements were made during the latter part of run-in and intervention feeding, the RZ-BPs used in this analysis are computed using the final one, two, three, four, and five baseline and intervention measurements for each participant. For ABPM, we present data for the change in mean 24-h and mean waking ambulatory BPs.

Estimates of treatment effects (eg, the effect of the DASH diet on blood pressure, net of control) were computed using analysis of variance, adjusting for clinical centre.[25] Equivalence of estimated treatment effects for different outcomes (eg, for RZ-BP *vs* mean 24-h ambulatory BP) was assessed using a simple *t*-test constructed using individual paired differences between the two outcomes for each individual (see Appendix A).

Results

In all, 54% of the sample were men, and the mean age for both men and women participants was 45 years. A total of 56% of the participants were of African American descent—73% of women and 42% of men. Another 6% derived from other minority backgrounds. Overall, 29% of the participants had mean baseline BPs in the hypertensive range—36% of women and 23% of men.

Approximately 45% of participants reported that wearing the monitor interfered with their usual home and work activities 'not at all', and another roughly 45% reported that wearing the monitor interfered with their usual activities only 'somewhat'. Only 5–10% of participants reported that the monitor interfered with their usual activities 'a lot', although between 15 and 20% of participants reported that the monitor interfered with their sleep 'a lot'. Nonetheless, 96% of participants reported that, except for showers, they wore their monitor for the full 24-h period, which is consistent with the completeness of the study data.

Table 1 compares, for hypertensives and nonhypertensives, the treatment effects for ABPM *vs* RZ-BPs. The treatment effects are expressed net of control, thus adjusting for any artifacts, such as regression to the mean, that may be present in the data. The use of 4-day RZ-BP estimates minimizes the variability of the estimates while still limiting the baseline measurements to those taken during run-in. We observed no consistent trend for the treatment effect estimates measured by ABPM to be

Table 1 Net impact of the DASH diet on changes in ambulatory and random zero blood pressures (mmHg) among participants in the DASH trial[a]

	4-day RZ-BP[b]	*Waking ABPM*	*24-h ABPM*
Nonhypertensive participants			
ΔSBP	−3.4 (−5.4, −1.3)	−1.9 (−4.1, 0.3)	−2.5 (−4.4, −0.5)
ΔDBP	−1.0 (−2.8, 0.8)	−1.4 (−3.2, −0.4)	−1.9 (−3.4, −0.3)
Hypertensive participants			
ΔSBP	−11.8 (−16.4, −7.3)	−10.4 (−15.0, −5.9)	−10.3 (−14.5, −6.2)
ΔDBP	−5.1 (−8.2, −1.9)	−5.2 (−8.5, −1.8)	−5.6 (−8.6, −2.6)

None of the estimated treatment effects differed between ABPM and RZ-BP or between waking and 24-h ABPM.
[a]Blood pressure changes are net of control and expressed as mean (95% conf. interval).
[b]Change from the four run-in blood pressures to final four end-of-study blood pressures.

Table 2 Standard deviation of blood pressure change (mmHg) in the DASH trial by number and type of blood pressure measurements

	Nonhypertensive participants		*Hypertensive participants*	
	ΔSBP	*ΔDBP*	*ΔSBP*	*ΔDBP*
RZ-BP				
1 set of daily measurements	9.7 (9.0, 10.6)	7.9 (7.3, 8.5)	13.5 (11.7, 15.8)	9.1 (7.9, 10.7)
2 sets of daily measurements	7.7 (7.2, 8.4)	6.5 (6.0, 7.1)	10.3 (8.9, 12.1)	7.6 (6.6, 8.9)
3 sets of daily measurements	6.7 (6.2, 7.3)	5.8 (5.4, 6.3)	9.3 (8.1, 10.9)	6.8 (5.9, 8.0)
4 sets of daily measurements	6.2 (5.8, 6.8)	5.5 (5.1, 6.0)	8.7 (7.5, 10.2)	6.1 (5.3, 7.1)
5 sets of daily measurements	5.9 (5.4, 6.4)	5.2 (4.8, 5.6)	8.9 (7.7, 10.4)	5.8 (5.1, 6.8)
ABPM				
Waking (33 measurements)[a]	6.7 (6.2, 7.3)	5.5 (5.1, 6.0)	8.7 (7.5, 10.2)	6.3 (5.5, 7.5)
24-h (48 measurements)[a]	6.0 (5.6, 6.6)	4.8 (4.4, 5.2)	8.0 (6.9, 9.4)	5.7 (5.0, 6.7)

All estimates are pooled estimates of s.d. of blood pressure change based on ANOVA model adjusting for treatment and site effects. 95% confidence intervals for SD change are shown in parentheses.
[a]Numbers in parentheses represent nominal number of ABPM measurements recorded during each observation period.

systematically greater (or less) than those measured by RZ-BP, and none of the pairwise comparisons of treatment effects estimated via the three measurement techniques were statistically significant. However, while all of the treatment effects estimated using 24-h ABPM measurements were statistically significant, only three of the four treatment effects estimated using RZ-BP and waking ABPM were statistically significant.

Table 2 presents estimates of the standard deviation of BP change for RZ-BPs measured over varying numbers of days and for mean waking and 24-h ABPM measurements. The first two columns display results for nonhypertensive participants, while the third and fourth columns are restricted to persons with hypertension at baseline. The table shows the increased precision resulting from increasing the number of measurement days (RZ options) or from increasing the number of measurements in any given day (ABPM options). For the RZ-BP measurements, the standard deviation of SBP change decreased on the order of 20% between one and two sets of daily measurements and an additional 10–13% between two and three daily sets. The reduction in variability was on the order of 7 and 5% on adding fourth and fifth days of measurements, respectively. For the ABPM measures, the standard deviation of BP change dropped on the order of 10% between the 33 (waking) measurements and all 48 measurements. The standard deviations of change in 24-h ambulatory SBP (8.0 mmHg in hypertensives and 6.0 mmHg in nonhypertensives) were either comparable or lower than the corresponding standard deviations of change in RZ-SBP measured over 5 days (8.9 and 5.9 mmHg), while the standard deviations of change for mean waking ambulatory SBP (8.7 and 6.7 mmHg) were comparable to those seen for three to four sets of daily RZ-BP measurements. Analysis of changes in DBP produced qualitatively similar results.

Assuming that BPs measured by RZ sphygmomanometers and by ABPM result in comparable treatment effect estimates in clinical trials, we can use the data from Table 2 to determine the relative sample size requirements for trials involving either ABPM or varying numbers of RZ-BP measurements. Table 3 shows calculations for a simple two-treatment, parallel-group design with a power of 90%. The calculations show the striking efficiency gains for the RZ-BP measurements by increasing

Table 3 Sample size per group required to detect given effect sizes among hypertensive participants with power 90%

Treatment effect (mmHg)	*SBP*			*DBP*		
	3.0	*5.0*	*10.0*	*3.0*	*5.0*	*10.0*
RZ-BP						
1 set of daily measurements	426	154	39	194	70	18
2 sets of daily measurements	248	90	23	135	49	13
3 sets of daily measurements	202	73	19	108	39	10
4 sets of daily measurements	177	64	16	87	32	8
5 sets of daily measurements	185	67	17	79	29	8
ABPM						
Waking (33 measurements)	177	64	16	93	34	9
24-h (48 measurements)	150	54	14	76	28	7

Assumes two-sided alpha level of 0.05 and standard deviations as per Table 2.

from one to three sets of daily measurements, the flattening out thereafter, and the superiority of the ABPM measurements over RZ-BP measurements based on three or fewer sets of daily measurements.

Finally, in an effort to determine the extent to which ambulatory and office-based BP measurement techniques identified the same responders, we correlated the change in BP assessed with the two techniques for participants eating the DASH diet. For those who were initially hypertensive, changes in both 24-h and waking ABPM measurements correlated highly (Pearson correlation coefficients = 0.79 and 0.82, resp., for SBPs and 0.60 and 0.60 for DBPs, all P-values <0.0001) with change in RZ-BP based on four measurements. Comparable correlations for participants who were not initially hypertensive were lower (0.41 and 0.39 for SBPs and 0.33 and 0.33 for DBPs), although still statistically significant.

Discussion

ABPM (24-h) appears to be a viable alternative to the use of RZ (or normal office-based) BP measurements for clinical trials in which BP is the primary outcome. Mean waking BPs recorded over a 16-h period appear to provide a precision comparable to that obtained from four sets of daily RZ-BPs. Furthermore, acceptance of the ambulatory devices appears high.

Strengths of the present study include the use of detailed, standardized protocols for both types of BP measurements, the ability to directly compare both techniques in the same population, and the large, diverse population. In all, 46% of the participants included in this analysis were female, 56% were African American, and 29% had baseline BPs in the hypertensive range.

The primary limitation of the study may be the nature of the intervention. DASH was a controlled feeding trial, which distinguishes it from longer-term trials such as the Trials of Hypertension Prevention.[27] Follow-up, from the beginning of run-in feeding to the end of intervention feeding, was just 11 weeks and included daily contact with participants. By contrast, lifestyle change trials may last for several years and include only quarterly or semi-annual clinic visits by participants. As a result, the standard deviations of change of RZ-BP and ABPM measurements and their relationship with the number of measurements will likely differ from the results reported here. In addition, the much greater intensity of participant contact in controlled feeding trials than in lifestyle change trials may have important logistical implications for the choice of measurement device. For instance, in an 18-month lifestyle change trial, a participant may be willing to come in for a single office BP measurement but not for an ABPM measurement, which is more burdensome and requires a second visit to return the equipment.

Our results are likely more applicable to shorter-term trials, such as drug studies, as well as to other feeding trials. This study represents the largest and most comprehensive analysis of this issue to date. These results may conceivably extend to the comparison of ABPM measurements with those made using nonmercury devices, including aneroid and automated stationary devices, which are increasingly being used as hospitals ban the use of mercury-containing devices.[28] Assuming that the primary sources of variability for such devices, as with carefully controlled RZ measurements, relate more to day-to-day and longer-term variation in true individual BPs than to measurement error, we expect that, qualitatively if not quantitatively, the results presented here would be generally applicable to other BP-measuring devices.

Rosner *et al*[29] documented substantially greater day-to-day than within-day variability in BP measurements. Their findings would tend to support taking BP measurements over several days, rather than simply taking the same total number of

measurements on one day. Our results suggest that the range of BPs sampled over the course of an entire day (or at least over several hours) must mimic some of the longer-term variation that has been observed for office-based BP measurements. That would also help to explain how increasing the number of ABPM measurements from 33 to 48 could continue to have an impact in reducing variability. If one focuses on awake BP, as few as 6 h of monitoring with 2–3 readings/h may be sufficient,[30] although our data would suggest that the added measurements further reduce variability.

Although three small studies of nonpharmacological therapy reported that the BP change measured by ABPM tended to be smaller than that detected by standard measurements,[20–22] we found no evidence of this in the DASH trial. Thus, we conclude that the standard deviation of BP change appears to be the key factor in determining relative sample size requirements and computing statistical power.

Our results also indicate that, to the extent that it is possible to identify 'responders' to the DASH diet, both ambulatory and office-based techniques tend to identify the same individuals. Correlation coefficients for relating change in RZ-BP measurements were much greater for those with hypertension at baseline than in those without. We believe that this is reflected both by the greater range of responses seen among hypertensive individuals (thus providing more variation to be explained), and by the fact that there is presumably more random variability in the change scores of the nonhypertensive group. We would caution, however, that defining individual 'responders' to a dietary intervention based on a single change score is at best an imprecise process.[31]

Ultimately, the decision between ABPM and office-based measurements for a given clinical trial will need to reflect a variety of considerations, including study duration, anticipated frequency of participant contact, and the differential risk of missing data during follow-up (potentially worse for ABPM). Cost is also a consideration. Obtaining high-quality office-based measurements requires extensive staff training and ongoing quality control monitoring, even if mercury devices become unavailable. At that time, the debate will shift to the use of one-time ABPM measurements *vs* multiple BP measurements made using aneroid or automated stationary devices. These devices would eliminate some but not all of the biases traditionally associated with standard BP measurements. Still, caution is warranted: it has not yet been established that a given change in ABPM pressures has the same impact on long-term cardiovascular risk as an equivalent change in BP measured by standard methods, even though recent evidence supports the clinical utility of using ABPM to predict CVD risk.[32,33] Nonetheless, we believe that the results presented here can inform the choice of BP measurement protocols now and in the future.

Acknowledgements

We wish to thank the DASH participants for their contribution to this study, and the DASH staff and the members of the DASH Data Safety and Monitoring Board for all their hard work on behalf of the study. This work was supported by cooperative agreements #HL50968, HL50972, HL50977, HL50981, HL50982, HL02642, RR02635, and RR00722 from the National Heart, Lung, and Blood Institute, the Office of Research on Minority Health, and the National Center for Research Resources of the National Institutes of Health.

References

1 Mancia G *et al.* Ambulatory blood pressure monitoring in the design of studies on antihypertensive drug efficacy. *Am J Hypertens* 1993; **6**(Part 2): S233–S235.
2 Coats AJ. Benefits of ambulatory blood pressure monitoring in the design of antihypertensive drug trials. *Blood Press Monit* 1996; **1**: 157–160.
3 Staessen JA *et al.* Ambulatory blood pressure monitoring in clinical trials. *Blood Press Monit* 1996; **1**: 223–225.
4 Waeber B, Brunner HR. Clinical value of ambulatory blood pressure monitoring in the assessment of antihypertensive therapy. *Blood Press Monit* 1999; **4**: 263–266.
5 National High Blood Pressure Education Program Working Group. Report on ambulatory blood pressure monitoring. *Arch Intern Med* 1990; **150**: 2270–2280.
6 Appel LJ, Stason WB. Ambulatory blood pressure monitoring and blood pressure self-measurement in the diagnosis and management of hypertension. *Ann Intern Med* 1993; **118**: 867–882.
7 Waeber B *et al.* Are some hypertensive patients overtreated? A prospective study of ambulatory blood pressure recording. *Lancet* 1987; **2**: 732–734.
8 Weber MA, Cheung DG, Graettinger WF, Lipson JL. Characterization of antihypertensive therapy by whole-day blood pressure monitoring. *JAMA* 1988; **259**: 3281–3285.
9 Cheung DG, Gasster JL, Weber MA. Assessing duration of antihypertensive effects with whole-day blood pressure monitoring. *Arch Intern Med* 1989; **149**: 2021–2025.
10 Trazzi S *et al.* Reproducibility of non-invasive and intra-arterial blood pressure monitoring: implications for studies on antihypertensive treatment. *J Hypertens* 1991; **9**: 115–119.
11 Conway J, Coats A. Value of ambulatory blood pressure monitoring in clinical pharmacology. *J Hypertens* 1989; **7**(Suppl): S29–S32.
12 Appel LJ, Marwaha S, Whelton PK, Patel M. The impact of automated blood pressure devices on the efficiency of clinical trials. *Control Clin Trials* 1992; **13**: 240–247.
13 Conway J *et al.* The use of ambulatory blood pressure monitoring to improve the accuracy and reduce the numbers of subjects in clinical trials of antihypertensive agents. *J Hypertens* 1988; **6**: 111–116.

14 Prisant LM *et al.* Repeatability of automated ambulatory blood pressure measurements. *J Fam Pract* 1992; **34**: 569–574.
15 Mancia G *et al.* Limited reproducibility of hourly blood pressure values obtained by ambulatory blood pressure monitoring: implications for studies on antihypertensive drugs. *J Hypertens* 1992; **10**: 1531–1535.
16 Silagy CA *et al.* Components of blood pressure variability in the elderly and effects on sample size calculations for clinical trials. *Am J Hypertens* 1992; **5**: 449–458.
17 Staessen JA *et al.* Clinical trials with ambulatory blood pressure monitoring: fewer patients needed? Syst-Eur Investigators. *Lancet* 1994; **344**: 1552–1556.
18 Mancia G, Lanfranchi A, Turri C, Grassi G. Can good surrogate end-points predict the prognosis of hypertensive patients? *J Hypertens* 1998; **16**(Suppl 5): S3–S7.
19 Mancia G *et al.* Ambulatory blood pressure is superior to clinic blood pressure in predicting treatment-induced regression of left ventricular hypertrophy. *Circulation* 1997; **95**: 1464–1470.
20 Fotherby MD, Potter JF. Potassium supplementation reduces clinic and ambulatory blood pressure in elderly hypertensive patients. *J Hypertens* 1992; **10**: 1403–1408.
21 Fotherby MD, Potter JF. Effects of moderate sodium restriction on clinic and twenty-four-hour ambulatory blood pressure in elderly hypertensive subjects. *J Hypertens* 1993; **11**: 657–663.
22 Moore TJ, Malarick C, Olmedo A, Klein RC. Salt restriction lowers resting blood pressure but not 24-h ambulatory blood pressure. *Am J Hypertens* 1991; **4**(Part 1): 410–415.
23 Moore TJ *et al.* Effect of dietary patterns on ambulatory blood pressure: results from the Dietary Approaches to Stop Hypertension (DASH) Trial. DASH Collaborative Research Group. *Hypertension* 1999; **34**: 472–477.
24 Sacks FM *et al.* Rationale and design of the Dietary Approaches to Stop Hypertension trial (DASH). A multicenter controlled-feeding study of dietary patterns to lower blood pressure. *Ann Epidemiol* 1995; **5**: 108–118.
25 Appel LJ *et al.* A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med* 1997; **336**: 1117–1124.
26 Karanja NM *et al.* Descriptive characteristics of the dietary patterns used in the Dietary Approaches to Stop Hypertension Trial. DASH Collaborative Research Group. *J Am Diet Assoc* 1999; **99**(Suppl): S19–S27.
27 Hebert PR *et al.* Design of a multicenter trial to evaluate long-term life-style intervention in adults with high-normal blood pressure levels. Trials of Hypertension Prevention (phase II). Trials of Hypertension Prevention (TOHP) Collaborative Research Group. *Ann Epidemiol* 1995; **5**: 130–139.
28 Jones DW, Appel LJ, Sheps SG, Roccella EJ. Measuring blood pressure accurately: new and persistent challenges. *JAMA* 2003; **289**: 1027–1030.
29 Rosner B, Polk BF. Predictive values of routine blood pressure measurements in screening for hypertension. *Am J Epidemiol* 1983; **117**: 429–442.
30 Sheps SG, Bailey KR, Zchariah PK. Short-term (six hour), ambulatory blood pressure monitoring. *J Hum Hypertens* 1994; **8**: 873–878.
31 Eva's paper on measurement of sodium responders.
32 Clement DL *et al.* Prognostic value of ambulatory blood-pressure recordings in patients with treated hypertension. *N Engl J Med* 2003; **348**: 2407–2415.
33 Kario K *et al.* Morning surge in blood pressure as a predictor of silent and clinical cerebrovascular disease in elderly hypertensives: a prospective study. *Circulation* 2003; **107**: 1401–1406.

Appendix A

This appendix describes the procedure used to compare the various intervention effect size estimates in Table 1.

Let OM1 and OM2 be two correlated outcome measures (eg, RZ-BP and mean 24-h ambulatory blood pressure), and assume interest is focused on the effect of the DASH dietary pattern. For OM1, the effect of the DASH diet, net of control, can be expressed as $(\overline{\Delta OM1_{DASH}} - \overline{\Delta OM1_{cntl}})$, where $\overline{\Delta OM1_{DASH}}$ is the observed mean difference in OM1 from baseline to end-of-study for participants eating the DASH diet, and $\overline{\Delta OM1_{cntl}}$ is the comparable difference for those eating the control diet. In order to test whether the DASH diet effects measured by OM1 and OM2 are equivalent, we compute

$$\delta = (\overline{\Delta OM1_{DASH}} - \overline{\Delta OM1_{cntl}}) - (\overline{\Delta OM2_{DASH}} - \overline{\Delta OM2_{cntl}}),$$

which can be rewritten as

$$\delta = (\overline{\Delta OM1_{DASH}} - \overline{\Delta OM2_{DASH}}) - (\overline{\Delta OM1_{cntl}} - \overline{\Delta OM2_{cntl}}).$$

Once in this form, δ can be seen to be the simple difference in two means, $\bar{X}$ and $\bar{Y}$, where $X_1, \ldots, X_{n1}$ and $Y_1, \ldots, Y_{n2}$ are $n1 + n2$ mutually independent individual paired differences defined by $X_i = (\Delta OM1_{DASH,i} - \Delta OM2_{DASH,i})$ and $Y_i = (\Delta OM1_{cntl,i} - \Delta OM2_{cntl,i})$. If s_x^2 and s_y^2 are the estimated standard deviations of the X's and Y's, then

$$T = \frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{s_x^2}{n1} + \frac{s_y^2}{n2}}}$$

will have an approximate Student's t-distribution for sufficiently large $n1$ and $n2$, and can thus be used to test the null hypothesis that $\delta = 0$.