



Article

Is Surface Metastability of Today's Ceramic Bearings a Clinical Issue?

Alessandro Alan Porporati ^{1,2}, Laurent Gremillard ³ , Jérôme Chevalier ³, Rocco Pitto ⁴ and Marco Deluca ^{5,*}

¹ CeramTec GmbH, Medical Products Division, Ceramtec-Platz 1-9, 73207 Plochingen, Germany; a.porporati@ceramtec.de

² Department of Engineering and Architecture, University of Trieste, Via Alfonso Valerio 6/1, 34127 Trieste, Italy

³ Univ Lyon, INSA Lyon, UCBL, CNRS, MATEIS, UMR 5510, F-69621 Villeurbanne, France; laurent.gremillard@insa-lyon.fr (L.G.); jerome.chevalier@insa-lyon.fr (J.C.)

⁴ Department of Orthopaedic Surgery, Middlemore Hospital, University of Auckland, South Auckland Clinical Campus, Auckland 93311, New Zealand; r.pitto@auckland.ac.nz

⁵ Materials Center Leoben Forschung GmbH, Roseggerstr. 12, 8700 Leoben, Austria

* Correspondence: marco.deluca@mcl.at; Tel.: +43-3842-45922-530

Abstract: Recent studies on zirconia-toughened alumina (ZTA) evidenced that in vivo aged implants display a much higher monoclinic zirconia content than expected from in vitro simulations by autoclaving. At the moment, there is no agreement on the source of this discrepancy: Some research groups ascribe it to the effect of mechanical impact shocks, which are generally not implemented in standard in vitro aging or hip walking simulators. Others invoke the effect of metal transfer, which should trigger an autocatalytic reaction in the body fluid environment, accelerating the kinetics of tetragonal-to-monoclinic transformation in vivo. Extrapolations of the aging kinetics from high (autoclave) to in vivo temperature are also often disputed. Last, Raman spectroscopy is by far the preferred method to quantify the amount of monoclinically transformed zirconia. There are, however, many sources of errors that may negatively affect Raman results, meaning that the final interpretation might be flawed. In this work, we applied Raman spectroscopy to determine the monoclinic content in as-received and in vitro aged ZTA hip joint implants, and in one long-term retrieval study. We calculated the monoclinic content with the most used equations in the literature and compared it with the results of X-ray diffraction obtained on a similar probe depth. Our results show, contrary to many previous studies, that the long-term surface stability of ZTA ceramics is preserved. This suggests that the Raman technique does not offer consistent and unique results for the analysis of surface degradation. Moreover, we discuss here that tetragonal-to-monoclinic transformation is also necessary to limit contact damage and wear stripe extension. Thus, the surface metastability of zirconia-containing ceramics may be a non-issue.

Keywords: zirconia-toughened alumina; phase transformation; Raman spectroscopy



Citation: Porporati, A.A.; Gremillard, L.; Chevalier, J.; Pitto, R.; Deluca, M. Is Surface Metastability of Today's Ceramic Bearings a Clinical Issue? *J. Compos. Sci.* **2021**, *5*, 273. <https://doi.org/10.3390/jcs5100273>

Academic Editor: Francesco Tornabene

Received: 1 August 2021

Accepted: 8 October 2021

Published: 14 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The current trend in total hip arthroplasty (THA) is to gradually prefer ceramic-based implants over metallic implants due to their excellent biocompatibility [1], both in bulk and particulate form, and high long-term survival [2]. Nowadays, the ceramic of choice for THA is zirconia-toughened alumina (ZTA), often in the form of the BIOLOX[®]delta material, the most commercially successful material. BIOLOX[®]delta (Delta) was developed in the early 2000s by CeramTec GmbH and is composed of 17 vol.% yttria-stabilized tetragonal zirconia particles (Y-TZP) embedded into an alumina matrix. The function of Y-TZP is to improve the toughness of the prosthetic material, through a mechanism called phase transformation toughening: under stress, a fraction of the tetragonal zirconia grains may undergo a phase transition to a monoclinic phase (t-m transformation). This phase has

a larger volume and thus sets the surrounding material under compressive stress. This increases the crack propagation resistance of the material. The alumina phase, on the other hand, provides hardness and wear resistance [3]. Furthermore, Delta also contains 0.3 w.% chromia, which imparts the pink color to the material, and about 3 vol.% of strontium hexaaluminate platelets, which further increases the fracture toughness of the ceramic composite. The stability of alumina matrix material bearings in the body of patients is still being discussed, despite the best long-term clinical outcomes shown by the arthroplasty registries when compared to alternative bearings [4–8]. Before being introduced into the medical market, the Delta material was thoroughly tested according to the available ASTM F2345-03 standard for in vitro testing, which states that a one-hour exposure at 134 °C under 2 bars water steam (in autoclave) corresponds to two years in vivo [9]. Extrapolations based on those tests predicted a very slow transformation (less than 5% increase over the first 10 years) [10]. However, investigations on revised Delta implants after few years (in some cases, months) in vivo [3,11–14], where the causes of revision were unrelated to the ceramic material, revealed worn areas [12,14,15], the presence of metal transfer [11–14], and a much higher monoclinic content than expected according to in vitro simulations [11,13]. This monoclinic content was always necessarily reported as the difference between worn and non-worn areas [11,15] because the initial monoclinic phase amount cannot be measured prior to implantation, and this value may vary slightly due to the manufacturer's batch processing [11].

Many material research groups have attempted to clarify this discrepancy between in vitro and in vivo results, but up to now, no consensus on the real causes has been found. Pezzotti et al. proposed one hypothesis [11,16–18] that ascribes this discrepancy to the effects of metal transfer. In fact, in correspondence with metal transfer, high wear values and high monoclinic contents in zirconia were measured [11–13]. According to Pezzotti et al. [11], in the aqueous body fluid environment, the autocatalytic dissociation of water molecules is promoted at the transferred metal's surface; this apparently causes an annihilation of oxygen vacancies first in the alumina phase and then in the tetragonal zirconia phase of Delta ceramics. This latter oxygen vacancy reduction should then trigger the tetragonal-to-monoclinic transformation in zirconia. This conclusion was supported by cathodoluminescence and X-ray photoelectron spectroscopies [11,16], and similar mechanisms were also observed in yttria-stabilized zirconia [19]. Although this explanation may sound plausible, especially in the case of metal transfer, it is not clear whether the initial data from the retrievals are consistent: the measured high wear could be due to the surface asperities produced by metal smearing, and not the ceramic. In addition, more likely, the presence of metal affects the quantification of the monoclinic phase by Raman spectroscopy because of the influence of metal on the intensity of Raman peaks due to plasmonic effects [20]. A poor signal-to-noise ratio could also affect spectra collected in the presence of chemisorbed proteins [21], which highlights the need for a thorough surface cleanup before Raman analysis.

Another explanation has been proposed by Perrichon et al. [15,22,23], who ascribe the discrepancy to the effect of shocks due to microseparation, which are not implemented in the ASTM in vitro standard, or in hip walking simulators. Perrichon et al. were able to demonstrate that using a specific in vitro test route that includes hip walking simulations, shock tests, and environmental aging tests, the discrepancy with in vivo studies can be reduced [23]. In this case, the monoclinic content on worn areas was much higher than on non-worn zones, in agreement with observations on retrieved implants. This shows that the phase transformation toughening mechanism was activated under stresses to limit mechanical damage. In other words, t–m transformation due to shock is not related to degradation due to aging.

The last explanation is related to the model used to relate the simulated aging kinetics obtained in the autoclave to in vivo aging. It is postulated that aging follows an Arrhenius law, and thus knowledge of its activation energy enables establishing a time–temperature equivalence. However, two main limits exist for this model: First, data at a low temper-

ature take a long time to obtain and thus are hardly available; thus, it is not completely certain that the Arrhenian behavior is still valid around body temperature. Second, small uncertainties in the activation energy may lead to large variability in the time–temperature equivalence, while the ASTM standard tacitly considers a single activation energy for all zirconia-containing materials.

However, more importantly, it has been recently shown that the *in vivo* zirconia transformation in Delta does not affect the mechanical performance of the total hip arthroplasty components [24]. It follows that differences in the *t*–*m* transformation of zirconia in the Delta material between *in vitro* tests and *ex vivo* components may not be clinically relevant; therefore, solving the problem is conducted only for the sake of knowledge. In the following, we will try to explain the differences due to the measurement method.

Although it is clear that *in vitro* testing standards should be reviewed in order to better approach the conditions encountered by the implant inside the patient’s body, one aspect may be overlooked: comparisons between *in vitro* and *in vivo* studies are based mainly on the value of V_m , the volume fraction of monoclinic zirconia, which is measured by Raman spectroscopy. However, currently, there is neither a standard that exists regarding how V_m should be measured from Raman spectra, nor sufficient explanations reported in the previous literature about how data collection and treatment have been performed. Very likely, each research group uses a different procedure for V_m quantification, which (at least) makes the comparison of V_m results obtained by different groups questionable. In addition, some data analysis procedures might lead to severe artefacts, with the consequence that wrong values of V_m are calculated, leading to flawed interpretations.

In this work, we show that the use of the Clarke/Adar equation—deemed unreliable by some research groups—is indeed the best choice for our measurement setup for Delta ceramics. This is confirmed by a control procedure using X-ray diffraction (XRD). Furthermore, we demonstrate—by analyzing a retrieval affected by metal transfer—that the effect of metal on the monoclinic transformation is negligible. Lastly, we show that specific choices in data analysis, such as the use of the absolute/integrated intensity, the choice of a baseline, or the overall signal-to-noise ratio of the spectrum, have a large impact on the obtainable results in terms of V_m . These results clearly suggest that a standard for V_m quantification using Raman spectroscopy (including sample preparation, spectroscopic procedures, and data treatment) should be promptly put in place.

2. V_m Quantification by Raman Spectroscopy

Raman spectroscopy probes the inelastic light scattering from vibrational motions of atoms in a solid, and as such, it is sensitive to any change in the way atoms vibrate, as caused, for instance, by the presence of a different phase. In other words, the Raman spectrum is a fingerprint of the state a solid assumes. In zirconia ceramics, the tetragonal and monoclinic polymorphs present very different spectra [25], and in mixed phases, a superposition between those two spectra appears, the extent of which depends on the volume fraction of the monoclinic phase, V_m , in the investigated area. Various researchers have derived an expression to quantify V_m from the intensity of Raman peaks belonging to tetragonal and monoclinic phases, building upon equations already available for XRD analyses [26]. The equation has the following form [26,27]:

$$V_m = \frac{I_m^{181} + I_m^{190}}{k(I_t^{147} + \delta I_t^{265}) + I_m^{181} + I_m^{190}} \quad (1)$$

and it differs among the available approaches only for the values of the δ and k coefficients. $I_{m,t}^i$ is the intensity of Raman peaks (at position i —in cm^{-1}) belonging to monoclinic and tetragonal phases. The two most used equations for the determination of V_m in THA implants are the one derived by Clarke and Adar ($\delta = 1$, $k = 0.97$) [27] and the one derived by Katagiri et al. ($\delta = 0$, $k = 2.2 \pm 0.2$) [26], whereby the latter has been used in the majority of recent studies concerning the Delta material. Tabares and Anglada [26] recently carried out a systematic study with both Raman and XRD using bulk mixtures of tetragonal and

monoclinic zirconia powders with from 0% to 100% of monoclinic phase content. They calculated V_m using both the Clarke/Adar and Katagiri equations and demonstrated that while Katagiri's equation correctly reproduced the monoclinic content, the Clarke/Adar equation largely underestimated it. XRD results were also in better accordance with the Katagiri equation. They suggested that this discrepancy is related to the localization of the monoclinic polymorph (i.e., the V_m profile) in the material used for calibration by Clarke and Adar: fracture surfaces in ZrO_2/Y_2O_3 specimens, where the monoclinic phase is expected to be present only near the surface. In this case, the penetration depth of X-rays depends on the angle of incidence, and thus it can be suggested that the discrepancy in the V_m calculated with Clarke and Adar's formula is due to a different angle of incidence that Tabares and Anglada used for their XRD measurements, compared to the one used by Clarke and Adar [26]. Hence, according to Tabares and Anglada, the Katagiri equation seems to have universal validity because it has been obtained on bulk mixtures of tetragonal and monoclinic zirconia powders, where the monoclinic content is homogeneous across the whole probed volume by both Raman and XRD. Its validity, however, has neither been systematically demonstrated in materials where a sharp gradient in the monoclinic phase is present, nor in sintered materials. Both aspects clearly apply to *in vitro* and *in vivo* aged Delta [3,9,11,28].

Apart from the choice of the equation to calculate V_m , there are many other aspects that could lead to errors and discrepancies between the V_m values reported in the literature:

- Raman spectra need to be fitted (using mathematical expressions) to obtain intensity values of the respective tetragonal and monoclinic peaks. Spectra with different qualities (i.e., different signal-to-noise ratios, SNRs) might lead to different V_m values because of fitting errors. Factors influencing the quality of spectra are the optical system, the laser, the time used for collection/accumulation of spectra, and the quality of the investigated surface.
- Spectra are often affected by a background due to elastic scattering or to the presence of fluorescence (particularly true for Delta). In these cases, a baseline is generally subtracted to avoid the influence of the background on the final result [26]. However, the choice of the baseline might affect the final result as well.
- It is not clear whether the integrated or the absolute intensity of Raman peaks should be used in Equation (1). With Equation (1) being an intensity ratio, this question may seem unimportant; however, the absolute intensity might not fully represent the monoclinic content, especially by low V_m values [29].
- Each spectrometer used for Raman analysis has different characteristics (e.g., the focal length, the number of gratings, the confocal pinhole width) affecting the SNR ratio and the spectral resolution, which could lead to different results if the same material is probed by different equipment.

A comparison between the Clarke/Adar and Katagiri equations using both Raman and XRD on Delta has not yet been reported in the literature, and also a thorough analysis of the aforementioned error sources (even partly) has never been attempted.

3. Materials and Methods

3.1. *In Vitro* Aging Study Samples

Ten Delta heads and ten Delta cup inserts (CeramTec GmbH, Plochingen, Germany) were analyzed by both XRD and Raman spectroscopy in order to independently quantify the monoclinic content. The areas investigated corresponded to the head apex in the heads (polished), and to the center of the bottom (opposite of the cup) in the inserts (ground). The two different surface finishes were selected with the intent to attempt to cover as much of the V_m range, from 0% to 100%, as possible, this way mimicking non-wear and wear zones in real implants, respectively. The aforementioned total hip arthroplasty implant components were tested both before and after extreme hydrothermal aging in an autoclave at 134 °C and 2.2 bars for 150 h, which would correspond to more than 300 years *in vivo* according to the ASTM standard.

3.2. Retrieval Sample

The studied retrieval was constituted by a fully Delta ceramic-on-ceramic (CoC) bearing couple. The total hip replacement (THR) components were a DePuy Pinnacle cup 60 mm and a Summit stem size 3 high offset. The 28 mm Delta ceramic head taper had a +8.5 mm neck length offset.

The patient was informed that the data concerning the case would be submitted for publication, and she provided consent.

The patient underwent complex total hip arthroplasty in 2004 (diagnosis: secondary osteoarthritis following developmental dysplasia of the hip). After twelve years of pain-free normal function, the patient presented with periprosthetic joint infection (caused by *Klebsiella pneumoniae*). After failure of debridement/irrigation and two dislocation events (managed with closed reduction), the patient underwent two-stage revision surgery in late 2016. The retrieved Delta 28 mm CoC bearing appeared intact, with titanium metal stripes in the femoral head caused by recurrent dislocation events. Areas with metal transfer were investigated before and after a cleaning procedure to remove the metal, which consisted in a 10 h bath at 60 °C with 30% aqueous H₂SO₄ solution.

The patient was 50 years old at the time of index surgery; The body weight of the patient was 78 kg, and the patient's body mass index (BMI) was 31. The patient was a housewife with a part-time administration job, and she performed no sports activities.

3.3. Characterization Methods

X-ray diffraction (XRD) analyses were carried out on a Bruker AXS D8 advance diffractometer (Bruker, Karlsruhe, Germany) using the Bragg–Brentano configuration. The excitation of the tube was fixed at 30 KV and 20 mA, the slit was fixed at 0.6 mm, and the probe size was around 6 × 12 mm. A position-sensitive detector (LynxEye, Bruker) was used to collect the data between 10° and 70° (2θ), with a 0.015° step size and a 0.4 s/step acquisition speed. With this configuration, and considering the peaks of interest (the (−111) and (111) monoclinic peaks and the (101) tetragonal peak, respectively, located around 28.3°, 31.5°, and 30.1°), 90% of the XRD signal comes from the first 17 μm below the surface. The monoclinic fraction was determined from the integrated intensities of the XRD peaks after subtracting a linear baseline, using Garvie and Nicholson's equation [30].

Raman spectra were collected with a single spectrograph (Horiba Jobin Yvon LabRAM HR800) with a grating of 1800 gr/mm and Ar+ laser excitation at 514.5 nm wavelength. The laser power on the sample was maintained at ~2 mW with a 100× long-working distance objective to avoid excessive laser-induced heating. With the chosen optical configuration, the laser had a lateral resolution of ~1 μm and a penetration depth of 4.2 μm or 15 μm when the confocal pinhole was fixed at 100 μm or 1000 μm, respectively (intended as the depth from which 90% of the signal comes from). These values were determined following the procedure outlined in Pezzotti et al. [31]. On each specimen, three adjacent points (>10 μm apart) were measured. The collected spectra were fitted with Gaussian–Lorentzian functions after subtracting a linear baseline; the integrated intensity values of monoclinic and tetragonal peaks were used to calculate V_m with both the Clarke/Adar and Katagiri equations and compared with the results of the XRD analyses. An example of the fitted spectrum after baseline subtraction is shown in Figure 1.

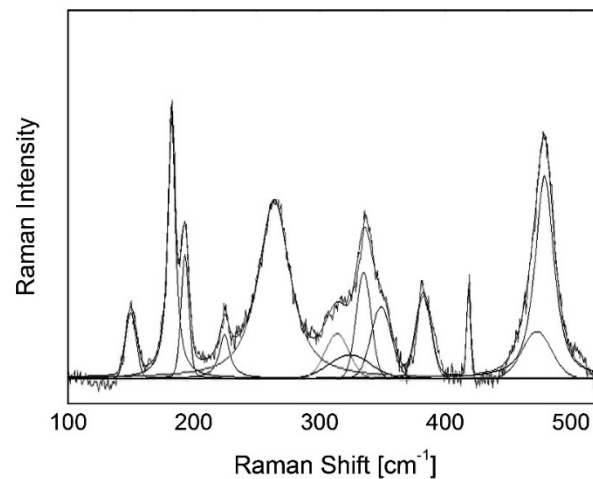


Figure 1. Example of a fitted Raman spectrum of Delta using the procedure followed in this paper. The spectrum was taken on a head specimen in a region with a high monoclinic fraction.

4. Results

4.1. In Vitro Aging Study

Figure 2 and Table 1 present the results of V_m measurements by Raman spectroscopy carried out on Delta femoral heads and inserts, both as received and after the aging procedure. The values of V_m were calculated from the integrated intensity of peaks belonging to the monoclinic and tetragonal phase after the fitting procedure described in Section 3. As it can clearly be seen, the values obtained with the through-focus configuration are smaller than those obtained with the confocal one; in the latter case, due to the smaller probe depth, the volume closer to the surface of the sample was analyzed. Hence, this result shows that the monoclinic fraction is higher in the vicinity of sample surfaces. Aged samples reveal a higher monoclinic content (up to a factor of 2 and higher), as expected, and the difference from the pristine state is larger near the sample surface. Moreover, inserts have a higher monoclinic content due to the raw (grinded backside) surface finish.

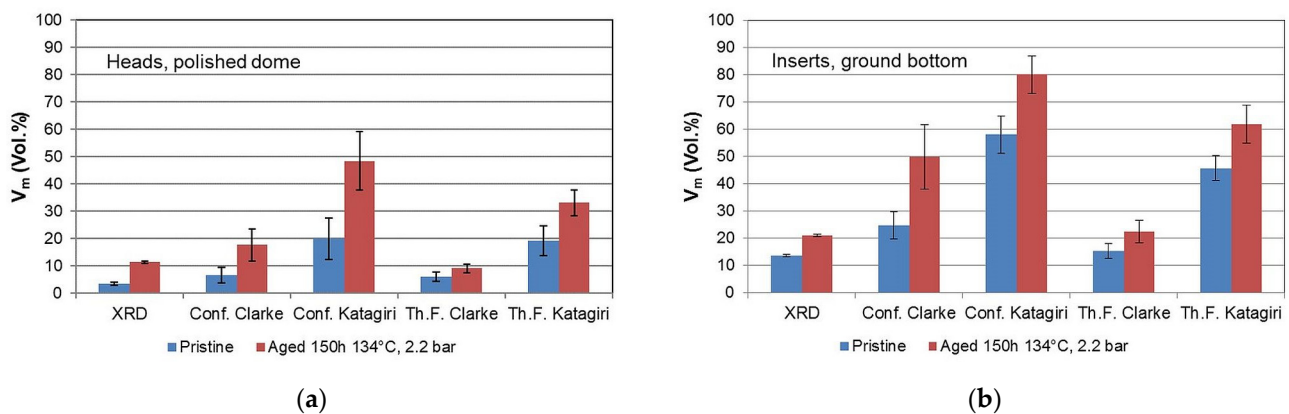


Figure 2. Values of V_m measured by Raman spectroscopy and XRD on both non-aged and in vitro aged Delta heads (a) and inserts (b). Conf. = confocal pinhole closed down to 100 μm (penetration depth of Raman signal: 4.2 μm); Th. F. = confocal pinhole fully open (1000 μm)—penetration depth of Raman signal: 15 μm .

Comparing values obtained with the Clarke/Adar and Katagiri equations, it is evident that a higher monoclinic content results from the Katagiri equation. This is in line with the findings of Tabares and Anglada [24], who concluded that the Clarke/Adar formula underestimated the monoclinic content for powders. However, a direct comparison between the through-focus Raman results and the XRD results (which have a very similar penetration depth of $\sim 15 \mu\text{m}$ and $17 \mu\text{m}$, respectively) shows that, indeed, it is the

Clarke/Adar equation that provides the best correspondence with the XRD measurements. This is valid on both sample types and for both pristine and aged specimens. This is more evident from Figure 3, where a direct comparison between V_m by XRD and Raman is provided for all samples both in the (a) confocal and (b) through-focus configurations. In Figure 3a, both equations overestimate the V_m by Raman, and this is due to the difference in the volume probed by the two techniques (with confocal Raman, the probe depth is much smaller). For the through-focus case (Figure 3b), where a direct comparison between Raman and XRD is more pertinent due to the very similar penetration depth, the Katagiri equation clearly overestimates (by a factor of 2.5) the monoclinic content, whereas the Clarke/Adar equation provides only a slightly lower V_m than XRD. This latter equation seems thus more suitable for the determination of V_m in the case of aged femoral heads, where the monoclinic content is not constant over the probed depth, keeping in mind that the obtained value is then a weighted average over 15 μm under the surface, as with V_m obtained by XRD.

Table 1. Values of V_m measured by Raman spectroscopy and XRD on both non-aged and in vitro aged Delta heads and inserts. Confocal Raman results belong to a depth up to $\sim 4 \mu\text{m}$ below the surface (pinhole diameter: 100 μm). Through-focus Raman data correspond to a fully opened confocal pinhole (1000 μm) and thus encompass a depth of 15 μm . The penetration depth of XRD is 17 μm .

	Raman CONFOCAL				Raman THROUGH-FOCUS				XRD	
	Clarke/Adar		Katagiri		Clarke/Adar		Katagiri		V_m	STD
	V_m	STD	V_m	STD	V_m	STD	V_m	STD		
Heads										
non-aged	7	5	20	13	6	2	19	7	4	1
aged	18	4	49	9	9	2	33	6	11	1
Inserts										
non-aged	25	8	58	8	15	3	46	6	14	1
aged	50	16	80	10	23	8	62	11	21	0

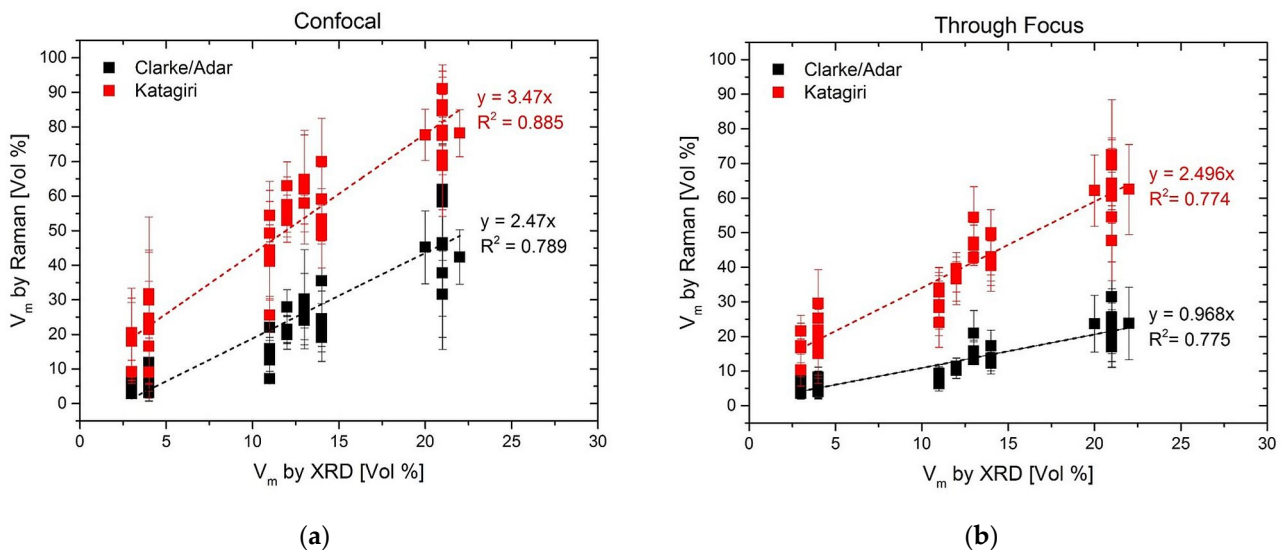


Figure 3. Comparison between all V_m values measured by XRD and Raman spectroscopy and on both non-aged and in vitro aged Delta heads and inserts. (a) Confocal = confocal pinhole closed down to 100 μm (penetration depth of Raman signal: 4.2 μm); (b) Through Focus = confocal pinhole fully open (1000 μm)—penetration depth of Raman signal: 15 μm . The Through Focus measurement mode (with the Clarke/Adar equation) best reproduces the XRD results due to the similar probed volume (penetration depth of XRD: 17 μm).

4.2. In Vivo Aging Study

The explanted femoral head presented significant metal transfer across the whole implant except on apex areas, caused by the two dislocation events with consequent closed reduction procedures. Roughness values measured before and after chemical attack revealed that in the metal transfer area, approximately half of the measured roughness (0.154 μm vs. 0.079 μm in the cleaned sample) was due to the metal smearing and not eventual ceramic surface wear nor scratches on the ceramic surface. Hence, the roughness results reported by other groups in metal transfer areas (without removing the metal) may be questioned [11]. A picture of the retrieved head before and after the cleaning and identification of zones is provided in Figure 4. Zones A, B, C, D, and E are defined as stripe wear, transition area, main wear, metal transfer, and no wear (control area).

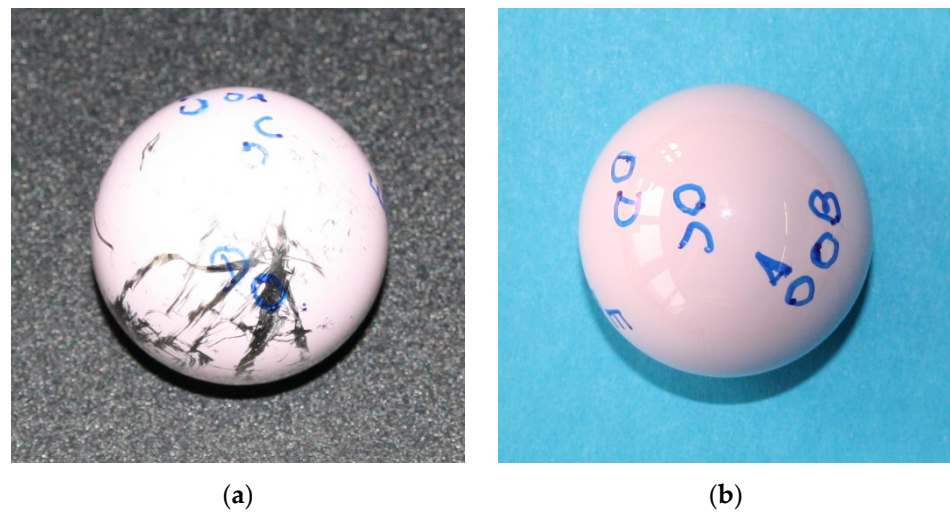


Figure 4. Picture of retrieved head before (a) and after (b) the cleaning and identification of zones.

No evidence was found of an increased monoclinic content led by metal transfer, whereas wear seemed to be more critically related to the monoclinic content: in wear areas, we found a higher monoclinic content (especially at the surface)—cf. Table 2. This result supports the interpretation that the discrepancy between in vitro and in vivo is related to shocks contributing to wear and stress-induced phase transformation rather than to metal transfer [15,22,23].

Table 2. Roughness values and values of V_m by XRD; Raman V_m values are determined as difference from E (non-wear case).

Area	Description	Roughness (μm)		V_m , Raman (%)			V_m , XRD (%)
		Metal	Metal Removed	Confocal (Clarke/Adar)	Confocal (Katagiri)	Through-Focus (Clarke/Adar)	Through-Focus (Katagiri)
A	Stripe wear	0.035	0.036	5	18	3	5
B	Transition area	0.012	0.017	16	36	8	19
C	Main wear	0.009	0.008	40	49	20	28
D	Metal transfer	0.154	0.079	20	35	7	13
E	No wear	0.008	0.007	-	-	-	-

5. Discussion

5.1. Use of Clarke/Adar and Katagiri Equations

Our study clearly demonstrates that in the investigated materials (both for in vitro and in vivo aged specimens), the Clarke/Adar equation, and not the Katagiri equation, produced results that are in better accordance with the XRD measurements. This contradicts the current trend in the literature and suggests that the validity of Raman data in the literature is questionable. There is, in particular, a discrepancy with Tabares and Anglada’s work [26], where on the basis of Raman and XRD measurements on several monoclinic/tetragonal powder mixtures, Katagiri’s equation was deemed more suitable,

whereas the Clarke/Adar equation underestimated the results. Tabares and Anglada explained this result with an intrinsic difference residing in the experimental procedure followed by Clarke and Adar: They used fracture surfaces of sintered samples in which the monoclinic phase was confined to a thin surface layer. Consequently, Clarke and Adar's specimens were affected by a concentration gradient in the depth direction, which caused the value of V_m measured by XRD to depend on the wavelength and the angle of incidence of the radiation. In other words, since Tabares and Anglada used different XRD settings for their calibration, the value of 0.97 for the k coefficient in Equation (1) is not valid in their case, and the Clarke/Adar equation underestimates V_m .

This, however, should also apply to our case. Interestingly, it is the Clarke/Adar equation that performs better in our case. One possible explanation is the fact that in our work, we carried out all measurements on sintered samples. It may be envisaged that the coefficient $k = 2.2$ obtained by Katagiri and confirmed by Tabares/Anglada for the Katagiri equation is valid only on powder mixtures, whereas the functional form including the tetragonal peak at $\sim 265 \text{ cm}^{-1}$ (and $k = 0.97$) has to be taken for a sintered material, which is the case of the calibration performed by Clarke and Adar. Another possible explanation is the fact that both Clarke/Adar and Tabares/Anglada worked on monolithic zirconia (thus with a much lower penetration depth for XRD: around $5 \mu\text{m}$).

A further proof that the Clarke/Adar equation, and not the Katagiri equation, has to be used for our setting is provided in Figure 5. The upper (blue) spectrum in Figure 6 belongs to an area (named area A) with a low V_m located at the apex of non-aged polished Delta head domes, whereas the lower spectrum (red) corresponds to regions with a high V_m (named area B) at the center of the ground bottom of aged heads and inserts. The spectrum in area A is associated with a V_m of 10.2% or 30.5% if calculated with the Clarke/Adar or Katagiri equation, respectively. The spectrum in area B corresponds to a V_m of 66.7% (Clarke/Adar) or 90.1% (Katagiri). Such a high monoclinic content as obtained from the Katagiri equation seems unlikely given the still very strong intensity of the tetragonal peak at $\sim 265 \text{ cm}^{-1}$. In a fully monoclinic material, the 265 cm^{-1} peak is, in fact, absent [18].

The main intrinsic limitation of the Katagiri equation is evident from its functional form displayed in Equation (1): For the calculation of V_m , it considers two monoclinic peaks and only one tetragonal peak. Consequently, if the coefficient k is not correct for the investigated material, the contribution of the monoclinic peaks is disproportionately high. Very likely, for the investigated sintered material, the coefficient k should be higher. Based on a comparison with the V_m obtained here using the Clarke/Adar equation, a coefficient of $k = 4.7$ for the Katagiri equation is probably more realistic in the present case. The coefficient k is probably not only dependent on the materials used for the calibration but also on the type of Raman spectrometer used and on the depth profile of the monoclinic fraction. A careful analysis of the available literature, in fact, suggests that the Katagiri equation performs better on triple spectrometers [11,24,32], whereas the Clarke/Adar equation performs better on single spectrographs [15,21]. This might be explained by differences in the measured relative intensities by the different equipment.

5.2. Spectral Quality and Fitting

Further aspects that could lead to differences in the values of V_m published by various research groups are (i) the overall quality (in terms of the SNR) of the collected spectra and (ii) the procedure used for data regression. Let us first investigate the former aspect. Figure 6 reports two spectra collected on the same polished spot of a Delta head. One spectrum was taken with shorter acquisition times and less repetitions in order to obtain two spectra with very different SNRs. The low-SNR spectrum mimics the case in which a spectrum was taken focusing through the metal in an area affected by metal transfer on a retrieved implant (cf. Figure 3c in [13]). The high-SNR spectrum (black line) corresponds to a V_m of 14% or 34% (with the Clarke/Adar or Katagiri equation, respectively), whereas the low-SNR spectrum corresponds to a V_m of 19% or 41% (Clarke/Adar or Katagiri equation, respectively). Therefore, despite those spectra belonging to the same area, a difference of

~20% was obtained. In other words, using spectra with a low SNR (such as the ones taken in a metal transfer area without removing the metal) may produce an overestimation of the monoclinic content of about 20%.

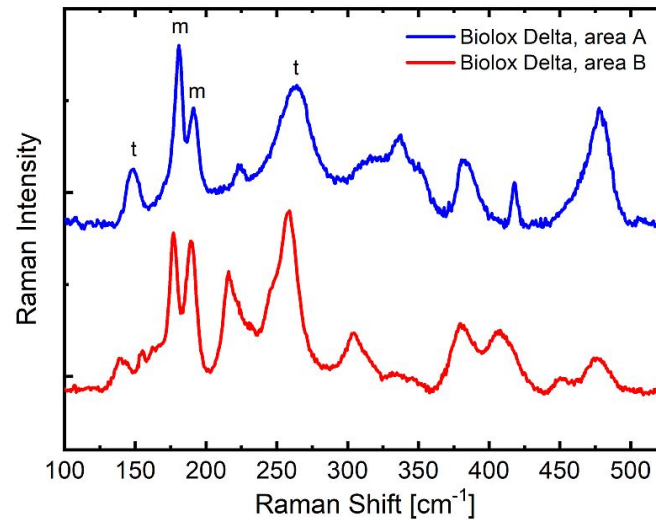


Figure 5. Comparison of Raman spectra of Delta that underwent a low monoclinic transformation (blue line—area A) and a high monoclinic transformation (red line—area B). Peaks belonging to the monoclinic (m) and tetragonal (t) phases of the area used in the analysis are labeled on the upper (area A) spectrum. The area A spectrum is associated with $V_m = 10\%$ and 31% with the Clarke/Adar and Katagiri equations, respectively. In area B, the monoclinic content amounts to 67% and 90% (according to the Clarke/Adar and Katagiri equations, respectively). Such a high monoclinic content as obtained from the Katagiri equation seems unlikely given the still very strong intensity of the tetragonal peak at $\sim 265\text{ cm}^{-1}$.

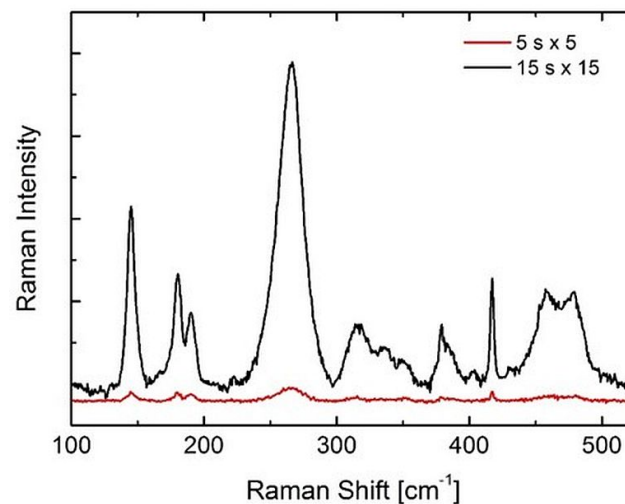


Figure 6. Raman spectra collected on Delta with different acquisition times in order to obtain different SNRs. The spectra were collected on the same point of a polished specimen surface, but the low-SNR spectrum mimics the case of a spectrum collected through a metal layer in correspondence with metal transfer.

Another issue that is often overlooked in the literature is the use of absolute or integrated intensities. In general, integrated intensities should be more suitable in low- V_m cases [27]. Nevertheless, the use of absolute intensities may seem attractive in cases in which a large fluorescence background is present. According to our analysis, using absolute instead of integrated intensities causes an overestimation of the monoclinic content

amounting to 26% for the Katagiri equation and up to 60% for the Clarke equation. Hence, use of integrated intensities is mandatory.

The latter result highlights the intrinsic weakness of the Clarke/Adar equation with respect to variations in the overall background of the spectrum, such as in the case of fluorescent emission. To highlight this aspect, we carried out a study in which we fitted the spectra using two different baselines, both in inserts and heads belonging to the investigated Delta implant components. Figure 7 shows the different baselines used in spectra collected on the rear of an insert (a, b) and on the apex of a head (c, d). These cases will be called Cups A and B and Heads A and B, henceforth. From Cup A, values of V_m of 53% and 84% were measured with the Clarke/Adar and Katagiri equations, respectively. The V_m for Cup B amounted to 64% with the Clarke/Adar equation and 84% with the Katagiri equation. Concerning the head, spectrum A had a V_m of 10% with the Clarke/Adar equation and 35% with the Katagiri equation. Head B produced a V_m of 15% with the Clarke/Adar equation and 34% with the Katagiri equation. Hence, choosing a different baseline brings about an error as high as 33% for the Clarke/Adar equation, whereas the Katagiri equation is more stable (maximum error ~3%). For the sake of clarity, we mention that in all spectra used for comparison in Figure 3 above, we used a 15 s acquisition time, 15 repetitions, and the same baseline used for background subtraction as that reported in Figure 7a.

Despite performing better than the Katagiri equation in the investigated samples, the Clarke/Adar equation thus seems to be more prone to errors. The reason resides in the use of the large tetragonal peak at $\sim 265\text{ cm}^{-1}$, which is strongly influenced by the background and is largely affected by changes in the choice of the baseline for data regression. Hence, Katagiri's choice of excluding this peak from the analysis is not at all wrong; however, we demonstrate here that in this case, the coefficient k has to be recalibrated every time a new material (e.g., sintered instead of powders) or a new instrument is used. Moreover, any modification in both the data collection and treatment procedures risks introducing sources of errors that are non-negligible even in the case of the Katagiri equation. This suggests that the V_m values obtained by different research groups using different equipment and different data treatments can hardly be compared. The only way out of this issue is to define a standard procedure for the analysis of the monoclinic content in zirconia via Raman spectroscopy.

5.3. Proposed Standard Procedure

In the following, we suggest a standard procedure that would allow obtaining V_m values on implants which can be compared between different laboratories:

- First, a series of standard, sintered zirconia samples with a large span of monoclinic content should be prepared in a single batch by the same laboratory or company. These samples should serve as a reference for the calibration of all Raman equipment worldwide.
- Each laboratory should carry out a defined calibration procedure on the standard samples in order to determine the value of the coefficient k for the Katagiri equation that is valid in that specific laboratory.
- The procedure for data treatment, including a minimum SNR, a defined baseline subtraction, and a fitting procedure, should be defined.
- A standard procedure for cleaning the surface of retrievals on all areas, in order to obtain spectra with comparable SNRs over the whole implant, should be defined.

A standardization procedure of this type should be attempted and defined within a round robin study with the participation of a large number of scientific institutions worldwide. However, it should be kept in mind that such a standard procedure will have the limitation—intrinsic within both the Raman and XRD techniques—that the measurement result is the average monoclinic content at several micrometers depth and thus reveals neither the bulk composition nor the spatial distribution of the monoclinic phase. The penetration depth can be varied in Raman by changing the width of the confocal pinhole,

and in XRD by modifying the angle of incidence of X-rays. Since spectra and diffractograms will be different for each modification of those parameters (for example, with a different SNR), the standard procedure should be repeated for each penetration depth selected.

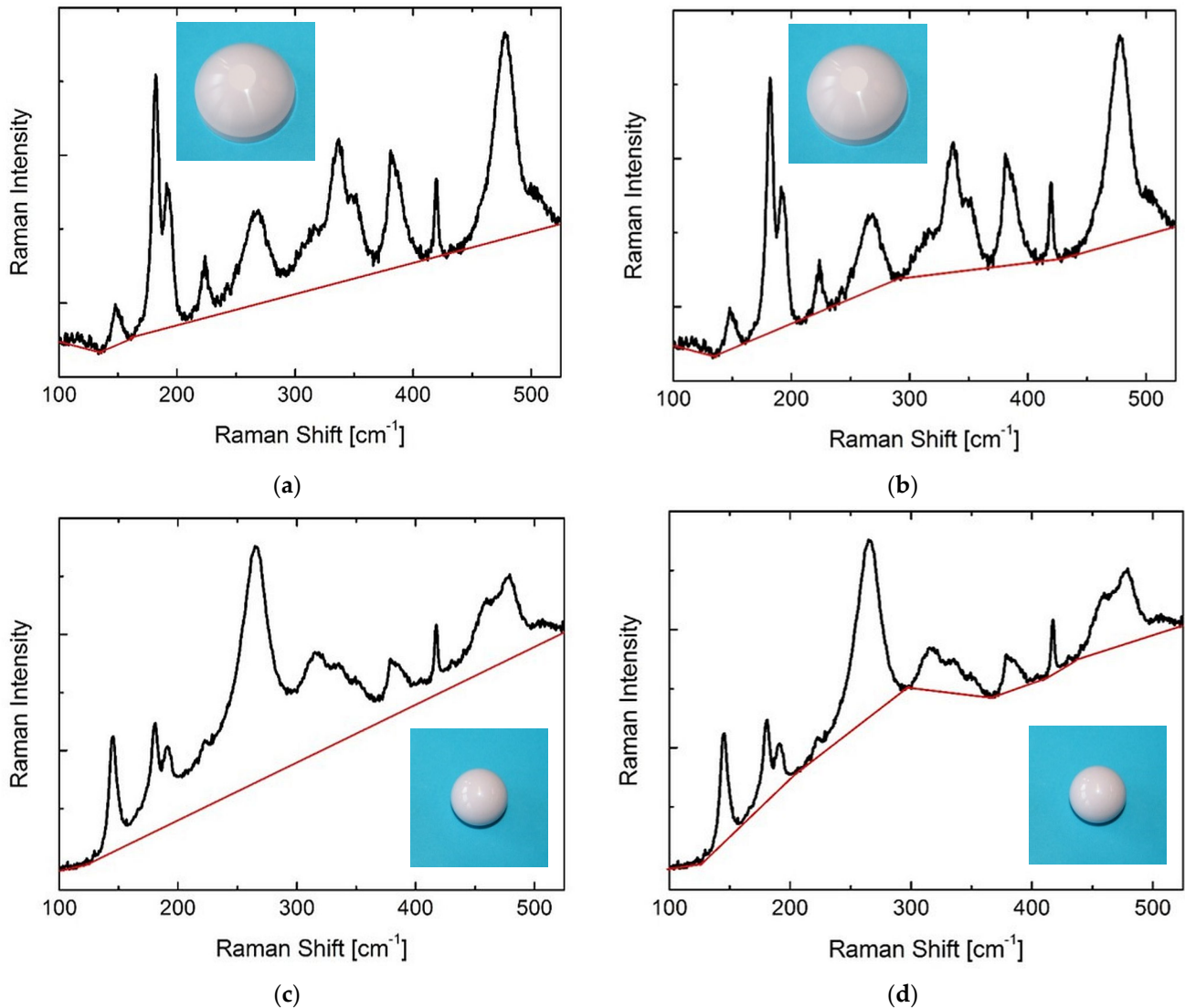


Figure 7. Different baseline choices on Raman spectra of Delta with (a,b) high and (c,d) low monoclinic contents. The spectrum shown in (a,b) was collected on the rear of an insert (rough surface), whereas the spectrum shown in (c,d) was collected on the apex of a head (polished surface).

A noteworthy alternative to the use of equations is the use of hyperspectral imaging and related statistical analyses (e.g., principal component analyses), as recently applied in surface-enhanced Raman spectroscopy [33].

5.4. Significance of V_m

Section 5.3 proposes a standard procedure for the evaluation of V_m from Raman spectrometry data. With this tool, one can now properly assess the amount of monoclinic phase on the surface of zirconia-toughened alumina hip prosthesis components. However, to determine whether the measured V_m has an influence on the performances of the components, one must also consider the origin of the transformation. Indeed, the origin can be twofold. First, the monoclinic phase can be formed by hydrothermal aging, after a spontaneous t-m transformation due to the presence of water. In this case, the t-m transformation is, in itself, a degradation mechanism. Second, the t-m transformation

can occur as a response to high mechanical stresses (phase transformation toughening), as can occur in or around a wear stripe [15], or during shocks. In the latter case, the t–m transformation is necessary to limit the damage. This is visible, for example, from the smaller width of wear stripes measured on ZTA than on alumina (that presents a comparable hardness but no phase transformation toughening) in vitro [34–36].

Stress-induced phase transformation is therefore required to obtain good crack and wear resistance, and the monoclinic content per se should not be considered as an indicator of degradation.

6. Conclusions

In our work, by measurements on both in vitro and in vivo aged BIOLOX[®]delta specimens, we determined that the Clarke/Adar equation is the most suitable equation to quantify the monoclinic content in the investigated material with the used experimental setup. Furthermore, we confirmed that metal transfer is not necessarily related to a high monoclinic content; previous studies showing the contrary might be affected by measurement artefacts leading to an exaggerated monoclinic content. This suggests that the discrepancy between in vitro and in vivo aged implants is rather ascribable to the effect of shocks than to the influence of metal transfer. Moreover, it must be considered that metastability of the tetragonal phase, to a certain extent, is necessary to guarantee good mechanical properties.

In addition, we demonstrated that Raman spectroscopy is a delicate procedure that is very much prone to errors. Critical aspects are associated with the used equation for the calculation of V_m , and with the definition of the related calibration coefficients. Other important issues are related to the spectral quality and data regression procedures. Our study demonstrates that there is a lack of standards concerning the quantification of the monoclinic content in zirconia by Raman spectroscopy. Such standards should be promptly put in place in order to avoid misinterpretations that could ultimately affect the well-being of THA patients.

Author Contributions: Conceptualization, A.A.P. and M.D.; methodology, A.A.P. and M.D.; software, M.D.; validation, A.A.P., M.D. and L.G.; formal analysis, M.D. and L.G.; investigation, A.A.P., M.D. and R.P.; resources, A.A.P., M.D. and R.P.; data curation, M.D. and L.G.; writing—original draft preparation, M.D. and A.A.P.; writing—review and editing, A.A.P., J.C., L.G. and R.P.; visualization, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: A.A.P. is an employee of CeramTec GmbH (Plochingen, Germany). M.D., L.G., J.C. and R.P. declare no conflict of interest.

References

1. Rahaman, M.N.; Yao, A.; Bal, B.S.; Garino, J.P.; Ries, M.D. Ceramics for prosthetic hip and knee joint replacement. *J. Am. Ceram. Soc.* **2007**, *90*, 1965–1988. [CrossRef]
2. Bierbaum, B.E.; Nairus, J.; Kuesis, D.; Morrison, J.C.; Ward, D. Ceramic-on-ceramic bearings in total hip arthroplasty. *Clin. Orthop. Relat. Res.* **2002**, *405*, 158–163. [CrossRef]
3. Taddei, P.; Pavoni, E.; Affatato, S. Raman and Photoemission Spectroscopic Analyses of Explanted Biolox[®] Delta Femoral Heads Showing Metal Transfer. *Materials* **2017**, *10*, 744. [CrossRef]
4. National Joint Registry for England, Wales, Northern Ireland and the Isle of Man, 17th Annual Report 2020. Surgical Data to 31 December 2019, 2020. Available online: <https://reports.njrcentre.org.uk/Portals/0/PDFdownloads/NJR%2017th%20Annual%20Report%202020.pdf> (accessed on 22 July 2021).
5. Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR). *Hip, Knee & Shoulder Arthroplasty: 2018 Annual Report*; Data Period 1 September 1999–31 December 2017; AOA: Adelaide, Australia, 2018; Available online: <https://aoanjrr.sahmri.com/annual-reports-2020> (accessed on 22 July 2021).
6. Peters, R.M.; Van Steenberg, L.N.; Stevens, M.; Rijk, P.C.; Bulstra, S.K.; Zijlstra, W.P. The effect of bearing type on the outcome of total hip arthroplasty. *Acta Orthop.* **2018**, *89*, 163–169. [CrossRef]

7. Report of R.I.P.O. Regional Register of Orthopaedic Prosthetic Implantology Overall Data Hip, Knee and Shoulder Arthroplasty in Emilia-Romagna Region (Italy) 2000–2018. 2020. Available online: <https://ripo.cineca.it/authzssl/Reports.html> (accessed on 22 July 2021).
8. Sharplin, P.; Wyatt, M.C.; Rothwell, A.; Frampton, C.; Hooper, G. Which is the best bearing surface for primary total hip replacement? A New Zealand Joint Registry study. *Hip. Int.* **2018**, *28*, 352–362. [[CrossRef](#)]
9. Arita, M.; Takahashi, Y.; Pezzotti, G.; Shishido, T.; Masaoka, T.; Sano, K.; Yamamoto, K. Environmental stability and residual stresses in zirconia femoral head for total hip arthroplasty: In vitro aging versus retrieval studies. *Biomed. Res. Int.* **2015**, *2015*, 638502. [[CrossRef](#)] [[PubMed](#)]
10. Kuntz, M. Live-Time Prediction of BIOLOX[®]delta. In *Bioceramics and Alternative Bearings in Joint Arthroplasty*, 1st ed.; Chang, J.-D., Billau, K., Eds.; Steinkopff: Darmstadt, Germany, 2007; pp. 281–288.
11. Pezzotti, G.; Affatato, S.; Rondinella, A.; Yorifuji, M.; Marin, E.; Zhu, W.; McEntire, B.; Bal, S.B.; Yamamoto, K. In vitro versus in vivo phase instability of zirconia-toughened alumina femoral heads: A critical comparative assessment. *Materials* **2017**, *10*, 466. [[CrossRef](#)]
12. Affatato, S.; Ruggiero, A.; De Mattia, J.S.; Taddei, P. Does metal transfer affect the tribological behaviour of femoral heads? Roughness and phase transformation analyses on retrieved zirconia and Biolox[®] Delta composites. *Compos. Part B Eng.* **2016**, *92*, 290–298. [[CrossRef](#)]
13. Zhu, W.; Pezzotti, G.; Boffelli, M.; Chotanaphuti, T.; Khuangsirikul, S.; Sugano, N. Chemistry-driven structural alterations in short-term retrieved ceramic-on-metal hip implants: Evidence for in vivo incompatibility between ceramic and metal counterparts. *J. Biomed. Mater. Res. B Appl. Biomater.* **2017**, *105*, 1469–1480. [[CrossRef](#)] [[PubMed](#)]
14. Elpers, M.; Nam, D.; Boydston-White, S.; Ast, M.P.; Wright, T.M.; Padgett, D.E. Zirconia phase transformation, metal transfer, and surface roughness in retrieved ceramic composite femoral heads in total hip arthroplasty. *J. Arthroplast.* **2014**, *29*, 2219–2223. [[CrossRef](#)] [[PubMed](#)]
15. Perrichon, A.; Reynard, B.; Gremillard, L.; Chevalier, J.; Farizon, F.; Geringer, J. Effects of in vitro shocks and hydrothermal degradation on wear of ceramic hip joints: Towards better experimental simulation of in vivo ageing. *J. Tribol. Int.* **2016**, *100*, 410–419. [[CrossRef](#)]
16. Pezzotti, G.; Bal, B.S.; Zanocco, M.; Marin, E.; Sugano, N.; McEntire, B.J.; Zhu, W. Reconciling in vivo and in vitro kinetics of the polymorphic transformation in zirconia-toughened alumina for hip joints: III. Molecular scale mechanisms. *Mater. Sci. Eng. C* **2017**, *71*, 552–557. [[CrossRef](#)] [[PubMed](#)]
17. Pezzotti, G.; Bal, B.S.; Zanocco, M.; Marin, E.; Sugano, N.; McEntire, B.J.; Zhu, W. Reconciling in vivo and in vitro kinetics of the polymorphic transformation in zirconia-toughened alumina for hip joints: I. Phenomenology. *Mater. Sci. Eng. C* **2017**, *72*, 252–258. [[CrossRef](#)] [[PubMed](#)]
18. Pezzotti, G.; Bal, B.S.; Zanocco, M.; Marin, E.; Sugano, N.; McEntire, B.J.; Zhu, W. Reconciling in vivo and in vitro kinetics of the polymorphic transformation in zirconia-toughened alumina for hip joints: II. Theor. *Mater. Sci. Eng. C* **2017**, *71*, 446–451. [[CrossRef](#)] [[PubMed](#)]
19. Zhu, W.; Nakashima, S.; Marin, E.; Gu, H.; Pezzotti, G. Microscopic mapping of dopant content and its link to the structural and thermal stability of yttria-stabilized zirconia polycrystals. *J. Mater. Sci.* **2019**, *55*, 524–534. [[CrossRef](#)]
20. Otto, A.; Mrozek, I.; Grabhorn, H.; Akemann, W. Surface-enhanced Raman scattering. *J. Phys. Condens. Matter.* **1992**, *4*, 1143. [[CrossRef](#)]
21. Rufaqua, R.; Vrbka, M.; Hemzal, D.; Choudhury, D.; Rebenda, D.; Krupka, I.; Hartl, M. Analysis of Chemisorbed Tribo-Film for Ceramic-on-Ceramic Hip Joint Prostheses by Raman Spectroscopy. *J. Funct. Biomater.* **2021**, *12*, 29. [[CrossRef](#)]
22. Perrichon, A.; Liu, B.; Chevalier, J.; Gremillard, L.; Reynard, B.; Farizon, F.; Liao, J.-D.; Geringer, J. Ageing, shocks and wear mechanisms in ZTA and the long-term performance of hip joint materials. *Materials* **2017**, *10*, 569. [[CrossRef](#)] [[PubMed](#)]
23. Perrichon, A.; Reynard, B.; Gremillard, L.; Chevalier, J.; Farizon, F.; Geringer, J. A testing protocol combining shocks, hydrothermal ageing and friction, applied to Zirconia Toughened Alumina (ZTA) hip implants. *Mech. Behav. Biomed. Mater.* **2017**, *65*, 600–608. [[CrossRef](#)]
24. Nguyen, T.M.; Weitzler, L.; Esposito, C.I.; Porporati, A.A.; Padgett, D.E.; Wright, T.M. Zirconia phase transformation in zirconia-toughened alumina ceramic femoral heads: An implant retrieval analysis. *J. Arthroplast.* **2019**, *34*, 3094–3098. [[CrossRef](#)]
25. Kim, B.K.; Hahn, J.W.; Han, K.R. Quantitative phase analysis in tetragonal-rich tetragonal/monoclinic two phase zirconia by raman spectroscopy. *J. Mater. Sci. Lett.* **2017**, *16*, 669–671. [[CrossRef](#)]
26. Munoz Tabares, J.A.; Anglada, M.J. Quantitative analysis of monoclinic phase in 3Y-TZP by raman spectroscopy. *J. Am. Ceram. Soc.* **2010**, *93*, 1790–1795. [[CrossRef](#)]
27. Clarke, D.R.; Adar, F. Measurement of the crystallographically transformed zone produced by fracture in ceramics containing tetragonal zirconia. *J. Am. Ceram. Soc.* **1982**, *65*, 284–288. [[CrossRef](#)]
28. Pezzotti, G.; Yamada, K.; Sakakura, S.; Pitto, R.P. Raman spectroscopic analysis of advanced ceramic composite for hip prosthesis. *J. Am. Ceram. Soc.* **2008**, *91*, 1199–1206. [[CrossRef](#)]
29. Vega, M.M.; Bonifacio, A.; Lughi, V.; Sergio, V. Low-Level Monoclinic Content Detection in Zirconia Implants Using Raman Spectroscopy. In *Nano-Structures for Optics and Photonics. NATO Science for Peace and Security Series B: Physics and Biophysics*, 1st ed.; Di Bartolo, B., Collins, J., Silvestri, L., Eds.; Springer: Dordrecht, The Netherlands, 2015; pp. 539–540.
30. Garvie, R.C.; Nicholson, P.S. Phase analysis in zirconia systems. *J. Am. Ceram. Soc.* **1972**, *55*, 303–305. [[CrossRef](#)]

31. Pezzotti, G.; Munisso, M.C.; Lessnau, K.; Zhu, W. Quantitative assessments of residual stress fields at the surface of alumina hip joints. *J. Biomed. Mater. Res. Part B Appl. Biomater.* **2010**, *95*, 250–262. [[CrossRef](#)]
32. Tateiwa, T.; Marin, E.; Rondinella, A.; Ciniglio, M.; Zhu, W.; Affatato, S.; Pezzotti, G.; Bock, R.M.; McEntire, B.J.; Bal, B.S.; et al. Burst Strength of BIOLOX[®] delta Femoral Heads and Its Dependence on Low-Temperature Environmental Degradation. *Materials* **2020**, *13*, 350. [[CrossRef](#)]
33. Vega, M.M.; Alois, B.; Vanni, L.; Sergo, V. Fine determination of monoclinic phase in zirconia-based implants: A surface-enhanced Raman spectroscopy (SERS) study. *J. Nanosci. Nanotechnol.* **2020**, *20*, 2430–2435. [[CrossRef](#)]
34. Uribe, J.; Geringer, J.; Gremillard, L.; Reynard, B. Degradation of alumina and zirconia toughened alumina (ZTA) hip prostheses tested under microseparation conditions in a shock device. *Tribol. Int.* **2013**, *63*, 151–157. [[CrossRef](#)]
35. Piconi, C.; Porporati, A.A.; Streicher, R.M. Ceramics in THR bearings: Behavior under off-normal Conditions. *Key Eng. Mater.* **2014**, *631*, 3–7. [[CrossRef](#)]
36. De Fine, M.; Terrando, S.; Hintner, M.; Porporati, A.A.; Pignatti, G. Pushing Ceramic-on-Ceramic in the most extreme wear conditions: A hip simulator study. *Orthop. Traumatol. Surg. Res.* **2021**, *107*, 102643. [[CrossRef](#)] [[PubMed](#)]