Reliability and Agreement of Three Devices for Measuring Implant Stability Quotient in the Animal Ex Vivo Model

Monica Blazquez-Hinarejos 1, Constanza Saka-Herrán 1, Victor Diez-Alonso 1, Raul Ayuso-Montero 2,3, *, Eugenio Velasco-Ortega 4 and Jose Lopez-Lopez 3,5

Abstract: Resonance frequency analysis (RFA) is the most extended method for measuring implant stability. The implant stability quotient (ISQ) is the measure obtained by different RFA devices; however, inter- and intra-rater reliability and agreement of these instruments remain unknown. Thirty implants were placed in three different pig mandibles. ISQ was measured parallel and perpendicular (lingual) to the peg axis with Osstell® Beacon, Penguin® and MegaISQ® by two different investigators and furthermore, one performed a test-retest. Intraclass correlation coefficient was calculated to assess the intra- and inter-rater reliability. Pearson correlation coefficient was used to assess the agreement. Intraclass correlation coefficients ranged from 0.20 to 0.65 for the Osstell® Beacon; 0.57 to 0.86 for the Penguin®, and 0.01 to 0.60 for the MegaISQ®. The highest ISQ values were obtained using Penguin® (66.3) in a parallel measurement; the lowest, using the MegaISQ® (60.1) in a parallel measurement. The highest correlation values with the other devices were obtained by MegaISQ® in a parallel measurement. Osstell® Beacon and MegaISQ® showed lower reliability than Penguin®. Osstell® had good agreement for measuring ISQ both in parallel and perpendicular, and MegaISQ® had the best agreement for measuring ISQ in parallel.

Keywords: resonance frequency analysis; implant stability quotient; reliability; agreement

1. Introduction

Implant stability is critical in implant therapy and varies during the osseointegration process, reflecting bone/implant interface changes [1,2]. Low levels of implant micromotion are necessary to avoid implant failure and to achieve successful osseointegration [3–5].

Several existing methods have addressed measuring implant stability, including theoretical [6] and experimental modal analysis [7]. Among these, Periotest, insertion torque value (ITV), and implant stability quotient (ISQ) using resonance frequency analysis [8] are the ones widely used clinically. Periotest is a damping method that requires to strike the implant abutment [9], ITV measures the newton centimeters used to screw the implant into the bone [5], and ISQ sensors register the response of the electromagnetic stimulation of an abutment fixed to the implant called transducer peg, measuring the implant stability [10]. Periotest and ISQ are considered modal analysis methods based on the displacement signal secondary to an external impulse force [7]. Successfully integrated implants involve low implant micromotion levels, which usually correspond to low Periotest values and high ITV and ISQ values. These ITV and ISQ values are inversely correlated with low implant
micromotion. However, the relationship between ITV and implant micromotion becomes exponential for higher ITV values [5]. Besides, ITV only measures implant stability at the moment of the insertion [8]. For these reasons, ISQ is usually the preferred method to measure implant stability.

Devices used to measure ISQ return a quotient value ranged from 0 to 100 corresponding to minimum and maximum vibration, respectively [11]. According to existing literature, a minimum ISQ value of 57 corresponds to a maximum implant micromotion of 150 µm. Clinically, this micromotion represents implant stability and it is required to maintain osseointegration [2,5].

There are several studies analyzing the reliability of existing devices for measuring the ISQ. These analyses have focused on two devices: Osstell® (W&H, Göteborg, Sweden) and Penguin® (Integration Diagnostics Sweden AB, Göteborg, Sweden) [12–14]. For instance, Buyukguclu et al. [12] reported better reliability for Osstell®, while Romanos et al. [13] reported that both devices were sensitive and reliable. Bural et al. reported excellent reliability for Penguin®, but they did not report the reliability of Osstell® [14]. Norton et al. [15] reported the agreement between the measurements obtained by two different Osstell® versions and Penguin®, and the authors considered these differences not clinically relevant. Therefore, reliability results reported so far are contradictory.

Osstell® Beacon is the wireless version of Osstell® that can be connected online to the Osstell® database for statistical analysis. The smartpegs for Osstell® are not autoclavable. On the other hand, the multipegs for Penguin® are autoclavable and the device is wireless. MegaISQ® (Megagen Implant CO, Daegu, Korea) is a portable, but not wireless device that uses the Osstell® smartpegs. To the best of our knowledge, inter- and intra-rater reliability and agreement among Osstell® Beacon, Penguin® and MegaISQ® have not yet been investigated.

The goal of this study was to determine and compare the inter- and intra-rater reliability of Osstell® Beacon and MegaISQ® versus Penguin® as control. This study also aimed to explore the agreement between these devices for ISQ measurement. This study was conducted in vitro by two investigators to obtain the inter- and intra-rater reliability and the agreement level among these three devices.

2. Materials and Methods

In this in vitro study, 30 BioHorizons® Internal implants (BioHorizons, Birmingham, AL, USA) were inserted in fresh pig mandibles (Figure 1). The manufacturer drilling protocol was used to place 10 implants (4.6 mm diameter, 12 mm height) in 10 different positions of 3 different mandibles (Figure 1). Considering an alpha error of 0.05 and a beta error of 0.2, in a two-sided test, a minimum of 20 samples were necessary from each group to identify a statistically significant difference greater than or equal to two units. Based on a recent study [16], standard deviation was assumed to be 5 and the correlation coefficient measurement was 0.9.

ISQ was measured using three different devices: Osstell® Beacon (W&H, Göteborg, Sweden), Penguin® (Integration Diagnostics Sweden AB, Göteborg, Sweden) and MegaISQ® (Megagen Implant CO, Daegu, Korea). The transducer peg for each device (smartpeg for Osstell and MegaISQ and multipeg for Penguin) was inserted on each implant according to the manufacturer’s instructions, and two measurements were recorded parallel and perpendicular (lingual) to the longitudinal axis of the peg. Due to the reliability provided by previous studies [12–15], Penguin® was used as the control.

The smartpeg for Osstell®, the multipeg for Penguin® and the smartpeg for MegaISQ® were removed between each measurement, and the stability of the 30 implants placed in the three different mandibles was evaluated with the three devices.
Figure 1. Implant locations in a representative mandible with a Penguin® multipeg inserted in one of the implants.

All procedures were repeated by two different experienced and calibrated investigators (MB and RA) after the implant insertion, in order to assess the inter-rater reliability, and one operator (RA) repeated the procedures 5 min later in order to perform a test-retest check for measuring the intra-rater reliability in the same conditions. While it was not possible to blind the device used, the order in which implants were measured and which device was used was randomized. The measurements were coded by these two operators in order to blind the statistical analysis.

Statistical Analysis

Shapiro–Wilks and Levene tests were respectively used for assessing criteria of normality and homogeneity of variances (Supplementary Table S1). Test-retest was used to calculate the intraclass correlation coefficient (ICC) using a mixed model with a random effect on the individual in order to assess the intra- and inter-rater reliability. ICC values were classified as poor-moderate-good according to Koo et al. criteria. [17]. Absolute ISQ values obtained using each method were reported as mean (95%CI). The agreement between the devices was assessed by means of the Pearson correlation coefficient. To establish the level of agreement and the correlation between the different devices a mean value from the two operators was calculated. All analyses were performed using the IBM Statistics for Windows v24.0 software package (IBM Corp., New York, NY, USA) (p < 0.05).

3. Results

Table 1 shows the implant stability measurements obtained by each device. As shown, using Penguin in parallel measurement yields the highest ISQ values; in contrast, using the MegaISQ® in a perpendicular measurement results in the lowest values.
Table 1. Mean values (95% CI) of implant stability quotient (ISQ) according to the device and the orientation.

| Device      | Technique | ISQ       | Difference with Mean ISQ Values |
|-------------|-----------|-----------|---------------------------------|
| OSSTELL® Beacon | parallel   | 62.2 (59.5 to 64.9) | 0.1 (−0.1 to 0.1) |
|             | perpendicular | 60.6 (58.0 to 63.3) | −1.5 (−1.7 to −1.15) |
| PENGUIN®    | parallel   | 66.3 (62.4 to 70.1) | 4.2 (3.25 to 5.1) |
|             | perpendicular | 63.1 (60.0 to 66.3) | 1 (0.85 to 1.3) |
| MEGAISQ®    | parallel   | 60.1 (57.5 to 62.6) | −2 (−2.4 to −1.65) |
|             | perpendicular | 60.2 (57.5 to 62.8) | −1.9 (−2.2 to −1.65) |

ISQ—implant stability quotient. 95% CI—95% confidence interval.

The difference between the values obtained with each device and technique and the mean ISQ values is shown in Figure 2. The mean of these differences was 0.13 (95% CI: −5.79 to 6.05) for the Osstell® Beacon in a parallel measurement; −1.45 (95% CI: −7.78 to 4.88) for the Osstell® Beacon in a perpendicular measurement; 4.2 (95% CI: −6.20 to 14.60) for the Penguin® in a parallel measurement; 1.03 (95% CI: −7.40 to 9.47) for the Penguin® in a perpendicular measurement; −2.02 (95% CI: −7.30 to 3.27) for the MegaISQ® in a parallel measurement and −1.09 (95% CI: −13.03 to 9.23) for MegaISQ® in a perpendicular measurement.

Figure 2. Cont.
Figure 2. Bland–Altman plot with the difference between the measurements of each device and technique and the mean ISQ values.

Table 2 shows the reliability scores for the three devices. The highest inter- and intra-rater reliability was obtained by Penguin® when measuring in parallel. The lowest inter-rater reliability was obtained by MegaISQ® measuring perpendicularly, Osstell® Beacon measuring perpendicularly, and MegaISQ® measuring in parallel. The lowest intra-rater reliability was obtained by Osstell® Beacon.

Table 2. Reliability (ICC; 95% CI) of the three devices used to measure ISQ.

| Reliability |
|-------------|
| INTER-RATER |
| ICC (95% CI) |
| OSSTELL® B parallel 0.37 (0.40 to 0.64) poor |
| OSSTELL® B perpendicular 0.20 (−0.17 to 0.52) poor |
| PENGUIN® parallel 0.86 (0.72 to 0.93) good |
| PENGUIN® perpendicular 0.57 (0.26 to 0.77) moderate |
| MEGAISQ® parallel 0.26 (−0.11 to 0.57) poor |
| MEGAISQ® perpendicular −0.01 (−0.38 to 0.36) poor |

| ICC Classification |
|---------------------|
| INTER-RATER |
| ICC (95% CI) |
| 0.65 (0.38 to 0.81) moderate |
| 0.47 (0.13 to 0.71) poor |
| 0.85 (0.70 to 0.92) good |
| 0.78 (0.56 to 0.89) good |
| 0.60 (0.26 to 0.79) moderate |
| 0.57 (0.27 to 0.77) moderate |

A matrix with the Pearson correlation coefficients is shown in Table 3. The highest correlation value with the other devices was obtained by MegaISQ® measuring parallel; however, this device obtained the lowest correlation value when measuring perpendicular. Osstell® Beacon obtained high correlation values with the other devices measuring either in parallel or perpendicular. Penguin® obtained correlation values lower than Osstell® Beacon, both measuring parallel and perpendicular, but higher than MegaISQ® when measuring perpendicular.
Table 3. Matrix of Pearson correlation coefficients for the different devices used to measure ISQ.

|                | Osstell® Beacon Parallel | Osstell® Beacon Perpend. | PENGUIN® Parallel | PENGUIN® Perpend. | MEGAISQ® Parallel | MEGAISQ® Perpend. | TOTAL |
|----------------|--------------------------|--------------------------|-------------------|-------------------|-------------------|-------------------|-------|
| OSSTELL® Beacon parallel | 1                        | 0.723 **                 | 0.667 **          | 0.555 **          | 0.766 **          | 0.405 **          | 4.12  |
| OSSTELL® Beacon perpend. | 0.723 **                 | 1                        | 0.575 **          | 0.652 **          | 0.653 **          | 0.525 **          | 4.13  |
| PENGUIN® parallel  | 0.667 **                 | 0.575 **                 | 1                 | 0.760 **          | 0.691 **          | 0.298 *           | 3.99  |
| PENGUIN® perpend. | 0.555 **                 | 0.652 **                 | 0.760 **          | 1                 | 0.675 **          | 0.404 **          | 4.05  |
| MEGAISQ® parallel  | 0.766 **                 | 0.653 **                 | 0.691 **          | 0.675 **          | 1                 | 0.449 **          | 4.23  |
| MEGAISQ® perpend. | 0.405 **                 | 0.525 **                 | 0.298 *           | 0.404 **          | 0.449 **          | 1                 | 3.08  |

* p < 0.01; ** p < 0.001; perpend.—perpendicular.

4. Discussion

Based on the ICC scores, our results suggest that Osstell® Beacon and MegaISQ® exhibited lower reliability than Penguin®. The reliability of Penguin® was good in parallel measurements, and between moderate to good when measuring perpendicularly. Osstell® Beacon presented a poor to moderate reliability when measuring parallel and poor when measuring perpendicular, and MegaISQ® obtained poor to moderate reliability when measuring both parallel and perpendicular. The lower reliability obtained by Osstell® Beacon and MegaISQ® compared to Penguin® can be attributed to differences in the electromagnetic functioning, since these devices use the same smartpeg from Osstell®, and Penguin® uses a magnetized multipeg. From these results, Penguin® should be used to monitor the implant micromotion and the evolution of osseointegration.

Our ICC scores were lower than a recent study [14]; however, in this study, the mean of the perpendicular and parallel ISQ values was considered as the final ISQ of each implant, then the differences of each device measuring parallel or perpendicular could be not detected. High ICC scores were also reported for both Penguin® and Osstell®, but only when the implant surrounding material was stiff [12]. The differences in bone density between studies could explain the different results obtained.

The inter-rater ICC values obtained in our study were mostly lower than the intra-rater ICC values for the majority of devices and techniques. This observation suggests that the values obtained from these devices can be operator dependent. All three devices presented higher ICC values when measuring parallel than perpendicular. One study reported increased variability and reduced reliability when measuring buccolingual [14]. These results suggest that the clinical evaluation of implant micromotion by means of parallel ISQ could be recommended.

Osstell® Beacon showed lower values than previous studies using different types of Osstell® (Osstell® ISQ and Osstell® IDX) [14,18]. These differences can be attributed to the bone density and the device version. In our study, the highest ISQ values were obtained using Penguin® and the lowest values were obtained using MegaISQ®. This observation can lead the clinician to overestimate the implant stability with Penguin®. The MegaISQ® values were the lowest, suggesting that MegaISQ® tends to underestimate the implant stability. However, the difference between the MegaISQ® values and the mean was twice as low as with Penguin, then the underestimation of MegaISQ® has not reached the magnitude of the Penguin® overestimation. These differences could be clinically relevant when the ISQ measure is around 57 corresponding to the minimum threshold of osseointegration. This
value can be interpreted as a correct osseointegration of a failed implant with Penguin® or as a failed implant with a correct osseointegration with MegaISQ®.

Comparing the correlation between each instrument, MegaISQ® measuring parallel had the higher correlation to the others (considering every instrument and technique). However, the same instrument obtained the lowest correlation when measuring perpendicular (lowest Pearson correlation coefficient and widest difference of agreement with the mean ISQ values in the Bland–Altman plot). On the other hand, Osstell® Beacon obtained good correlation for measuring both parallel and perpendicular (and the narrowest difference of agreement with the mean ISQ values), and Penguin® had similar correlation values with the other methods measuring parallel and perpendicular.

This study has some limitations. The bone density can affect ISQ values [19] and no previous evaluation of the different bone locations where implants were placed was done. However, some aspects that could affect ISQ values, such as implant length and diameter were controlled using the same implant size for all measurements. Another limitation was the manual tightening of the transducers, but this technique was previously reported to be objective and reliable [20]. Finally, it was not possible to blind the investigators within the instrument used, and the study was performed in an animal model. Therefore, further research is needed to clinically assess in vivo the behavior of these devices.

5. Conclusions

Within the limitations of this study, Osstell® Beacon and MegaISQ® showed a larger deviation in the measurements than Penguin®; Penguin® exhibited moderate to good inter-rater reliability and good intra-rater reliability for measuring the implant micromotion; Osstell® Beacon had good agreement for measuring ISQ both parallel and perpendicular and MegaISQ® had the best agreement for measuring ISQ parallel, but not for measuring perpendicular.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/app11083453/s1, Table S1. Normality and homogeneity of variances.

Author Contributions: Conceptualization, J.L.-L., R.A.-M. and E.V.-O.; Methodology, R.A.-M., J.L.-L., C.S.-H. and E.V.-O.; Software, R.A.-M. and C.S.-H.; Validation, M.B.-H., C.S.-H., V.D.-A., J.L.-L., R.A.-M. and E.V.-O.; Formal Analysis, R.A.-M. and C.S.-H.; Investigation, V.D.-A., M.B.-H., J.L.-L. and R.A.-M.; Resources, R.A.-M. and J.L.-L.; Data Curation, V.D.-A., M.B.-H. and R.A.-M.; Writing—Original Draft Preparation, R.A.-M. and J.L.-L.; Writing—Review and Editing, M.B.-H., R.A.-M. and E.V.-O.; Visualization, R.A.-M., J.L.-L.; Supervision, J.L.-L., R.A.-M. and E.V.-O.; Project Administration, M.B.-H., R.A.-M. and J.L.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from Biohorizons-Camlog® Ibérica with the grant number 018582/2019, and from the Faculty of Medicine and Health Sciences, University of Barcelona.

Data Availability Statement: Data of obtained from this study is not available publicly. Information regarding these data should be requested to the corresponding author.

Acknowledgments: The authors gratefully acknowledge Jordi Martinez-Gomis for the statistical analysis support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Turkyilmaz, I.; Sennery, L.; Tumer, C.; Yenigul, M.; Avci, M. Stability and marginal bone level measurements of unsplinted implants used for mandibular overdentures: A 1-year randomized prospective clinical study comparing early and conventional loading protocols. *Clin. Oral Implant. Res.* 2006, 17, 501–505. [CrossRef] [PubMed]

2. Huwiler, M.; Pjeturson, B.E.; Bosshardt, D.D.; Salvi, G.E.; Lang, N.P. Resonance Frequency Analysis (RFA) in relation to jaw bone characteristics during early healing. *Clin. Oral Implant. Res.* 2007, 18, 275–280. [CrossRef] [PubMed]

3. Trisi, P.; Perfetti, G.; Baldoni, E.; Berardi, D.; Colagiovanni, M.; Scogna, G. Implant micromotion is related to peak insertion torque and bone density. *Clin. Oral Implant. Res.* 2009, 20, 467–471. [CrossRef] [PubMed]
4. Winter, W.; Klein, D.; Karl, M. Micromotion of dental implants: Basic mechanical considerations. *J. Med. Eng.* 2013, 2013, 1. [CrossRef] [PubMed]

5. Brizuela-Velasco, A.; Alvarez-Arenal, A.; Gil-Mur, F.J.; Herrero-Climent, M.; Chávarri-Prado, D.; Chento-Valiente, Y.; Dieguez-Pereira, M. Relationship between torque and resonance frequency measurements, performed by resonance frequency analysis, in micromobility of dental implants: An in vitro study. *Implant Dent.* 2015, 24, 607–611. [CrossRef] [PubMed]

6. Zanetti, E.M.; Ciaramella, S.; Calì, M.; Pascoletti, G.; Martorelli, M.; Asero, R.; Watts, D.C. Modal analysis for implant stability assessment: Sensitivity of this methodology for different implant designs. *Dent. Mater.* 2018, 34, 1235–1245. [CrossRef] [PubMed]

7. Atsumi, M.; Park, S.H.; Wang, H.L. Methods used to assess implant stability: Current status. *Int. J. Oral Maxillofac. Implant.* 2007, 22, 743–754.

8. Aparicio, C.; Lang, N.P.; Rangert, B. Validity and clinical significance of biomechanical testing of implant/bone interface. *Clin. Oral Implant. Res.* 2006, 17 (Suppl. S2), 2–7. [CrossRef] [PubMed]

9. Lee, D.H.; Shin, Y.H.; Park, J.H.; Shim, J.H.; Shin, S.W.; Lee, J.Y. The reliability of anycheck device related to healing abutment diameter. *J. Adv. Prosthodont.* 2020, 12, 83–88. [CrossRef] [PubMed]

10. Brizuela-Velasco, A.; Chávarri-Prado, D. The functional loading of implants increases their stability: A retrospective clinical study. *Clin. Implant Dent. Relat. Res.* 2019, 21, 122–129. [CrossRef] [PubMed]

11. Santamaría-Arrieta, G.; Brizuela-Velasco, A.; Fernández-González, F.J.; Chávarri-Prado, D.; Chento-Valiente, Y.; Solaberrieta, E.; Chávarri-Prado, D.; Chento-Valiente, Y.; Solaberrieta, E. Biomechanical evaluation of oversized drilling technique on primary implant stability measured by insertion torque and resonance frequency analysis. *J. Clin. Exp. Dent.* 2016, 1, 307–311. [CrossRef]

12. Buyukguclu, G.; Ozkurt-Kayahan, Z.; Kazazoglu, E. Reliability of the Osstell implant stability quotient and Penguin resonance frequency analysis to evaluate implant stability. *Implant Dent.* 2018, 27, 429–433. [CrossRef] [PubMed]

13. Romanos, G.E.; Bastardi, D.J.; Kakar, A.; Moore, R.; Delgado-Ruiz, R.A.; Javed, F. In vitro comparison of resonance frequency analysis devices to evaluate implant stability of narrow diameter implants at varying drilling speeds in dense artificial bone blocks. *Clin. Implant Dent. Relat. Res.* 2019, 21, 1023–1027. [CrossRef] [PubMed]

14. Bural, C.; Dayan, C.; Geçkili, O. Initial stability measurements of implants using a new magnetic resonance frequency analyzer with titanium transducers: An ex vivo study. *J. Oral Implantol.* 2020, 46, 35–40. [CrossRef] [PubMed]

15. Norton, M.R. Resonance Frequency Analysis: Agreement and correlation of implant stability quotients between three commercially available instruments. *Int. J. Oral Maxillofac. Implant.* 2018. [CrossRef] [PubMed]

16. Lee, J.; Pyo, S.W.; Cho, H.J.; An, J.S.; Lee, J.H.; Koo, K.T.; Lee, Y.M. Comparison of implant stability measurements between a resonance frequency analysis device and a modified damping capacity analysis device: An in vitro study. *J. Periodontal. Implant Sci.* 2020, 50, 56–66. [CrossRef] [PubMed]

17. Koo, T.K.; Li, M. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* 2016, 15, 155–163. [CrossRef] [PubMed]

18. Becker, W.; Hujoel, P.; Becker, B.E. Resonance frequency analysis: Comparing two clinical instruments. *Clin. Implant Dent. Relat. Res.* 2018, 20, 308–312. [CrossRef] [PubMed]

19. Sim, C.P.; Lang, N.P. Factors influencing resonance frequency analysis assessed by Osstell mentor during implant tissue integration: I. Instrument positioning, bone structure, implant length. *Clin. Oral Implant. Res.* 2010, 21, 598–604. [CrossRef] [PubMed]

20. Kästel, I.; de Quincey, G.; Neugebauer, J.; Sader, R.; Gehrke, P. Does the Manual Insertion Torque of Smartpegs Affect the Outcome of Implant Stability Quotients (ISQ) During Resonance Frequency Analysis (RFA)? *Int. J. Implant Dent.* 2019, 5, 42. [CrossRef] [PubMed]