File ID	uvapub:102476
Filename	Chapter 4: The internal responsiveness of the Oral Health Impact Profile-14 to detect differences in clinical parameters related to surgical third molar removal
Version	unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):
Type	PhD thesis
Title	Some fundamental issues in oral health-related quality of life research
Author(s)	J.M. Kieffer
Faculty	ACTA
Year	2012

FULL BIBLIOGRAPHIC DETAILS:
> http://hdl.handle.net/11245/1.372260

CHAPTER 4
The internal responsiveness of the Oral Health Impact Profile-14 to detect differences in clinical parameters related to surgical third molar removal

INTRODUCTION

Nowadays, impacts of oral conditions on patients' well-being and treatment efficiency are a central focus in dental research. The Oral Health Impact Profile (OHIP-49) is the most commonly used instrument to assess impacts of oral conditions. This instrument is based on Locker's adaptation of the 'Classification of Impairments, Disabilities and Handicaps' developed by the 'World Health Organisation' (Locker 1988, Slade et al. 1994, John et al. 2004), and contains seven hierarchically ordered dimensions. A derivative of this instrument, the OHIP-14, was developed later and captures the same dimensions as the original. It consists of 14 items, including two items out of every dimension of the OHIP-49 (Slade 1997a). A higher score on the OHIP-14 (or 49) indicates a greater impact of oral conditions. Both instruments are generally used in cross-sectional and longitudinal studies and are aimed to evaluate the physical, psychological and social impacts of oral conditions [5, 6].

Even though short-form instruments are more convenient compared to the original to administer in certain settings (such as clinical settings), it is recognized that short-form instruments are more prone to reliability and validity issues (Slade 1997a). There are some studies which resulted in some attention towards the reliability and validity of the OHIP-14, but even so, the focus is mostly on cross-sectional validity and test-retest reliability. However, when it comes to treatment efficiency in longitudinal studies, responsiveness - the ability for an instrument to detect clinical changes – is ever so important (Locker et al. 2004). However, there seems to be some confusion in the literature as to what a responsive measure actually is. One review study (Husted et al. 2000) suggests that there are to two main aspects when it comes to responsiveness; external and internal responsiveness. External responsiveness is the extent to which changes over time match up with an external standard. Internal responsiveness is defined as the ability of a measure to detect change before and after treatment. In this study, responsiveness refers to internal responsiveness.

There are different studies with regard to the responsiveness of the OHIP-49 and the OHIP-14. The OHIP 49 has been shown responsive to treatment among edentulous patients (Awad et al. 2000, Allen et al. 2001) and to tooth whitening among adolescents (McGrath et al. 2005). With respect to the OHIP-14, one study found the OHIP-14 to be modestly responsive to change when used to evaluate a dental care program for the elderly (Locker et al. 2004) and it has even been shown that the responsiveness of the OHIP-49 can be maintained while reducing the number of items (Awad et al. 2008).

However, is the OHIP-14 responsive to the surgical removal of third molars? Although relatively common, third-molar surgery is considerably invasive and is most often performed in an outpatient setting on relatively young people who are expected to be healthy and therefore have rarely had prior experience with surgery (Litt et al. 1995). Furthermore, it is realistic to assume that a large part of the people who have had third molar surgery will experience postoperative pain. This pain can be exacerbated by clinical parameters, such as preoperative or postoperative complications and the number of molars removed (Van Wijk et al. 2009). Thus, the side-effects of third molar surgery can have a great impact on patients' well-being, and should therefore be detected by the short-form OHIP-14. One study indicated the OHIP-14 to be responsive to clinical changes in status of impacted third molars in a Scottish general dental practice (Fernandes et al. 2006). As with this study, the present study also focuses on the responsiveness of the OHIP-14 with

regard to third molar surgery, which is one of the most frequently performed oral surgery procedures (Worrall et al. 1998).

Thus the aim of the present study is to assess the internal responsiveness of the OHIP-14 with respect to third molar surgery. It has been hypothesised that OHIP scores are higher in the first few days after surgery than in the preoperative state. Furthermore, this study will explore the OHIP-14's ability to differentiate between patients with and without preoperative and postoperative complaints, between patients with a different preoperative status and other clinical variables.

MATERIALS & METHODS
Participants
Patients who were referred to the Department of Oral & Maxillofacial Surgery of the Academic Medical Centre (AMC) in Amsterdam by their dentist for surgical removal of their impacted third molars and who were 18 years or older, were eligible to participate in this study. Inclusion criteria were the ability to read, understand and fill in questionnaires and willingness to participate. Only impacted mandibular third molars were included and surgery was performed unilaterally (one side only). The study was approved by the medical ethics committee of the Academic Medical Centre Amsterdam, University of Amsterdam. The study was performed with the understanding and written consent of each patient and according to the ethical principles described in the Helsinki Declaration.

Impacts of oral conditions
The OHIP-14 (Slade 1997a) was used to assess impacts of oral conditions. A Dutch OHIP-14 was constructed by using the relevant items from the original questionnaire (49 items) that had recently been translated into Dutch (Van der Meulen et al. 2008). The OHIP-14 consists of seven dimensions, namely: functional limitations, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap. Before surgery, patients were asked for each item of the OHIP-14 how often in the past 4 weeks they had experienced a certain problem regarding their teeth, mouth or dentures. They responded on a Likert-type scale, which was coded as follows: 5, very often; 4, fairly often; 3, sometimes; 2, hardly ever and 1, never. Thus, higher scores indicate more impacts of oral conditions. The total score ranged from 14 – 70. The same instrument was used for the eight follow-ups: directly after the surgery, every day thereafter for a six-day period, and one month after the surgery. The time period of the OHIP- items was changed for the first seven follow-ups in which the items were introduced with "Today" instead of "During the past four weeks".

Clinical variables
Patient characteristics were measured by;
- Smoking (yes or no)
- Mandibular third molar that was surgically removed (left or right).

Preoperative status was measured by:
- Preoperative complaints (1 = no complaints, 2 = pain, 3 = other)

- Mucosal coverage (1 = none, 2 = partial, 3 = covered): Mucosal coverage is the extent to which soft tissue (the gingiva) covers the third molar. As such, more mucosal coverage requires more removal of more soft tissue.
- Classification of angulations (1.Vertical, 2.Mesioangular, 3.Distoangular, 4.Horizontal, 5.Inverted, 6.Normal): How the molar is situated in the bone. An angulation deviating from a normal one (vertical) is more difficult to remove.
- Molar classification = a combination of class (I, II or III) with position (A, B or C). Higher classes reflect less space in the jawbone, and increased risk of damaging a nerve and thus the occurrence of postoperative complication. Position reflects the degree of impaction in the bone; higher positions will require more bone removal to access the third molar (Pell et al. 1942).

Procedural variables that were taken into account:
- Type of alveotomy (whether splitting the molar was required for removal from the bone: yes or no)
- duration of surgery (in minutes)

Postoperative complications (1 = abscess [infection], 2 = alveolitis [infection of the jawbone], 3 = other) were assessed within one week and four weeks after surgery.

Statistical analyses
Sample size was based on time and on availability of patients and it was determined that a one year period of testing would yield approximately 100 eligible patients. Associations between categorical variables were analysed using the Chi^2 test. Independent mean scores were compared using the independent samples t-test. The responsiveness of the OHIP was assessed by using repeated measures ANOVA, followed by paired-samples t-tests when appropriate. Effect size estimates for within subject effects were calculated by Partial eta-squared (η^2), and effects size estimates for between subject effects (ES) were t-based calculations. These estimates are interpreted in terms of the ranges of magnitude as follows: < 0.2 (small difference), through 0.2–0.8 (moderate difference), to > 0.8 (large difference) (Cohen 1977). The level of significance was set at alpha = 0.05.

RESULTS
Participants
A total of 107 patients were eligible to be included in the study. As a result of incomplete data (there where the majority of the postoperative week measurements were not filled in, and/or in combination with no one-month postoperative measurement), 10 patients were excluded from the analyses. The resulting sample consisted of 45 male (mean age = 26.2, SD. = 6.6) and 52 female patients (mean age = 25.0, SD = 4.7) who did not differ with respect to age.

The majority of patients did not experience preoperative complaints (73.2%), preoperative pain was experienced by 24.7% and 2 patients (2.1%) had another pathological condition (a distal carious lesion of the second molar, and periodontal bone loss distal from the second molar). Smokers (20.6%) seemed to report preoperative pain

more often than non-smokers, however, the Chi^2-test was only marginal significant, $X^2(1)$ = 3.57, p = 0.06.

The right mandibular third molar was removed more often (62.9%) than the left mandibular third molar. The surgical procedure lasted, on average, for 16.1 minutes (SD = 5.7) with a range of 8 to 45 minutes. The majority of impacted third molars (87.6%) required splitting of the tooth in order to be removed. In 48.5% of the patients, their third molar was partially covered by mucosa, the remaining molars were covered completely (51.5%). A limited number of angular classifications were found. The most common position was horizontal (43.3%) followed by mesioangular (36.1%), distoangular (12.4%) and vertical (8.2%). The distribution of the molar classification was as follows: The majority (54.6%) had a 3B classification, the next most common was 2B (22.7%) followed by 3C (13.4%). The remaining 9.3% was distributed across 1B, 2A, 2C and 3A. Only three patients developed some kind of postoperative complication within 4 weeks.

Responsiveness of the OHIP

The internal responsiveness of the OHIP-14 was assessed by using ANOVA for repeated measures. Data were analysed twice: (1) for the entire sample and (2) for patients without any pre- or postoperative complaints. Results are presented in Table 1. For the entire sample, a significant effect over time was found, $F(8. 87) = 21.6$, $p < 0.001$, $\eta^2 = 0.67$. This effect resulted from a significant increase in mean score on the first day of the postoperative week, relative to the pre-test measure. In addition, all mean scores were significantly different relative to each other, with exception of the mean difference between the preoperative score and the one-month postoperative score. For the sample of patients without pre- or postoperative complaints a nearly identical result was found. A significant effect for time was found, $F (8, 61) = 15.8$, $p < 0.001$, $\eta^2 = 0.67$. All mean scores were significantly different from each other with the exception of the mean scores on day 6 and 7. After one month the mean OHIP-score did not differ from the preoperative score. These results show that the OHIP-14 is able to differentiate between the preoperative day, nearly all days within the postoperative week and one month postoperatively, and can therefore be considered internally responsive to changes in impacts of oral conditions as a result of surgical third molar removal.

Table 1. Mean OHIP-14 score (standard deviation) preoperatively, on each postoperative day for a week and one month postoperatively

	Pre-	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Month
Group 1[a]	16.85	28.47	25.71	23.53	22.17	20.98	19.92	18.62	16.34^
(N = 95)	(5.35)	(9.32)	(8.89)	(8.27)	(8.59)	(8.15)	(7.70)	(6.57)	(4.54)
Group 2[b]	16.26	28.33	25.45	22.98	21.69	20.25	19.40	18.14	16.13^
(N = 69)	(3.89)	(9.33)	(8.66)	(7.52)	(7.76)	(6.80)	(7.28)	(6.40)	(4.17)

[a] Entire sample
[b] Subsample without patients experiencing preoperative or postoperative complications
^ Differences in the preoperative score and one-month postoperative score are *not* significant at α= 0.05

The previous analyses have shown that all repeated measurements differed significantly (expect for pre- and one-month postoperatively) on the total score. In the following subscale analyses, those points in time are reported where the observed differences were largest. That is, the preoperative measurement, the first postoperative day (on which the

difference relative to the preoperative measurement was largest) and the seventh postoperative day (which is still higher than the preoperative measurement but lower than on the first preoperative day). In other words, these results were specifically selected since they provide the reader with the maximum amount of information with the least amount of space. Table 2 represents the mean OHIP-14 subscale scores on the preoperative day, the first day postoperatively and one week postoperatively. Results show that, relative to the other subscales, patients score significantly higher on the Physical pain subscale $F (6, 91) = 7.27$, $p < 0.001$, $\eta^2 = 0.32$, preoperatively (most likely resulting from the patients that reported pain as preoperative complaint). On the first postoperative day, patients scored even higher on Physical pain, $F (6, 90) = 66.16$, $p < 0.001$, $\eta^2 = 0.82$. Also, there are still significant differences after one week $F (6, 90) = 13.36$, $p < .001$, $\eta^2 = 0.47$, which can still be attributed to the Physical pain subscale. After correcting for multiple tests (seven subscales so $0.05 / 7 = 0.007$) the results described above remain significant.

Table 2. Mean OHIP-14 subscale scores and standard deviations for the preoperative and postoperative period (first day and after one week).

	Preoperative	Postoperative Day 1	Postoperative Day 7
	Mean (SD)	Mean (SD)	Mean (SD)
Functional limitations	2.07 (0.33)	3.74 (1.63)	2.39 (0.72)
Physical pain	2.94 (1.53)	6.25 (1.77)	3.53 (1.70)
Psychological discomfort	2.59 (1.22)	3.34 (1.57)	2.45 (1.06)
Physical disability	2.33 (1.05)	4.39 (2.00)	2.77 (1.45)
Psychological disability	2.47 (1.11)	3.39 (1.45)	2.45 (0.92)
Social disability	2.29 (0.79)	3.93 (1.66)	2.57 (1.18)
Handicap	2.17 (0.72)	3.45 (1.90)	2.42 (0.94)

SD, standard deviation

Clinical variables

Patients without preoperative complaints scored significantly lower (mean = 16.2, SD = 3.9) than patients with preoperative complaints (i.e. pain; mean = 18.9, SD = 8.1) on the mean OHIP-14 score preoperatively, $t (93) = -2.2$, $p < 0.03$, ES = 0.24.

Furthermore, patients with partial and complete mucosa coverage were compared across time on the OHIP-14 score. Besides an expected within-patients effect from time $F (8, 86) = 21.9$, $p = < 0.001$, $\eta^2 = 0.67$, a trend was found between mucosa and time, $F (8, 86) = 1.99$, $p < 0.06$, $\eta^2 = 0.16$. Exploratory inspection of the means plotted(?) over time revealed that the interaction resulted from a difference in the change from the preoperative-mean score to the first postoperative-mean score. The group with partially covered mucosa had a higher preoperative score (mean = 17.79, SD = 6.81) than the completely covered group, but showed a smaller increase of impact on oral health (postoperative mean = 27.38, SD = 7.73), while the completely covered group started out lower (mean = 15.94, SD = 3.17), but ended up higher (mean = 29.53, SD = 10.63). This was related to the mean surgery time, which was longer for patients with full mucosa coverage (mean = 18.1 min. SD = 6.6) than for patients with partial mucosa coverage (mean = 14.1 min. SD = 3.6), $t (95) = 3.65$, $p < 0.001$, ES = 0.38.

ANOVA was used to compare the mean OHIP-14 score between the different angular classifications on each day of measurement, but no significant differences could be shown at any point in time. This analysis was repeated using molar classification as the

independent variable. Only the three largest groups were analysed, the 3B, 2B and 3C position. A significant difference between groups was found on mean OHIP-14 score on the first postoperative day, $F_{(2, 84)} = 3.22$, $p = 0.045$, $\eta^2 = 0.07$. Next, the two largest groups (3B and 2B) were analysed using ANOVA for repeated measures on the mean OHIP-14 scores of the preoperative and first postoperative day. Results showed a significant interaction between position and time, $F_{(1, 72)} = 4.24$, $p = 0.043$, $\eta^2 = 0.06$, resulting from a higher increase in mean OHIP-14 score for the molars in the 3B position. A logical result, since more than half the crown is impacted in the mandibular ramus. Chi^2-analysis indeed shows a strong association, $X^2_{(2)} = 18.35$, $p < 0.001$, between type of removal (with and without splitting) and degree of impaction (2A, 3B and 3C), since nearly all molars (51 vs. 2) with a 3B degree of impaction, and all molars with a 3C degree of impaction, required alveotomy using splitting.

Patients who needed splitting (n=85) of the third molar were compared to patients who did not require splitting of the molar (n= 12) on OHIP-14 score across time. The following results should be considered preliminary given the small sample size of the subgroups. Results again show the time effect, $F_{(8, 86)} = 7.91$, $p < 0.001$, $\eta^2 = 0.42$ but no interaction with type of alveotomy (with/without splitting), $F_{(8, 86)} = 0.59$, $p < 0.78$, $\eta^2 = 0.05$. Nevertheless, the group that required splitting scored consistently higher than the other group, except for preoperatively, where they scored lower. In addition, the group without splitting was rather small, i.e. n = 12, resulting in low power to detect possible differences, if present. Surgical removal without splitting lasted significantly shorter (mean = 12.4, SD = 3.5) than surgery that did require splitting the molar (mean = 16.7, SD = 5.8), $t_{(95)} = -2.48$, $p < 0.02$, ES = 0.25.

DISCUSSION

As was hypothesized, patients scored higher on the OHIP in the first few days after surgery, which can be explained by the impact of surgery. Furthermore, by day seven and one month after surgery a large proportion of the patients scored even below their preoperative level. In addition, the OHIP-14 differentiates well between patients with or without pre-operative complaints, and between patients differing in other clinical variables (mucosa coverage, position and extent of surgery). Patients without preoperative complaints reported less oral impacts than patients with preoperative complaints. Although the difference was relatively small, it does suggest that the OHIP-14 is sensitive enough to distinguish between patients with and without preoperative complaints. Also, patients with partially covered mucosa reported more oral impacts after surgery than patients with completely covered mucosa. This could be explained because partially covered mucosa is perhaps more prone to infections than completely covered mucosa. However, patients with completely covered mucosa, although starting out with less oral impacts, ended up reporting more oral impacts right after surgery. This may result from the fact that more tissue needs to be removed and /or a longer surgery time is necessary for patients with full mucosa coverage. Although no significant difference was found, the time of surgery for patients who required splitting was also longer and resulted in higher scores on the OHIP-14.

Whether these results are clinically significant is a matter of perspective. There is no specific guideline for determining clinical significance, because small differences in

mean scores can be statistically significant. Nevertheless, statistical significance is not equivalent to clinical significance (Crosby et al. 2003). To give some support to significance testing and to be able to interpret these outcomes, effect-size estimates can be used, but these estimates do not quantify clinical significance or relevance because that depends highly on the disease or the condition under consideration (Addy et al. 2005).

One important point of discussion concerns the comparison of the OHIP-14 score that was filled in for different time frames (i.e. oral impacts in the last four weeks or today). The authors like to point out that, in the original instruction, the 4-week time interval is used to yield a more or less stable estimate. That is, not all oral complaints are present continuously, but may vary across time. That is why the 4-week interval is used in the original OHIP. However, suppose a healthy patient undergoes third molar surgery. On the second postoperative day the OHIP-14 is filled in, asking for the impact of oral health status on the last 4 weeks. We expect that the patient will answer the questions based on the discomfort felt in the last two days. We do not expect that the patient will 'calculate' an average impact based on 26 days without discomfort and 2 days with discomfort. Given this reasoning, we feel that comparing the preoperative, one-week postoperative and one-month postoperative is suitable despite differences in instruction. Related to this, we found that mean scores on the OHIP of the first post-operative day corresponded with postoperative scores on the OHIP that were taken in another study (Van Wijk et al. 2009) one week post-operatively. In some way, it is remarkable that approximately the same score is found on the first postoperative day and one week postoperatively. A difference between the two studies is that in the present study patients were asked to rate the impact of oral condition on normal functioning TODAY, while in the other study (Van Wijk et al. 2009) patients were asked to rate the impact of oral condition in the previous WEEK. Nevertheless, mean scores were about the same. This suggests that responses are influenced by that point in time where problems and/or pain are felt the most. This is an interesting finding and definitely worth further investigation.

Limitations regarding the different recall-periods are apparent, and the comparison with another study does not exclude the possibility of recall-effects. Convincing evidence could be provided by repeatedly evaluating patient's oral impacts while their oral health is stable, or by applying different recall-periods to groups of patients with similar oral health conditions. For future reference, perhaps a global transition judgment could also be included to record the change over time retrospectively, which can be compared to the changes found between the before and after judgments. Also the lack of a control group that has not undergone surgery makes it difficult to conclude whether the effects found are actually due to the improvement of oral health or due to regression to the mean. Furthermore, social desirability should also be considered as a potential liability to effect the results of this study.

CONCLUSION
Overall, this study shows that the OHIP-14 is responsive to changes as a result of third molar surgery, and is able to differentiate between patients on a number of clinical variables. However, considering the limitations, more research is needed.