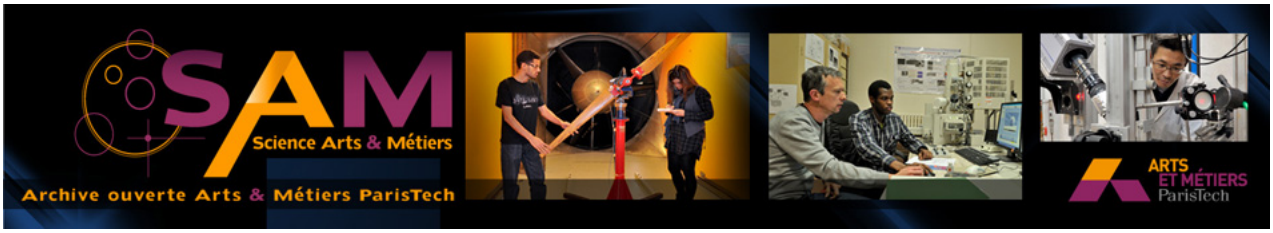

An expert system to characterise the surfaces morphological properties according to their tribological functionalities: The relevance of a pair of roughness parameters

M. Bigerelle [a,b,*], D. Najjar [c,d,e], T. Mathia [f], A. Iost [c,d,e], T. Coorevits [c,d,e], K. Anselme [g]

[a] Laboratoire Roberval, UMR 6253, UTC/CNRS, Centre de Recherches de Royallieu, BP20529, 60205 Compiègne, France
[b] Laboratoire TemPo/LAMIH UMR 8201, Université de Valenciennes et du Hainaut Cambrésis, Le Mont Houy, 59313 Valenciennes, France
[c] Université Lille Nord de France, F-59000 Lille, France
[d] Arts et Metiers ParisTech, Laboratoire de Mécanique de Lille, F-59000 Lille, France
[e] CNRS, UMR 8107, F-59650 Villeneuve d'Ascq, France
[f] Laboratoire de Tribologie et Dynamique des Systèmes, UMR 5513, Ecole Centrale de Lyon, 36 Av Guy de Collongue, 69134 Ecully Cedex, France
[g] Institut de Chimie des Surfaces et Interfaces, UPR CNRS 9069, 15 rue Jean Starcky, BP2488, 68057, Mulhouse Cedex, France

Keywords:
Roughness
Discriminant analysis
Femoral head
Third body abrasive mechanism

A B S T R A C T

Knowing that a surface or profile can be characterized by numerous roughness parameters, the objective of this investigation was to present a methodology which aims to determine quantitatively and without preconceived opinion the most relevant pair of roughness parameters that describe an abraded surface. The methodology was firstly validated on simulated fractal profiles having different amplitudes and Hölder exponents and it was secondly applied to characterise different worn regions of a retrieved metallic femoral head articulated against an ultra-high molecular weight polyethylene (UHMWPE) acetabular cup containing an embedded metallic fibber into its surface. The methodology consists in combining the recent Bootstrap method with the usual discriminant analysis. It was validated on simulated fractal profiles showing that, among more than 3000 pairs tested, the total amplitude R_t and the fractal dimension Δ is the most relevant pair of roughness parameters; parameters corresponding to the variables modulated in the analytical expression of the fractal function. The application of this methodology on a retrieved metallic femoral head shows that the most relevant pair of parameters for discriminating the different investigated worn regions is the arithmetic roughness parameter R_a paired with the mean peak height R_{pm}. This methodology finally helps in a better understanding of the scratch mechanism of this orthopedic bearing component.

1. Introduction

The topography of machined surfaces plays a key role with regard to their properties and therefore to the potential applications of the manufactured products. However, the characterization of a surface topography made through the estimation of roughness parameters remains a large subject of debate. Indeed, because of the various industrial and scientific interests, a proliferation of roughness parameters, termed by Whitehouse the "parameter rash" [1], has been triggered to describe the different kinds of natural, manufactured and modeled surface topographies. Probably running into hundreds, these roughness parameters include amplitude, frequency and hybrid parameters.

This multiparameter representation of surface roughness has been outlined in various works and some efforts have been done in previous ones to define a method for selecting relevant parameters [2–6]. However, only a limited set of parameters was systematically analyzed in these works. In fact, deciding which parameter is the most relevant to describe a surface topography with regard to a specific application remains a difficult task of paramount importance. That is why we developed some years ago a quantitative methodology to select, without preconceived opinion and among a high number of roughness parameters, the most relevant one with regard to different applications [7–11]. Based on the combination of usual statistical methods (least squares linear regression or analysis of variance depending on the application) with the Computer Based Bootstrap Method, the aim of this methodology is to define a numerical indicator of relevance associated with a statistical confidence interval for each roughness parameter under investigation. The higher the value of this indicator, the more relevant is

* Corresponding author at: Laboratoire TemPo/LAMIH UMR 8201, Université de Valenciennes et du Hainaut Cambrésis, Le Mont Houy, 59313 Valenciennes, France.
E-mail address: maxence.bigerelle@utc.fr (M. Bigerelle).

the related roughness parameter to describe the surface topography with regard to the application considered.

For sake of simplicity and by analogy with most of the works dealing with the characterization of topography of profiles or surfaces, our purpose was focused only on the relevance of a single roughness parameter in the aforementioned studies. Nevertheless, because a roughness profile is usually described physically by frequency components of different amplitudes, the aim of the present investigation is to present a new methodology designed to select the most relevant pair of roughness parameter to be used to describe the topography of a machined surface. Based on the combination of the Discriminant Analysis and the Computer Based Bootstrap Method, this methodology is calibrated on simulated fractal profiles for which the scaling amplitude factor and the Hölder exponent have been intentionally modulated.

2. Problematic characterization of surfaces by a pair of roughness parameters

The relevance of characterization of surfaces by two roughness parameters are illustrated by scheme described in Fig. 1a. To characterise both frequency and amplitude components, two roughness parameters are considered. These roughness parameters are an amplitude parameter (total amplitude, R_t) and a frequency parameter (mean distance between profile elements, S_m,). As it could be graphically observed, surfaces A and B get the same S_m value but different R_t values as well as surfaces C and D. On the contrary, surfaces A and C get the same R_t value with different S_m values, such as B and D surfaces. By plotting values of parameters (S_m, R_t) of the four surfaces (Fig. 1b), it becomes obvious that it is impossible to characterise profiles considering a single roughness parameter; either R_t or S_m parameter: R_t does not differentiate A and C surfaces (or B and D) as well as S_m does not differentiate A and B surfaces (or C and D). Consequently, in this schematic example two parameters are necessarily required to characterise the four surfaces.

3. Creation and characterization of simulated roughness profiles

3.1. Creation of fractal profiles

To validate the methodology that allows us to find the most relevant pair of roughness parameters for discriminating different topographies, simulated profiles have been created using the analytical expression of a fractal function depending on two parameters: a scaling amplitude factor A and the Hölder exponent $H \in [0-1]$; exponent related to the fractal dimension Δ of the profiles by the relation $\Delta = 2 - H$. In fact, the fractal function considered in this investigation was created earlier by the authors to simulate profiles resulting from a random grinding process of a metallic surface by abrasive particles having a hemispherical shape [10]. Called the Stochastic Fractal Circle Function (SFCF), it is defined as follows:

$$F_{SFCF}(x,p) = A \sum_{n=0}^{\infty} \Psi_n 2^{-(nH/p)} g(2^{n/p}x + \phi_n) \quad (1)$$

with $g(x)$ an elementary term to which is associated an half-circle representation simulating the shape of the grooves generated during the grinding process. This elementary function is of period 1 and is defined on the $[0-1]$ interval as follows:

$$g(x) = \sqrt{0.5^2 - (x-0.5)^2} \quad x \in [0-1] \quad (2)$$

In Eq. (1), p is an integer higher than unity, Ψ_n are positive Gaussian random numbers that physically simulate the stochastic variation of penetration depth of abrasive particles during the grinding process, and φ_n are uniform random numbers that simulate the random disorientation between the grooves. In this investigation, this function is used to take benefit from these stochastic components which enable to introduce some variability in our methodology.

Fig. 2 shows 16 roughness profiles simulated with different values for A and H. As a consequence, it is obvious that a single roughness parameter cannot be sufficient to differentiate these profiles intentionally created by modulating two roughness parameters: an amplitude roughness parameter and the fractal dimension. To create a databank with a view to assessing the relevance of a high number of roughness parameters, 100 profiles of 10,000 points have been simulated for the different pairs (A, H) under study. For each profile, the least squares mean line has been selected for use as reference datum.

3.2. Computation of roughness parameters

In this investigation, 78 roughness parameters have been computed and assessed for each profile thanks to a personal computer program [9]. Some of the most important roughness parameters considered in this investigation are listed in Table 1. From a general point of view, these roughness parameters are amplitude ones like the total amplitude R_t, the arithmetic roughness R_a or the root mean square roughness R_q, frequency ones like the mean distance between profile elements, S_m and, finally, hybrid ones like the average curvature radius of peaks R_{wz}, the mean slope of the profile D_a and the fractal dimension Δ estimated by different methods (like the oscillation [14], the structure [14], the spectrum [15] and the ANAM [16] methods). In summary, this computational procedure enables to create a databank containing a set of $4 \times 4 \times 100 \times 78 = 124{,}800$ roughness parameters values available for a subsequent statistical analysis.

Fig. 3 shows some examples of paired roughness parameters for each pair (A, H) called a class hereafter. It can be observed on the graph (Δ_{ANAM}, R_t) that the data related to each class seem to be clustered with the clusters being significantly spaced each other. In other words, this means that, as expected, the pair of parameters 'total amplitude R_t–fractal dimension estimated by the ANAM method Δ'_{ANAM} seems relevant to differentiate each class.

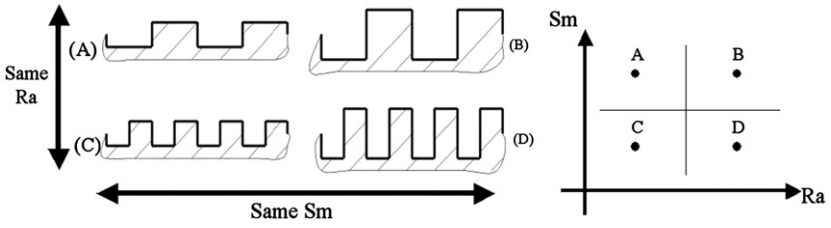

Fig. 1. Example of the necessity to differentiate four surfaces noted A, B, C, D by a pair of roughness parameters (S_m, R_t).

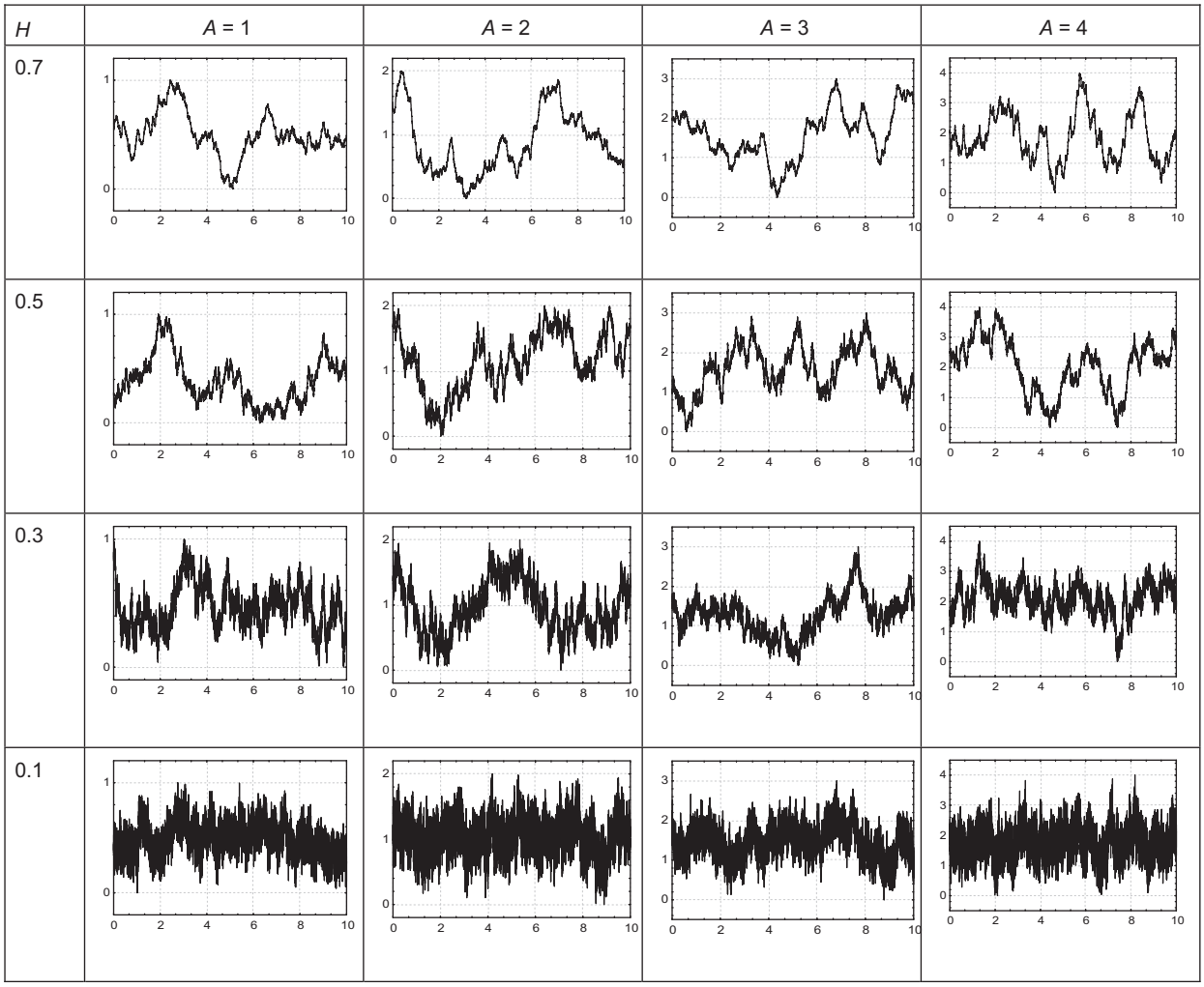

Fig. 2. Simulated profiles obtained by the stochastic fractal circle function for different values of amplitude *A* and Hölder exponent *H*.

On the contrary, in the case of the pair (*Sk*, R_t), only the total amplitude of the profiles R_t seems appropriate to differentiate the data of the different classes regardless of the value of the skewness *Sk*. If parameters are correlated, like the pair (R_a, R_q), it becomes more difficult to make a clear distinction between the data of each class. Finally, the graph (*Sk*, *Ek*) shows that any of these parameters can be used to distinguish the different classes. In fact, this is completely logical since neither the skewness *Sk* nor the kurtosis *Ek* have been considered in the analytical formulation of the simulated profiles.

4. Statistical treatment by Discriminant Analysis

4.1. General information on Discriminant Analysis and limits

The main problem is now to define a statistical methodology that allows finding the most relevant paired roughness parameters for discriminating the simulated profiles related to the different classes (*A*, *H*) under consideration. To address this issue, we decided to retain the Discriminant Analysis [17–19] that, computationally speaking, is very similar to analysis of variance (ANOVA). In practice, the usual discriminant analysis is computed in two parts. In the first part, the total variance/covariance matrix and the within-groups variance/covariance matrix are computed for all the tested variables. In the second part, a forward or a backward stepwise discriminant analysis is run. In the former case, the program will choose at each step the variable for entry which has the largest *F* value (higher than a threshold *F* enter value decided by the user). In the latter case, the program will choose at each step the variable for exclusion from the model which has the smallest *F* value (lower than a threshold *F* remove value decided by the user). This usual approach presents major disadvantages with regard to the present investigation which considers a high number of roughness parameters:

- In the case of a high number of roughness parameters, often the most relevant pair obtained in the forward stepwise discriminant analysis is rarely the same as the one obtained in the backward stepwise discriminant analysis.
- The result depends on parameters selected by the user in the retained algorithm.
- The probability to find the right most relevant pair of roughness parameters that distinguish the different classes drastically decreases as the number of roughness parameters increases.
- The measure of discriminating power is not so intuitive since it is based on the *F* Snedecor variable. A more accurate indicator to test the superiority of a pair of roughness parameters could be the percentage of well classified cases after algorithm computation.
- This technique does not allow us to estimate the pairs of roughness parameters that get the same discriminating power.

Table 1

Roughness parameter	Name	Standard
R_a	Arithmetic roughness	ISO 4287 [11]
R_t	Total amplitude	ISO 4287
R_{pm}	Mean peak height	ASME B46 [12]
Δ	Fractal dimension	**No norm**
S_m	Mean distance between profile elements	ISO 4287
R_q	Root mean square roughness	ISO 4287
R_{wz}	Average curvature radius of peaks	No norm
$R_{\Delta a}$	Mean slope of profile	ASME B46
S_k	Skewness	ISO 4287
E_k	Kurtosis	ISO 4287
$R_{\Delta q}$	Root mean square of the slope of profile	ISO 4287
L_0	Developed Profile Length	**No norm**
L_R	Relative Developed Profile Length	**No norm**
R_k	Roughness core profile	ISO 13565-2 [13]
R_{pk}	Reduced peak height	ISO 13565-2
R_{vk}	Reduced valley	ISO 13565-2
Mr_1	Upper material portion	ISO 13565-2
Mr_2	Lower material portion	ISO 13565-2
A_1	Upper area relative to peaks	ISO 13565-2
A_2	Lower area relative to valleys	ISO 13565-2

ISO 4287:1997. Geometrical Product Specifications (GPS)—Surface texture: Profile method—Terms, definitions and surface texture parameters.
ISO 13565-2:1996. Geometrical Product Specification (GPS)—Surface texture: Profile method; Surfaces having stratified functional properties—Part 2: Height characterization using the linear material ratio curve.
ASME B46.1:2009. Surface Texture, Surface Roughness, Waviness and Lay.

- Finally, this technique does not propose a ranking of the relative discriminating powers of the different tested pairs of roughness parameters.

4.2. Alternative method for the classification of paired roughness parameters

Thanks to the power of modern computers, we shall propose an alternative and original method that both avoids all the precited disadvantages and gives some visual indicators intended to help the operator in the research of the most relevant discriminating pair of roughness parameters. The basic idea is to perform the discriminant analysis for all pairs of roughness parameters, which represents $p=n(n-1)/2+n$ discriminant analyses where n is the number of roughness parameters. For each of the p pairs of roughness parameters, thanks to the use of discriminating functions, the number of well-classified data in each class indiced by c_k is computed by $n_w(q_i, q_j, c_k)$ where q_i represents the ith roughness parameters and the suffix w means "well classified". On the contrary, $n_b(q_i, q_j, c_k)$ denotes the number of "badly classified" with $n(q_i, q_j, c_k)=n_w(q_i, q_j, c_k)+n_b(q_i, q_j, c_k)$ the number of data that belong to the class c_k in original data set. Hereafter, we also shall note $n(q_i, q_j, c_k, c_l)$ the number of classified data originally associated to the class c_k and finally classified in the class c_l with $n_w(q_i, q_j, c_k)=n(q_i, q_j, c_k, c_k)$ and $n_b(q_i, q_j, c_k)=\sum_{\substack{l\in C \\ l\neq k}} n(q_i, q_j, c_k, c_l)$ where C is the number of classes. With these

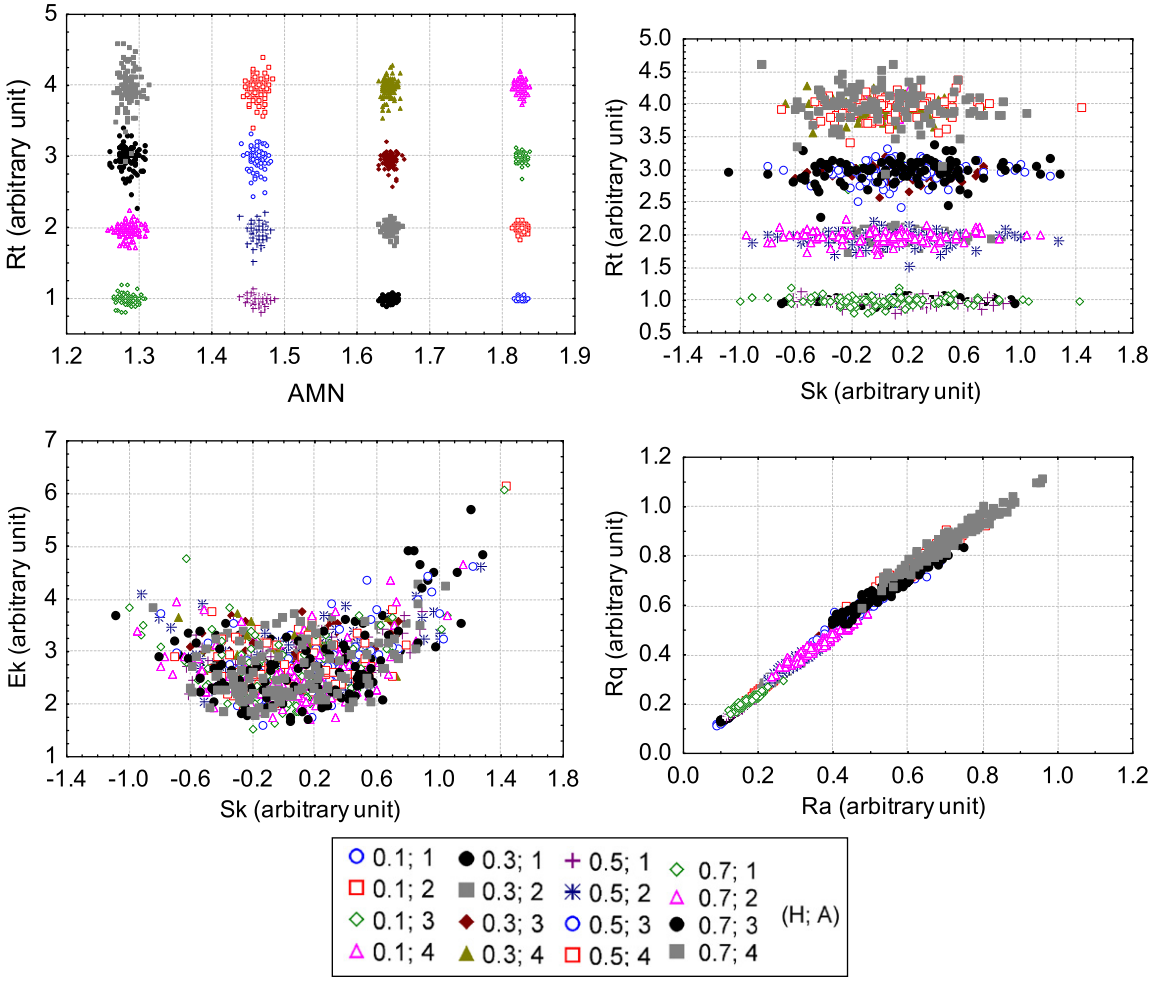

Fig. 3. Four examples of pairs of roughness parameters obtained for 100 profiles simulated in each class. The pairs of roughness parameters are (Δ_{ANAM}, R_t), (Sk, R_t), (R_a, R_q), (Sk, Ek).

notations, the percentage of well classified data in the c_k class can be expressed as:

$$p_w(q_i, q_j, c_k) = 100n(q_i, q_j, c_k, c_k)/\sum_{l \in C} n(q_i, q_j, c_k, c_l) \quad (3)$$

Finally, the average percentage of well classified data in the different classes is retained as the best classification indicator. Its analytical expression is

$$p_w(q_i, q_j) = 100/C\sum_{k \in C} \left(n(q_i, q_j, c_k, c_k)/\sum_{l \in C} n(q_i, q_j, c_k, c_l) \right) \quad (4)$$

Fig. 4 shows the well and badly classified data obtained when applying the Discriminant Analysis to pairs of roughness parameters already considered in Fig. 3. The related average percentages of well classified data equal 99% for the pair (Δ_{ANAM}, R_t), 50% for the pair (R_a, R_q), 26% for the pair (Sk, R_t) and 14% for the pair (Sk, Ek). As expected, the results of the Discriminant Analysis confirm that the pair of roughness parameter (Δ_{ANAM}, R_t) seems really appropriate to distinguish the different classes related to profiles generated in Section 2. On the contrary, the pair of roughness parameters (Sk, Ek) seems absolutely not relevant.

4.3. Bootstrap protocol

In the previous section, two subtle limits must be however highlighted:

The first limit is that no confidence interval is provided with the $p_w(q_i, q_j)$ values. If we assume that the selected scanning length is sufficient to avoid a scaling effect on the estimation of the roughness parameters (fractal concept); the longer the scanning length (for the same sampling rate), the more precise the value of the roughness parameters and the $p_w(q_i, q_j)$ estimation. Besides, increasing the number of profiles will also increase the classical estimator used in the discriminant analysis, which will decrease the $p_w(q_i, q_j)$ variances. These two kinds of variability lead to the result where, at a given confidence level, it is not possible to affirm that many pairs of parameters possess different discrimination powers.

The second limit is that it is not possible to affirm that any pair of parameters (q_i, q_j) is relevant as long as a threshold value of $p_w(q_i, q_j)$ (denoted $p_{w,H_0}(q_i, q_j)$) and its confidence interval have not been defined; threshold value for which the pair of roughness parameter is therefore no longer relevant.

For these reasons, we shall introduce a recent technique called the Bootstrap which is a resampling technique [20,21]. The basic idea of the bootstrap is to create a new dataset by randomly sampling with replacement data from the original data set and then performing the same statistical analysis as was carried out on the original data set. This analysis enables to obtain a new set of values for the variable $p_w(q_i, q_j)$ denoted $\hat{p}^1_w(q_i, q_j)$ where $\hat{p}$ represents the estimator of p and subscript^1 indicates the first bootstrap sample. Fig. 5 shows an example of Bootstrap sample for the relevant pair of roughness parameters (Δ_{ANAM}, R_t). A high number of bootstrap samples B are processed again to obtain a set of B estimators $\hat{p}^1_w(q_i, q_j)$, $\hat{p}^2_w(q_i, q_j)$,…, $\hat{p}^B_w(q_i, q_j)$. From this set that constitutes an empirical probability density of the variable $p_w(q_i, q_j)$,

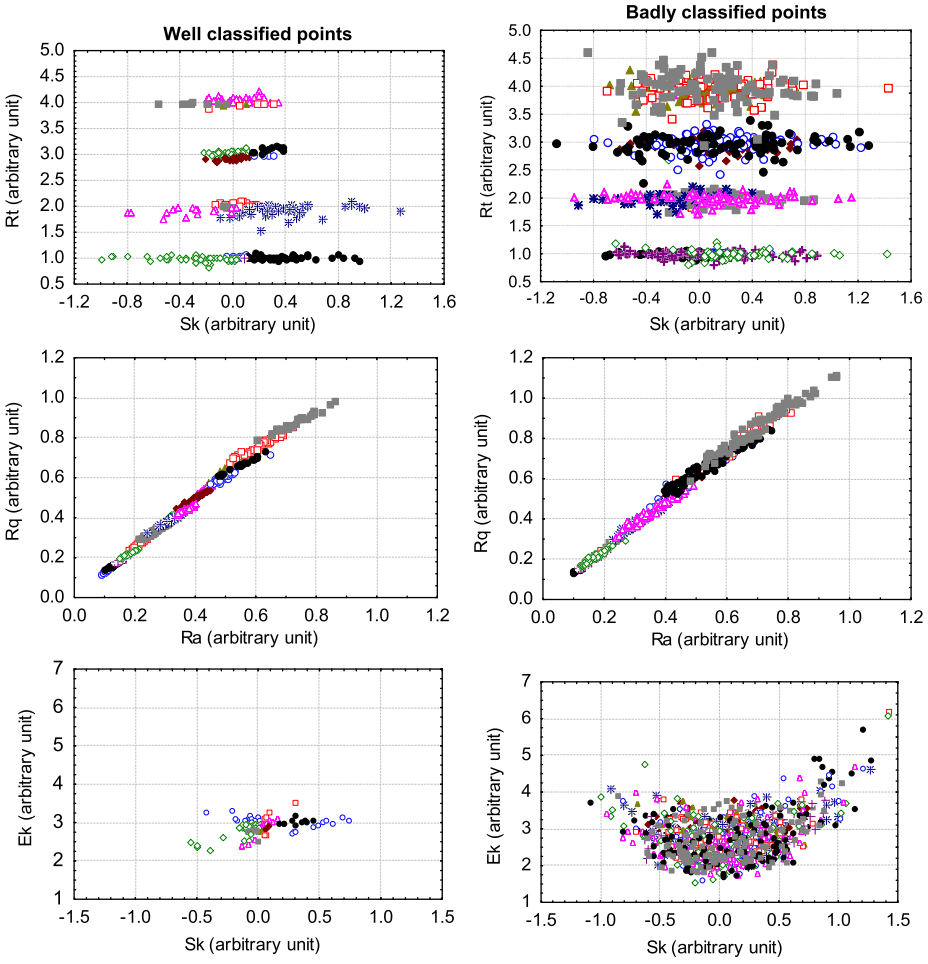

Fig. 4. Classification of the three pairs of roughness parameters (Sk, R_t), (R_a, R_q) and (Sk, Ek). On the left, classification of the well classified points and on the right the badly classified points obtained after the discriminant analysis of the original data.

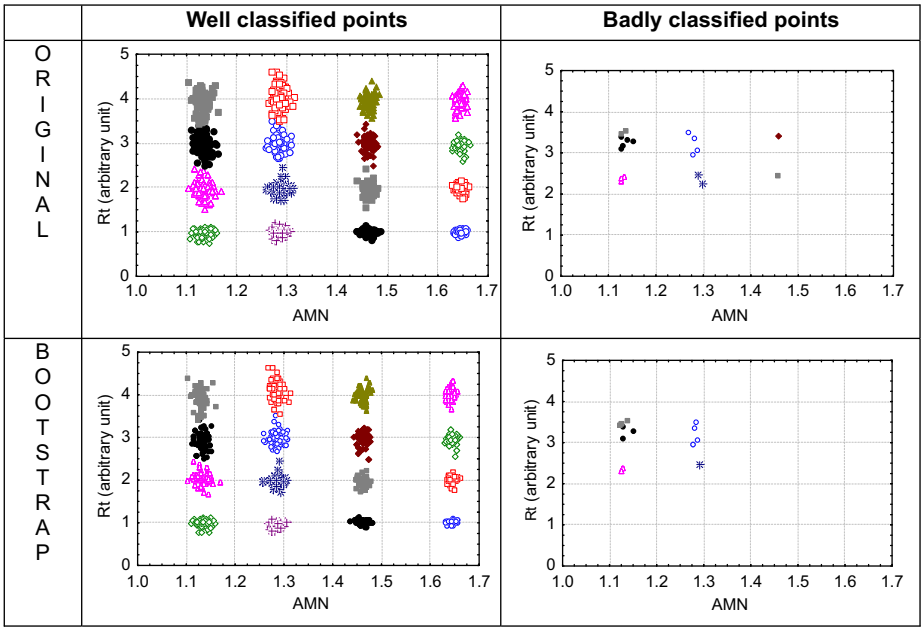

Fig. 5. Classification of the pair of roughness parameters (Δ_{ANAM}, R_t). On the left, the well classified points and on the right the badly classified points obtained from both the original data and a bootstrap sample.

usual descriptive statistics can be computed like mean, mode, median, quartile, confidence intervals and so on… Applied to the discriminant analysis, for one class, we randomly get N pairs of (q_i, q_j) with replacement where N is the number of pairs of roughness coefficients on this class calculated from a profile. This procedure is reproduced for each class of the original data set to finally constitute the bootstrap data set.

To obtain the $p_{w,\,H_0}(q_i, q_j)$ distribution, we applied the bootstrap protocol except that a preliminary operation has to be done. If we assume that roughness parameters (q_i, q_j) do not play any role (what is called the H_0 hypothesis), it means that classes have no influence on the $p_w(q_i, q_j)$ values or equivalently that $p_w(q_i, q_j, c_k) = p_w(q_i, q_j, c_{k'})$ for $k \neq k'$ This means also that the elements of the class c_k can be permuted with those of the class $c_{k'}$ without affecting statistically the value of $p_w(q_i, q_j)$ which is then denoted $p_{w,\,H_0}(q_i, q_j)$. Thus, for the tested pair (q_i, q_j), the preliminary step consists in mixing all the elements without taking account of their original class and then randomly sampling with replacement to affect them into the different new classes. This means that the new class c_k consists now of elements originating from all the classes $c_1, c_2, ..., c_k, ..., c_C$. The same protocol as described in the preceding paragraph is then reproduced to construct the $\hat{p}^1_{w,H_0}(q_i, q_j)$, $\hat{p}^2_{w,H_0}(q_i, q_j), \ldots, \hat{p}^B_{w,H_0}(q_i, q_j)$ bootstrap set on which an empirical probability density of the $p_{w,H_0}(q_i, q_j)$ is built. We shall call this method the **B**ootstrapped **P**aired **D**iscriminant **A**nalysis (**BPDA**). The Fig. 6 sum up the algorithm of the whole two pair analyses.

5. Results

5.1. Searching for the most relevant paired roughness parameters

Because of the analytical expression used to generate the simulated fractal profiles under consideration, two roughness parameters are expected to distinguish the different classes defined by the different studied pairs (A, H). Fig. 7 represents the real, the mean and the 95% centile and the 5% centile percentage of well classified data by decreasing order obtained by applying our BPDA algorithm with 100 bootstraps. As can be observed, classification decreases from 99% to 10% of well classified data. This clearly means that some pair of roughness parameters distinguish the different classes of profiles better than many others. As can be expected, the fractal dimension Δ and the total amplitude R_t values is the most relevant discriminating pair of parameters proving that our method is really well adapted to obtain a quantitative ranking of the relative relevance of the different pairs of roughness parameters tested. Let us note that, as the profile is straightened up, the total amplitude R_t is not perfectly equal to the A values and gets an intrinsic variance. For information, the pairs of roughness parameters (R_a, R_q), (Sk, R_t) and (Sk, Ek) rank in 1157th, 2427th and 3073th positions. Fig. 5 allows us to appreciate graphically the ability of the most relevant pair of roughness parameters 'total amplitude R_t–fractal dimension estimated by ANAM Δ'_{ANAM} to well distinguish the different classes (A, H) under study.

As already mentioned, several methods have been developed to estimate the fractal dimension of a profile and it is really interesting to note that the four first places of the ranking are occupied by a pair of roughness parameters consisting of the total amplitude R_t combined with the fractal dimension Δ estimated by using the oscillation and the structure methods [14] as well as two authors' methods [16]. Besides, the pair of roughness parameter total amplitude R_t–fractal dimension Δ estimated by the spectrum method only ranks in the 335th position confirming that this method is not satisfactory for evaluating the fractal dimension of a profile [14]. To make sure that all the results obtained are statistically significant, a bootstrap under H_0 hypothesis has been finally processed and leads to a percentage of well classified data of about $15\% \pm 5\%$ for the 10 most relevant pairs of roughness parameters (Fig. 8). This means that many pairs of roughness parameters tested are relevant; however the pair 'total amplitude R_t–fractal dimension Δ'_{ANAM} is the most relevant of them all.

5.2. Searching for most relevant single roughness parameter

From a simple computational point of view, searching for a single relevant roughness parameter is equivalent to extracting from the general BPDA results those only corresponding to the

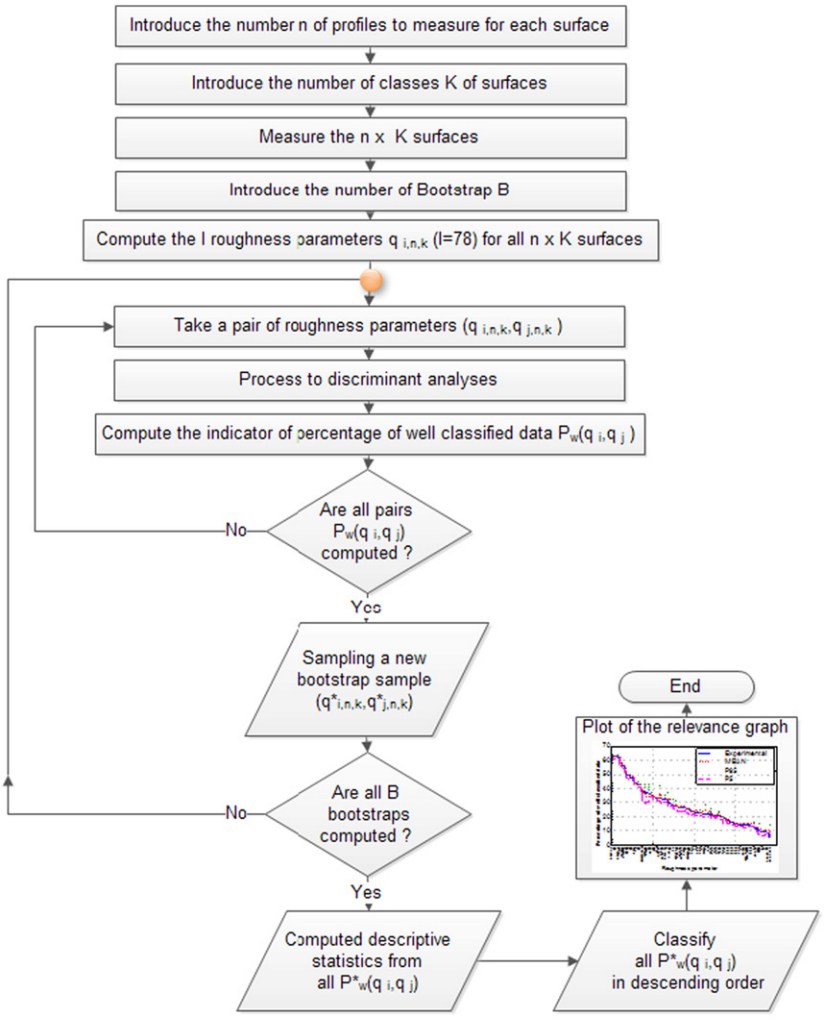

Fig. 6. Flowchart of the algorithm used to classify the pairs of roughness parameters.

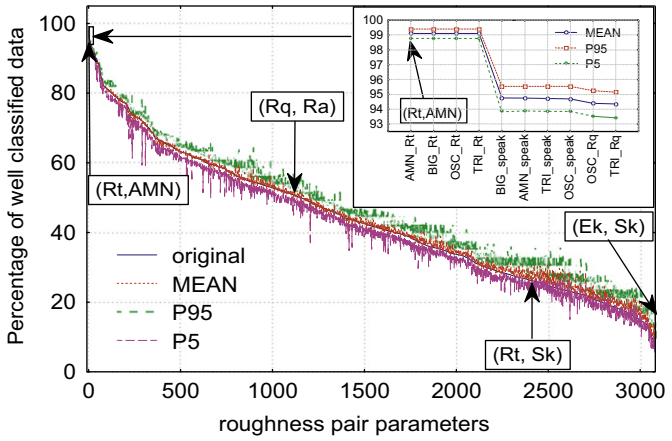

Fig. 7. Ranking of the pairs of roughness parameters with respect to their average percentage of well classified data and with the mean, 5% and 95% centiles obtained from 100 bootstrap simulations. More than 3000 pairs of roughness parameters were considered.

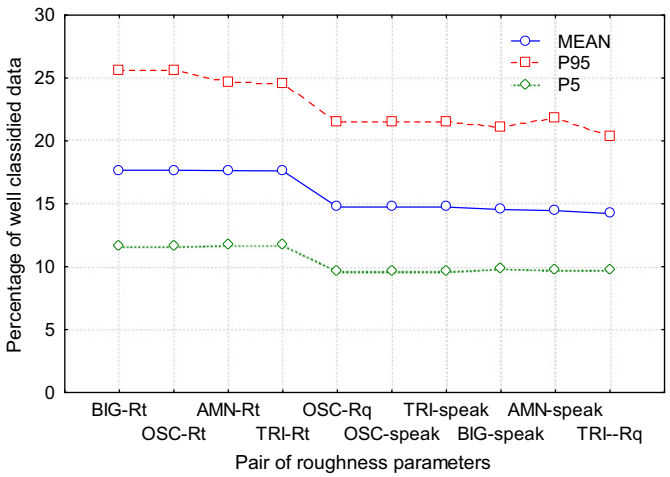

Fig. 8. Percentage of the well classified data for the ten best ranked pairs of parameters and with the mean, 5% and 95% centiles obtained from 1000 bootstrap simulations under the H_0 hypothesis (meaning that no parameters can distinguish any class).

pairs for which $q_i=q_j$; i.e. the 78 pairs (q_i, q_i). It is really interesting to notice that, neither the total amplitude R_t nor the fractal dimension Δ; which are the variables used in the analytical expression of the studied profiles, are relevant when they are considered individually in the overall ranking. Indeed, the pairs (R_t, R_t) and (Δ_{ANAM}, Δ_{ANAM}), respectively, rank in 2410th and 1792th positions with related percentages of well classified data only equal to 26% and 37%. In fact, as can be seen in Fig. 9, the five most relevant single parameters are hybrid parameters depending on both the amplitude and the fractal dimension of the

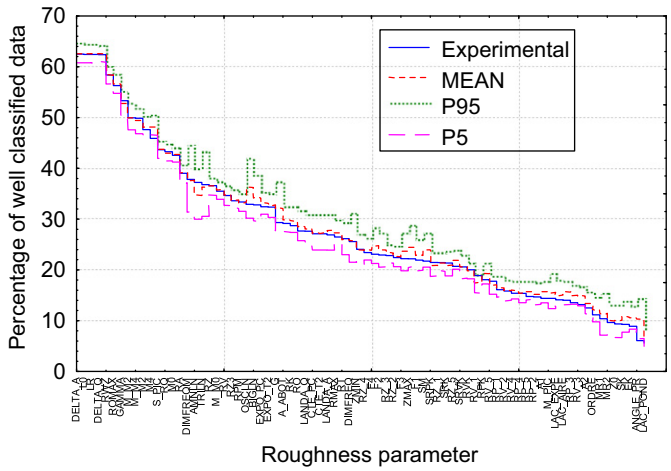

Fig. 9. Ranking of the 78 identical pairs of roughness parameters with respect to their average percentage of well classified data and with the mean, 5% and 95% centiles obtained from 100 bootstrap simulations.

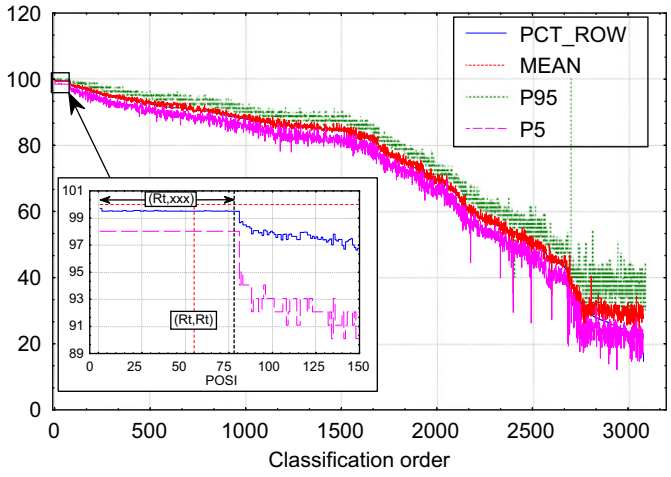

Fig. 10. Ranking of the pairs of roughness parameters with respect to their average percentage of well classified data and with the mean, 5% and 95% centiles obtained from 100 bootstrap simulations. More than 3000 pairs of roughness parameters were considered.

profiles. These parameters are the mean slope of profile D_a, the developed length of the profile L_0, the relative developed length L_R, the root mean square of the slope of profile Δ_q and the average curvature radius of peaks R_{wz}. It is interesting to mention that Berglund and al. noted a similarity between the mean slope and the developed length of profile in their work [6]. Anyway, this result from Fig. 9 means that, if a single roughness parameter had to be selected in a blind analysis for characterising these profiles defined by two roughness parameters, the least wrong choice would have to retain a hybrid parameter. However, even the most relevant of them all (i.e. the mean slope of profile Δ_a) only ranks in 491th position with a percentage of well classified data equals to 62%. In other words, single hybrid roughness parameter rank very far behind many pairs of roughness parameters and particularly the pair 'total amplitude R_t–fractal dimension estimated by ANAM Δ'_{ANAM} which is the most relevant of them all.

Let us now deliberately simulate a set of profiles for which only one roughness parameter is discriminating by fixing for example the value of $H=0.1$ and by varying the amplitude of profiles A from 1 to 4 in Eq. (1). If the BPDA is applied on this new series of profiles, then the first 78 most relevant pairs of roughness parameters all include the total amplitude R_t we know in our simulation to be the only relevant parameter. In particular, the pair (R_t, R_t) ranks in 55th position. However, it can be observed in Fig. 10 that this particular pair has the same discriminating power than the 77 other ones since the confidence intervals on the mean percentage of well classified data are overlapped for the overall paired roughness parameters. This clearly means, as expected, that the total amplitude R_t is the only one relevant roughness parameter since no more information is obtained and the ranking remains unchanged when combining this roughness parameter with any other non relevant one. Based on simulated profiles, all these results proves the efficiency of the BPDA methodology we developed to select without preconceived opinion either a single or a pair of relevant roughness parameters for discriminating different classes of profiles or surfaces. If in a blind study the first same bootstrap confidence intervals of most classified parameters contain an identical pair parameter (q_i, q_i), then only one parameter is required to characterise classes of profiles. On the contrary, if no identical pair parameter belongs to this interval, a pair of parameters (q_i, q_j) is required.

6. Application of the BPDA methodology to the characterization of the surface topography of a retrieved metallic femoral head

6.1. Presentation of the general problem

Despite the growing success of the total hip prostheses surgery, limiting the wear of UHMWPE components remains a key issue to improve the long-term performance of the most implanted Charnley type prostheses. In this part, the BPDA methodology previously described is applied on a retrieved metallic femoral head coming from a total hip prosthesis of Charnley type to specify which roughness parameters are relevant for assessing the surface topography. Knowing that the scratching of metallic femoral head is widely thought to be a major factor affecting the wear of the Ultra-High Molecular Weight Polyethylene (UHMWPE) counterface component [22–31] and that third bodies are increasingly suspected of being the major cause of the scratch generation at the surface of the metallic femoral heads of Charnley type total hip prostheses [22–26,32,33], the titanium alloy (TA6V) femoral head retained for this study was especially selected because, after revision surgery, a visual inspection revealed that this bearing component was severely scratched and that a centimeter-length metallic fiber was embedded into the surface of the UHMWPE counterface component. It should be mentioned that the scratching mechanism of this retrieved metallic femoral head has been already studied in details by the authors in two previous studies focused on the characterization of its surface topography by means of visual inspection, optical microscopy and three dimensional tactile profilometry. These studies particularly emphasize the deleterious role of the embedded centimeter-length metallic fiber on the surface degradation of the articulation components through a third body abrasive mechanism [34,35].

The topography of a surface is usually quantified thanks to roughness parameters. However, it should be pointed out that, among the various studies carried out for the last twenty years, no roughness parameter has universally been admitted to describe the topography of orthopedic bearing surfaces. Even if the arithmetic roughness parameter R_a is the one parameter which has systematically been used in these studies, the root mean square roughness R_q, the total amplitude R_t, the peak height R_p, the mean peak height R_{pm} and the skewness S_k of the Amplitude Distribution Function (ADF) have been also considered but in a lesser extent [22,23,26,31]. Consequently to an analysis of the

limits of the aforementioned roughness parameters, Hall et al. claimed that considerable work still needs to be undertaken in terms of specifying which relevant roughness parameter should be used in assessing surface topography of femoral heads in relation to the wear of UHMWPE components and that investigation of the parameters derived from the bearing area curve (BAC) is urgently required [23]. That is to say the roughness core profile R_k, the reduced peak height R_{pk}, the reduced valley R_{vk}, the upper material portion Mr_1, the lower material portion Mr_2, the upper area relative to peaks A_1 and the lower area relative to valleys A_2.

Specifying which relevant roughness parameter should be used is all the more difficult since, as mentioned in the introduction of this paper, more than an hundred of roughness parameters can be used from a general point of view to characterise a surface topography whereas only a restricted set of roughness parameters have been considered by the authors of the aforementioned studies related to the characterization of the surface topography of orthopedic bearing surfaces. Moreover, the results obtained from the analysis of simulated fractal profiles presented more specifically in the previous parts of this paper show that it should be more relevant to take into consideration a pair of roughness parameters to characterise the topography of a surface rather than a single one. The aim of this part is to determine without preconceived opinion the most relevant pair of roughness parameters that should be used for discriminating the different investigated worn regions of the retrieved metallic femoral head considered in this study.

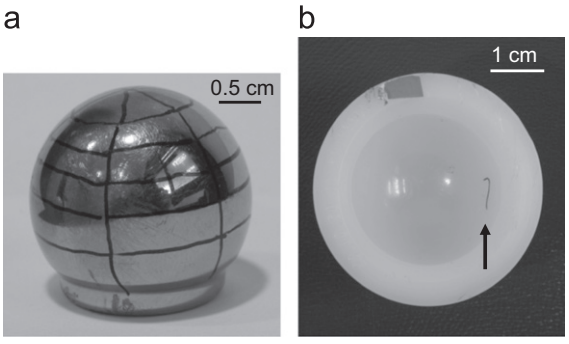

Fig. 11. Photographs showing the retrieved components: (a) the titanium based femoral head (TA6V), (b) the associated UHMWPE acetabular liner with an embedded metallic fiber marked with a black arrow.

6.2. Clinical information and surface topography characterization of the retrieved femoral head

The articulating components under investigation were retrieved after 4 years and 5 months because of the detection of an osteolysis phenomenon (without loosening) on the survey radiographs of the patient. During the period of implantation, the UHMWPE acetabular liner was inserted in a Harris-Galante Mark I metallic shell and the 28 mm diameter metallic femoral head was mounted on a cemented femoral stem. After revision surgery, a centimeter-length metallic fiber coming from the titanium fibermesh deposited on the Harris-Galante metallic cup during the fabrication process was detected by visual inspection into the surface of the UHMWPE counterface component (Fig. 11). It must be pointed out that polyethylene and metallic debris were detected by histological analyses on the peri-prosthetic tissues. As far as the degradation of the titanium alloy femoral head is concerned, three different kinds of regions were observed by visual inspection (Fig. 12).

Regions covering about 30% of the entire surface, containing only few isolated scratches and having a bright finish meaning that the femoral head has only suffered a light surface degradation in these regions. These regions noted LS (for lightly scratched) are white-colored on the schematic representation of the retrieved femoral head shown in the center of Fig. 12.

Regions covering between 10% and 20% of the entire surface and visually having a low brightness level. Two regions of this type were detected along a meridian; a region located near the polar region and another one located near the equatorial region. Both these regions suffered a severe surface degradation and contain a high density of large, deep and multidirectional scratches. These regions respectively noted SSPR (for severely scratched polar region) and SSER (for severely scratched equatorial region) are dark gray-colored on the schematic representation of the retrieved femoral head shown in the center of Fig. 12. It should be mentioned that some material build-up can be observed on either side of the large scratches crossing the region SSER whereas it is not the case for the region SSPR (Fig. 13).

Regions covering more than 50% of the entire surface, containing a high density of small and multidirectional scratches, and visually having an intermediate brightness level. These regions noted MS (for moderately scratched) are gray-colored in on the schematic representation of the retrieved femoral head shown in the center of Fig. 12.

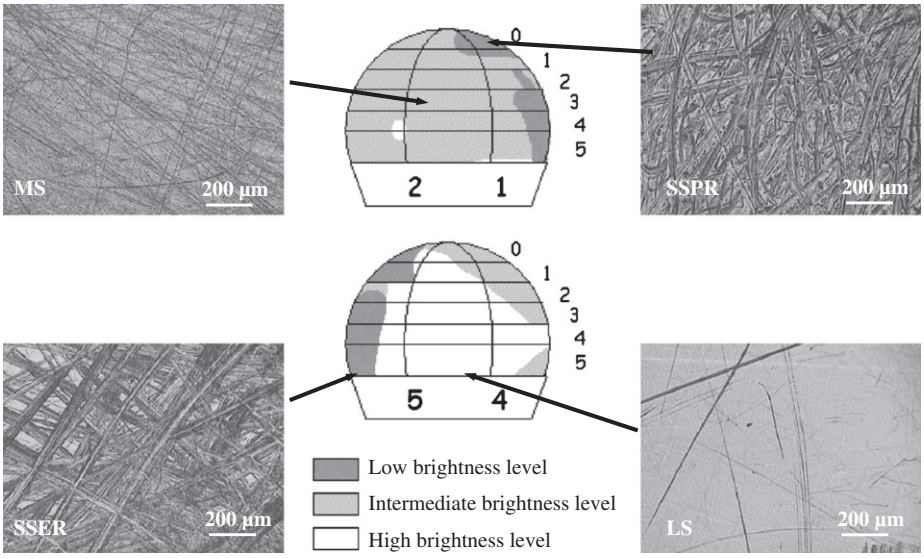

Fig. 12. Summary of the observations made by visual inspection combined with an optical microscope.

Combined to these optical observations, quantitative roughness measurements have been carried out by means of a three dimensional contacting profilometer (KLA Tencor P10) having a 2 μm stylus radius to obtain information about the magnitude of damage produced by scratching in these four characteristic regions. In each region, six measurements have been performed. These selected experimental conditions correspond to a 1 μm horizontal resolution along a scanning trace, a $600 \times 600\ \mu m^2$ size for each scanned area and a number of measured and treated profiles equals to 1200 in each analyzed region.

6.3. Experimental results and discussion

In a previous work focused on the surface characterization of the same retrieved femoral head as that considered in this investigation, only the roughness parameters commonly used to describe the topography of orthopedic bearing surfaces were analyzed in a first time (parameters R_a, R_q, R_t, R_p, R_{pm}, S_k but also the parameters R_k, R_{pk}, R_{vk}, Mr_1, Mr_2, A_1, A_2 derived from the Abbott–Firestone curve) [32]. Fig. 14 illustrates the evolution of the values recorded in the four worn regions studied in the particular case of the arithmetic roughness parameter R_a paired the mean peak height R_{pm}. In fact, it was shown in the previous work that not only the values of the arithmetic roughness parameter R_a but the values of all the amplitude parameters R_α ($\alpha = a, q, t, p, pm, k, pk, vk$) tend to increase with the severity of surface degradation if all the worn regions are considered reflecting the physical fact that the height of the peaks and the depth of the valleys simply increase with the severity of surface degradation. Combining the Analysis of Variance (ANOVA) with the Computer-Based Bootstrap Method (CBBM), it was shown that the average amplitude roughness parameters are more relevant than the other ones for discriminating the different topographies of the four investigated worn regions; the most relevant of them all being the arithmetic roughness parameter R_a.

However, when regions SSER and SSPR were considered separately from the two others regions, this average amplitude parameter as well as the other average amplitude parameter failed to differentiate the topographies of the two severely scratched regions whose patterns are still quite different. Indeed, as it can be noticed as an example on figure in the particular case of the arithmetic roughness parameter R_a paired with the mean peak height R_{pm}, the values of the average amplitude parameters R_α ($\alpha = a, q, k$) were of the same order for regions SSER and SSPR whereas the values of the extreme-value parameters R_α ($\alpha = p, pm, pk, t$) were quite different as well as the values of the roughness parameters Mr_1 and A_1; the values of all these parameters being always higher in the region SSER. Moreover, the values of the roughness parameters R_{vk}, Mr_2 and A_2 were of the same order for the two regions. These results physically mean that, while the regimes related to the core roughness and to the valleys lying below the core roughness are similar for the two severely scratched regions, the regime related to the peaks are significantly different; the peaks being higher in the region SSER than in the region SSPR. Combining the ANOVA with the CBBM, it was shown that the most relevant roughness parameter for discriminating the topographies of these two latter regions were the roughness parameters sensitive to the height and the density of peaks; the most relevant of them all being the mean peak height R_{pm}. These experimental results related to a single retrieved femoral head highlight the difficulty to find a universal and single roughness parameter to describe the topography of worn surfaces in the case of bearing components.

In the present investigation, more than an hundred of roughness parameters (including the aforementioned ones) were considered to find without preconceived opinion the most relevant pair of roughness parameters for discriminating the four worn regions by applying the BPDA methodology. Fig. 15 presents the evolution of the percentage of well classified data for all the pairs of roughness parameters under consideration in this study. In this figure, the results of the analysis are presented for the original data with a 90% level of confidence determined using the centiles 5% and 95% extracted from the empirical probability density function (PDF) generated by means of 100 bootstrap simulations for each pair of roughness parameter under study. Fig. 16 shows the empirical PDF obtained for the pair (R_a, R_{pm}) for which the highest percentage of well classified data has been recorded using

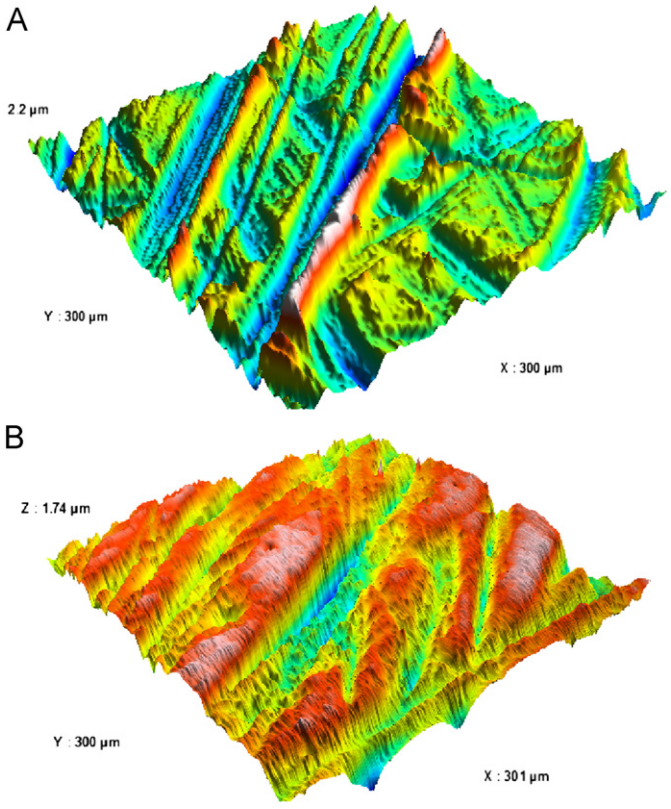

Fig. 13. Isometric plots of typical areas having a low brightness level and containing a high density of large, deep and multidirectional scratches: (A) with build-up material to their sides in the equatorial region SSER, (B) without build-up material to their sides in the polar region SSPR.

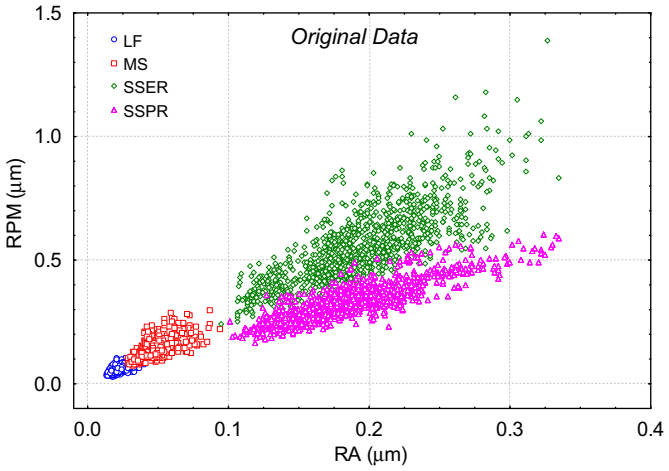

Fig. 14. (R_a, R_{pm}) values recorded in the four worn regions denoted LS, MS, SSER and SSPR.

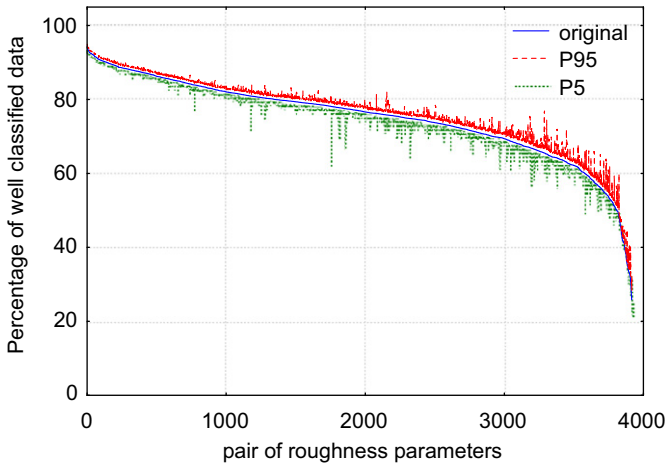

Fig. 15. Results of the Bootstrapped Paired Discriminant Analysis applied on the retrieved metallic femoral head showing the ranking of the pairs of roughness parameters with respect to their average percentage of well classified data and with the original data, 5% and 95% centiles obtained from 100 bootstrap simulations.

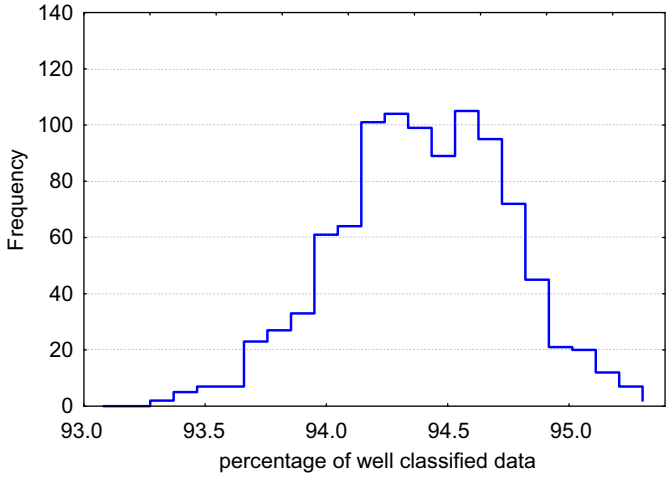

Fig. 16. Empirical PDF of the percentage of well classified data obtained for the pair (R_a, R_{pm}) using the Bootstrapped Paired Discriminant Analysis.

the BPDA methodology while more than an hundred of roughness parameters have been considered in this statistical treatment. It should be noted that this result is statistically significant since the same procedure has been applied under H_0 hypothesis performing 1000 bootstrap simulations and leads to a fortuitous percentage of well classified data of about 28% $\pm$ 7% for the pair (R_a, R_{pm}). Fig. 17 shows the results obtained by means of Discriminant Analysis and reflected in terms of well classified data and wrongly classified data for these two relevant roughness parameters when applied on the original data related to the four worn regions under consideration; original data already presented in Fig. 14.

The statistical results obtained by the BPDA methodology are therefore really interesting since they are fully in agreement with the conclusions proposed by the authors in other recent works about the relevance of roughness parameters with regard to the topography related the scratch mechanism of the retrieved femoral head under consideration. Combined with those obtained in the aforementioned recent works, the results presented in this paper enables a better understanding of the overall degradation mechanism which can be summarized as follows. After migration from the fibermesh of the Harris-Galante Mark I metallic shell in which the UHMWPE was fixed, the centimeter-length metallic

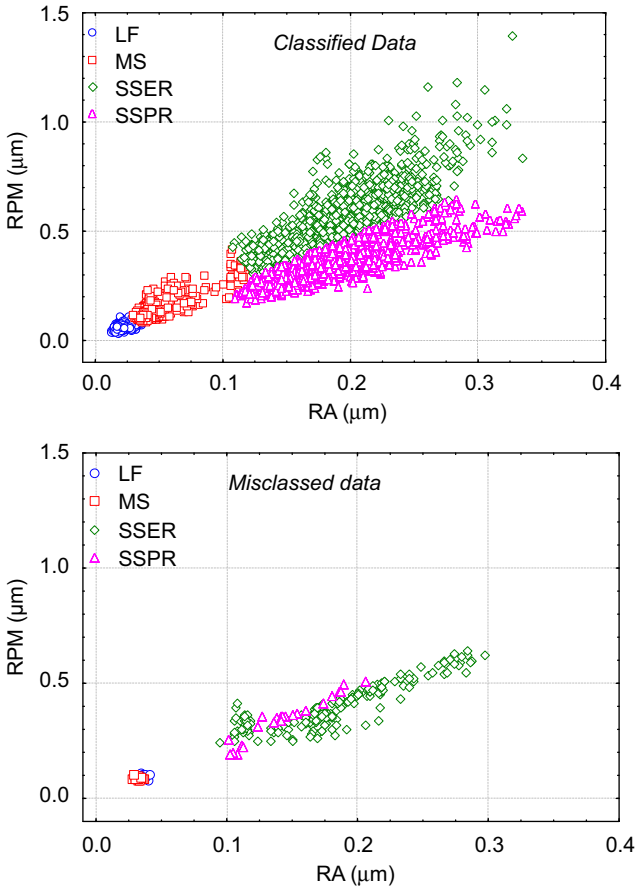

Fig. 17. Results using Discriminant Analysis and showing the well classified data and the wrongly classified data for the pair (R_a, R_{pm}) when applied on the original data related to the four worn regions of the retrieved femoral head under consideration.

fiber, which can be observed in figure, is thought to have entered in the equatorial region SSER from which it progressively and randomly moved towards the polar region SSPR. During its random movements into the joint space, this fiber acted as a large-sized third body that is likely to have generated in a first time the numerous large scratches accompanied material build-up on either sides; this region showing consequently the highest values both for the arithmetic roughness parameter R_a and the mean peak height R_{pm}. After being embedded and fixed in its definitive position relatively to the UHMWPE acetabular liner, this hard third body is suspected to have participated for a second time to a continuously acting wear process due to its multiple passages over the region SSPR. In this region, this continuously acting wear process might have consequently eroded the material build-up (highest peaks) while leaving unaffected the deepest valleys of the large scratches previously generated during the random movement of the fiber into the joint space. Consequently, in comparison with the region SSER, the value of arithmetic roughness parameter R_a remains approximately the same whereas that of the mean peak height R_{pm} decreases significantly. The effect of such an abrasive process in this region on the evolution of the roughness parameters has been confirmed by means of simulation in [34]. This metal/metal friction between the centimeter-sized fiber and the femoral head occurring both in the regions SSER and SSPR is thought to have generated numerous metallic micrometer-sized wear debris. These debris could have acted themselves as small hard third bodies at the origin of the numerous or isolated small scratches, respectively found in the regions MS and LS. Consequently, in comparison with the severely

scratched regions SSER and SSPR, both the values of the arithmetic roughness parameter R_a and the mean peak height R_{pm} decreases significantly; these values being of course higher in the region MS than in the region LS.

Since scratching of metallic femoral heads of Charnley type prostheses is increasingly suspected to be caused by hard third bodies such as particles of bone, cement and/or metal, which have often been found embedded in acetabular cups and in the periprosthetic tissues [22–26,32,33], there is therefore an utility in exploring which sites of head scratching tend to be the most deleterious in terms of increasing UHMWPE wear, with a view to devising steps to impede third body access to these most deleterious head sites [24]. However, due to the stochastic aspects related to the generation, the migration and the characteristics of these hard third bodies, the locations of these deleterious head sites and the textural variation of the surface topography of the worn regions are difficult to predict as well as the time in service life at which roughening occurred [24]. Taking into consideration such a variable and unpredictable clinical situations, the hope to find a single and universal relevant roughness parameter for assessing the surface degradation of femoral heads in relation to the wear of UHMWPE components can be questioned from a general point of view.

7. Conclusion

In this paper, we proposed a new and original methodology designed to select without preconceived opinion either a single or a pair of roughness parameters relevant for discriminating different topographies with regard to a specific application. Two statistical methods were combined in this methodology we calibrated on simulated fractal profiles: the usual discriminant analysis which enabled to define and estimate a quantitative indicator of performance (the average percentage of well classified data) for each roughness parameter and the recently developed and powerful Computer Based Bootstrap Method which enabled to define and estimate a 90% confidence interval on the value of this indicator.

Among more than 3000 pairs of roughness parameters tested, the results of this methodology revealed that, as expected, the pair 'total amplitude R_t–fractal dimension estimated by ANAM Δ_{ANAM}' was the most relevant to be selected for discriminating the topography of different simulated fractal profiles which roughness was intentionally modulated by varying a scaling amplitude factor and the value of the Hölder exponent.

This methodology has been applied on a retrieved metallic femoral head which surface was severely scratched due to the abrasive action of a centimeter-length metallic fiber embedded into the surface of the UHMWPE counterface component. For this retrieved femoral head, it is shown that the most relevant pair of roughness parameters for discriminating the different investigated worn regions is the arithmetic roughness parameter paired with the mean peak height. This methodology helps in a better understanding of the scratch mechanism of this orthopedic component by a centimeter-length metallic fiber that entered into the join space and participated to a third body abrasive mechanism.

It should be recalled that knowledge of the mechanism behind femoral head scratching is not yet fully understood in clinical situations, especially in the case of contemporary constructs with modular interfaces and/or porous metal coatings likely to liberate aggressive particulate metal debris in unpredictable conditions. However, the knowledge of this scratch mechanism through the characterization of femoral heads surface topography by means of roughness parameters is of paramount importance to improve the long term clinical performance of the Charnley total hip prostheses. As far as foreign third bodies are frequently concerned in the overall degradation process of articulating components of contemporary medical devices, it should be interesting to perform similar investigations on other retrieved femoral heads, firstly, to improve our knowledge of the characteristics (location, size and topography) of the most deleterious sites with regard to the surface degradation and, secondly, to specify which roughness parameters are relevant to assess the surface topography in relation to the wear of UHMWPE components.

Acknowledgment

We wish to thank Véronique Hague for her assistance in English. This work was supported by the Region Picardie and FEDER found in the Project "FoncRug3D *Nouvelles méthodes d'analyze des états de surfaces: de la caractérisation à la recherche de paramètres pertinents*" and by the Région Nord-Pas de Calais in the Project *"Fédération des Biomatériaux et Dispositifs Médicaux Fonctionnalisés"*.

References

[1] Whitehouse DJ. The parameter rash—Is there a cure? Wear 1982;83:75–8.
[2] Nowicki B. Multiparameter representation of surface roughness. Wear 1985;69.
[3] Scott RS, Ungar PS, Bergstrom TS, Brown CA, Grine FE, Teaford, Walker A. Dental microwear texture analysis within-species diet variability in fossil hominins. Nature 2005;436(4):693–5.
[4] Narayan P, Hancock B, Hamel R, Bergstrom TS, Brown CA. Differentiation of the surface topography of various pharmaceutical excipient compacts. Materials Science and Engineering 2006;A430(1–2):79–89.
[5] Jordan SE, Brown CA. Comparing texture characterization parameters on their ability to differentiate ground polyethylene ski bases. Wear 2006;261:398–409.
[6] Berglund J, Brown CA, Rosen BG, Bay N. Milled die steel surface roughness correlation with steel sheet friction. CIRP Annals Manufacturing Technology 2010;59(1):577–80.
[7] Bigerelle M, Najjar D, Iost A. Relevance of roughness parameters for description and modelling of machined surfaces. Journal of Materials Science 2003;38(11):2525–36.
[8] Najjar D, Bigerelle M, Iost A. The Computer Based Bootstrap Method as a tool to select a relevant surface roughness parameter. Wear 2003;254(5–6):450–60.
[9] Bigerelle M, Najjar D, Iost A. Description d'une nouvelle méthode de corrélation entre la rugosité et une propriété de surface, application à la brillance de tôles skin-passées. Revue de Métallurgie 2002;5:467–79.
[10] Bigerelle M, Najjar D, Iost A. Multiscale functional analysis of wear: a fractal model of the grinding process. Wear 2005;258:232–9.
[11] ISO 4287:1997. Geometrical Product Specifications (GPS)—Surface texture: Profile method — Terms, definitions and surface texture parameters.
[12] ASME B46.1:2009. Surface texture, surface roughness, waviness and lay.
[13] ISO 13565-2:1996 Geometrical Product Specification (GPS)—Surface texture: Profile method; Surfaces having stratified functional properties—Part 2: Height characterization using the linear material ratio curve.
[14] Dubuc B, Quiniou JF, Roques-Carnes C, Tricot C, Zucker SW. Evaluating the fractal dimension of profiles. Physical Review A 1989;39(3):1500–12.
[15] Majumdar A, Bhushan B. Role of fractal geometry in roughness characterization and contact mechanics of surfaces. ASME Journal of Tribology 1980;112:205–16.
[16] Bigerelle M, Iost A. Calcul de la dimension fractale d'un profil par la méthode des autocorrélations moyennées normées. Comptes Rendus De L Academie Des Sciences Serie b 1996;323:669–75.
[17] Cooley WW, Lohnes PR. Multivariate data analysis. New York: John Wiley & Sons; 1971.
[18] Klecka W. Discriminant analysis. quantitative applications in the social sciences, series no. 19. Thousand Oaks, CA: Sage Publications; 1980.
[19] Lachenbruch PA. Discriminant analysis. NY: Hafner; 1975.
[20] Efron B, Tibshirani RJ. An introduction to the Bootstrap. New York: Chapman and Hall; 1993.
[21] Hall P. The Bootstrap and the edgeworth expansion. New York: Springer-Verlag; 1992.
[22] Hall RM, Unsworth A, Siney P, Wroblewski BM. The surface topography of retrieved femoral heads. Journal of Materials Science: Materials in Medicine 1996;7:739–44.
[23] Hall RM, Siney P, Unsworth A, Wroblewski BM. The effect of surface topography of retrieved femoral heads on the wear of UHMWPE sockets. Medical Engineering and Physics 1997;19:711–9.

[24] Brown TD, Stewart KJ, Nieman JC, Pedersen DR, Callaghan JJ. Local head roughening as a factor contributing to variability of total hip wear: a finite element analysis. Journal of Biomechanical Engineering 2002;124:691–8.
[25] McNie CM, Barton DC, Ingham E, Tipper JL, Fischer J, Stone MH. The predictions of polyethylene wear rate and debris morphology produced by microasperities on femoral heads. Journal of Materials Science: Materials in Medicine 2000;11:163–74.
[26] Tipper JL, Ingham E, Hailey JL, Besong AA, Fisher J, Wroblewski BM, Stone MH. Quantitative analysis of polyethylene wear debris, wear rate and head damage in retrieved Charnley hip prostheses. Journal of Materials Science: Materials in Medicine 2000;11:117–24.
[27] Dowson D, El-Hady Diab MM, Gillis BJ, Atkinson JR. Influence of counterface topography on the wear of UHMWPE under wet and dry conditions. In: Lee LH, editor. Polymer wear and its control, 287. American Society Series; 1985. p. 171–87.
[28] Dowson D, Taheri S, Wallbridge NC. The role of counterface imperfections in the wear of polyethylene. Wear 1987;119:277–93.
[29] Wang A, Essner A, Polineli VK, Stark C, Dumbleton JH. Lubrication and wear of UHMWPE in total joint replacements. Tribology International 1998;31: 17–33.
[30] Elfick APD, Smith SL, Green SM, Unsworth A. The quantitative assessment of UHMWPE wear debris produced in hip simulator testing: the influence of head material and roughness, motion and loading. Wear 2001;249:517–27.
[31] Bowsher JG, Shelton JC. A hip simulator study of the influence of patient activity level on the wear of crosslinked polyethylene under smooth and roughened femoral conditions. Wear 2000;259:167–79.
[32] McNie CM, Barton DC, Ingham E, Tipper JL, Fischer J, Stone MH. Modeling of damage to articulating surfaces by third body particles in total joint replacements. Journal of Materials Science: Materials in Medicine 2000;11: 569–78.
[33] Najjar D, Behnamghader A, Iost A, Migaud H. Influence of a foreign body on the wear of metallic femoral heads and polyethylene acetabular cups of total hip prostheses. Journal of Materials Science 2000;35:4583–8.
[34] Najjar D, Bigerelle M, Migaud H, Iost A. Identification of scratch mechanisms on a retrieved metallic femoral head. Wear 2005;258:240–50.
[35] Najjar D, Bigerelle M, Migaud H, Iost A. About the relevance of roughness parameters used for characterizing worn femoral heads. Tribology International 2006;39:1527–37.