Acta Odontologica Scandinavica, 2007; 65: 109–118

informa
healthcare

ORIGINAL ARTICLE

Validation and improvement of a predictive model for treatment outcome in patients with temporomandibular disorders

BERTIL SUNDQVIST[1], BENGT WENNEBERG[2] & TOMAS MAGNUSSON[3]

[1]*Department of Stomatognathic Physiology, Postgraduate Dental Education Centre, Örebro, Sweden* [2]*Department of Stomatognathic Physiology, Göteborg University, Göteborg, Sweden and* [3]*Department of Stomatognathic Physiology, The Institute for Postgraduate Dental Education, Jönköping, Sweden*

Abstract

Objective. To examine whether it was possible to improve individual prediction of treatment outcome in patients with temporomandibular disorders (TMD) through continuous quality improvement registrations and implementation of changes in clinical routines. ***Material and methods.*** All 5777 patients referred to a specialist clinic for stomatognathic physiology; 2187 TMD patients started treatment. The patients were divided into Muscle or Mainly TMJ symptom groups. Actual treatment outcome was predicted as Good or Dubious based on patient history and clinical findings. The degree of improvement of initial symptoms was graded using a numeric rating scale 0–100. Improvement in initial complaints of 50% or more was judged as a clinically important difference (CID); 989 patients (76%) completed treatment 1992–1998 (Sample 1) and 769 (86%) treatment 1999–2004 (Sample 2). ***Results.*** For patients with Muscle symptoms in Sample 1, CID was reached by 90% of those predicted Good and by 56% of those predicted Dubious. In Sample 2, the figures were Good 93% and Dubious 57%. In subgroup Mainly TMJ symptoms, CID was reached by 94% of those predicted Good and by 88% if predicted Dubious in Sample 1. In Sample 2, the figures were the same for those predicted Good, i.e. 94%, and Dubious, 73%. ***Conclusion.*** Continuous quality improvement registrations and implementation of changes in clinical routines improved the ability to predict individually the actual treatment outcome between patients predicted Good or Dubious, respectively. Rheumatic disease turned out not to be a negative predictor for treatment outcome.

Key Words: *Clinical trial, craniomandibular disorders, occlusal adjustment, occlusal splints, quality control*

Introduction

The possibility to predict individual treatment outcome in patients diagnosed with temporomandibular disorders (TMD) treated mainly with interocclusal appliance and occlusal adjustment has been reported recently [1,2]. With a focus on mistakes and pitfalls, i.e. all patients predicted Good, reaching an objective treatment goal, but not the predicted treatment outcome, and patients predicted Dubious but repeatedly reporting no improvement, prediction and evaluation of individual treatment outcome may be regarded as a history of errors in TMD treatment.

TMD has been determined as a major cause of non-dental pain in the orofacial region [3]. There is controversy among clinicians and researchers concerning its cause as well as how to treat TMD patients, and the efficacy of occlusal treatment for managing TMD is also still controversial [4].

Predictability can be defined as "a clinician's probability of achieving a successful outcome of treatment" [5] or as a method clinically identifying moderators or mediators of actual treatment outcome to be tested in well-controlled experiments [6]. Shewhart [7] stated that: "A phenomenon will be said to be controlled when, through the use of past experience, we can predict, at least within limits, how the phenomenon may be expected to vary in the future." He later added that there are three important components of knowledge: (a) the data of experience in which the process of knowing begins, (b) the prediction in terms of data that one would expect to get if performing certain experiments in the future, and (c) the degree of belief in the prediction based on the original data or some summary thereof as evidence [8]. A prediction is a statement of how, for example a process, is expected to perform.

Correspondence: Bertil Sundqvist, Department of Stomatognathic Physiology, Postgraduate Dental Education Centre, Box 1126, SE 70371 Örebro, Sweden. Tel: +46 19 602 40 70. Fax: +46 19 602 40 55. E-mail: bertil.sundqvist@orebroll.se

(Received 9 May 2006; accepted 7 November 2006)

ISSN 0001-6357 print/ISSN 1502-3850 online
DOI: 10.1080/00016350601106082

Implicit in every case is a prediction that the change will result in future improvement. Test results that do not concur with predictions may be one reason for rethinking the theory grounding the prediction [9]. "Prediction" has to be separated from "prognosis", defined as "a forecast as to the probable outcome of an attack of disease, the prospect as to recovery from a disease as indicated by the nature and the symptoms of the case" [10]. A suggestion for a new definition of prediction is: "a clinician's possibility to predict an in advance defined subjective treatment outcome at an objective treatment goal".

Quality improvement reporting is a new kind of research through which authors describe improvement projects that others can learn from. The science of quality improvement is intended to generate knowledge concerning the prediction and control of health-care systems for producing better treatment outcomes [11]. Quality improvement in health care is a process for producing outcomes using evidence-based medicine (EBM) and best clinical skills to meet the needs and expectations of patients. Core activities are: 1) to seek an understanding of the sources of systematic as well as unwanted and unnecessary variation, 2) to implement cost-effective strategies to reduce unwanted variation, and 3) to produce organization-wide knowledge on structured approaches to change process and improve outcomes. Understanding variation and raising outcomes to a higher level requires thinking from the point of view of the system. Effective and efficient implementation of continuous improvement requires data-driven methods and ongoing evaluation. Implementation in everyday practice is the single most important factor distinguishing quality improvement from traditional evaluate research [12]. Some authors have described this process as real-time science. Quality improvement reports are said to broaden the view of EBM [13].

The main aim in quality improvement research is assessing whether a study intervention, imposed to change a process, produces an improvement in outcome. Quality improvement research is commonly based on the PDSA cycle, developed by theoreticians in economy. The PDSA model advocates: formation of a hypothesis for improvement, stating the objectives of the plan, making predictions (Plan), treating, documenting and analyzing data (Do), comparing data to predictions (Study), and implementing changes (Act). Conducting a scientific study is the most important concept of PDSA quality improvement that is fundamental to iterative learning [9].

The hypothesis of the present study was whether the possibility to predict individual treatment outcome in patients diagnosed with TMD, treated mainly with interocclusal appliance and/or occlusal adjustment, may improve over time as a result of continuous quality improvement registration and implementation of changes in clinical routines according to a model similar to the PDSA cycle. As far as we are aware, this is the first quality improvement report about the treatment of TMD.

Material and methods

The original patient sample comprised all 5777 patients who had been referred to the specialist clinic of stomatognathic physiology in Sollefteå/Sundsvall, County Council of Västernorrland, Sweden, during the period 1992–2004. Eleven percent (612) of the subjects rejected an examination, in most cases owing to spontaneous remission of symptoms or other ongoing treatment at the time when they were called for examination. All the other 5165 patients were subjected to a clinical examination carried out by one and the same TMD specialist or by specially trained general practitioners [2]. Patients without the co-existence of subjective symptoms and clinical signs of TMD ($n=1563$) were excluded from the investigation.

Out of a total of 3602 patients diagnosed with TMD, 46 were excluded because the treatment outcome was predicted Poor [1]. Another 595 patients were sent back to their ordinary dentist with a treatment plan and a further 511 patients were treated by specially trained general practitioners [2].

The remaining 2450 patients had both signs and symptoms of TMD and were offered treatment performed by a TMD specialist. Of those, 263 rejected the suggested treatment for varying reasons. The remaining 2187 patients commenced treatment (Figure 1).

Inclusion criteria: examined, predicted, treatment planned, and treated by a TMD specialist, signs and symptoms of TMD, individual predicted treatment outcome Good or Dubious, patients expressing a demand for treatment, and treatment outcome evaluated when a stable occlusion in retruded contact position (RCP), measured by double-folded 12μ thin plastic foil, had been established on the interocclusal appliance and/or in the dentition and had been kept that way for at least 2 months.

The patients were divided into two groups: *Group 1* – Patients who commenced, discontinued and/or completed treatment initiated by the TMD specialist during the period 1992–1998 ($n=1297$). Patients who completed treatment will henceforth be called Sample 1 ($n=989$). *Group 2* – Patients who commenced, discontinued and/or completed treatment initiated by the TMD specialist during the period 1999–2004 ($n=890$). Patients who completed treatment will henceforth be called Sample 2 ($n=769$).

The examinations performed have been described in detail previously [1,2]. Based on signs and

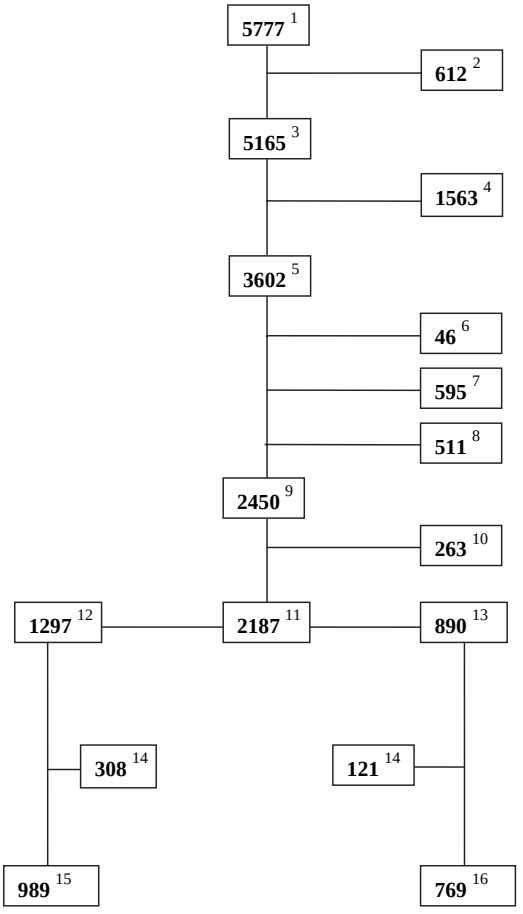

Figure 1. Loss of participants. 1: All referred to the clinic. 2: Rejected examination. 3: Examined. 4: Patients without signs or symptoms of TMD. 5: Patients with signs and symptoms of TMD. 6: Patients with TMJ clicking not offered treatment because of poor prediction of treatment outcome. 7: Patients sent back to ordinary dentist with treatment plan. 8: Patients who commenced treatment by specially TMD -trained general practitioner. 9: Patients fulfilling inclusion criteria. 10: Patients rejecting suggested treatment. 11: Patients who commenced treatment performed by TMD specialist. 12: Patients who commenced treatment 1992–1998. 13: Patients who commenced treatment 1999–2004. 14: Patients who discontinued treatment. 15: Patients who completed treatment 1992–1998 16: Patients who completed treatment 1999–2004.

symptoms, the patients were allocated to one of two subgroups using principally the same criteria as defined by Okeson [3]: *Muscle group* – symptoms and signs of myogenous origin; *Mainly TMJ group* – symptoms and signs of mainly arthrogenous origin. Patients with fatigue, stiffness and/or pain in the masticatory muscles and/or tension-type headache, and who exhibited palpatory tenderness in the masticatory muscles, were allocated to the *Muscle group*. Patients with joint pain and/or locking at rest and/or during function and palpatory tenderness in the region of the TMJ, other clinical and/or radiological signs of internal derangement, arthritis or arthrosis in the TMJ, were allocated to the *Mainly TMJ* group, regardless of whether these patients had muscular signs and symptoms or not.

All patients were also subjected to an individual prediction of treatment outcome based on all data collected at the first examination, and, in some cases, information from radiographic examinations. A correct prediction of treatment outcome was defined as an improvement of 50% or more, judged as a clinically important difference (CID) [14]. The patients were unaware whether the cut-point for judging a treatment outcome was positive or not.

The possibility of reaching a CID was determined as *Good*, *Dubious*, or *Poor* [15]. Patients were determined *Good* if the TMD symptoms were known to have a good long-term treatment outcome; pain in the face and jaws, difficulty opening the mouth wide, pain on movement of mandible, chewing difficulties, locking or dislocation of mandible [16], or patients with tension-type headache and clinical signs of TMD [17]. Patients determined *Dubious* were classified within subgroups.

Subgroup 1 – patients suffering from symptoms with no significant improvement at a 4-year follow-up [16]: Feeling of fatigue in the jaws or cheek, tinnitus/impaired hearing, dizziness, tongue pain, swallowing difficulties/globus in the throat, extensive wear of teeth, and/or TMJ crepitation.

Subgroup 2 – patients with TMD symptoms with prediction Good but also rheumatoid arthritis [18], ankylosing spondylitis [19], psoriatic arthritis [20], fibromyalgia [21], or whiplash injury [22], diseases/illnesses all known to affect masticatory muscles and/or TMJs.

Subgroup 3 – patients with clinical and/or radiological signs that might affect treatment outcome; TMD symptoms but no clinical signs corresponding to the symptoms, i.e. bilateral symptoms and unilateral signs or vice versa, or patients with gross structural changes in the TMJs due to general diseases or previous fractures, making it impossible to reproduce the RCP [23].

Patients determined *Poor* were: TMJ clicking without pain and/or locking and no other signs or symptoms of TMD [24] or orofacial pain without clinical signs of TMD. Patients predicted Poor were not offered any treatment.

The TMD specialist made individual treatment plans for all 2187 patients. The individualized treatment followed a strict protocol that has been described in detail in a previous article [1]. In short, simple and mostly conservative treatment methods have been utilized, such as interocclusal appliances and/or selective occlusal grinding. In cases where interocclusal appliances were used, full coverage, stabilization appliances of the Michigan type were the most common. When selective occlusal grinding was performed, it followed the principles described by Ash & Ramfjord [25]. The treatment goal was defined as a stable occlusion in RCP on the interocclusal appliance and/or in the dentition measured with double-folded 12μ thin plastic foil in a clamping

tweezer. Treatment outcome was evaluated when the occlusion had been stable for at least 2 months.

At treatment goal, all the patients were asked to grade the degree of improvement of their initial symptoms. The improvement was recorded in percent using a numeric rating scale ranging from 0 to 100 [26]. To evaluate the treatment outcome, a 2-grade scale based on the numeric ratings was used. Grade 1: improved 50% or more; Grade 2: improved less than 50%, no change or impairment of initial symptoms.

During the period covered by the investigation, new predictors were identified in two ways: (1) Patients with a general illness/disease that might affect muscles and/or joints, and (2) patients predicted Good who reached the treatment goal but did not have a CID. Furthermore, the patients' records were examined regarding past and present history, with a focus on psychosocial and psychological factors, previous or existing illness/disease that might influence treatment outcome. Identified hypothetical negative predictors have been used as keywords in a PubMed search. Patients later identified with these hypothetical negative predictors have been predicted Dubious. In patients predicted Dubious, who during the study period reached treatment goal but repeatedly failed to have a CID, the prediction has been changed to Poor. After identification, such cases have not been offered any treatment. All identified predictors have continuously been implemented into clinical routines (Figure 2).

Subgroup 4: Identified new predictors: Psychological/psychosocial factors such as patients living in a chronic pain family [27], and patients with secondary gain of pain [28], psychiatric diagnoses, e.g. schizophrenic patients [29], orofacial symptoms for more than 30 years [30], scuba divers with internal derangement [31], narcotic drug abuse [32].

General illness/disease such as gout [33], facial paralysis [34], epileptic spasm in masticatory muscles [35], radiation therapy in jaw, face or head [36], neurological diseases with possible influence in jaws, face or head [37], systemic lupus erythematosus (SLE) [38], mixed connected tissue disease (MCTD) [39], Ehlers-Danlos syndrome [40].

Local factors such as complete dentures and crista flaccida, previous trauma in the region of symptom [41], tension-type headache not localized to the temples and/or forehead, tension-type headache described by the patient as emanating from neck, shoulders and/or backhead, and localized in the temples and/or forehead. Factors such as incomplete past history, e.g. language problems, refusal to answer questions, and, in patients with internal derangement, sleeping position on the stomach [42] and/or nail biting.

Identified predictors of *Poor* treatment outcome: patients with tinnitus, impaired hearing, and/or dizziness, without any TMD symptoms known to have good long-term treatment outcome [16]. This group is no longer offered treatment.

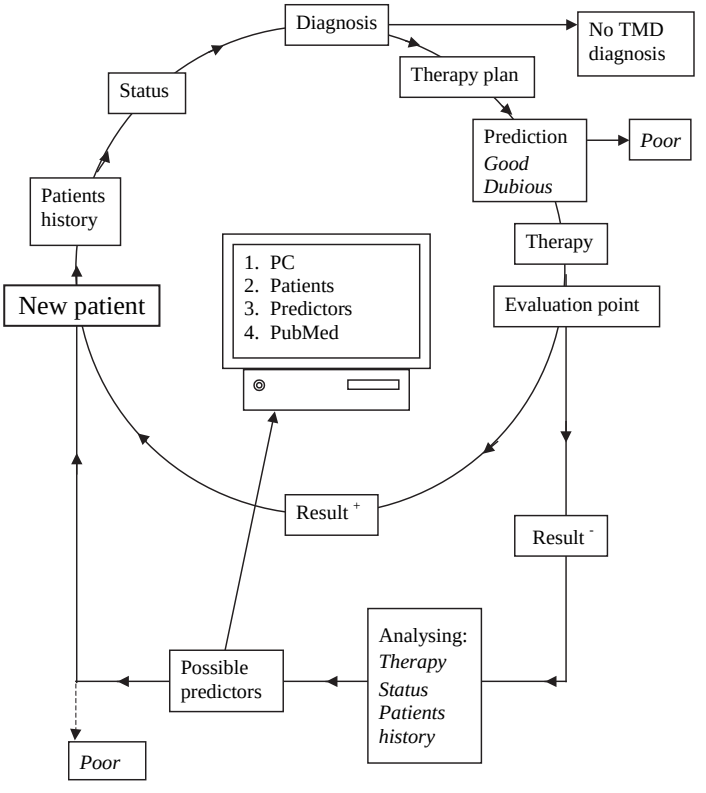

Figure 2. The hypothesis generator.

The agreement between predicted and actual treatment outcome in the 989 patients in Sample 1 who completed treatment (870 predicted Good and 119 predicted Dubious) was evaluated and compared with the results obtained in the 769 patients in Sample 2 (410 predicted Good and 359 predicted Dubious) (Figure 1). An improvement in the possibility to individually predict treatment outcome was defined as an increased difference in actual treatment outcome between patients predicted Good or Dubious and an increased number of patients predicted Good reaching a CID.

In Sample 2, we also made an attempt to evaluate whether there was a difference in actual treatment outcome when comparing patients predicted Dubious in the four different subgroups of predictors with all patients predicted Good. Furthermore, in Sample 2 we evaluated whether there was a cumulative effect with increasing number of negative predictors through comparing patients with 1 predictor and those with 2 or more predictors.

Statistics

To test for differences between groups, the chi-square test has been used [43]. The levels of significance used were: $p \geq 0.05$ N.S. (not significant), $0.01 \leq p < 0.05$, $0.001 \leq p < 0.01$, $p < 0.001$.

Results

In Sample 1, 308 patients (24%) discontinued the treatment. The corresponding figure for Sample 2 was 121 (14%). The difference was statistically significant ($p < 0.001$). Reasons for discontinuing treatment were, for example, inability to wear the appliance, not coming to scheduled appointments, and moved out of the district.

In Sample 1, 630 out of 989 patients (64%) were allocated to subgroup *Muscle* and 359 (36%) to subgroup *Mainly TMJ*. In Sample 2, the figures were much the same: 475 (62%) and 294 (38%), respectively. The distribution of the patients according to subgroup, gender and age is presented in Table I. The age distribution between men and women was much the same in the two samples in both the *Muscle* and *Mainly TMJ* groups, but patients with *Muscle* symptoms were statistically significantly older compared to those with *Mainly TMJ* symptoms in both samples (Sample 1; *Muscle*; mean age: 43 years, range 5–87, *Mainly TMJ*; mean age: 37, range 10–82, respectively; $p < 0.001$. Sample 2; *Muscle*; mean age: 44 years, range 5–85, *Mainly TMJ*; mean age: 39, range 10–87, respectively; $p < 0.001$).

In Sample 1, 537 out of 630 patients (85%) of the subjects in subgroup *Muscle* and 333 out of 359 patients (93%) in subgroup *Mainly TMJ* were predicted Good ($p =$ N.S.). The corresponding figure for Sample 2 was 215 out of 475 (45%) and 195 out of 294 (66%), ($p < 0.01$). When comparing Samples 1 and 2 in respect of patients predicted Good, there was a statistically significant decrease in both subgroup *Muscle* and subgroup *Mainly TMJ* ($p < 0.001$ and $p < 0.001$, respectively).

The different treatment modalities performed are presented in Table II. One significant difference in the treatment panorama could be found comparing the two samples; more extensive occlusal therapy, such as onlays, new dentures, fixed partial dentures (FPDs), orthodontics and/or orthognathic surgery, had decreased from 18% to 5% ($p < 0.001$).

When comparing prediction Good and Dubious with actual treatment outcome, there was a statistically significant difference in subgroup *Muscle* in both samples ($p < 0.001$ and $p < 0.001$, respectively). In subgroup *Mainly TMJ* there was no statistically significant difference between prediction Good or Dubious and actual treatment outcome in Sample 1 ($p =$ N.S), while in Sample 2 a statistically significant difference was found ($p < 0.001$). Comparing prediction Good or Dubious and actual treatment outcome in Sample 1 with Sample 2 there were no statistically significant differences for either of the two subgroups ($p =$ N.S.). Comparing prediction Good and actual treatment outcome in subgroup *Muscle* with subgroup *Mainly TMJ* there was no statistically significant difference either in Sample 1 or Sample 2. However, when comparing prediction Dubious and actual treatment outcome between the two subgroups *Muscle* or *Mainly TMJ* symptoms, there was a statistically significant difference of Dubious cases reaching CID in both samples ($p < 0.01$ and $p < 0.01$, respectively, Table III).

Table I. Comparison of patients with Muscle or Mainly TMJ symptoms according to gender and age in Samples 1 and 2. Percentage distribution

	Muscle Sample 1			Muscle Sample 2			Mainly TMJ Sample 1			Mainly TMJ Sample 2		
Age	Female $n=479$	Male $n=151$	Total $n=630$	Female $n=381$	Male $n=94$	Total $n=475$	Female $n=289$	Male $n=70$	Total $n=359$	Female $n=245$	Male $n=49$	Total $n=294$
0–19	9	10	9	6	9	7	22	17	21	21	35	23
20–39	29	32	30	34	35	34	33	43	35	33	31	33
40–59	44	42	44	45	38	44	33	23	31	28	22	27
60–	18	16	17	14	18	15	12	17	13	18	12	17

Table II. Main treatment/treatments for patients in Samples 1 and 2. Percentage distribution

Treatment	Sample 1			Sample 2		
	Muscle $n=630$	Mainly TMJ $n=359$	Total $n=989$	Muscle $n=475$	Mainly TMJ $n=294$	Total $n=769$
Only acute treatment	3	3	3	0	1	1
Appliance night wear	39	26	34	38	33	36
Appliance and occlusal adjustment	37	38	37	46	46	46
Only occlusal adjustment	7	10	8	11	13	12
Onlays full time	2	6	4	2	2	2
New complete or partial dentures	4	4	4	1	1	1
FPDs, orthodontics and/or orthognathic surgery	8	13	10	2	3	2

A correct prediction of treatment outcome (Good $\geq$50% improvement, Dubious <50% improvement, respectively) in subgroup *Muscle* was made in 524 out of 630 patients in Sample 1, and in subgroup *Mainly TMJ* symptoms the corresponding figures were 316 out of 359 patients (83% and 88%, respectively; $p<0.05$). The figures in Sample 2 were 314 out of 475 patients in subgroup *Muscle* and 210 out of 294 patients in subgroup *Mainly TMJ* (66% and 71%, respectively; $p=$N.S.). Comparing the figures for correct prediction between Samples 1 and 2 in subgroup *Muscle* and *Mainly TMJ*, respectively, the figures for a correct prediction had decreased in both subgroups ($p<0.001$, $p<0.001$, respectively; Table IV).

The actual treatment outcome for all patients predicted Dubious in the four different subgroups ($n=22$, $n=188$, $n=26$, $n=123$, respectively) was compared with those predicted Good in Sample 2 ($n=410$). In all four subgroups, the actual treatment outcome was statistically significantly less favorable compared to patients initially predicted Good ($p<0.001$, $p<0.001$, $p<0.001$ and $p<0.001$, respectively) (Table V).

In Subgroup 2, there was a sufficient number of individuals for each of the three subdiagnoses to allow separate analyses (Table VI). Patients with whiplash injury or fibromyalgia had a statistically significantly poorer treatment outcome compared to those predicted Good ($p<0.001$ and $p<0.001$, respectively), while patients with some rheumatic disorder (rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis) had a treatment outcome comparable to patients predicted Good ($p=$N.S.).

Of all patients in Sample 2 who had been predicted Dubious, 316 had 1 negative predictor while 43 had 2 or more negative predictors. Sixty-three percent of those with one negative predictor reported an improvement of 50% or more, while only 49% of those with more than one negative predictor reported such an improvement. The difference was close to statistically significant ($p=0.08$, Table VII).

Discussion

Of all patients examined, 1563 (30%) were not diagnosed with any TMD and were not offered any treatment. This figure is similar to that in a recent study [44]. If these patients had symptoms from teeth, gums, the oral mucosa, tongue or other oral structures they were either recommended to contact their ordinary dentist or referred to other specialist dentists. If they had orofacial pain not emanating from the teeth, gums, oral mucosa, tongue, masticatory muscles and/or TMJs they were referred to, or recommended to contact, a physician.

Forty-six patients had TMJ clicking without pain and/or locking and no other signs or symptoms of TMD. They were not offered any treatment because of Poor prediction of treatment outcome [24]. The number of patients in this group had decreased. During the period 1992–1998, the number of patients with only TMJ clicking was 33, while only 13 such patients were referred during the period 1999–2004. The reason for this decrease is probably information from the TMD clinic to dentists in the region that we do not treat TMJ clicking without any other TMD symptom.

Table III. Comparison of patients in Sample 1 or 2 according to predicted and actual treatment outcome in the Muscle and Mainly TMJ group. Percentage distribution within parentheses

Improved $\geq$50%	Muscle group		Mainly TMJ group	
	Sample 1 $n=630$	Sample 2 $n=475$	Sample 1 $n=359$	Sample 2 $n=294$
Good	483 (90%)	201 (93%)	313 (94%)	183 (94%)
Dubious	52 (56%)	147 (57%)	23 (88%)	72 (73%)
	$p<0.001$	$p<0.001$	N.S.	$p<0.001$

Table IV. Distribution of the patients who fulfilled treatment in Samples 1 and 2 according to predicted and actual treatment outcome in subgroup Muscle and Mainly TMJ symptoms. Percentage distribution within parentheses

	Muscle group		Mainly TMJ group		Muscle group		Mainly TMJ group	
	Sample 1	$n=630$	Sample 1	$n=359$	Sample 2	$n=475$	Sample 2	$n=294$
	Prediction		Prediction		Prediction		Prediction	
Degree of improvement	Good	Dubious	Good	Dubious	Good	Dubious	Good	Dubious
≥50%	483 (90%)	52 (56%)	313 (94%)	23 (88%)	201 (93%)	147 (57%)	183 (94%)	72 (73%)
<50%	54 (10%)	41 (44%)	20 (6%)	3 (12%)	14 (7%)	113 (43%)	12 (6%)	27 (27%)

The age and sex distribution of the patients, and the proportion between patients with *Muscle* or *Mainly TMJ* symptoms, was in accordance with a recent study [44], with a dominance for females (78%) and for muscular symptoms (61%).

In Sample 1, 24% discontinued treatment and in Sample 2, 14%. The reason for the decrease might be that the difference between prediction Good, Dubious or Poor has been more distinct over time, which may result in fewer tentative treatments and more motivated patients.

In the subgroup *Muscle* the difference in actual treatment outcome between prediction Good and Dubious was obvious already in Sample 1 and difficult to improve further. In subgroup *Mainly TMJ* the difference in actual treatment outcome between patients predicted Good and Dubious had increased in Sample 2 compared to Sample 1. There was a small improvement of actual treatment outcome for patients in subgroup *Muscle* in Sample 2 compared to Sample 1 irrespective of prediction Good or Dubious. For those with *Mainly TMJ* symptoms, the actual treatment outcome was equal for patients predicted Good in both samples, while fewer patients in Sample 2 predicted Dubious reached CID compared to the corresponding patients in Sample 1. This means that overall more patients had a correct prediction of actual treatment outcome, especially among patients with *Mainly TMJ* symptoms and a Dubious prediction in Sample 2 compared to Sample 1.

On the other hand, a correct prediction of treatment outcome Good and Dubious, respectively, had decreased in both subgroup *Muscle* and in subgroup *Mainly TMJ*, from 83% to 66% and from 88% to 71%, respectively. This might seem contradictory to our statement above, claiming that more patients in Sample 2 had a correct prediction of actual treatment outcome compared to Sample 1. The reason is that the proportion of patients predicted Dubious had increased from 12% in Sample 1 to 47% in Sample 2. Since a substantial number of patients predicted Dubious reached the CID (thus wrongly predicted), an increased proportion of cases predicted Dubious will erroneously result in poorer figures for the overall prediction of treatment outcome.

How is it that the proportion of patients predicted Dubious increased over time? The reasons for this are several. The increase can be explained to some extent by the fact that more negative predictors had been included in the hypothesis generator. Furthermore, an increased awareness by the TMD specialist of the importance of finding predictors, resulting in a more attentive listening to the patient's past history, leads to an increased number of predictors being found. However, the most important factor is that a larger proportion of "simple TMD cases" are now being handled by their ordinary dentists. At the same time, the number of difficult cases being referred to the specialist clinic, especially from physicians, has gradually increased over time.

There was a significant difference between actual treatment outcome for those predicted Dubious when comparing the subgroup *Muscle* with subgroup *Mainly TMJ*. In both samples, more patients predicted Dubious in the group with *Mainly TMJ* symptoms reached CID compared to those predicted Dubious in the group with *Muscle* symptoms. One probable explanation for this is the overall good

Table V. Comparison in Sample 2 between predicted and actual treatment outcome in all patients predicted Good and all predicted Dubious, divided into four subgroups of negative predictors. Subgroup 1: symptoms with no significant improvement at a 4-year follow-up. Subgroup 2: rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, fibromyalgia or whiplash injury. Subgroup 3: TMD symptoms but no clinical signs corresponding to the symptoms or impossible to reproduce the RCP. Subgroup 4: identified new predictors. Percentage distribution within parentheses

Degree of improvement	Prediction Good $n=410$	Subgroup 1 $n=22$	Subgroup 2 $n=188$	Subgroup 3 $n=26$	Subgroup 4 $n=123$
≥50%	384 (94%)	7 (28%)	122 (65%)	19 (73%)	72 (59%)
<50%	26 (6%)	18 (72%)	66 (35%)	7 (27%)	51 (41%)
		$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$

Table VI. Comparison in Sample 2 between predicted and actual treatment outcome in patients predicted Good or Dubious in three subgroups Whiplash trauma (WAD), fibromyalgia (Fibro), or rheumatic disorders in the TMJs (RA). Percentage distribution within parentheses

Degree of improvement	Prediction Good $n=410$	WAD $n=107$	Fibro $n=53$	RA $n=28$
≥50%	384 (94%)	69 (64%)	28 (53%)	25 (89%)
<50%	26 (6%)	38 (36%)	25 (47%)	3 (11%)
		$p<0.001$	$p<0.001$	$p=$N.S.

prognosis when treating *Mainly TMJ* symptoms compared to treatment of *Muscle* symptoms [45]. It is also likely that some of the negative predictors might influence treatment outcome more in the subgroup *Muscle* compared to subgroup *Mainly TMJ*.

The present results have also identified one false negative predictor. The presence of a rheumatic disease, most common in subgroup *Mainly TMJ*, has been judged as a negative predictor but turned out to have no negative effect for the alleviation of local symptoms in the masticatory system.

A tendency for a cumulative effect of negative predictors was also found, since only half of those with more than one negative predictor reached CID, compared to two out of three with only one negative predictor.

Common treatment modalities for TMD are jaw exercises, interocclusal appliances, and selective occlusal adjustment. This was also true in the present material. The number of patients with more extensive occlusal therapy had decreased from 18% to 5%. The reason is that in Sample 1 all patients with severe malocclusions or loss of molar support were routinely recommended permanent occlusal stabilization. The need for extensive permanent occlusal therapy in treatment of TMD has been questioned [23]. Furthermore, extensive occlusal therapies are expensive and time consuming. Consequently, fewer patients in Sample 2 have been subjected to such therapies.

Psychological and psychosocial factors have been suggested as possible predictors of treatment outcome [46]. In line with others, we have suggested two psychological or psychosocial negative predictors for treatment outcome, namely patients living in a chronic pain family [27] and secondary gain of pain [28].

Table VII. Comparison in Sample 2 between predicted and actual treatment outcome in patients predicted Dubious with 1 or 2 or more negative predictors of actual treatment outcome. Percentage distribution within parentheses

Degree of improvement	1 predictor $n=316$	2 or more predictors $n=43$
≥50%	198 (63%)	21 (49%)
<50%	118 (37%)	22 (51%)
	$p=$N.S.	

In this investigation, an improvement of 50% or more was defined as a positive treatment outcome. Evaluating the reasons for a positive treatment effect is never easy, because so many different factors interact, except from a therapeutic action from a specific treatment. In randomized trials, estimates of treatment effect may be exaggerated by up to 40% by the patients [47]. The role of spontaneous improvement in patients, i.e. regression to the mean [48], should not be forgotten when interpreting the results. It is well known that the prediction in itself may increase the placebo effect [49], and also that a convinced therapist, with an explanation that makes sense to the patients, can improve the placebo effect [50]. Furthermore, we have tried to give a high quality of "caregiver communication" using a professional attitude as defined by Holm [51]. However, in patients with orofacial pain the quality of caregiver communication predicts only 10% to 14% of the variance in outcomes [52]. Finally, no one has ever been able to find a reliable way of predicting who is going to respond to inert treatment and who is not, and no one has been able to predict placebo [53].

The method used, i.e. a hypothesis generator, is scientifically biased through its continuous change of inclusion and exclusion criteria, but this is also the clinical strength of the method. Improving the daily practice of medicine or dentistry requires making changes in processes of care. The most powerful way to make such changes is to conduct small local tests according to the PDSA cycle, in which one learns from taking action [9]. The findings in a quality improvement study should add knowledge to the field and the organization. The lessons learned may be both about what didn't work and what was successful [54].

A limitation of this report is that the method has been developed and tested at one clinic. Whether the results are generalizable is not known, but as Codman declared as long ago as in 1914: "we must formulate some method of hospital report showing as nearly as possible what are the results of treatment obtained at different institutions. This report must be made out and published by each hospital in a uniform manner, so that comparison will be possible. With such a report as a starting point, those interested can begin to ask questions as to management and efficiency" [55]. We welcome other

institutions testing the method and comparing the results.

In conclusion, continuous quality improvement registrations and implementation of changes in clinical routines improved the ability to predict actual treatment outcome individually between patients predicted Good and Dubious, respectively. Rheumatic diseases turned out not to be a negative predictor for treatment outcome.

Acknowledgment

This study was financially supported by Folktandvården, County Council of Västernorrland and Örebro, Sweden.

References

[1] Sundqvist B, Magnusson T. Individual prediction of treatment outcome in patients with temporomandibular disorders. Swed Dent J 2001;25:1–11.

[2] Sundqvist B, Magnusson T, Wenneberg B. Comparison between predicted and actual treatment outcome in patients with temporomandibular disorders treated by TMD-trained general dental practitioners. Swed Dent J 2003;27:131–41.

[3] Okeson JP. Differential diagnosis and management considerations of Temporomandibular Disorders. In: Okeson JP, editor. Orofacial pain. Guidelines for assessment, diagnosis, and management. Carol Stream: Quintessence; 1996. p. 113–84.

[4] Al-Ani Z, Gray RJ, Davies SJ, Sloan P, Glenny AM. Stabilization splint therapy for the treatment of temporomandibular myofascial pain: a systematic review. J Dent Educ 2005;69:1242–50.

[5] Newman M, Kornman K, Holtzman S. Association of clinical risk factors with treatment outcomes. J Periodontol 1994;65:489–97.

[6] March JS, Curry JF. Predicting the outcome of treatment. J Abnorm Child Psychol 1998;26:39–51.

[7] Shewhart WA. The economic control of quality of manufactured product. London: Macmillan; 1931.

[8] Shewhart WA. Statistical method from the viewpoint of quality control. Washington D.C.: Graduate School, Department of Agriculture; 1939.

[9] Langley GJ, Nolan KM, Nolan TW, Norman CL, Provost LP. The improvement guide. San Francisco: Jossey-Bass; 1996.

[10] Dorland's Illustrated Medical Dictionary, 29th edn. Philadelphia: W. B. Saunders; 2000.

[11] Speroff T, O'Connor GT. Study designs for PDSA quality improvement research. Qual Manag Health Care 2004;13: 17–32.

[12] Speroff T, James BC, Nelson EC, Headrick LA, Brommels M. Guidelines for appraisal and publication of PDSA quality improvement. Qual Manag Health Care 2004;13:33–9.

[13] Berwick DM. Broadening the view of evidence-based medicine. Qual Saf Health Care 2005;14:315–6.

[14] Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149–58.

[15] Rizzi DA. Medical prognosis – some fundamentals. Theor Med 1993;14:365–75.

[16] Wedel A, Carlsson GE. A four-year follow-up, by means of a questionnaire, of patients with functional disturbances of the masticatory system. J Oral Rehabil 1986;13:105–13.

[17] Forssell H. Mandibular dysfunction and headache [Thesis]. Turku: University of Turku, Finland; 1985.

[18] Tegelberg Å. Temporomandibular joint involvement in rheumatoid arthritis. A clinical study [Thesis]. Malmö: University of Lund; 1987.

[19] Wenneberg B. Inflammatory involvement of the temporomandibular joint. Diagnostic and therapeutic aspects and a study of individuals with ankylosing spondylitis [Thesis]. Göteborg and Malmö: Universities of Lund and Göteborg; 1983.

[20] Könönen M. Craniomandibular disorders in psoriasis. Community Dent Oral Epidemiol 1987;15:108–12.

[21] Eriksson PO, Lindman R, Stål P, Bengtsson A. Symptoms and signs of mandibular dysfunction in primary fibromyalgia syndrome (PFS) patients. Swed Dent J 1988;12:141–9.

[22] Burgess J. Symptom characteristics in TMD patients reporting blunt trauma and/or whiplash injury. J Craniomandib Disord Oral Pain 1991;5:251–7.

[23] Okeson JP. Management of temporomandibular disorders and occlusion, 5th edn. St. Louis: Mosby; 2003.

[24] Clark GT. A critical evaluation of orthopedic interocclusal appliance therapy: effectiveness for specific symptoms. J Am Dent Assoc 1984;108:364–8.

[25] Ash MM, Ramfjord S. Occlusion, 4th edn. Philadelphia: W. B. Saunders; 1995.

[26] Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain 1986;27: 117–26.

[27] Payne B, Norfleet MA. Chronic pain and the family: a review. Pain 1986;26:1–22.

[28] Fishbain DA, Rosomoff HL, Cutler RB, Rosomoff RS. Secondary gain concept: a review of the scientific evidence. Clin J Pain 1995;11:6–21.

[29] Velasco-Ortega E, Monsalve-Guil L, Velasco-Ponferrada C, Medel-Soteras R, Segura-Egea JJ. Temporomandibular disorders among schizophrenic patients. A case-control study. Med Oral Patol Oral Cir Bucal 2005;10:315–22.

[30] Emshoff R. Clinical factors affecting the outcome of occlusal splint therapy of temporomandibular joint disorders. J Oral Rehabil 2006;33:393–401.

[31] Taddey JJ. Scuba diving and TMD. Cranio 1993;11:73–4.

[32] Winocur E, Gavish A, Volfin G, Halachmi M, Gazit E. Oral motor parafunctions among heavy drug addicts and their effects on signs and symptoms of temporomandibular disorders. J Orofac Pain 2001;15:56–63.

[33] Barthelemy I, Karanas Y, Sannajust JP, Emering C, Mondie JM. Gout of the temporomandibular joint: pitfalls in diagnosis. J Craniomaxillofac Surg 2001;29:307–10.

[34] Heir GM, Fein LA. Lyme disease: considerations for dentistry. J Orofac Pain 1996;10:74–86.

[35] Sinclair NA. Unusual mandibular fracture. Aust Dent J 1976;21:517–9.

[36] Nicholls DW, Lowe N. Use of a modified distraction appliance to treat radiation-induced trismus. J Oral Maxillofac Surg 2003;61:972–4.

[37] Okeson JP. Bell's orofacial pains, 5th edn. Chicago: Quintessence; 1995.

[38] Grinin VM, Nasonova VA, Maksimovskii Iu M, Ternovoi SK, Sinitsyn VE, Smirnov AV. The differential diagnosis of avascular necrosis of the temporomandibular joint in systemic lupus erythematosus]. Stomatologiia 2000;79:23–5.

[39] Konttinen YT, Tuominen TS, Piirainen HI, Kononen MH, Wolf JE, Hietanen JH, et al. Signs and symptoms in the masticatory system in ten patients with mixed connective tissue disease. Scand J Rheumatol 1990;19:363–73.

[40] Hagberg C, Korpe L, Berglund B. Temporomandibular joint problems and self-registration of mandibular opening capacity among adults with Ehlers-Danlos syndrome. A questionnaire study. Orthod Craniofac Res 2004;7:40–6.

[41] Hansson PT, Lacerenza M, Marchettini P. Aspects of clinical and experimental neuropathic pain: the clinical perspective. In: Hansson PT, Fields HL, Hill RG, Marchettini P, editors. Neuropathic pain: pathophysiology and treatment. Seattle: IASP Press; 2001. p. 1–18.

[42] Elfving L, Helkimo M, Magnusson T. Prevalence of different temporomandibular joint sounds, with emphasis on disc-displacement, in patients with temporomandibular disorders and controls. Swed Dent J 2002;26:9–19.

[43] Siegel S. Nonparametric statistics. New York: McGraw-Hill; 1956.

[44] Anastassaki A, Magnusson T. Patients referred to a specialist clinic because of suspected temporomandibular disorders: a survey of 3194 patients in respect of diagnoses, treatments, and treatment outcome. Acta Odontol Scand 2004;62:183–92.

[45] Scholte AM, Steenks MH, Bosman F. Characteristics and treatment outcome of diagnostic subgroups of CMD patients: retrospective study. Community Dent Oral Epidemiol 1993;21:215–20.

[46] Grossi ML, Goldberg MB, Locker D, Tenenbaum HC. Reduced neuropsychologic measures as predictors of treatment outcome in patients with temporomandibular disorders. J Orofac Pain 2001;15:329–39.

[47] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. J Am Med Assoc 1995;273:408–12.

[48] Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. J Am Med Assoc 1994;271:1609–14.

[49] Gracely RH. Charisma and the art of healing: can non-specific factors be enough? In: Devor M, Rowbotham MC, Wiesenfeld-Hallin Z, editors. Proceedings of the 9th World Congress on Pain, Progress in Pain Research and Management. Seattle: IASP Press; 2000. p. 1045–67.

[50] Moerman D. Meaning, medicine and the "placebo effect". Cambridge: Cambridge University Press; 2002.

[51] Holm U. Empathy in the doctor–patient relationship. A theoretical and empirical analysis. Stockholm: Almqvist & Wiksell; 1985.

[52] Riley JL, Myers CD, Robinson ME, Bulcourf B, Gremillion HA. Factors predicting orofacial pain patient satisfaction with improvement. J Orofac Pain 2001;15:29–35.

[53] Fisher S. The placebo reactor: thesis, antithesis, synthesis, and hypothesis. Dis Nerv Syst 1967;28:510–5.

[54] Smith R. Quality improvement reports: a new kind of article. They should allow authors to describe improvement projects so others can learn. Br Med J 2000;321:1428.

[55] Codman EA. The product of a hospital. 1914. Arch Pathol Lab Med 1990;114:1106–11.