HAL
open science

A semi-automated Computer Tool for the Analysis of Retinal Vessel Diameter Dynamics

Guillaume Euvrard, Olivier Genevois, Isabelle Rivals, Pascale Massin, Amélie Collet, José-Alain Sahel, Michel Paques

▶ **To cite this version:**

Guillaume Euvrard, Olivier Genevois, Isabelle Rivals, Pascale Massin, Amélie Collet, et al.. A semi-automated Computer Tool for the Analysis of Retinal Vessel Diameter Dynamics. Computers in Biology and Medicine, 2013, 43, pp.513-523. hal-00805371

HAL Id: hal-00805371

https://espci.hal.science/hal-00805371v1

Submitted on 27 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A semi-automated Computer Tool for the Analysis of Retinal Vessel Diameter Dynamics

Guillaume Euvrard[1,3], Olivier Genevois[2], Isabelle Rivals[3], Pascale Massin[4], Amélie Collet[5], José-Alain Sahel[5] and Michel Paques[5]

December 2012

[1]Anagos, 55 rue des petites écuries, 75010 Paris, France

[2]Centre Hospitalier Charles Nicolle, 1 rue Germont, 76000 Rouen, France

[3]Équipe de Statistique Appliquée, ESPCI ParisTech, 10 rue Vauquelin, 75005 Paris, France

[4]Centre Hospitalier Lariboisière, Université Paris 7, 2 rue Ambroise Paré, 75010 Paris, France

[5]Clinical Investigation Center 503, Centre Hospitalier National des Quinze-Vingts, INSERM & Université Pierre et Marie Curie-Paris 6, Paris, France.

Correspondence should be addressed to:

Guillaume Euvrard, Anagos, 55 rue des petites écuries, F75010 Paris, France

Phone: +33 148016590

Email: guillaume.euvrard@wanadoo.fr

Abstract:

Retinal vessels are directly accessible to clinical observation. This has numerous potential interests for medical investigations. Using the Retinal Vessel Analyzer, a dedicated eye fundus camera enabling dynamic, video-rate recording of micrometric changes of the diameter of retinal vessels, we developed a semi-automated computer tool that extracts the heart beat rate and pulse amplitude values from the records. The extracted data enabled us to show that there is a decreasing relationship between heart beat rate and pulse amplitude of arteries and veins. Such an approach will facilitate the modeling of hemodynamic interactions in small vessels.

Keywords:

Retinal Vessel Analyser; Pulse Amplitude; Heart Beat Rate, Signal Processing.

1 Introduction

The retinal vascular network is the only microcirculatory network that can be thoroughly observed in a noninvasive fashion. In most subjects, spontaneous pulsation of retinal arteries and/or veins can be detected on fundus examination. These pulsations result from the cyclic nature of cardiac output which propagates a pulse wave throughout the vascular tree at each systole. Passive vessel pulse associates longitudinal pulse (i.e. variation in length and/or tortuosity, (1)) and transversal pulse (i.e. variations in diameter). In retinal vessels, it is likely that the systolodiastolic variations of vessel diameter, termed here *pulse amplitude* (PA), result from the net effect of the interaction of variations of transmural pressure (which is proportional to the difference between intraluminal and intravitreal pressure) and retinal vessels compliance. Hence, it can be hypothesized that measuring the PA is likely to provide potential cues about several parameters related to ocular or general circulation, in particular regarding vessel compliance and intraluminal pressure. For instance, arteriolosclerosis increases the stiffness of arterioles; the latter should shift to the right the pressure-diameter relationship and the question remains opened, whether the macrocirculation compensates or not such changes.. Also, it is conceivable that luminal pressure modulates the PA. This may explain why venous pulse is not detectable in most patients with increased intracranial

pressure (2). Yet, the quantitative relationship between the PA and hemodynamic factors remains to be investigated.

Initial studies of retinal vessel pulse used fundus photographs synchronized with the electrocardiogram and manual determination of vessel diameter (3). Recently, an automated, on-line approach of the PA of retinal vessel was made possible through computerized analysis of videorecordings of the fundus (4). The Retinal Vessel Analyzer® (RVA; Imedos, Jena, Germany) indeed records the temporal evolution of the vascular segment diameter. In order to perform a quantitative study of the PA, we have developed a computer tool that extracts the PA and *heart beat rate* (HBR) sequences from the RVA signal.

Several studies used the RVA signal to investigate the vessel stiffness and the HBR. In (5), the authors suggest stimuli to apply to the patient eyes during the measurements, like flicker stimulation or systemic hyperoxia, with the goal to analyze their influence on the resulting signals. Another approach is to assess the pulse delay between vein and artery signals in order to estimate the retinal pulse wave velocity, as a measure of the vessel rigidity. Using this approach, it has been possible to correlate the vessel rigidity with glaucoma damage (6), vasospastic propensity (7) or blood pressure (8). The pulse delay assessment implies a Fourier signal decomposition whose first peak corresponds to the HBR. Furthermore, in (8), the authors considered the mean and scattering of the PA of each vessel and tested their correlations with other parameters. But to our knowledge, the HBR and PA have only been considered as two aggregate values for a whole signal. Here, our approach results in assessing their temporal evolution, allowing to investigate whether their temporal variations are correlated or not.

Section 2 is a presentation of the computer tool and of its algorithms, while section 3 is a description of some experimental results we obtained from a PA/HBR relationship investigation.

2 Methods

2.1 Subjects

This study followed the principles of the Declaration of Helsinki and was approved by an Ethics Committee. Sixteen healthy subjects older than 18 years with at least 20/20 vision and a normal fundus were considered for the study. Each of them received full oral and written information and gave written consent prior to inclusion. Before RVA examination, participants were instructed to restrain from coffee, cigarette and alcohol consumption, as

well as from physical exercise for 12 hours. Then, routine ophthalmological examination was performed including a medical history, best-corrected visual acuity testing with manifest refraction, intraocular pressure (IOP) measure by applanation tonometry, slit lamp biomicroscopy and fundoscopy. Arteriovenous (AV) ratio was calculated using the built-in software of the RVA. For pupil dilation, topical tropicamide was applied on the examined eye.

2.2 Signal acquisition

Details on the RVA technology have been published earlier, by other authors ((4), (5)). Schematically, it consists in the association of a classical fundus camera (FF450, Carl Zeiss GmbH, Germany) which records the fundus image through a videorecorder, and a built-in software which performs the image analysis. The illumination light of the fundus camera is reflected by the different layers of the retina and by the retinal vessels before reaching the camera (charge-coupled device). The basic principle relies on the specific optical properties of the hemoglobin within the retinal vessels, which absorbs light at a maximum wavelength of 400-620 nm whereas the surrounding tissues mostly reflect it. Thus, when a green filter is inserted between the white light source and the retina, the images are very contrasted, with a high brightness difference between the hemoglobin and the surrounding tissues (figure 1).

The build-in software of the RVA makes use of this contrast to locate the vascular vessels on the images. It allows the user to select vascular segments and then tracks these segments in the successive images (25 images per second) and assesses their diameters. The successive diameter measurements are then exported as Excel datasheets (Microsoft Corporation, Redmond, WA) for off-line analysis.

Since the image scale of each eye is unknown, the diameter values are expressed in relative units (RU). If the examined eye has the dimensions of the normal Gullstrand eye, this unit corresponds to the micrometer.

In our study, the selected vascular segments were an arterial and a venous segment, approximately 500 µm long and 100 µm large (figure 1), whose diameter temporal evolution has been recorded during one or two minutes for each subject. We focused on the spontaneous changes of HBR and PA and did not use external stimulation such as flicker light that has been suggested by other authors (5) and is made available by the RVA.

2.3 Signal decomposition

An individual record contains two signals, corresponding to the two selected vessel segments. Figure 2 shows an example of observed record: both artery and vein signals clearly exhibit cardiac pulsations. The tool that we present analyses each record and divides it into a succession of pulsations, for which a duration (and hence a corresponding heart beat rate HBR) and a pulse amplitude PA for both vessels are estimated (figure 3). Thus, each individual record results in a sequence of cycle HBR, arterial PA and venous PA.

A signal can be represented as a sequence of values (t_i, y_i) for the time and the corresponding vessel diameter. It has some missing and spurious data. The process leading to the signal division runs through three steps: first, the spurious data are identified and removed; second, the signal is decomposed into multi-resolution signals; and third, this decomposition is used to delimit the cardiac cycles.

2.4 Artifact identification

The first task is to detect and remove artifacts from the signal. Indeed, the RVA recorder generates spurious data, for example during loss of fixation by the patient. Furthermore, these artifacts are sometime grouped: their occurrence then results in a succession of outliers.

The algorithm we developed to detect artifacts is based on a basic idea: around each signal data (t_i, y_i), we build a local model of the signal and estimate the local residual standard deviation σ, which characterizes the measurement noise. Then we compare the model residual at point t_i to σ: if it is significantly greater than σ, we decide that the observation (t_i, y_i) is spurious (figure 4).

More precisely, for a data point (t_i, y_i), we build a local model to fit the training examples $\{(t_j, y_j), j \in J\}$ of the set

$$J = \left\{ j / r \le \left| t_j - t_i \right| \le R \right\},$$

where r and R are predefined parameters. The constraint that $|t_j-t_i|\le R$ ensures that the resulting model is local, while the constraint that $|t_j-t_i|\ge r$ ensures that the i'th point and its neighbors are unused in the model construction. Indeed, if the i'th point is an artifact, the neighbors are likely to be also artifacts and should therefore be excluded as well. The local model is a function $f(t, \boldsymbol{\theta})$ whose parameter vector $\boldsymbol{\theta}$ is set to the least squares value $\boldsymbol{\theta}_{\mathbf{ls}}$, the one the minimizes the total squared error

$$L(\boldsymbol{\theta}) = \sum_{j \in J} \left(y_j - f(t_j, \boldsymbol{\theta}) \right)^2 \qquad (1)$$

In our program, the functional form is a two-phase linear function:

$$f(t,\boldsymbol{\theta}) = \left| \begin{array}{ll} a_1 + b_1 t & \text{if} \quad t \le t_c \\ a_2 + b_2 t & \text{if} \quad t > t_c \end{array} \right.$$
$$\text{with} \quad \boldsymbol{\theta} = \begin{pmatrix} a_1 & b_1 & a_2 & b_2 \end{pmatrix}^t \tag{2}$$
$$t_c = -\frac{a_2 - a_1}{b_2 - b_1}$$

This model is piecewise linear in its parameters and the parameter vector $\boldsymbol{\theta}_{\mathbf{ls}}$ is obtained from the ordinary least squares formula: we call **y** the column vector of the training example outputs $\mathbf{y} = (y_j, j \in J)$; we show in the appendix how to build an experience matrix **x** from the training inputs $(t_j, j \in J)$ such that the model output to the training examples is $\mathbf{x}\boldsymbol{\theta}$. Then the total squared error $L(\boldsymbol{\theta})$ is the quadratic norm $||\mathbf{y}-\mathbf{x}\boldsymbol{\theta}||^2$ of the residual vector and its minimum is found by setting its gradient with respect to $\boldsymbol{\theta}$ to zero. When the matrix **x** is full rank, it results in:

$$\boldsymbol{\theta}_{\mathbf{ls}} = \left(\mathbf{x}^{\mathbf{t}}\mathbf{x}\right)^{-1}\mathbf{x}^{\mathbf{t}}\mathbf{y} \tag{3}$$

We can now define a confidence interval for Y_i in order to decide whether the example (t_i, y_i) follows the same distribution as the examples $\{(t_j, y_j)\ ,\ j \in J\}$ (see (9) and the illustration on figure 4). In order to distinguish this interval for a *random* variable Y_i, from the classical confidence interval for a *non random* value such as a mathematical expectation or a standard deviation, we call it a "prediction interval" (10).

If the model contains the "real" regression function and if the noise of Y is Gaussian with a uniform variance σ^2, that is if there exists a parameter value $\boldsymbol{\theta}_{\mathbf{0}}$ and a normal random variable W with zero mean and unit variance such that

$$\begin{array}{l} \forall t / r \le |t - t_i| \le R, \quad Y = f(t, \boldsymbol{\theta}_{\mathbf{0}}) + \sigma W \\ W \to N(0,1) \end{array} \tag{4}$$

then the model value $f(t_i,\boldsymbol{\theta}_{\mathbf{ls}})$ is a normal random variable whose expectation and variance are

$$\begin{array}{lcl} E\left(f(t_i, \boldsymbol{\theta}_{\mathbf{ls}})\right) & = & f(t_i, \boldsymbol{\theta}_{\mathbf{0}}) \\ Var\left(f(t_i, \boldsymbol{\theta}_{\mathbf{ls}})\right) & = & \sigma^2 \mathbf{x_i}\left(\mathbf{x}^{\mathbf{t}}\mathbf{x}\right)^{-1}\mathbf{x_i^t} \end{array} \tag{5}$$

Since Y_i does not belong to the training set, it is independent from $f(t_i,\boldsymbol{\theta}_{\mathbf{ls}})$ and their difference $Y_i - f(t_i, \boldsymbol{\theta}_{\mathbf{ls}})$ is normally distributed with a mean equal to the difference of the means and a variance equal to the sum of the variances. Therefore, the following random variable is normally distributed with zero mean and unit variance:

$$\frac{Y_i - f(t_i, \boldsymbol{\theta}_{\mathbf{ls}})}{\sigma\sqrt{1 + \mathbf{x_i}\left(\mathbf{x}^{\mathbf{t}}\mathbf{x}\right)^{-1}\mathbf{x_i^t}}} \to N(0,1) \tag{6}$$

Furthermore, the estimator S^2 of the variance σ^2 is χ^2 distributed:

$$S^2 = \frac{1}{n-p}\sum_{j\in J}\left(Y_j - f(t_j, \boldsymbol{\theta}_{\mathbf{ls}})\right)^2$$
$$\frac{(n-p)S^2}{\sigma^2} \rightarrow \chi^2(n-p) \tag{7}$$

where n is the number of training examples (the size of J) and p the number of independent parameters of the model. Considering the independence of these two random variables, their ratio defines a Student variable with n-p degrees of freedom:

$$\frac{Y_i - f(t_i, \boldsymbol{\theta}_{\mathbf{ls}})}{S\sqrt{1+\mathbf{x_i}\left(\mathbf{x^t x}\right)^{-1}\mathbf{x_i^t}}} \rightarrow Student(n-p) \tag{8}$$

We can hence compute a 1-α prediction interval:

$$Y_i = f(t_{i,}\boldsymbol{\theta}_{\mathbf{ls}}) \pm tinv_{n-p}\left(1-\frac{\alpha}{2}\right)\times S\sqrt{1+\mathbf{x_i}\left(\mathbf{x^t x}\right)^{-1}\mathbf{x_i^t}} \tag{9}$$

where $tinv_{n-p}$ is the inverse of the *Student(n-p)* cumulative distribution function. If the value y_i of Y_i lies in the interval, then the hypothesis that Y_i has the distribution (4) is accepted. Conversely, if y_i lies outside the prediction interval, the hypothesis is rejected with a risk of error α and the point (t_i, y_i) is considered as a spurious point.

The local model and the artifact identification are applied iteratively: first, all the signal points are tested and, for each individual test, all neighboring points are used to build the local model and the prediction interval. Then the spurious points are removed, and the algorithm is reiterated: only the remaining points are tested and, for each individual test, only the remaining points around are used as training examples to build the local model and the prediction interval. The points that are identified as spurious are removed and the process is reiterated until there is no residual spurious point.

All the points identified as artifacts are removed from the signal: the next steps of the analysis use signals without any of these points.

2.5 Multiresolution signal decomposition

The next step is to decompose the signals at the appropriate time scale. Indeed, the signal variations have other causes than heart beats, for example vasomotion or noise. In order to extract the cardiac pulsations, we use the signal dynamic properties and remove its variations that are too slow or too fast as compared to the HBR order of magnitude. We do this by using a multiresolution decomposition inspired by (11). The two timescales T_1 and T_2 are predefined

parameters such that $T_1 >> T_2$, (3 and 0.1 sec. as default values) and the signal is decomposed into these two scales and a residual according to:

$$\begin{array}{lll} \text{Scale 1} & \text{decomposition term} & d_1(t) = mean\left(y(\tau) \middle| |\tau - t| \le \frac{T_1}{2} \right) \\ & \text{scale 1 resolution signal} & y_1(t) = d_1(t) \\ & \text{residual term} & r_1(t) = y(t) - y_1(t) \\ \text{Scale 2:} & \text{decomposition term} & d_2(t) = mean\left(r_1(\tau) \middle| |\tau - t| \le \frac{T_2}{2} \right) \\ & \text{scale 2 resolution signal} & y_2(t) = y_1(t) + d_2(t) \\ & \text{residual term} & r_2(t) = y(t) - y_2(t) = r_1(t) - d_2(t) \end{array} \tag{10}$$

The decomposed form of the signal is (see figure 5)

$$y(t) = \underbrace{\overbrace{d_1(t)}^{y_1(t)} + d_2(t)}_{y_2(t)} + \underbrace{\overbrace{d_2(t) + r_2(t)}^{r_1(t)}}_{r_2(t)} \tag{11}$$

The first scale decomposition term $d_1(t)$ contains slow variations that can be seen at a time scale T_1 (3 sec.) or more; those are long term movements that are not caused by cardiac beat. The second scale resolution signal $d_2(t)$ contains the variations which, being faster, cannot be seen at the time scale T_1, but are still visible at the scale T_2 (0.1 sec.). This is the case of the heart beat signal. At least, the residual $r_2(t)$ is a short term movement that cannot be seen neither at scale T_1 nor T_2. It is mainly made of noise. This decomposition enables to isolate the second decomposition scale $d_2(t)$, which is the part of the signal we use to delimit the cycles.

2.6 Cardiac cycle delimitation and validation

After the artifacts have been removed and the signal decomposed at the appropriate timescales, the cardiac cycles can be delimited and validated, on both vein and artery signals (figure 3). As can be seen on figures 2 and 3, the cycles are much more noticeable on the venous than on the arterial signal. For this reason, we first set the periods of the first signal, to which we synchronize the second signal division, with a flexible time advance. For the same reason, the HBR of the period is estimated with the venous period duration, while the arterial division is only used to estimate the arterial PA.

The vein signal period delimitation is performed with the second scale decomposition term $d_2(t)$ (formula (12)). Two new parameters are defined, T_{min} and T_{max}, which are the minimum

and maximum values for a cycle duration. To start with, the period ending points are set as the local minima of $d_2(t)$: they are the time points t_{lim} such that

$$\forall t / |t - t_{\lim}| \leq T_{\min}, d_2(t) \geq d_2(t_{\lim}) \quad (12)$$

and the periods are the intervals between two such successive points. Then the algorithm tries to concatenate these periods. Each one is grouped with the next one, and a "shape validation procedure" evaluates whether the enlarged period is likely to correspond to a cardiac cycle or not. If yes, the enlarged period is accepted and the intermediate delimitation point is removed.

The arterial period synchronization implies another parameter, dt_{max} (0.32 sec. as default value), which is the maximal time delay of one signal on the other. For every vein period with ending point t_{lim}, we define a corresponding arterial period whose ending point is the minimum of the arterial second scale decomposition term $d_2(t)$ on the interval $[t_{lim}- dt_{max}, t_{lim}]$.

The last step is the period validation. Each period is evaluated: first, its duration is compared with the parameters T_{min} and T_{max}; second, the "shape validation procedure" is run. The periods that are not validated as cardiac cycles remain in the signal division, but are not taken into account for further statistics: their invalidation is generally the consequence of a locally too noisy or too flat signal, and their HBR and PA assessment would not be trustworthy.

The basic idea of the "shape validation procedure" is that a signal interval, when it corresponds to a cardiac cycle, consists more or less of an increasing phase followed by a decreasing one, with limited oscillations around this skeleton (figure 6). To implement it, the procedure models the second scale decomposition $d_2(t)$ as a two-phase linear regression (the model construction is described in the appendix), checks that the model has the expected variation signs and that its error is small enough as compared to the d_2 standard deviation. The procedure also checks that the second scale residual $r_2(t)$ has small variations as compared to $d_2(t)$.

Let $P=[t_{begin}, t_{end}]$ denote the evaluated period, $f(t,\boldsymbol{\theta}_{\mathbf{ls}})$ the two-phase linear model of the signal in the period and t_c the model changeover point. Let $min_P(y(t))$, $max_P(y(t))$, $mean_P(y(t))$, and $std_P(y(t))$ denote the extreme values, mean and standard deviation of signal $y(t_j)$ on the period P. The procedure validates the period if the following conditions are all fulfilled:

$$\begin{cases} C1: & \dfrac{std_P\left(r_2(t)\right)}{max_P\left(d_2(t)\right) - min_P\left(d_2(t)\right)} \le noise_max \\ C2: & \dfrac{\sqrt{mean_P\left(\left(d_2(t) - f(t,\theta_{ls})\right)^2\right)}}{std_P\left(d_2(t)\right)} \le error_max \\ C3: & \left[\begin{array}{l} f(t_c, \boldsymbol{\theta}_{\mathbf{ls}}) > f(t_{begin}, \boldsymbol{\theta}_{\mathbf{ls}}) \\ f(t_c, \boldsymbol{\theta}_{\mathbf{ls}}) < f(t_{end}, \boldsymbol{\theta}_{\mathbf{ls}}) \end{array} \right. \end{cases} \tag{13}$$

Condition C1 controls the data noise as compared to the pulse amplitude. Condition C2 controls how close the two-phase linear model is to the second scale decomposition term $d_2(t)$ and condition C3 checks that the changeover point t_c corresponds to a maximum. The bounds *noise_max* and *error_max* are parameters of the algorithm whose default values are 35% and 50%.One should notice that this procedure is invariant to homothetic transformation of the signal according to the time t and/or to the vessel width y.

Once the periods are set and validated, their HBR and PA can be estimated: the HBR of period P is inverse to the venous period duration t_{end}-t_{begin}, and the PA of each vessel is the amplitude $max_P(y(t))$-$min_P(y(t))$ of the second scale decomposition term during the period.

2.7 User defined modification

The algorithm presented above is automatic. Once its parameters are set, it directly gives a structure to the RVA signal: it removes the spurious data, decomposes the remaining signal at the appropriate time scales, divides it into periods and selects the ones which are likely to correspond to cardiac cycles. The algorithm parameters are recalled in table 1.

However, the user may disagree with some of the algorithm results and the program allows him to modify them manually. These corrections comprise changing the artifact status of any point, adding or deleting a period delimitation and forcing a cycle validation or invalidation. In particular, the artifact detection inevitably leads to false positives: if the test level is α, a proportion α of the non spurious points will be wrongly identified as spurious. Furthermore, the proportion of false negative is unknown but is non-zero and some spurious points may remain undetected. The manually modified result can then be saved into a file and reloaded in further sessions using the signal.

To limit the subjectivity induced by allowing to modify the algorithm result or not, we restricted ourselves to some artifacts undetected by the algorithm and to some period invalidation. The experimental results presented in the next section were obtained with only a few user defined modifications, and those respected the following rules:

- We felt free to change the artifact status of any point. But once the artifact list had been set, we kept it constant, whatever the results of the signal decomposition,
- The periods delimitations remained unmodified: no delimitation has been added, deleted nor shifted,
- However, we permitted ourselves a few periods invalidation: when there was a significant change of duration of two successive cycles, both have been rejected if there was uncertainty about their delimitation.

3 Experimental results

The tool described above has been applied to the individual RVA records obtained for the 16 subjects. The algorithm hence produced 16 series of validated pairs *(HBR, PA)* for both the arteries and the veins. The signal properties are not homogenous: the cyclic variations due to the heart beats are more easily identified for the veins than for the arteries, resulting in more validated periods (table 2).

We used these divisions to investigate whether the PAs and HBR are related or not. Each of the 16 records has been studied statistically for both vessels.

3.1 Individual modeling

To start with, we built individual models: for each record we computed:

- The linear correlations between the PAs and HBR,
- A p-value testing the absence of correlation between the variables (Pearson's test) against the alternative hypothesis of their negative correlation,
- a linear regression of the PAs on HBR and an estimation of its standard error.

The results are displayed on table 3 and the regressions are plotted on figure 7. For most subjects there was a negative correlation between PA and HBR for both arteries and veins, yet statistical significance ($p \leq 0.05$) was observed in only 1 subject.

A point worth mentioning from figure 7 is the disparity of the HBR and PA domains, which differ significantly from an individual to another. This observation suggests that PA is not a simple function of HBR, but also depends on individual specificities. The model standard

errors are shown on figure 8. When computed on all the records together, they equal 1.01 RU for the arteries and 1.14 RU for the veins. Each individual regression has 2 parameters, hence the overall modeling (one model per record) has 32 independent parameters for each vessel type.

3.2 Global modeling

In building the individual models, a significant negative correlation between PA and HBR appeared only for one individual. A reason may have been the small number of validated periods per individual, as compared to the correlation order of magnitude. To overcome this, we have concatenated all the records and built two large datasets, one for the arteries and one for the veins. In both cases, we computed the linear correlations, tested the correlation, built 2-degree regressions of PA on HBR and estimated their standard errors. Table 4 shows the results, figure 7 displays the two data sets and the corresponding second degree regressions and figure 8 shows the model standard errors.

These results show a significant negative correlation between the HBR and PA. However, the polynomial regressions are not monotonic and their standard errors are much larger than those obtained with the individual models. This may be consistent with the number of independent parameters, which is now of only 3 instead of 32 for each vessel type, but this model reduction does not take the individual specificities into account. For example, it can be seen from figure 7 (middle right) that some vein data points have a low HBR and a high PA. But it can also be seen, from figure 7 (top right), that they may correspond to records whose PA is globally very high. It is therefore difficult to know whether these high observed values have to be attributed to low HBR values or to individual effect.

3.3 Linear mixed effect modeling

In order to include both the individual and the HBR effects in the modeling, we have built, for each vessel type, a linear mixed effect (LME) model (12). A LME model has a random effect, here the individual record, and a fixed effect, here the HBR. In our case, for a data point *i* of the record *r*, the model output is

$$PA_{r,i} = a_r + b \times HBR_{r,i} + \varepsilon_{r,i} \tag{14}$$

The coefficient a_r is specific to the record *r*, it is the individual effect which is considered as random; the fixed effect *b* is common to all the records and the residuals $\varepsilon_{r,i}$ are independent identically distributed random variables.

Fitting the models to the data, by using the “restricted maximum likelihood” method (12), resulted in the parameters of table 5. The model plots are shown on figure 7 and their standard errors on figure 8.

The standard model errors are 1.03 RU for the arteries and 1.14 RU for the veins. They are significantly lower than the global model standard errors (1.72 RU and 2.53 RU), and nearly the same than the individual model standard errors (1.02 RU and 1.14 RU). The number of independent parameters lies between the two former modeling sizes: there is one common parameter b and a specific a_r for each record, that is 17 parameters for each vessel type. They have to be compared with 32 parameters in the case of individual modeling, and with 3 parameters in the global modeling. The LME modeling has less parameters than the individual modeling with the same performance, while the global modeling has less parameters than both LME and individual, but this gain is paid by an important loss of performance. The best modeling is clearly the LME. It is rich enough to take the individual disparities into account, and still allows a *HBR* term that is common to all.

Both for the arteries and for the veins, the multiplicative coefficient b of the *HBR* term is negative. Furthermore, we evaluated its significance by testing the null hypothesis $b=0$ (i.e. zero correlation) against the alternative hypothesis $b\neq0$. As shown in the last row of table 5, the p-value is 0.0006 for the arteries and 0.0002 for the veins: in both vessel types, the null hypothesis is clearly rejected and the coefficient b is significantly negative.

4 Discussion

We found a negative correlation between the heart beat rate and the pulse amplitude, in particular the venous one, i.e. the lower the HBR, the higher the pulse amplitude. The most important component of the variability of the HBR is the diastole, that it, the time during which the heart is at reast between two constrictions (the systoles). Hence, a lower heart beat rate means a longer diastole. During diastole, the venous diameter progressively decreases to reach a nadir immediately after the beginning of the systole. It is assumed indeed that the venous flow is continuous, and may even lead to a complete emptiness of the vein. The latter can indeed be observed around the optic disc. The depth of the nadir is therefore linked to the length of the diastole. Since the venous diameter during systole does not vary significantly, any change in length will modify the depth of the venous nadir and hence the pulse amplitude at the next systole. The fact that the vein may become completely empty during diastole indicates that venous flow does not depend only on the "pushing effect" of arteriolar flow, but that intraocular pressure probably also contributes to venous outflow. Future works on the

relationship between PA and HBR could lead to algorithms that could provide an estimate of intraluminal venous pressure, an important parameter in cardiology. For instance, it is expected that an increased venous pressure would weaken the relationship between PA and HBR. This will require to improve the quality of data and in particular adress the issue of the frequent artifacts.

5 Conclusion

We have proposed a method for semi-automated analysis of the variations of retinal vessel diameters which decomposes the RVA signal into a succession of pulsations corresponding to cardiac cycles. It allows to display the temporal changes of the pulse amplitude and heart beat rate. Using this tool, we have shown that, for healthy patients, there is a significant negative correlation between these two variables. However, significant individual disparities suggest that the pulse amplitude also depends on other hemodynamic factors which remain to be identified. The construction of a model linking hemodynamic factors, intraocular pressure and vessel diameters, may allow inferring hemodynamic values such as intraluminal pressure and/or vessel compliance. This emphasizes the potential medical importance of our approach.

6 References

1. **Kapuscinski, W.** Observations on the pulse and retinal arterial pressure. *Br. J. Ophtalmol.* 1948, 32, pp. 273-300.

2. **Levin, B E.** The clinical significance of spontaneous pulsations of the retinal vein. *Archives of Neurology.* 1978, 35, pp. 37-40.

3. **Chen, H C, et al., et al.** Vessel diameter changes during the cardiac cycle. *Eye.* 1994, 8, pp. 97-103.

4. **Seifert, B U and Vilser, W.** Retinal Vessel Analyser (RVA) - Design and Function. *Biomed Tech.* 2002, Vol. 47, Suppl. 1.

5. **Garhofer, Gerhard, et al., et al.** Use of retinal vessel analyzer in ocular blood flow research. *Acta Ophthalmologica.* 2010, Vol. 88, pp. 717-722.

6. **Oettli, Alexander, et al., et al.** Rigidity of Retinal Vessel in Untreated Eyes of Normal Tension Primary Open-angle Glaucoma Patients. *Journal Of Glaucoma.* 2011, Vol. 20, 5, pp. 303-306.

7. **Gugleta, Konstantin, et al., et al.** On Pulse-Wave Propagation in the Occular Circulation. *Investigative Ophthalmology & Visual Science.* 2006, Vol. 47, 9, pp. 4019-4025.

8. **Kotliar, Konstantin E., et al., et al.** Does Increased Blood Pressure Rather Than Aging Influence Retinal Pulse Wave Velocity? *Investigative Ophthalmology & Visual Science.* 2012, Vol. 53, 4, pp. 2119-2126.

9. **Seber, G. A. F.** *Linear regression analysis.* s.l. : John Wiley & Sons, 1977.

10. **Personnaz, L. and Rivals, I.** *Réseaux de neurones formels pour la modélisation, la commande et la classification.* s.l. : CNRS Editions, 2003.

11. **Mallat, Stéphane.** *A Wavelet Tour of Signal Processing.* 2nd. s.l. : Academic Press, 1999.

12. **Pinheiro, José C. and Bates, Douglas M.** *Mixed-Effects Models in S and S-PLUS.* New York : Springer Verlag, 2000.

Appendix: construction of a two-phase linear regression model

First of all we recall the least squares formula. Let **y** denote a $N\times 1$ vector, **x** a $N\times p$ matrix and $\boldsymbol{\theta}$ a $p\times 1$ unknown vector. If $p<N$ and if the rank of **x** is p, then the minimization in $\boldsymbol{\theta}$ of the quadratic function

$$\begin{aligned} L(\mathbf{x},\mathbf{y},\boldsymbol{\theta}) &= \sum_{j=1}^{N}\left(y_j - \sum_{i=1}^{p} x_{ji}\theta_i\right)^2 \\ &= \left\|\mathbf{y}-\mathbf{x}\boldsymbol{\theta}\right\|^2 \end{aligned} \tag{A1}$$

is obtained for the least squares value $\boldsymbol{\theta}_{\mathbf{ls}}$:

$$\boldsymbol{\theta}_{\mathbf{ls}} = \left(\mathbf{x}^{\mathbf{t}}\mathbf{x}\right)^{-1}\mathbf{x}^{\mathbf{t}}\mathbf{y} \tag{A2}$$

This solution is found by setting the gradient of (A1) with respect to $\boldsymbol{\theta}$ to $\mathbf{0}$.

Now, let (t_j, y_j), $j=1,\ldots,N$ be a set of training points such that

$$t_1 \le t_2 \le t_3 \le \ldots \le t_N \tag{A3}$$

and let $f(t,\boldsymbol{\theta})$ be the two-phase linear model

$$\begin{aligned} f(t,\boldsymbol{\theta}) &= \left| \begin{matrix} a_1 + b_1 t & \text{if} \quad t \le t_c \\ a_2 + b_2 t & \text{if} \quad t > t_c \end{matrix} \right. \\ \text{with} \quad \boldsymbol{\theta} &= \begin{pmatrix} a_1 & b_1 & a_2 & b_2 \end{pmatrix}^t \\ t_c &= -\frac{a_2 - a_1}{b_2 - b_1} \end{aligned} \tag{A4}$$

We look for the least squares parameter value $\boldsymbol{\theta}_{\mathbf{ls}}$ which minimizes the quadratic loss function $L(\boldsymbol{\theta})$:

$$\begin{aligned} L(\boldsymbol{\theta}) &= \sum_{j=1}^{N}\left(y_j - f(t_j,\boldsymbol{\theta})\right)^2 \\ \boldsymbol{\theta}_{\mathbf{ls}} &= \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, L(\boldsymbol{\theta}) \end{aligned} \tag{A5}$$

In this appendix, we show how to compute the least square parameter value $\boldsymbol{\theta}_{\mathbf{ls}}$. The basic idea is that, since the model $f(t,\boldsymbol{\theta})$ is piecewise linear in its parameter $\boldsymbol{\theta}$, the loss function $L(\boldsymbol{\theta})$ is piecewise quadratic with the form (A1). In particular, there is a parameter dimension p and a matrix **x** of size $N\times p$ such that $\boldsymbol{\theta}_{\mathbf{ls}}$ is the minimum of the quadratic function (A1) and can be expressed by (A2).

The changeover point t_c is either in the set $S^{(0)} = (-\infty, t_t] \cup [t_N, +\infty)$, or in a set $S^{(k)} = [t_k, t_{k+1}]$. We solve (A5) analytically on each set $S^{(k)}$ and denote $p^{(k)}$, $\boldsymbol{\theta}_{\mathbf{ls}}^{(\mathbf{k})}$, $\mathbf{x}^{(\mathbf{k})}$ and $L^{(k)}$ the corresponding parameter dimension, least square parameter, experience matrix and loss function value.

i. <u>Solution on $S^{(0)}$</u>

If $t_c \leq t_1$ or $t_c \geq t_N$, then the model is linear on $[t_1, t_N]$ (single-phase model):

$$\begin{aligned} f(t,\boldsymbol{\theta}) &= a + bt \\ &= \mathbf{x}\boldsymbol{\theta} \\ \text{with} \quad \mathbf{x} &= \begin{pmatrix} 1 & t \end{pmatrix} \\ \boldsymbol{\theta} &= \begin{pmatrix} a & b \end{pmatrix}^t \end{aligned} \tag{A6}$$

The loss function is

$$\begin{aligned} L(\boldsymbol{\theta}) &= \sum_{j=1}^{N} \left(y_j - f(t_j, \boldsymbol{\theta}) \right)^2 \\ &= \left\| \mathbf{y} - \mathbf{x}^{(0)} \boldsymbol{\theta} \right\|^2 \end{aligned} \tag{A7}$$

where $\mathbf{y}$ is the vector of the (y_j) and $\mathbf{x}^{(0)}$ is:

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_N \end{pmatrix} \tag{A8}$$

This loss function has the form (A1) and its minimum is hence given by (A2):

$$\boldsymbol{\theta}_{\mathbf{ls}}^{(0)} = \left(\mathbf{x}^{(0)t} \mathbf{x}^{(0)} \right)^{-1} \mathbf{x}^{(0)t} \mathbf{y} \tag{A9}$$

The optimal loss value $L^{(0)}$ is deduced from (A7) with $\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathbf{ls}}^{(0)}$ and the number $p^{(0)}$ of independent parameters is 2.

ii. <u>Solution on $S^{(k)}$ *with* $k \neq 0$</u>

If $t_k \leq t_c \leq t_{k+1}$, then the model is

$$\begin{aligned} f(t,\boldsymbol{\theta}) &= \mathbf{x}\boldsymbol{\theta} \\ \text{with} \quad \mathbf{x} &= \left| \begin{matrix} \begin{pmatrix} 1 & t & 0 & 0 \end{pmatrix} & \text{if } t \leq t_c \\ \begin{pmatrix} 0 & 0 & 1 & t \end{pmatrix} & \text{if } t > t_c \end{matrix} \right. \\ \boldsymbol{\theta} &= \begin{pmatrix} a_1 & b_1 & a_2 & b_2 \end{pmatrix}^t \end{aligned} \tag{A10}$$

The following constrained optimization has to be solved:

$$\begin{aligned} &\underset{\boldsymbol{\theta}}{Min} \quad L(\boldsymbol{\theta}) \\ &\text{subject to} \quad \left| \begin{matrix} t_c = \dfrac{a_1 - a_2}{b_2 - b_1} \geq t_k \\ t_c = \dfrac{a_1 - a_2}{b_2 - b_1} \leq t_{k+1} \end{matrix} \right. \end{aligned} \tag{A11}$$

First, the minimization is solved without constraint, and we check whether the solution satisfies the constraints or not. If yes, it is the solution of (A11). If not, the optimization is performed after saturation of one constraint and the other (both cannot be saturated simultaneously), and the solution with the lowest loss value is selected.

- Unconstrained problem

The loss function takes the form (A1) with the experience matrix $\mathbf{x}^{(k)}$

$$\mathbf{x}^{(k)} = \begin{pmatrix} 1 & t_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_k & 0 & 0 \\ 0 & 0 & 1 & t_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & t_N \end{pmatrix} \tag{A12}$$

Its solution is therefore obtained by applying (A2):

$$\boldsymbol{\theta}_{\mathbf{ls}}^{(k)} = \left(\mathbf{x}^{(k)t}\mathbf{x}^{(k)}\right)^{-1}\mathbf{x}^{(k)t}\mathbf{y} \tag{A13}$$

If it fulfills the constraints, it is the solution of (A11). The number $p^{(k)}$ of independant parameters is 4, the experience matrix is given by (A12) and the loss value by (A1).

- Constraint saturation

If the parameter value given by (A13) does not fulfill the constraints, the latter have to be saturated. Both cases $t_c = t_k$ and $t_c = t_{k+1}$ are considered, and the solution with the lowest error is kept. Since both are solved in the same way, we now explicit only the first case. The constraint implies

$$a_2 = a_1 - t_k(b_2 - b_1) \tag{A14}$$

By injecting (A14) in (A10), we get

$$\begin{aligned} f(t,\boldsymbol{\theta}) &= \mathbf{x}\boldsymbol{\theta} \\ \text{with} \quad \mathbf{x} &= \left| \begin{matrix} \begin{pmatrix} 1 & t & 0 \end{pmatrix} & \text{if } t \le t_k \\ \begin{pmatrix} 1 & t_k & t - t_k \end{pmatrix} & \text{if } t \ge t_k \end{matrix} \right. \\ \boldsymbol{\theta} &= \begin{pmatrix} a_1 & b_1 & b_2 \end{pmatrix}^t \end{aligned} \tag{A15}$$

The loss function then takes the form (A1) with the experience matrix $\mathbf{x}^{(k)}$

$$\mathbf{x}^{(k)} = \begin{pmatrix} 1 & t_1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & t_k & 0 \\ 1 & t_k & t_{k+1} - t_k \\ \vdots & \vdots & \vdots \\ 1 & t_k & t_N - t_k \end{pmatrix} \tag{A16}$$

and its solution is given by (A2):

$$\boldsymbol{\theta}_{\mathbf{ls}}^{(\mathbf{k})} = \left(\mathbf{x}^{(\mathbf{k})\mathbf{t}}\mathbf{x}^{(\mathbf{k})}\right)^{-1}\mathbf{x}^{(\mathbf{k})\mathbf{t}}\mathbf{y} \tag{A17}$$

This parameter vector, completed by the a_2 value from (A14), gives the solution of (A11) with the first constraint saturated. The number $p^{(k)}$ of independant parameters is 3, the loss value $L^{(k)}$ is deduced from (A1) with the experience of (A16). The same procedure is used to solve (A11) with the second constraint saturated, and the best of the two is the solution of (A11).

Note that the solution of (A11) on $S^{(k)}$ under the *second* constraint saturation may be used to solve (A11) on $S^{(k+1)}$ under the *first* constraint saturation.

iii. <u>Global solution</u>

The global solution of (A5) is merely obtained by considering all the solutions $\boldsymbol{\theta}_{\mathbf{ls}}^{(\mathbf{k})}$ on the subsets $S^{(k)}$ (k=0,…,N-1), and by selecting the one with the lowest loss value $L^{(k)}$. The parameter dimension p and experience matrix $\mathbf{x}$ are the corresponding $p^{(k)}$ and $\mathbf{x}^{(\mathbf{k})}$.

	Parameter	Description	Default value
Spurious point detection	r and R	Temporal radius for training examples selection	0.10 and 0.40 sec.
	α	Statistical test level	2%
Multiresolution decomposition	T_1 and T_2	Long and short term timescale	3.0 and 0.1 sec.
Period delimitation	T_{min} and T_{max}	Extrema of cycle duration	0.5 and 1.71 sec. (correspond to HBR extrema of 35 and 120 bpm)
	dt_{max}	Maximum time delay between arterie and vein	0.32 sec.
Period shape validation	*noise_max*	Maximum noise/PA ratio	35%
	error_max	Maximum model error/std ratio	50%

Table 1: synthetic view of the algorithm parameters.

Record	Artery				Vein			
	Periods	Validated periods	Mean HBR	SD HBR	Periods	Validated periods	Mean HBR	std HBR
1	116	99	72.5	7.0	102	101	72.2	7.5
2	118	25	77.5	13.5	124	62	72.7	12.6
3	91	31	80.2	12.6	90	43	78.0	14.2
4	124	97	65.0	6.6	115	98	65.8	6.8
5	112	29	60.6	6.7	108	104	60.3	5.7
6	133	26	84.8	12.5	134	53	84.9	12.0
7	42	5	54.4	2.3	42	42	53.6	3.8
8	63	19	87.2	11.5	60	36	87.7	11.0
9	55	28	72.7	8.6	55	44	73.6	10.0
10	49	13	71.0	9.8	53	51	69.4	6.8
11	44	18	62.7	5.4	43	43	62.2	5.9
12	68	5	90.6	4.7	66	60	90.6	7.3
13	78	35	86.1	5.7	75	72	86.0	5.8
14	51	24	62.4	4.8	47	47	62.8	4.7
15	108	14	102.1	10.9	106	101	99.2	9.0
16	36	14	56.3	13.5	40	30	55.1	11.0

Table 2: synthetic view of the 16 records (Mean HBR and SD HBR in beats per minute).

Record	Artery				Vein			
	Validated periods	Corr HBR/PA	p	Model std error	Validated periods	Corr HBR/PA	p	Model std error
1	99	-0.35	**<0.01**	0.81	101	-0.23	**0.02**	0.78
2	25	-0.32	0.12	1.09	62	-0.12	0.37	1.04
3	31	-0.02	0.92	0.63	43	-0.09	0.57	1.11
4	97	-0.07	0.47	0.64	98	-0.04	0.68	0.69
5	29	-0.35	0.06	0.83	104	-0.00	0.99	0.79
6	26	0.15	0.45	0.79	53	-0.05	0.72	1.24
7	5	0.20	0.75	0.50	42	-0.15	0.35	0.87
8	19	0.04	0.88	1.78	36	-0.17	0.33	1.41
9	28	-0.07	0.71	1.44	44	-0.08	0.59	1.12
10	13	-0.18	0.55	1.45	51	-0.11	0.46	1.90
11	18	-0.41	0.09	0.64	43	-0.24	0.13	1.00
12	5	-0.21	0.74	0.34	60	-0.12	0.37	1.36
13	35	-0.28	0.10	0.60	72	-0.15	0.21	0.57
14	24	-0.01	0.96	1.21	47	-0.23	0.12	1.14
15	14	0.09	0.75	1.92	101	-0.14	0.15	1.15
16	14	-0.47	0.09	2.29	30	-0.17	0.36	2.64

Table 3: correlations between HBR and PA, p-values of the independence test and standard error (in RU) of the individual linear regressions.

	Arteries	Veins
Regression formula	PA = 4.26 - 0.03 (HBR-72.76) + 0.0013 $(HBR-72.76)^2$	PA = 4.85 - 0.07 (HBR-74.45) + 0.0023 $(HBR-74.45)^2$
Regression standard error	1.72	2.53
Correlation	-0.09	-0.31
p	0.04	<0.0001

Table 4: results for the global modeling (HBR in beats per minute, PA and regression error in RU).

Record	Arteries		Veins	
	Model formula (*PA*:RU, *HBR*: bpm)	Standard error (RU)	Model formula (*PA*:RU, *HBR*: bpm)	Standard error (RU)
1	PA = 6.32 -0.019 *HBR*	0.82	PA = 5.12 -0.016 *HBR*	0.79
2	PA = 5.99 -0.019 *HBR*	1.09	PA = 4.57 -0.016 *HBR*	1.05
3	PA = 4.75 -0.019 *HBR*	0.67	PA = 5.32 -0.016 *HBR*	1.12
4	PA = 4.58 -0.019 *HBR*	0.64	PA = 4.05 -0.016 *HBR*	0.69
5	PA = 5.00 -0.019 *HBR*	0.85	PA = 5.68 -0.016 *HBR*	0.80
6	PA = 4.68 -0.019 *HBR*	0.87	PA = 5.91 -0.016 *HBR*	1.25
7	PA = 3.80 -0.019 *HBR*	0.55	PA = 11.62 -0.016 *HBR*	0.87
8	PA = 6.88 -0.019 *HBR*	1.80	PA = 6.61 -0.016 *HBR*	1.41
9	PA = 7.67 -0.019 *HBR*	1.44	PA = 6.93 -0.016 *HBR*	1.12
10	PA = 8.17 -0.019 *HBR*	1.45	PA = 12.51 -0.016 *HBR*	1.90
11	PA = 4.83 -0.019 *HBR*	0.66	PA = 7.86 -0.016 *HBR*	1.01
12	PA = 3.70 -0.019 *HBR*	0.39	PA = 7.07 -0.016 *HBR*	1.36
13	PA = 4.74 -0.019 *HBR*	0.60	PA = 4.28 -0.016 *HBR*	0.57
14	PA = 8.25 -0.019 *HBR*	1.21	PA = 9.82 -0.016 *HBR*	1.16
15	PA = 8.99 -0.019 *HBR*	1.96	PA = 6.46 -0.016 *HBR*	1.15
16	PA = 8.83 -0.019 *HBR*	2.49	PA = 11.21 -0.016 *HBR*	2.65
All	PA = a_r -0.019 *HBR*	1.03	PA = a_r -0.016 *HBR*	1.14
p-value(*b*=0)	0.0006		0.0002	

Table 5: the LME models and their performances. For each vessel type, the additive coefficients are specific to each record and only the coefficient *b* of the HBR term is common to all records. The last row is the result of testing the hypothesis *b*=0.

Figures

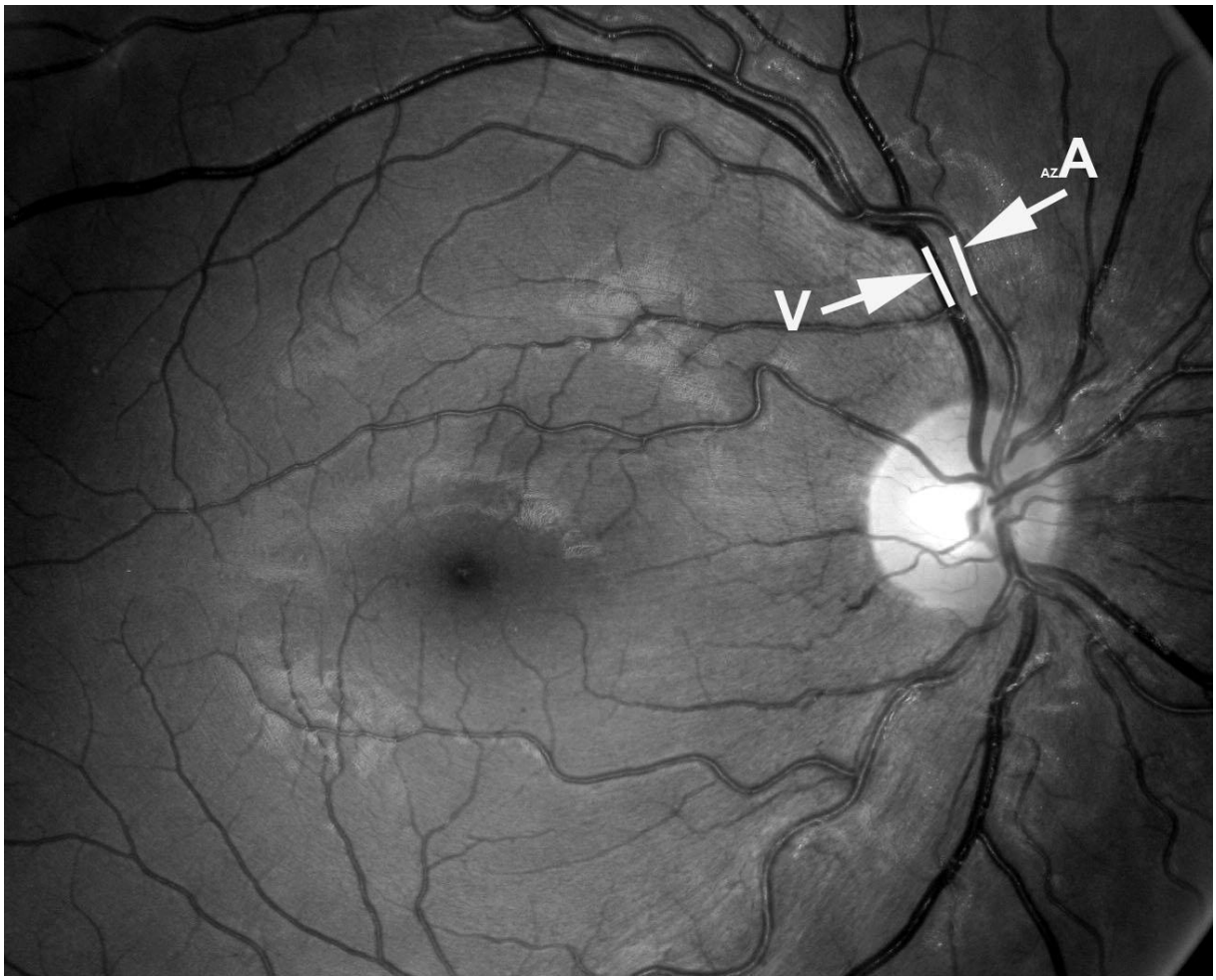

Figure 1: eye fundus image showing an example of two selected vascular segments whose diameters are measured by the RVA.

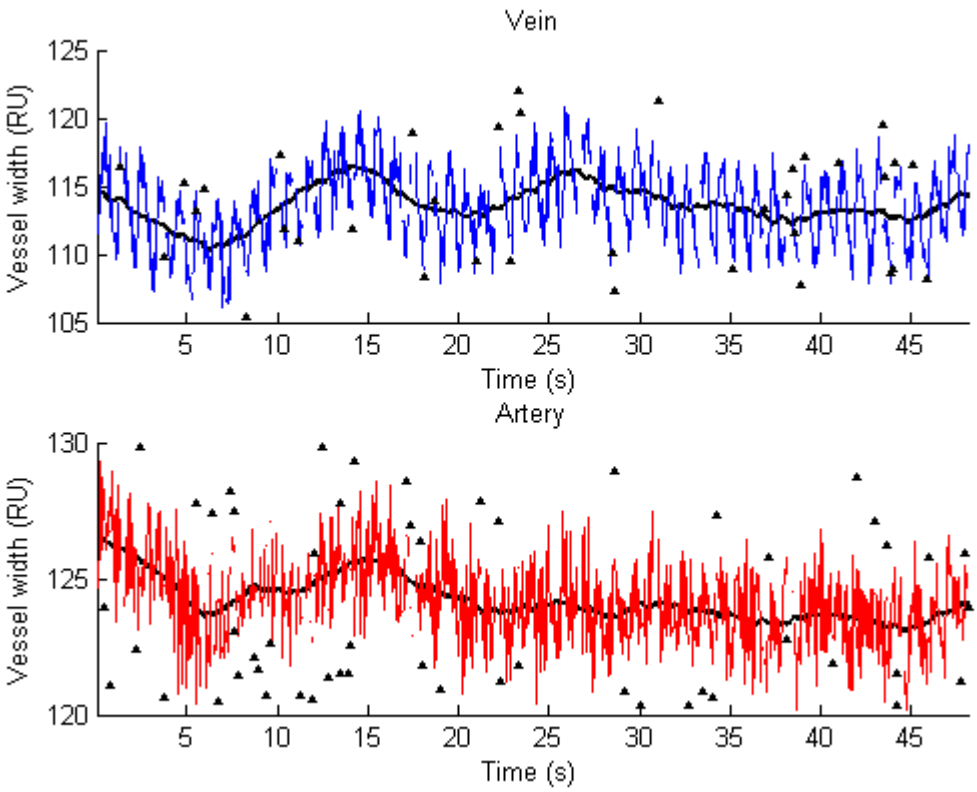

Figure 2: an example of RVA record, with a venous and an arterial signal. Spurious points (▲) are removed and a baseline (bold black line) is calculated. The heart beat pulsations appear clearly around the baseline.

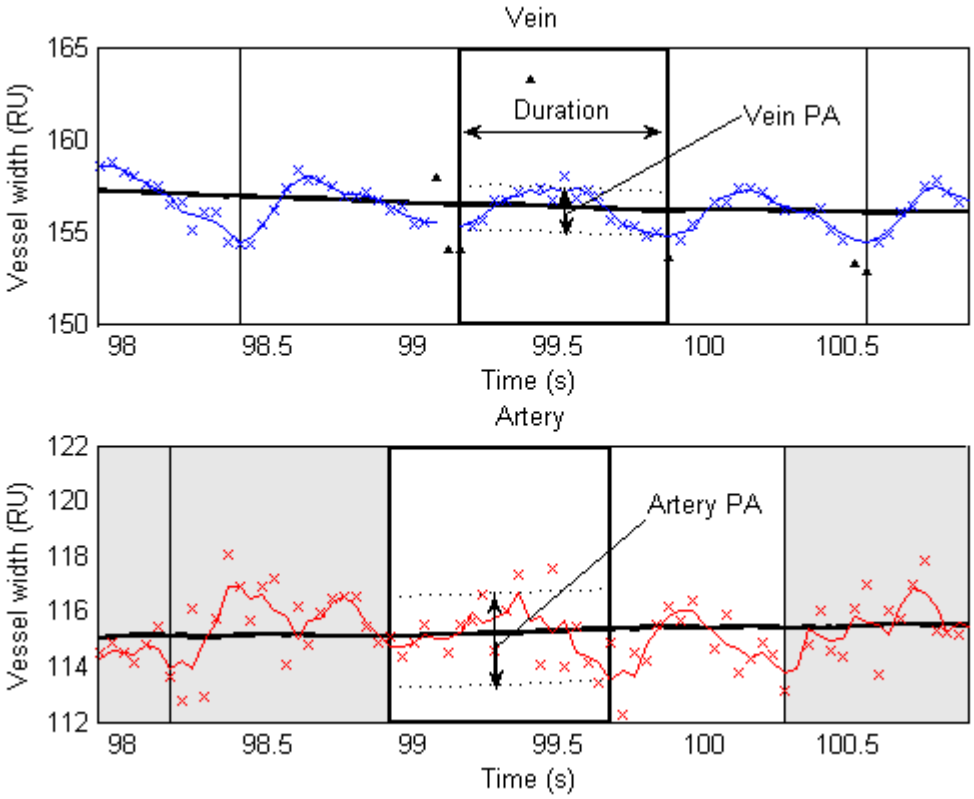

Figure 3: the RVA signal is decomposed into a succession of periods. Some of them, like the central one, in the bold rectangle, are likely to be cardiac cycles; a duration is estimated on the vein signal, and two pulse amplitudes are then estimated, one for the artery and one for the vein. Other periods may be rejected, for one or both vessels, because of the lack of regularity of the signal (grey rectangles). This decomposition requires first to identify spurious (▲) and correct (×) data, second to compute a multi-resolution decomposition (plain lines) of the signal, and third to delimit and validate the cardiac cycles.

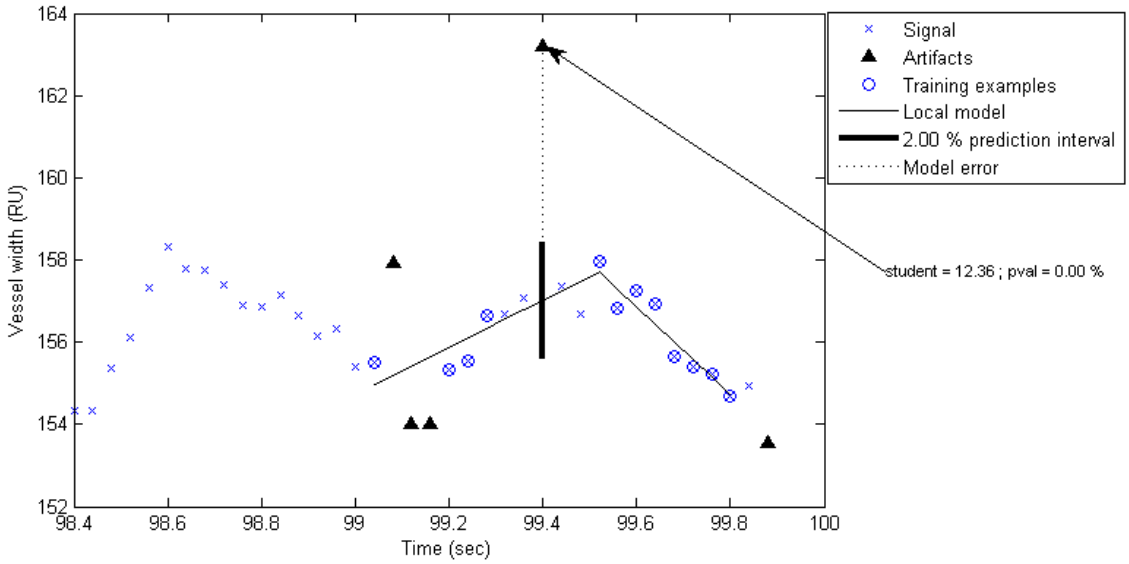

Figure 4: illustration of spurious points identification. The tested point is shown by the arrow (t=99.40 sec, y=163.19 RU). The data points whose distance to this point is between 0.1 and 0.4 sec. are used as training examples to build a local model (a two-phase linear regression). From the dispersion of these training examples around the local model, the procedure estimates the local noise σ and infers a 2% prediction interval for the tested point. Here, the tested point is outside the prediction interval and is hence identified as spurious.

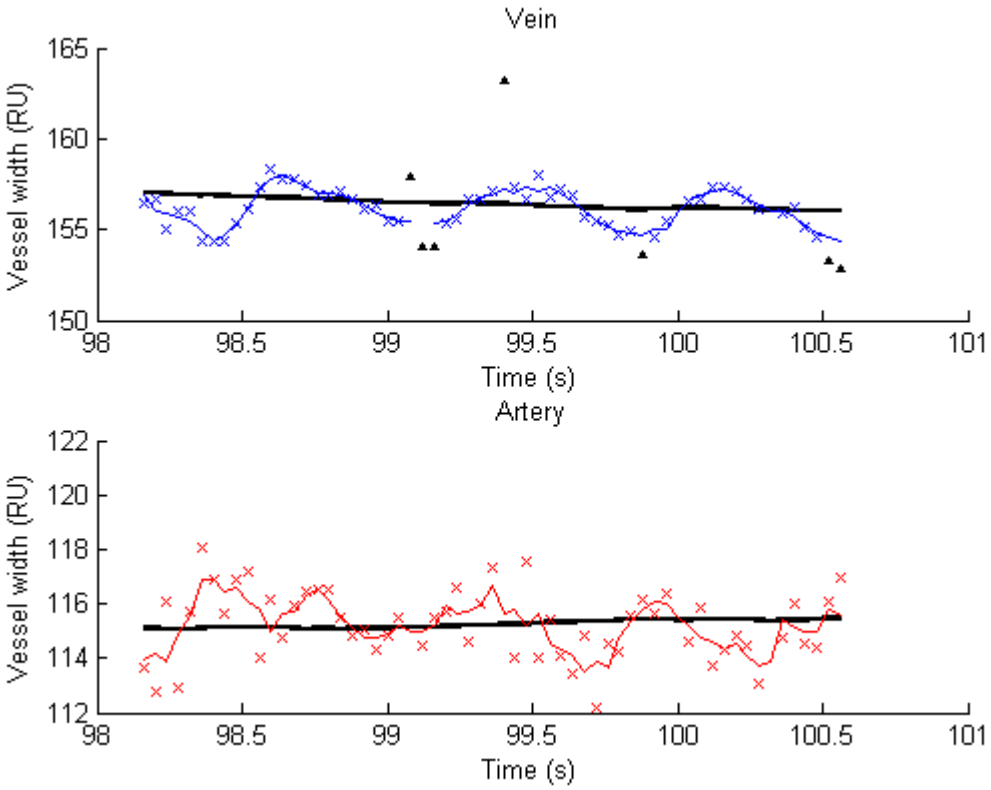

Figure 5: the signal is decomposed into a long term resolution signal $y_1(t)$ **(**T_1 = 3 sec, in bold line), a mid term resolution signal $y_2(t)$ (T_2 = 0.1 sec, in thin line) and the original signal $y(t)$ (in ×).

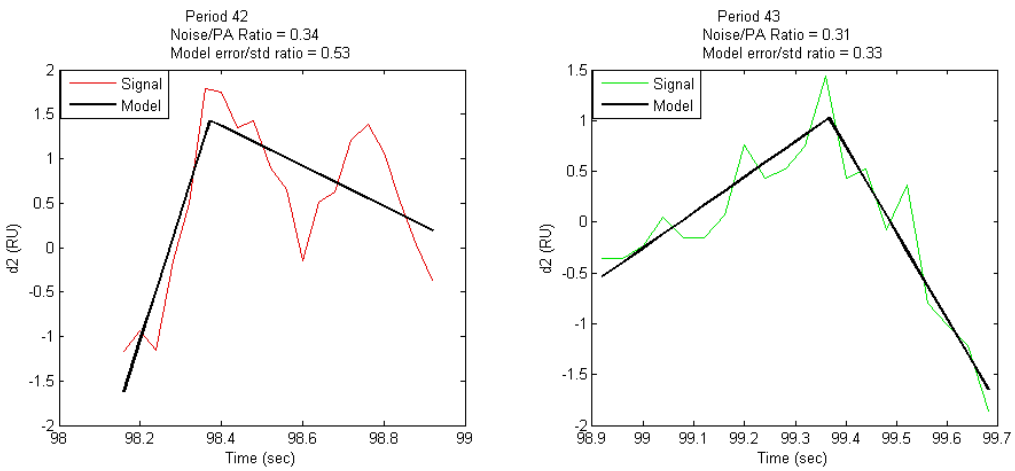

Figure 6: illustration of the «shape validation procedure». The evaluated intervals are the left and central arterial periods of the figure 3. A two-phase linear regression model is built to approximate the second scale decomposition term $d_2(t)$. In the first example (left), the model error, as compared to the $d_2(t)$ standard deviation, is too important (ratio of 0.53>0.50) and the interval is not validated. The second interval (right), fulfilling all the conditions, is validated.

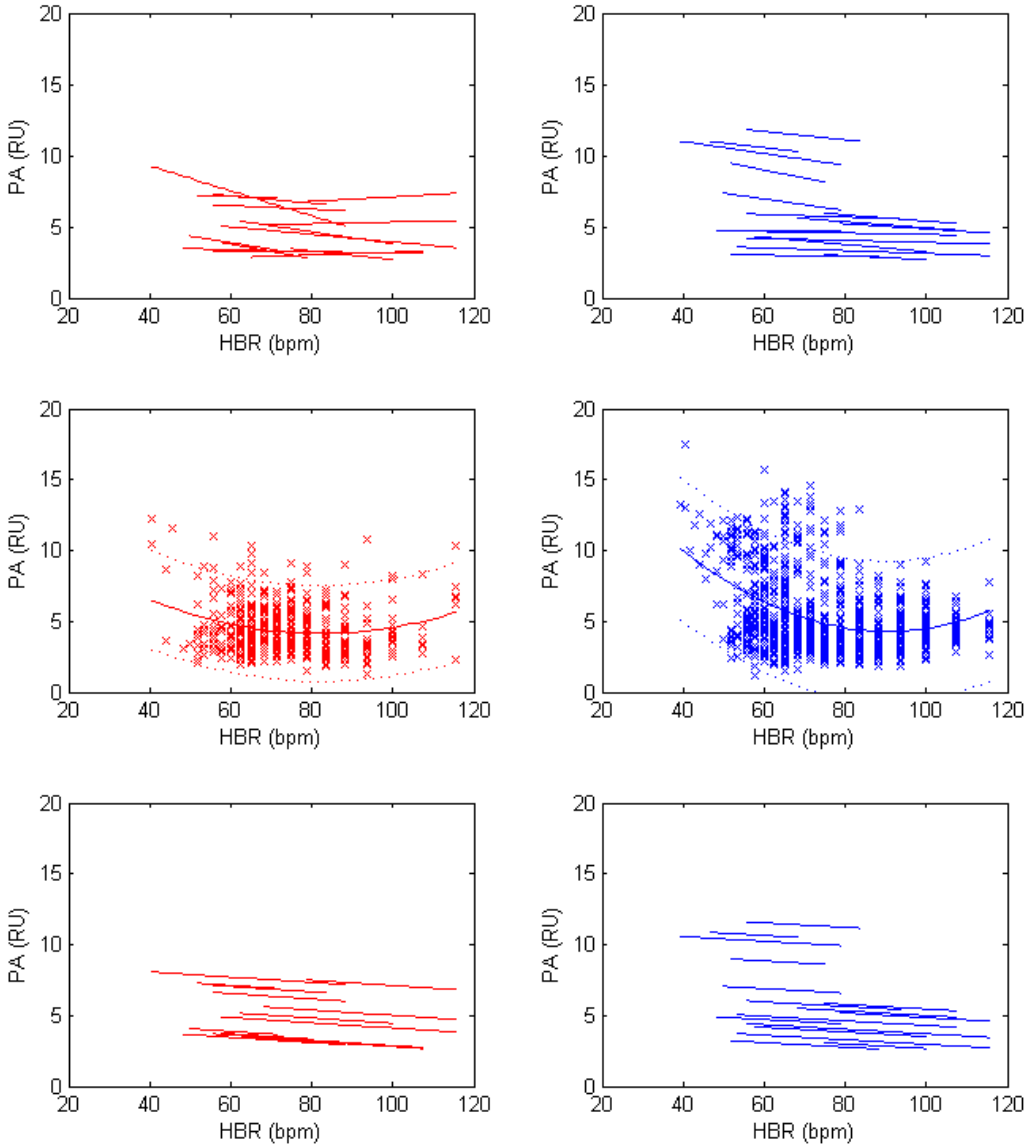

Figure 7: the individual linear regressions (top), global 2-degree regressions (middle) and linear mixed models (down) for the arteries (left) and the veins (right). The global models are shown with the data sets and a 0.95 prediction interval.

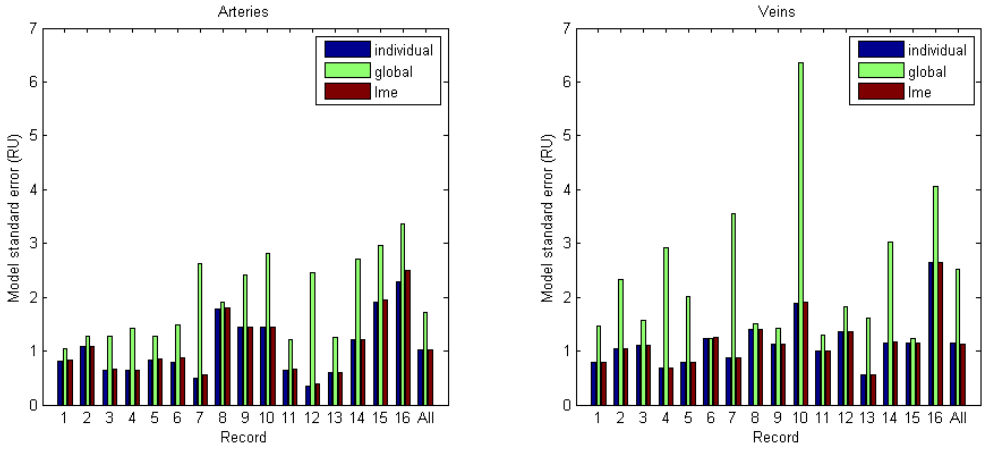

Figure 8: standard errors of the 3 modeling approaches for the arteries (left) and veins (right). The standard errors are displayed record by record and then for all records together.