The individual RDH index: a novel vector index for statistical assessment of antihypertensive treatment reduction, duration, and homogeneity

Mateo Aboy^{a,b}, José R. Fernández^a, James McNames^c and Ramón C. Hermida^a

We propose a new vector index for the statistical assessment of antihypertensive treatment duration and homogeneity from ambulatory blood pressure monitoring. We termed this approach for evaluating and comparing blood pressure coverage offered by antihypertensive drugs over 24 h as the reduction-duration-homogeneity index. The reduction-duration-homogeneity index is a three-component vector index that incorporates information about the reduction, duration, and homogeneity of antihypertensive treatment, as well as their statistical significance. The advantages of the reduction-duration-homogeneity index are demonstrated by several comparative examples. *Blood Press Monit* 11:69–78 © 2006 Lippincott Williams & Wilkins.

Keywords: ambulatory blood pressure monitoring, antihypertensive drug efficacy, antihypertensive duration of action, antihypertensive therapy, blood pressure variability, trough-peak ratio, smoothness index

^aBioengineering and Chronobiology Laboratories, Signal Theory and Communications Department, ETSIT, University of Vigo, Vigo, Spain, ^bDepartment of Electronics Engineering Technology at Oregon Institute of Technology and ^cBiomedical Signal Processing Laboratory at Portland State University, Portland, Oregon, USA

Correspondence and requests for reprints to Mateo Aboy E-mail: mateoaboy@ieee.org

Sponsorship: This research was supported in part by grants from Xunta de Galicia (PGIDIT03-PXIB-32201PR) and Vicerrectorado de Investigacion, University of Vigo.

Received 23 February 2005 Revised 8 June 2005 Accepted 8 August 2005

Blood Pressure Monitoring 2006, 11:69-78

Introduction

The duration and homogeneity of antihypertensive drugs are commonly quantified by computation of the trough: peak ratio (TP) and the smoothness index (SI) [1–7]. Currently, both indices are computed from ambulatory blood pressure monitoring (ABPM) recordings. ABPM has drastically improved the ability to assess the efficacy of antihypertensive treatment in clinical studies and in medical practice [8–11]. ABPM offers a number of advantages over clinic sphygmomanometric readings obtained casually in the examiner's office. The ABPM is characterized by higher reproducibility, is not subject to observer bias and white-coat effect, enables us to test the effectiveness of a given antihypertensive drug in daily life conditions, and can be used to estimate the pharmacodynamics of antihypertensive drugs [2,12,13].

The TP and the SI aim at quantifying the efficacy of the antihypertensive drug by taking into account the duration and homogeneity of its antihypertensive effect. The underlying idea is based on the contention that optimal control and management of blood pressure (BP) should be based upon therapeutic strategies that consistently reduce BP in a homogeneous or smooth manner.

The relative advantages and limitations of these two different indices for assessing the efficacy and duration of the action of antihypertensive drugs have been addressed in several studies [3,4,7,14]. Despite their widespread use, their implementation in a clinical setting is still mainly restricted to research, and at the moment inferences on the clinical superiority of a particular treatment over another based on a higher TP or SI are still considered speculative in nature [6]. Both indices have intrinsic limitations. Specifically, the TP and the SI indices are based on the BP reduction, but do not assess the statistical significance of these reductions. Additionally, there are intrinsic limitations to any single number (scalar) index for the task of assessing the antihypertensive treatment drug effect, as it is impossible to capture the complexity of antihypertensive treatment by means of a single number.

The significance of assessment of antihypertensive treatment and the need for a new index are well established in the literature [6,7,15]. In order to correct one of the main limitations of the SI, we have recently proposed a simple correction factor for the SI, which prevents the index from reaching erroneous high values in situations when the reduction in BP is inadequate but very homogeneous. We have termed this corrected SI index as the SIn (normalized SI) [16].

In this paper, we propose a novel vector index for assessment of antihypertensive treatment duration and homogeneity from ABPM. We termed this approach for evaluating and comparing BP coverage offered by antihypertensive drugs over 24 h as the reduction– duration–homogeneity index (RDH). The RDH is a three-component index which can be used to characterize

1359-5237 © 2006 Lippincott Williams & Wilkins

and evaluate the effect of antihypertensive treatment more precisely than using the TP, SI, or SIn.

In the following subsections of the Introduction, we give an overview of the TP, SI, and SIn methodologies. In the next section, we present the proposed index and methods. In the subsequent section, we describe the participants and data used in this study and we show several examples illustrating and comparing the TP, SI, SIn, and RDH indices and discuss their relative advantages and limitations. In the last section, we provide some concluding remarks.

The trough-peak ratio

The TP was introduced after the US Food and Drug Administration (FDA) suggested that an antihypertensive agent should retain most of its peak effect at trough. In 1988, the FDA published guidelines indicating that a minimum TP of 50-60% was required for the efficacy of an antihypertensive drug to be considered satisfactory in relation to its proposed dosage interval [17]. These guidelines were established with two objectives: (1) to define a simple index to assess the duration of action and homogeneity of an antihypertensive drug and (2) to prevent the use of inappropriate high doses of drugs with the objective of extending the duration of action. The second objective was aimed at detecting drugs with short duration of action, which could achieve an optimal trough effect at the expense of using a high dose and causing a large BP drop at peak [1]. Following the FDA definition, the TP is obtained by dividing the BP reduction at the end of the between-dose interval (trough) by the BP at the time of maximal drug effect (peak). Despite this simple definition, the application of TP to ABPM recordings raises several methodological questions. These include how to select the optimal time window for the calculation of BP changes at peak and trough, the exact definition of the peak and trough effects, whether to limit the TP calculation only to responders, and other methodological considerations caused by the lack of more specific methodological requirements for its calculation [3]. As a consequence, the application of the TP has resulted in inconsistencies and discrepancies caused by the use of different methodologies for its calculation. These have called into question the validity of the index altogether. In general, to obtain the TP, peak changes are calculated by averaging the values of (1) the hour in which the BP reduction compared with the corresponding baseline value was maximal within the first 8h after drug intake and (2) the adjacent hour in which the BP fall was greater. Trough changes are calculated by averaging the differences between baseline and after treatment BP values over the last 2-h of the interdose period [3]. In addition to these methodological issues, other concerns about the potential theoretical limitations of the index have also been expressed. The most significant of these limitations are as follows: (1) it makes use of only a small portion of the 24-h ABPM recording; (2) even though individual peak and trough changes in ABPM follow a normal distribution with narrow scatter, TPs exhibit pronounced scatter and do not follow a normal distribution; (3) TP has no relation to variability of BP; (4) there is a small placebo effect at peak; (5) it has limited reproducibility and clinical value [5,6]. Nonetheless, although the FDA no longer requires to provide its value, the TP is still widely used, and there is convincing evidence that it is a useful index when employed in an appropriately conducted study [4] and interpreted correctly.

The smoothness index

The SI is defined as the ratio between the average of the 24 hourly reductions in BP induced by treatment and the standard deviation of these hourly reductions [15,18]. The index was proposed in an attempt to overcome some of the limitations of the TP, especially the fact that the TP ratio does not take advantage of all the 24-h ABPM data available [5]. The SI is derived from the analysis of 24-h BP profiles obtained before and during drug treatment, and takes into account all 24-h BP differences between baseline and treatment, which enables researchers to obtain a more comprehensive evaluation of the homogeneity of 24-h therapeutic coverage. Currently, the SI index is the most popular index used in the evaluation of antihypertensive treatment. This is due, in part, to the fact that SI has been shown to have clinical validity, as SI has been reported to be associated with changes in left ventricular mass and with changes in carotid artery wall thickness under antihypertensive treatment [19,20]. Additionally, the SI has been found to have a normal distribution with narrow scatter [5,6] and exhibits greater reproducibility in comparison between mediumterm and long-term treatment when compared with TP [4,5,18].

The normalized smoothness index

We have recently proposed the SIn in an attempt to overcome one of the documented limitations of the SI [16]. According to the standard definition, the SI is calculated as the ratio between the mean hourly reductions and the standard deviation of these. The main limitation of this definition is that a drug with a negligible BP reduction could still have a very high SI, provided that the reduction is very homogeneous. The SI tends to infinity regardless of the BP reduction as the standard deviation of the mean hourly reductions tends to zero for any non-zero BP mean reduction. To overcome this limitation, we proposed a normalized index (SIn) defined as the ratio of the mean hourly reductions to the standard deviation of these plus one. Adding a one to the denominator makes the index robust to the problem and improves interpretability of results. As the standard deviation tends to zero, the SIn tends to the mean BP reduction.

Methodology Notation and definitions

In this section, we introduce notation and precise definitions that will be used in subsequent sections to define the RDH index.

Given two individual ABPM recordings, let us denote the vector containing the individual raw BP values before treatment (baseline) as x and the vector containing the individual raw BP values after treatment as y. In general, these time series are not uniformly sampled. ABPM monitors are commonly configured so that the sampling frequency is higher during daytime than at night, and it can also happen that the patient may need to take out the monitor for some period of time. Let $k_{k=1}^{K}$ denote the *K* time containers (categories), where k = 1 corresponds to the first time (i.e. wake up, drug intake time). For the purposes of this paper, we will assume we have 24 categories (K = 24). Let $x_{k,j}$ denote the *j*th BP sample belonging to the *k*th category. Using this notation, let us define a set of 24 ABPM vectors at baseline as

$$\begin{array}{rcl}
x_1 &=& (x_{1,1}, x_{1,2}, \dots, x_{1,L_1}) \\
x_2 &=& (x_{2,1}, x_{2,2}, \dots, x_{2,L_2}) \\
\vdots && , \\
x_{24} &=& (x_{24,1}, x_{24,2}, \dots, x_{24,L_{24}})
\end{array}$$
(1)

where L_k denotes the number of BP samples in the *k*th class at baseline. In general, the dimension of vectors from different classes is not equal, that is, $L_k \neq L_j$, where *k* and *j* denote the index of the *k*th and *j*th classes.

The set of 24 ABPM vectors after treatment is defined analogously,

$$y_{1} = (y_{1,1}, y_{1,2}, \dots, y_{1,L'_{1}})
y_{2} = (y_{2,1}, y_{2,2}, \dots, y_{2,L'_{2}})
\vdots
y_{24} = (y_{24,1}, y_{24,2}, \dots, y_{24,L'_{24}})$$
(2)

where L_{k}' denotes the number of BP samples in the *k*th class after treatment. In general, $L_{k} \neq L'_{k}$, that is, the dimension of the vector before treatment corresponding to the *k*th class, $\mathbf{x}_{k} = (x_{k,1}, x_{k,2}, \dots, x_{k,L_{k}})$, is not necessarily equal to the dimension of the vector after treatment corresponding the same class $\mathbf{y}_{k} = (y_{k,1}, y_{k,2}, \dots, y_{k,L'_{k}})$. Let us denote the sample mean of the ABPM vector corresponding to the *k*th class before and after treatment as \bar{x}_{k} and \bar{y}_{k} , respectively

$$\bar{x}_k = \sum_{i=1}^{L_k} \frac{x_{k,i}}{L_k},$$
(3)

$$\bar{y}_k = \sum_{i=1}^{L'_k} \frac{y_{k,i}}{L'_k},\tag{4}$$

and define two vectors of 24 components each, with components of sample means corresponding to each class

before and after treatment:

$$\overline{\mathbf{x}} = (\overline{x}_1, \overline{x}_2, \dots, \overline{x}_{24}), \tag{5}$$

$$\overline{\mathbf{y}} = (\overline{y}_1, \overline{y}_2, \dots, \overline{y}_{24}). \tag{6}$$

Analogously, let us define two vectors with components corresponding to the estimated standard errors of the means for each class, before and after treatment,

$$\widehat{se}_x = (\widehat{se}_{x,1}, \widehat{se}_{x,2}, \dots, \widehat{se}_{x,24}), \tag{7}$$

$$\widehat{se}_{\gamma} = (\widehat{se}_{\gamma,1}, \widehat{se}_{\gamma,2}, \dots, \widehat{se}_{\gamma,24}), \tag{8}$$

where the standard error for a given class k before treatment is given by $\hat{se}_{x,k} = \sqrt{s_x^2/L_k}$ and after treatment by $\hat{se}_{y,k} = \sqrt{s_y^2/L_k}$, and s_x^2 and s_y^2 are the sample variance of class k before and after treatment.

For a given patient, we have one vector containing the BP sample means for each class before treatment, $\overline{\mathbf{x}}$, and another with the means after treatment, $\overline{\mathbf{y}}$. The vector containing the class-by-class differences is denoted as

$$d = \overline{\mathbf{x}} - \overline{\mathbf{y}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{24}) - (\bar{y}_1, \bar{y}_2, \bar{y}_{24}) = (d_1, d_2, \dots, d_{24}).$$
(9)

Smoothness index, normalized smoothness index and trough-peak ratio

According to the standard definition, the SI is calculated as the ratio between the mean of the hourly reductions and the standard deviation of these, that is,

$$SI = \frac{\bar{d}}{s_d},$$
 (10)

where \overline{d} denotes the sample mean of the class-by-class differences (reductions) and s_d is the sample standard deviation

$$\bar{d} = \sum_{i=1}^{24} \frac{d_i}{24},$$

$$s_d = \sqrt{\sum_{i=1}^{24} \frac{(d_i - \bar{d})^2}{24 - 1}}.$$
(11)

The SI incorporates two effects in a single number, namely, the mean BP reduction \overline{d} and its homogeneity s_d . The main limitation of this definition is that a drug with a negligible BP reduction could still have a very high SI, provided that the reduction is very homogeneous. In fact, the index tends to infinity regardless of the BP reduction as the standard deviation s_d tends to zero for any non-zero BP mean reduction.

To overcome this limitation, we have recently proposed a SIn defined as

$$\operatorname{SIn} = \frac{\overline{d}}{1 + s_d}.$$
 (12)

Unlike SI, SIn does not tend to infinity as s_d tends to zero, instead it tends to \bar{d} . The best SIn possible is \bar{d} , which is reached only when the reduction is constant ($s_d = 0$). This correcting factor also has the benefit that it removes outliers (i.e. very high SI values because of $s_d \approx 0$).

The TP is calculated as the ratio of the mean BP reduction at the end of the between-dose interval (trough) to the mean BP reduction at the time of maximum drug effect (peak). The exact method for calculating the trough and peak differs among researchers. One of the most commonly used methods was proposed by Omboni *et al.* [3,5]. The peak effect is calculated by considering the interval between the second and eighth hour after drug intake. The average is computed over 2-h time-windows as follows:

$$TP = \frac{T}{P},$$

$$T = \frac{d_{23} + d_{24}}{2},$$

$$P = \frac{d_{\max} + \max\{d_{\max-1}, d_{\max+1}\}}{2},$$
(13)

where T is the average of the BP differences over the last 2 h of the interdose period, P is the average around the peak effect, $d_{\text{max}} = \max\{d_2, d_2, \dots, d_8\}$ is the maximum BP difference in the time interval between the second and eighth hour after drug intake, and $d_{\text{max}-1}$ and $d_{\text{max}+1}$ are the BP differences in the adjacent classes to d_{max} .

Parametric RDH index

TP, SI, and SIn are obtained from analysis of the mean hourly BP reductions d. In contrast, the proposed RDH index is based on the idea that what is important is not the absolute BP differences (i.e. reductions d in mmHg), but whether or not these reductions are significant when we take into account the intrinsic variability of the BP at each of the 24 categories before and after treatment. The objective is to perform a hypothesis test for each of the 24 categories (hours) in order to test whether or not the reduction is significant.

Given the vector of the mean BP differences **d** (reductions) and the vectors containing the estimated standard errors of the mean before and after treatment, \widehat{se}_x and \widehat{se}_y , we can estimate the standard error of the mean BP reductions $\mathbf{d} = \overline{\mathbf{x}} - \overline{\mathbf{y}}$ as

$$\widehat{se}_{d_k} = \sqrt{\left(\widehat{se}_{x,y}\right)^2 + \left(\widehat{se}_{y,k}\right)^2},\tag{14}$$

$$\widehat{\mathbf{se}}_d = (\widehat{se}_{d_1}, \widehat{se}_{d_2}, \dots, \widehat{se}_{d_{24}},$$
(15)

as the variance of the difference of two independent random variables is the sum of their variances.

Having the vector of reductions **d** and the vector of estimated standard errors \widehat{se}_d , we can determine whether

or not the reduction is significant or just due to chance by calculating the ratio of the two

$$\mathbf{r} = \frac{d}{\widehat{\mathbf{s}}\widehat{\mathbf{e}}_d} = \left(\frac{d_1}{\widehat{s}\widehat{e}_{d_1}}, \frac{d_2}{\widehat{s}\widehat{e}_{d_2}}, \dots, \frac{d_{24}}{\widehat{s}\widehat{e}_{d_{24}}}\right).$$
(16)

The vector **r** enables us to assess the statistical significance of each of the 24-h reductions. Given an estimate $\hat{\theta}$ and an estimate of the standard error \hat{se} , the 90% confidence interval (CI) for θ assuming a normal distribution is $\hat{\theta} \pm 1.645 \cdot \hat{se}$. Thus, components of **r** greater than 1.645 can be considered statistically significant assuming normality (note that in this case a 90% two-sided test is equivalent to a 95% one-sided test). This threshold may be increased according to the *t*-distribution.

The RDH index is based on reductions in \hat{s} units, that is, on the vector **r** as opposed to raw reductions **d**. This vector **r** can be plotted instead of the **d** to illustrate the reductions for each of the 24 categories in \hat{s} units.

An indicator function \mathscr{I} operates on **r** to generate a sequence of ones and zeros corresponding to significant and non-significant reductions, respectively,

$$\mathbf{r}_T = \mathscr{I}(\mathbf{r} > 1.645). \tag{17}$$

The RDH index is defined in terms of the \mathbf{r}_T as a triplet

$$RDT = (c_1, c_2, c_3),$$
 (18)

where c_1 denotes the total number of ones in \mathbf{r}_T (significant reductions), c_2 is the maximum number of consecutive significant reductions, and c_3 is the maximum number of consecutive non-significant reductions. For the c_2 and c_3 , it is also useful to provide the time information. For instance, an RDH = (15, 8, 4) in an ABPM with 24 categories would provide the following information (assuming a c_2 corresponding to the time 0000–0800 h and c_3 to 1900–2300 h): (1) there were significant reductions in 15 out of the 24 categories; (2) there were a maximum of eight consecutive hours with significant reductions corresponding to the period between 0000 and 0800 h; (3) there were 4 h without coverage (no significant treatment efficacy, corresponding to the period between 1900 and 2300 h). Thus, the RDH provides information about the reduction (i.e. whether each reduction is significant or not), homogeneity (i.e. what is the number of categories with significant reductions), and duration and coverage (i.e. what is the maximum number of consecutive significant reductions, and what is the maximum number of consecutive nonsignificant reductions, corresponding to the coverage and lack of coverage of an antihypertensive drug).

Non-parametric RDH

In the development presented in the previous section, the RDH quantifies the number of significant reductions induced by an antihypertensive drug. We intentionally referred to these reductions as significant as opposed to statistically significant, as they can only be interpreted as statistically significant under the assumption of normality. In this section, we describe a methodology to perform non-parametric hypothesis testing based on bootstrap in order to determine statistically significant BP reductions and to generate *P*-values for each category without making any assumptions about the distribution of the BP differences [21].

We are interested in testing a hypothesis for the difference of means for each different class. The null and the alternative hypotheses are

$$\begin{aligned} H_0 : \mu_{\mathbf{y}_k} &= \mu_{\mathbf{x}_k} \Rightarrow \mu_{\mathbf{y}_k} - \mu_{\mathbf{x}_k} = 0, \\ H_1 : \mu_{\mathbf{y}_k} &< \mu_{\mathbf{x}_k} \Rightarrow \mu_{\mathbf{y}_k} - \mu_{\mathbf{x}_k} < 0, \end{aligned}$$
(19)

where $\mu_{\mathbf{y}_k}$ and $\mu_{\mathbf{x}_k}$ are estimated by \bar{y}_k and \bar{x}_k , respectively. Recall that \mathbf{x}_k is the mean of the vector corresponding to class k before treatment, $\mathbf{x}_k = (x_{k,1}, x_{k,2}, \dots, x_{k,L_k})$.

The process involves computing a bootstrap CI for the difference of the parameters of interest (mean at each category); equality of the parameters is assumed if zero is a possible value in the CI. The bootstrap technique was introduced in 1979 as a computer-based method for estimating the standard error of an estimator [21]. The natural example used as estimator was the sample mean, the estimator of the population mean.

We obtain a bootstrap sample $\mathbf{x}_{k}^{*} = (x_{k,1}^{*}, x_{k,2}^{*}, \dots, x_{k,L_{k}}^{*})$ by randomly sampling L_{k} times, with replacement, from the original BP data points $x_{k,1}^{*}, x_{k,2}^{*}, \dots, x_{k,L_{k}}^{*}$. For instance, with $L_{k} = 6$, a possible \mathbf{x}_{k}^{*} could be $\mathbf{x}_{k}^{*} = (x_{k,6}^{*}, x_{k,1}^{*}, x_{k,4}^{*}, x_{k,1}^{*}, x_{k,5}^{*}, x_{k,2}^{*})$. Analogously, for each category k, we can generate a large number of independent bootstrap samples $\mathbf{x}_{k}^{*^{1}}, \mathbf{x}_{k}^{*^{2}}, \dots, \mathbf{x}_{k}^{*^{B}}$, each of size L_{k} . The same process is followed to create bootstrap samples for the after-treatment vectors $\mathbf{y}_{k}^{*^{1}}, \mathbf{y}_{k}^{*^{2}}, \dots, \mathbf{y}_{k}^{*^{B}}$. For each replica, we calculate the sample mean for before and after treatment, \overline{x}_{k}^{*} and \overline{y}_{k}^{*} , respectively. Using these bootstrap sample means enables us to generate bootstrap differences for each category k,

$$\begin{array}{rcl}
d_{k}^{*^{1}} &=& (\bar{y}_{k}^{*} - \bar{x}_{k}^{*})^{1} \\
d_{k}^{*2} &=& (\bar{y}_{k}^{*} - \bar{x}_{k}^{*})^{2} \\
\vdots \\
d_{k}^{*B} &=& (\bar{y}_{k}^{*} - \bar{x}_{k}^{*})^{B}
\end{array}$$
(20)

Thus, for a given category k, we have B bootstrap differences $\{d_k^{*'}\}_{l=1}^B$. This enables us to define a vector of B components for each class

$$\mathbf{d}_{k}^{*} = (d_{k}^{*1}, d_{k}^{*2}, \dots, d_{k}^{*B}).$$
(21)

The histogram of \mathbf{d}_{k}^{*} is an estimate of the probability density function of the differences of mean for category k. The bootstrap CIs for the BP reduction in class k are

obtained as

$$\hat{d}_{klo} = 100 \cdot \alpha$$
th percentile of d_k^* 's distribution
 $\hat{d}_{kup} = 100 \cdot (1 - \alpha)$ th percentile of d_k^* 's distribution

If this interval contains zero, it cannot be assumed, with $(1-2\alpha)$ confidence, that the parameters of the two populations are statistically different. In order to perform a one-sided test like (19), the null hypothesis will be rejected, with $(1-\alpha)$ confidence, if \hat{d}_{kup} is lower than zero. Moreover, an approximation for the *P*-value of the test can be obtained by computing the relative position of zero in the sorted estimations of differences.

In order to have good estimates of CIs, $B \approx 1000$ [21]. Given a category k with L_k BP samples, the theoretical maximum number of replicas before repetition is given by

$$B_{max} = \begin{pmatrix} 2L_k - 1\\ L_k \end{pmatrix}.$$
 (22)

Thus, a minimum of four BP samples before and after treatment are needed for a given category in order to accurately generate a non-parametric CI for the BP reduction, as

$$\binom{7}{4} = 35, \text{ and } 35 \cdot 35 = 1225 (35 \text{ for } \mathbf{x}_k \text{ and } 35 \text{ for } \mathbf{y}_k).$$

The problem of having too few samples per category can be solved by overlapping categories as in the case of the TP ratio in which both the peak and trough are calculated by averaging two classes or by taking longer ABPM recordings (48 h instead of 24 h) and using the wellknown circadian variability of BP [22].

Illustrative examples

In this section, we present selected examples to illustrate the methodology presented earlier and to compare the proposed RDH index with the TP ratio, the SI, and the SIn.

Subjects and materials

We studied 90 white participants (30 men and 60 women), 49.0 \pm 14.3 years of age, with mild to moderate (grade 1 or 2) essential hypertension based on criteria of the European Society of Hypertension-–European Society of Cardiology guidelines [23] for conventional cuff BP measurements (systolic BP (SBP) between 140 and 179 mmHg or diastolic BP (DBP) between 90 and 109 mmHg) and corroboration by ABPM at the time of recruitment. A positive diagnosis of hypertension based on ABPM required that either the 24-h mean SBP/DBP be above 130/80 mmHg, the diurnal mean be above 135/85 mmHg, or the nocturnal mean be above 120/70 mmHg [24].

All the participants received their routine medical care at the Hypertension and Vascular Risk Unit, Hospital Clínico Universitario, Santiago de Compostela, Spain. They participated in a clinical trial on the antihypertensive efficacy of 160 mg/day valsartan, as reported previously [25]. Specific details on the participants and the design of this clinical trial have been provided previously [25]. The SBP and DBP, and heart rate (HR) of each participant were automatically measured every 20 min from 0700 to 2300 h and every 30 min during the night for 48 consecutive hours with a validated SpaceLabs 90207 device (SpaceLabs Inc., Issaquah, Washington, USA) [26]. Participants were studied by ABPM under baseline conditions when participants were free of medication before and again after 3 months of therapy. During 48-h ABPM, each participant wore a MiniMotionLogger actigraph (Ambulatory Monitoring Inc., Ardsley, New York, USA) on the dominant wrist to monitor physical activity every minute. This compact device, about half the size of a wrist watch, functions as an accelerometer. The internal clocks of the actigraph and the ABPM devices were synchronized through their respective interfaces by the same computer. The actigraphy data were used to determine the onset and offset times of diurnal activity and nocturnal sleep so as to accurately determine the diurnal and nocturnal BP means of each participant. The mean activity for the 5 min before each BP reading was then calculated for further statistical analysis on circadian variability of activity, according to previous studies on this area [27,28]. Each individual's clock hour BP and HR values were first re-referenced from clock time to hours after awakening from nocturnal sleep, according to the information obtained from wrist actigraphy. This transformation avoided the introduction of bias because of differences among participants in their sleep/activity routine [22]. BP and HR time series were then edited according to conventional criteria to remove measurement errors and outliers [29].

Results

We analyzed the participants described above by calculating the TP, SI, Sin, and RDH of the SBP. The following examples were selected to illustrate the methodological contributions presented in this paper.

In Fig. 1, we show an example of potential problems with the TP ratio and the SI. Figure 1a shows a plot of simulated BP reductions and the calculated TP ratio and SI. From the graph of the BP reductions, we can see how these are negligible BP reductions but very homogeneous ones. As the TP ratio is defined as a ratio of two BP reductions (trough over peak reduction), the index can, in principle, give very high values even though there is almost no reduction at all, as long as the reductions at peak and at trough are very close. This is precisely what happens in this example. The TP ratio is very close to one (TP = 0.943), as the reduction in BP at trough is very close to the reduction at peak. The SI also has this problem, as the index incorporates two effects (i.e. mean BP reduction and reduction homogeneity) in a single number. Given a high SI, we cannot determine whether it is due to a considerable BP reduction or whether the reduction is very homogeneous. From the definition of $SI(\overline{d}/s_d)$, we can see that in the case of a nearly constant reduction, SI would be very large, regardless of how significant the BP reduction is; Fig. 1a illustrates this problem. Even though the antihypertensive treatment has no reduction, the SI results in a very high value (SI = 13.463) because of the homogeneity of this reduction, indicating a good overall drug effect. Note, however, how the SIn provides a more accurate assessment of the drug efficacy (SIn = 0.480).

Figure 1b shows an example of good antihypertensive coverage (real ABPM data). In this case, the drug induces a significant BP reduction with 24 h coverage. According to the TP and the SI indices, the treatment efficacy in this case, however, would be lower than that for the treatment in Fig. 1a. The SIn indicates a considerable difference between both treatments, providing a better assessment of drug efficacy.

The above example illustrates a potential limitation of the TP ratio and the SI. These indices are sensitive to treatment efficacy but they are not specific. That is, they give high values in cases when the antihypertensive drug causes a considerable and homogeneous BP reduction (Fig. 1b), but they can also give high values in situations when the BP reduction is inadequate (Fig. 1a). A simple solution to correct this potential limitation is to normalize the SI (SIn), which results in a better estimate of the combined reduction/homogeneity effect.

Figure 2 shows an example that demonstrates the advantages of the RDH index. Figure 2a shows the BP profiles before and after treatment. The methodology used in the RDH calculation is shown in Fig. 2b. The upper part of this plot shows a bar graph of the BP reductions in units of standard errors, $\mathbf{r}(\mathbf{r} = \mathbf{d}/\widehat{\mathbf{se}}_d)$. This plot enables one to assess the significance of a given BP reduction. The plot also shows the threshold values to determine significance (1.645 and 1.960). The bottom plot shows a graph of $\mathbf{r}_T(\mathbf{r}_T = \mathscr{I}(\mathbf{r} > 1.645))$, which is used to assess the number and location of the significant BP reductions and to compute the RDH index. In this example, TP = 0.875 and SI = 0.833, which incorrectly indicate an adequate treatment efficacy. The RDH index, RDH = (4, 2, 12), however, provides a more appropriate and complete assessment of treatment efficacy. The RDH index tells us that there were significant reductions in only four out of the 24 h, that the maximum number of consecutive significant reductions was 2, and that there were 12 h with consecutive non-significant reductions. Figure 2d shows the non-parametric 90% CIs for the BP reductions computed using bootstrap and the non-parametric estimate of the RDH index $RDH^* = (6, 2, 8)$. The main advantage of the non-parametric RDH (RDH*) index computed based on bootstrap is that it



Comparison of the trough-peak ratio (TP), smoothness index (SI), and the proposed normalized SI (SIn). Negative *d* values indicate reduction in blood pressure (BP) (note the plot is showing -d). (a) Plot of the BP hourly reductions of an antihypertensive drug. Note that despite the fact that the BP reduction is negligible (d=0.4 mmHg), both the TP ratio and the SI result in high values (TP=0.943, SI=13.463), which indicates an adequate antihypertensive coverage. The SIn provides a more accurate assessment of the antihypertensive coverage (SIn=0.480). (b) Plot showing the BP reductions of an other antihypertensive drug. In this case, the reduction is considerable and relatively homogeneous; however, note that the TP ratio and the SI in (b) are lower than in (a). The SIn in (b) results in a much higher value than in (a), which is a more appropriate assessment of the antihypertensive effect.

allows us to establish statistical significance without making any assumptions about the distribution of the BP reductions. In cases where the BP reductions follow a normal distribution the RDH and RDH* coincide. In this study, we observed that both RDH and RDH* give similar results. This is due to the fact that the BP differences follow a close-to-normal distribution. Figure 2d shows an example of the distribution of the BP differences used in the computation of the non-parametric CI in a given category.

In Fig. 3, we show another example illustrating the methodology. In this case, the antihypertensive drug provides a good coverage, as can be seen from the graph of



Illustration of the reduction-duration-homogeneity (RDH) index and methodological comparison between the trough-peak (TP) ratio, smoothness index (SI), normalized SI (Sin), and RDH in the assessment of antihypertensive drug effect. (a) BP means $(\bar{x} \text{ and } \bar{y})$ for each hour before (light gray) and after treatment (dark gray). Time 0 corresponds to the awakening time and time of drug intake. (b) Illustration of the BP reductions in units of standard error. The upper plot shows a bar plot of the **r** vector ($\mathbf{r} = d/\widehat{se_d}$) and thresholds at 1.645 and 1.960. The graph at the bottom shows the vector $\mathbf{r}_T (\mathbf{r}_T = I(\mathbf{r} > 1.645))$, the times at which significant reductions took place. In this case, the TP=0.875, SI=0.833, SIn=0.652, and RDH=(4, 2, 12). On the basis of the TP and SI, we would conclude a good antihypertensive drug coverage; however, the RDH indicates that in only two out of the 24 h there was a significant reduction, that the maximum coverage was of 2 h, and that there were 12 h without coverage. (c) The upper plot shows the mean and bootstrap confidence intervals (CIs) for the reduction in BP. The bottom plot shows the results of performing a non-parametric RDH (RDH*) results (6, 2, 8). (d) Example of the histogram of the BP bootstrap reductions used to determine the CI corresponding to the 11th hour.

BP reductions. The TP ratio, however, is lower than that in Fig. 2. On the other hand, the RDH index appropriately establishes the much better antihypertensive coverage of the drug in this case, TP = 0.841, SI = 3.03, SIn = 2.66, RDH = (21, 14, 1), and RDH^* (23, 22, 1)).

Discussion

The TP and SI are indices commonly used to assess the duration of action of an antihypertensive drug. Both, however, have significant limitations. Specifically, both indices are based on the BP reduction comparing two profiles, but do not assess the statistical significance of these reductions [6]. Indices like TP or SI are usually computed for a given patient. An additional problem is how to extrapolate the obtained indices (TP, SI) to one population or group [5]. The intra-individual and

interindividual variability of the RDH should be studied as in the case of the TP [30].

In this methodological paper, we introduced the RDH, a new index for assessment of antihypertensive treatment reduction and homogeneity with ABPM for a given individual. The application of the RDH index to hypertensive populations to assess antihypertensive drug efficacy still requires further research.

We illustrated the potential usefulness of the SIn and the RDH indices using several examples. The RDH tests the statistical significance of the BP reduction by means of parametric or non-parametric techniques. The use of one or another method depends on the sample size for each category. The parametric test assumes normality of data



Illustration of the RDH index and methodological comparison between the TP ratio, SI, Sin, and RDH in the assessment of antihypertensive drug effect. (a) BP reductions in standard error units (top) and vectors indicating significant reductions (bottom). In this case, TP=0.841, SI=3.03, SIn=2.66, RDH=(21, 14, 1), and RDH* (23, 22, 1). Note the TP is less in this case than that for the treatment shown in Fig. 2. The RDH index clearly distinguishes between the two, providing a more accurate assessment of drug efficacy. (b) The mean and bootstrap CI for BP reductions (top) and plot of statistically significant reductions (bottom).

on each time class. ABPM sampling recommendations usually result in three or four data on each class with 24-h ABPM or 6-8 data in the case of 48-h ABPM. Sometimes one or more of these samples is missing, resulting in a smaller sample size. Thus, in most applications, the hypothesis of normality cannot be assumed. In these situations, we recommend using the RDH*. This method, based on bootstrap, does not assume normality or any other distribution. Having a small number of samples significantly widens the CI and it is therefore more difficult for a reduction to be statistically significant. Any statistically significant reduction obtained with a small number of samples, however, is valid and statistically sound. Having a small number of samples affects only the type II error and not the type I error. Thus, we can conclude that the RDH index is

conservative in this regard. If the sample size is very small (four samples or less), the number of possible replications could be insufficient to obtain appropriate CIs. To solve this problem, we recommend overlapping the time classes as currently done in the calculation of the TP, where both the peak and the trough are estimated by averaging two classes. For instance, by using 2-h classes with 1-h of overlap, the sample size on each category is duplicated. The overlapping procedure has another advantage. Two consecutive classes share half of the data, so the results of the statistical test will tend to be more similar. This gives some kind of continuity to the statistical significance of the tests along consecutive time classes. This prevents situations in which we obtain an isolated nonsignificant class within several significant classes. It is, however, important to emphasize that when classes are overlapped, the class-by-class statistical tests will not be independent. An adjustment/correction procedure for multiple comparisons may be employed in these situations.

The RDH provides information about the total number of classes in which BP reduction is statistically significant and about the maximum number of consecutive significant and non-significant reductions. Additional information can be incorporated into the index, such as the temporary location of those intervals of consecutive significant and non-significant reductions. These time intervals could be used, for instance, to evaluate times in which medication is effective or to determinate the lapse of time between the drug intake and the BP reduction. Moreover, these time intervals could be used to individualize the medication for each patient and to evaluate the patient's response to treatment. The most significant advantages of the RDH* are as follows: (1) it provides a complete description (graphical and numerical) of the reduction, homogeneity, and duration (coverage) of an antihypertensive drug; (2) it is based on nonparametric statistics; (3) it can be used to determine time periods without coverage; (4) it can be used to assess antihypertensive efficacy in situations when several antihypertensive drugs are combined (polytherapy); (5) it can be used to study the treatment effect in selected time intervals. The RDH should be viewed as a possible complementary index to SI and TP for evaluating BP coverage offered by antihypertensive drugs. The SI has been reported to be associated with changes in left ventricular mass, as well with changes in carotid artery wall thickness [19,20]. The potential correlations of RDH or SIn with organ damage should be investigated.

References

- Elliott H, Meredith PA. Methodological considerations in calculation of the trough:peak ratio. J Hypertens 1994; 12:S3–S7.
- 2 Lipicky R. Trough:peak ratio: the rationale behind the United States Food and Drug Administration recommendations. *J Hypertens* 1994; 12:S17–S19.

- 3 Omboni S, Parati G, Zanchetti A, Mancia G. Calculation of trough:peak ratio of antihypertensive treatment from ambulatory blood pressure: methodological aspects. *J Hypertens* 1995; 13:1105–1112.
- 4 Meredith PA. Trough:peak ratio and smoothness index for antihypertensive agents. *Blood Press Monit* 1999; 4:257–262.
- 5 Omboni S, Parati I, Mania G. The trough:peak ratio and the smoothness index in the evaluation of control of 24 h blood pressure by treatment in hypertension. *Blood Press Monit* 1998; **3**:201–204.
- 6 Zannad F, Radauceanu A, Parati G. Trough-to-peak ratio, smoothness index and morning-to-evening ratio: why, which and when? J Hypertens 2003; 21:851–854.
- 7 Stergiou GS, Efstathiou SP, Skeva II, Baibas NM, Roussias LG, Mountokalakis T. Comparison of the smoothness index, the trough:peak ratio and the morning:evening ratio in assessing the features of the antihypertensive drug effect. J Hypertens 2003; 21:913–920.
- 8 Sega R, Corrao G, Bombelli M, Beitrame L, Facchetti R, Grassi G. Blood pressure variability and organ damage in a general population: results from the PAMELA study. *Hypertension* 2002; **39**:710–714.
- 9 O'Brien E, Cox JP, O'Malley K. Ambulatory pressure measurement in the evaluation of blood pressure lowering drugs. *J Hypertens* 1989; 7:243–247.
- 10 Coats AJS, Radaelli A, Clark SJ, Conway J, Sleight P. The influence of ambulatory blood pressure monitoring. Design and interpretation of trials in hypertension. J Hypertens 1992; 10:385–391.
- 11 Mallion JM, Baguet JP, Mania G. European Society of Hypertension Scientific Newsletter: update on hypertension management. *Eur Soc Hypertens* 2003; 4:1–2.
- 12 Parati G, Ravogli A, Mutti E, Santucciu C, Omboni S, Mancia G. Ambulatory blood pressure monitoring in the evaluation of antihypertensive drugs. *J Hypertens* 1994; 12:S9–S15.
- 13 Coats AJ. Benefits of ambulatory blood pressure monitoring in the design of antihypertensive drug trials. *Blood Press Monit* 1996; 1:157–160.
- 14 Zanchetti A. Twenty-four-hour ambulatory blood pressure evaluation of antihypertensive agents. J Hypertens 1997; 15:S21–S25.
- 15 Omboni S, Fogari R, Palatini P, Rappelli A, Salvetti A, Mancia G. Reproducibility and clinical value of the trough-to-peak ratio of the antihypertensive effect: evidence from the SAMPLE study. *Hypertens* 1998; 16:733–738.
- 16 Aboy M, Fernandez R, Hermida RC. Methodological considerations in the evaluation of the duration of action of antihypertensive therapy using ambulatory blood pressure monitoring. *Blood Press Monit* 2005; 10:111–115.

- 17 FDA. Proposed guidelines for the clinical evaluation of antihypertensive drug products. FDA Division of Cardio-Renal Drug; 1988.
- 18 Parati G, Rizzoni D, Omboni S, Agabiti-Rosei E, Mancia G. Smoothness index but not trough:peak ratio estimates balanced 24-hour blood pressure control and predicts regression organ damage by antihypertensive treatment. J Hypertens 1997; 15:S7.
- 19 Parati G, Omboni S, Rizzoni D, Agabiti R, Mancia G. The smoothness index: a new, reproducible and clinically relevant measure of the homogeneity of the blood pressure reduction with treatment for hypertension. *J Hypertens* 1998; **16**:1685–1691.
- 20 Rizzoni D, Muiesan ML, Salvetti M, Castellano M, Bettoni G, Monteduro C. The smoothness index, but not the trough-to-peak ratio predicts changes in carotid artery wall thickness during antihypertensive treatment. J Hypertens 2001; 19:703–711.
- 21 Efron B, Tibshirani RJ. An introduction to the bootstrap. London: Chapman & Hall; 1993.
- 22 Hermida RC, Ayala DE, Fernández JR, Mojón A, Alonso I, Calvo C. Modeling the circadian variability of ambulatory monitored blood pressure by multiplecomponent analysis. *Chronobiol Int* 2002; **19**:461–481.
- 23 Committee Guidelines. Guidelines Committee. 2003 European Society of Hypertension–European Society of Cardiology guidelines. *Hypertens* 2003; 21:1011–1053.
- 24 O'Brien E, Staessen J. Normotension and hypertension as defined by 24-hour ambulatory blood pressure monitoring. *Blood Press* 1995; 4:266–282.
- 25 Hermida RC, Calvo C, Ayala DE, Domínguez MJ, Covelo M, Fernández JR, et al. Administration-time-dependent effects of valsartan on ambulatory blood pressure in hypertensive subjects. *Hypertension* 2003; 42:283–290.
- 26 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the SpaceLabs 90207 determined by the British Hypertension Society protocol. J Hypertens 1991; 9:573–574.
- 27 Mansoor GA, White WB, McCabe EJ, Giacco S. The relationship of electronically monitored physical activity to blood pressure, heart rate, and the circadian blood pressure profile. *Am J Hypertens* 2000; **13**:262–267.
- 28 Hermida RC, Calvo C, Ayala DE, Mojón A, López JE. Relationship between physical activity and blood pressure in dipper and non-dipper hypertensive patients. J Hypertens 2002; 20:1097–1104.
- 29 Staessen J, Fagard R, Lijnen P, Thijs L, van Hoof R, Amery A. Ambulatory blood pressure monitoring in clinical trials. J Hypertens 1991; 9:s13-s19.
- 30 Staessen J, Bieniaszewski L, Buntinx F, Celis H, O'Brien E, Hoof R, et al. The trough-to-peak ratio as an instrument to evaluate antihypertensive drugs. *Hypertension* 1995; 26:942–949.