Do orthodontic research outcomes reflect patient values? A systematic review of randomized controlled trials involving children

Aliki Tsichlaki[a] **and Kevin O'Brien**[b]
Manchester, United Kingdom

Introduction: The selection of appropriate outcomes is a fundamental part of the design of clinical trials. Orthodontic treatment aims to improve a person's dentofacial appearance, and research outcomes should therefore reflect the perspectives of both clinicians and patients. In this study, we aimed to identify which outcomes were measured in recent orthodontic trials and to explore whether any relevant outcome domains were underrepresented. **Methods:** Five electronic databases were searched to identify all randomized controlled trials of orthodontic treatment interventions in children published in the last 5 years. Abstracts and eligible full-text articles were screened independently and in duplicate by 2 reviewers. Outcome measures were identified and categorized into 6 predetermined outcome domains. **Results:** The search identified 650 abstracts, of which 244 eligible articles were retrieved in full. One hundred thirty-three studies met the inclusion criteria and were included. Morphologic features of malocclusion were measured in 84 studies (63%); health resource utilization in 43 (32%); adverse effects of orthodontic treatment in 43 (32%); quality of life in 12 (9%); functional status in 10 (8%); and physical consequences of malocclusion in 3 (2%). There was no consistency in the outcomes selected among the trials to measure these domains. **Conclusions:** Most of the outcomes used in orthodontic research are concerned with measuring morphologic changes of treatment and do not reflect patient perspectives. Five of the 6 domains were infrequently evaluated, and outcomes were heterogenous. A core set of outcomes for clinical trials of orthodontic treatment interventions would help to overcome these issues. (Am J Orthod Dentofacial Orthop 2014;146:279-85)

In this article, we intend to outline a study in which we evaluated whether outcome measures that have been used in orthodontic trials since 2008 are relevant to patients.

First, when considering the factors that should influence the selection of outcome measures for orthodontic research, a useful starting point is to consider the nature of malocclusion. Malocclusion is not a disease but, rather, a variation from an accepted societal norm that can lead to functional difficulties or concerns about dentofacial appearance for a patient.[1] As a result, malocclusion falls under the World Health Organization's framework of functioning, disability, and health, which considers the psychologic and sociologic in addition to the purely biologic aspects of disability.[2] Therefore, it can be suggested that malocclusion might be a chronic disability that is amenable to treatment that can render a patient back to a state of oral health.

There have been many definitions of oral health; arguably, the most accepted is that proposed by Dolan,[3] as "a comfortable and functional dentition which allows individuals to continue in their desired social role." Importantly, this takes into account the social and functional elements as fundamental aspects. It is implied that when we study the treatment of malocclusion, within the broader context of oral health, the measurement of perceptions and behaviors is as essential as the measurement of the "disability" itself. Hence, research outcomes should reflect this.

From the Department of Orthodontics, School of Dentistry, Manchester, United Kingdom.
[a]National Institute for Health Research academic clinical fellow and honorary specialist registrar in Orthodontics.
[b]Professor in Orthodontics.
Aliki Tsichlaki's post is funded by the NIHR UK.
All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported.
Address correspondence to: Aliki Tsichlaki, School of Dentistry, Coupland St, Manchester M15 6FH, United Kingdom; e-mail, aliki.tsichlaki@googlemail.com.
Submitted, February 2014; revised and accepted, May 2014.
0889-5406/$36.00

http://dx.doi.org/10.1016/j.ajodo.2014.05.022

Although the adoption of randomized clinical trial (RCT) methodology in orthodontic research is increasing,[4] it has been suggested that the reported outcomes appear to be mostly relevant to clinicians and not to our patients.[5] This is important because the relevance of a study is derived from the outcomes it reports, so concentrating on measures that are important only to clinicians might fail to consider pertinent issues. This has been succinctly stated by Sinha et al,[6] who evaluated the outcomes used in clinical trials for the treatment of childhood asthma and reported that "the selection of inappropriate outcomes can lead to wasted resources or misleading information that overestimates, underestimates, or completely misses the potential benefits of an intervention." The issues arising from poor outcome measurement and reporting in studies are therefore multifaceted. First, the time and resources invested in the research are wasted, with approximately 40% to 89% of published trials not replicable due to poor descriptions of their interventions and outcomes.[7] Second, valuable information regarding the effectiveness of an intervention can be overlooked by neglecting to measure outcomes important to patients. As a result, including patient values is at the core of evidence-based medicine, and integration of these values with clinical research evidence is necessary to enable decision making.[8]

Another frequently encountered problem is the difficulty in combining the results of trials into systematic reviews, because the selection of outcome measures in the trials has been inconsistent. This outcome heterogeneity was clearly demonstrated in the Cochrane review on the treatment of increased overjets,[9] where the included studies all used different cephalometric analyses to answer the same questions.[5]

Recently, there has been extensive work in the initial stages of development of an agreed standardized set of outcomes for health care. These are termed core outcome sets (COS). It is suggested that "these outcomes should be measured as a minimum in trials assessing effectiveness of interventions, and would help eliminate issues relating to outcome heterogeneity and outcome reporting bias while ensuring that the perspectives of both clinicians and patients are measured, thus enhancing the value of RCTs and systematic reviews."[10] At present, COS development in dentistry and in certain fields of medicine is still in its infancy. However, in others and most notably in rheumatology, such work has advanced greatly through the work by the OMERACT initiative.[11] This international collaboration used standardized consensus techniques to develop COS in clinical trials of rheumatology to reduce the discrepancies and inconsistencies in outcome measurement that mainly existed between United States and Europe.[6] Currently, there are no COS available for orthodontic trials.

The aims of this study were to (1) identify the outcomes measured in recent orthodontic trials, (2) classify them into various outcome domains, (3) consider whether they were relevant to patients, and (4) suggest whether any relevant outcome domains were underrepresented.

MATERIAL AND METHODS

The outcomes used in previous orthodontic research were evaluated by conducting a systematic review of the literature. Studies were considered eligible if they met the following inclusion criteria.

1. Study design: prospective RCTs. All parallel-group RCTs, including those of crossover or cluster design, were considered eligible for inclusion.
2. Participants: children up to age 16 years at the start of treatment, including those with nonsyndromic clefts, were eligible.
3. Interventions: any orthodontic treatment intervention with no restrictions placed on the control groups was eligible.
4. Outcome measures: all reported outcomes (primary and secondary) were to be identified.
5. Exclusions: studies involving solely adults, patients with syndromic conditions, surgical or pharmacologic interventions, and purely laboratory investigations were excluded.

The electronic search strategy was designed to include the relevant literature, published from January 1, 2008, through December 31, 2012. The following electronic databases were searched: MEDLINE via Ovid, EMBASE via Ovid, the Cumulative Index to Nursing and Allied Health Literature (CINAHL) via EBSCO, and the Cochrane Central Register of Controlled Trials (CENTRAL, Cochrane Library). The search strategy was informed by the identified PICO concepts and inclusion criteria, and was tailored to each database to ensure appropriate use of search terms and limits. Controlled vocabulary using appropriate subject headings (MeSH terms), as well as free text search terms, were used as necessary, before the identified terms and concepts were grouped together with Boolean operators. No language restrictions were applied.

In addition, the reference lists and trials identified in recently published Cochrane systematic reviews were crosschecked to ensure that no relevant studies were missed from the electronic search.

The abstracts of all studies identified by the searches were assessed independently and in duplicate by 2 reviewers (A.T. and K.O'B.). Full-text reports of studies

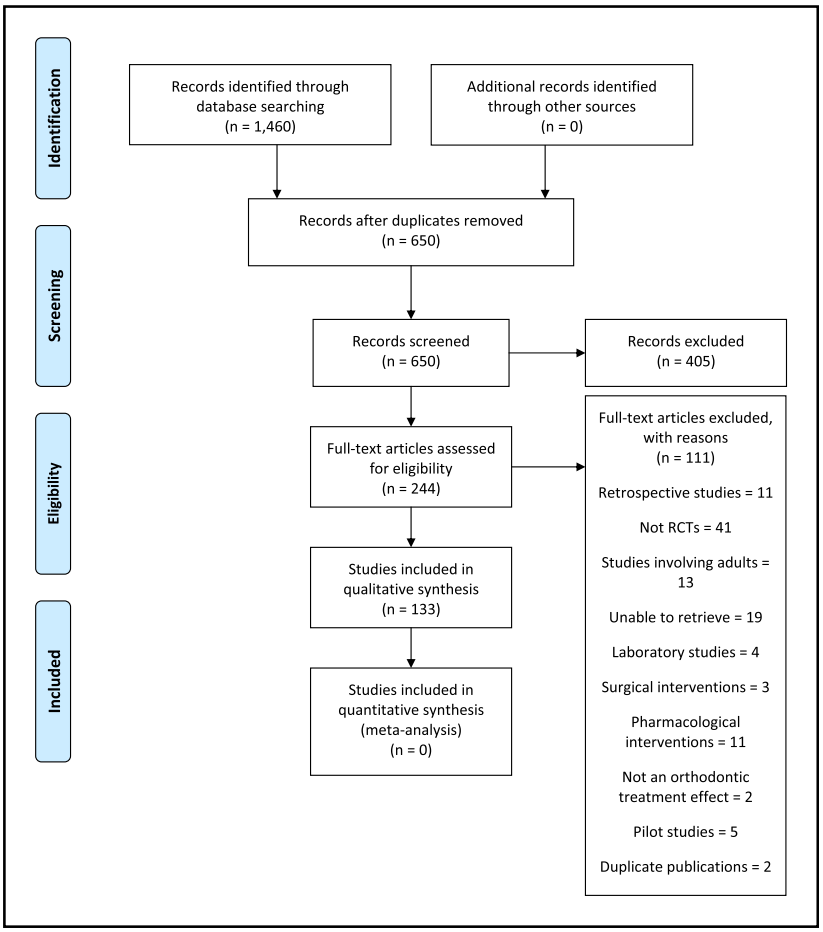

Fig 1. PRISMA flow diagram.

that appeared to meet the inclusion criteria and of studies for which there was insufficient information in the title or abstract to make a clear decision were obtained. The 2 reviewers independently assessed the full-text articles, and any disagreement regarding final inclusion was resolved by discussion until a consensus was reached.

The primary and any secondary outcomes were identified from the information stated by the authors. If this was not clear, the primary outcome was inferred from the aim of the study and any sample size calculation. Any subsequent outcomes reported in the results were also identified and recorded in a previously piloted data extraction spreadsheet as secondary outcomes. In the event of uncertainty as to which outcomes constituted the primary and secondary outcomes, all were recorded as primary outcomes, and a note of this was made in the data extraction sheet.

The reviewers independently categorized the outcomes into the following domains using the method adopted by Sinha et al[12]: (1) disease activity (morphologic features or changes of malocclusion), (2) physical consequence of malocclusion, (3) functional status, (4) social outcomes and quality of life, (5) health service resource utilization, and (6) adverse effects of treatment.

For the purpose of this review, there was no synthesis of outcome data; therefore, an overall methodologic quality appraisal of the studies was not necessary because the outcomes might not be reflected by the quality of the trials.

RESULTS

We identified 650 articles through the searches, after removing duplicates. After screening, we considered that 244 abstracts were eligible for inclusion, and full-text articles were sought. After careful examination of the remaining full texts, 133 studies met the inclusion criteria and were included in the review. The PRISMA flow diagram is shown in Figure 1.

Table I. Frequency with which outcome domains and outcomes were measured in studies (N = 133)

Domain	*Subdomain*	*Outcome*	*Number (%) of studies that measured the outcome reported as any outcome*	*Number (%) of studies that measured the outcome reported as the primary outcome*
Disease activity (features of malocclusion)	Study cast measurements	Tooth movement (mm and °)	17 (13)	7 (2)
		Dental alignment (Little's index/HLD index)	12 (9)	12 (9)
		Arch-width changes	15 (11)	9 (7)
		Occlusal outcome (PAR)	12 (9)	8 (6)
		Occlusal outcome (ICON)	2 (2)	0 (0)
	Bite registrations	Areas of contact and near contact	2 (2)	2 (2)
	Cephalometric measurements	Dental changes (mm and °)	33 (25)	25 (19)
		Skeletal changes (mm and °)	24 (18)	15 (11)
		Soft-tissue changes (mm and °)	9 (7)	8 (6)
	DPT measurements	Dentoalveolar changes	3 (2)	2 (2)
	CBCT/CT measurements	Dentoalveolar changes	5 (4)	5 (4)
	Ultrasound measurements	Skeletal (condylar) changes	1 (1)	1 (1)
	Clinical (intraoral) findings	Eruption of teeth	4 (3)	3 (2)
		Periodontal condition (BPE/GCF volume)	2 (2)	1 (1)
		AC of IOTN	1 (1)	0 (0)
		Dental alignment (Little's index)	1 (1)	0 (0)
	Notes/records review	Dental changes (overjet/crossbite correction)	3 (2)	2 (2)
Adverse effects of treatment	Radiographic findings (CT, PA)	Root resorption	5 (4)	4 (3)
	Laboratory findings	Plaque/bacterial accumulation	4 (3)	4 (3)
	Clinical findings	Plaque accumulation/periodontal condition	5 (4)	3 (2)
		Demineralization	4 (3)	2 (2)
		Complications	4 (3)	0 (0)
	Questionnaire	Pain/discomfort	22 (17)	11 (8)
		Patient expectations of treatment	5 (4)	3 (2)
Health resources utilization	Clinical findings	Attachment/appliance failure	22 (17)	16 (12)
		Treatment duration	19 (14)	10 (8)
		Cost analysis	2 (2)	0 (0)
		Patient motivation and/or compliance	2 (2)	1 (1)
	Questionnaire	Patient motivation and/or compliance	2 (2)	0 (0)
	Questionnaire	Clinician's preference	2 (2)	1 (1)
Quality of life	Questionnaire	Patient perceptions of occlusion/impact of malocclusion	5 (4)	0 (0)
		Patent acceptability of treatment	3 (2)	3 (2)
		Patient anxiety	2 (2)	1 (1)
	Interview	Patient perceptions of occlusion	1 (1)	0 (0)
Functional status	Clinical findings	Intraoral examination (TMJ)	3 (2)	1 (1)
	Self-reported questionnaire	TMJ/jaw impairment	2 (2)	1 (1)
	CT measurements	Condylar position	1 (1)	1 (1)
	Rastereography	Back/spine position	1 (1)	1 (1)
	Inclinometer	Head posture	1 (1)	1 (1)
	EMG	Muscle activity	2 (2)	2 (2)
Physical consequence of malocclusion	Clinical findings	Incisor injury	2 (2)	1 (1)
		Relapse	2 (2)	1 (1)

HLD, Handicapping Labio-Lingual Deviation Index; *PAR*, peer assessment rating; *ICON*, Index of Complexity, Outcome, and Need; *DPT*, dental panoramic tomogram; *CT*, computed tomography; *CBCT*, cone-beam CT; *BPE*, basic periodontal examination; *GCF*, gingival crevicular fluid; *AC*, aesthetic component; *PA*, peri-apical; *TMJ*, temporomandibular joint; *EMG*, electromyography.

When we found publications derived from the same trial but involving different outcome measures or different follow-up periods, we considered these to be separate studies. We did not find any RCTs on the treatment of children with clefts published during this period.

Most studies were carried out in single centers (n = 104, 78%), of which 23 (17%) used a split-mouth design. Twenty-one studies (16%) were multicentered, whereas for 8 (6%) the number of centers involved was unclear.

A sample size calculation was reported in half (n = 65, 49%) of the studies. In 5 studies (4%), it was not clear whether a sample size calculation had been performed. In 6 (5%) of those that reported a sample size calculation, the primary outcome on which this was based was not clearly stated. In 2 studies (2%), the outcome used for the sample size calculation was different from the authors' stated primary outcome. In these cases, both outcomes were noted as primary outcomes during data extraction.

Table I includes data on the identified outcomes and domains. This shows a wide variety of outcomes reported across the studies, even within the same domains.

When outcomes were grouped into the 6 domains, we found that most of the outcome domains were underrepresented in this cohort of studies, with the exception of the domain for disease activity, as shown in Table II. In this domain, which essentially covered outcome measures relating to features of malocclusion, the outcomes were varied, and many forms of measurement had been used (Table I).

Figure 2 illustrates the changes in outcome domain reporting over time. This shows that disease activity was the highest represented domain for the 5-year period. Importantly, all the other domains were represented in less than 50% of the studies, with health resource utilization showing the largest fluctuation. This change might have occurred because investigators are becoming more aware of collecting information on the process and cost of treatment, but this is only conjecture. Nevertheless, there has been a tendency to increase the measurements of adverse effects of treatment and quality of life over the past few years.

Table II. Outcome domain representation across studies (N = 133)

Outcome domain	*Number of studies measuring outcome domain (%)*
Disease activity (morphologic changes of malocclusion)	84 (63)
Health service resource utilization	43 (32)
Adverse effects of treatment	43 (32)
Quality of life/social outcomes	12 (9)
Functional status	10 (8)
Physical consequence of malocclusion	3 (2)

DISCUSSION

This study showed that most (63%) of the recently published RCTs of orthodontic treatment interventions were focused on measuring morphologic changes that result from treatment. Understandably, this is both necessary and important when assessing effectiveness of care, as clinicians need to know whether an intervention has worked. Unfortunately, other outcomes of comparable value, such as cost-benefit analysis, adverse effects of treatment, patient perceptions, and impact of and compliance with treatment, have remained largely unexplored. These outcomes are relevant to patients and provide essential information when operators and patients make shared decisions on care.

Our findings are comparable those of Sinha et al,[12] who conducted a systematic review of 159 trials to assess which outcomes had been measured in RCTs in children with asthma. They found that short-term disease activity was the most frequently measured outcome domain, with quality of life, functional status, and physical consequences of the disease largely underrepresented.

It was noteworthy that these 3 domains were measured infrequently in studies of orthodontic treatment. This is disappointing, since evaluation of the quality of life outcomes could provide insight in and articulate more clearly the relationship between disease, risks, prevention, and treatment with functional oral health and well-being.[13] For example, in the study by Kumar and Bansal,[14] which evaluated the effectiveness of Essix and Begg retainers, the authors also sought the perceptions of patients regarding their acceptability in daily life. Any differences in the perceived acceptability of the retainers would strongly suggest a superiority of one intervention over the other, even if both were effective at maintaining dental alignment.

Previous studies that have evaluated perceptions of orthodontic treatment benefits found that one purpose of seeking orthodontic treatment was to improve the patient's self-esteem and esthetics.[1,15,16] Yet, we found that patient perceptions, and self-esteem in particular, were rarely measured. Arguably, it is often difficult to ascertain whether changes in such psychosocial measures occur as a result of treatment or developmental changes, and perhaps the difficulty in accurately measuring these means that they are often omitted in studies. However, advances in the development and validation of scales and questionnaires to measure patient-reported outcomes should encourage the use of such outcome measures.[17] Indeed, a few such tools have

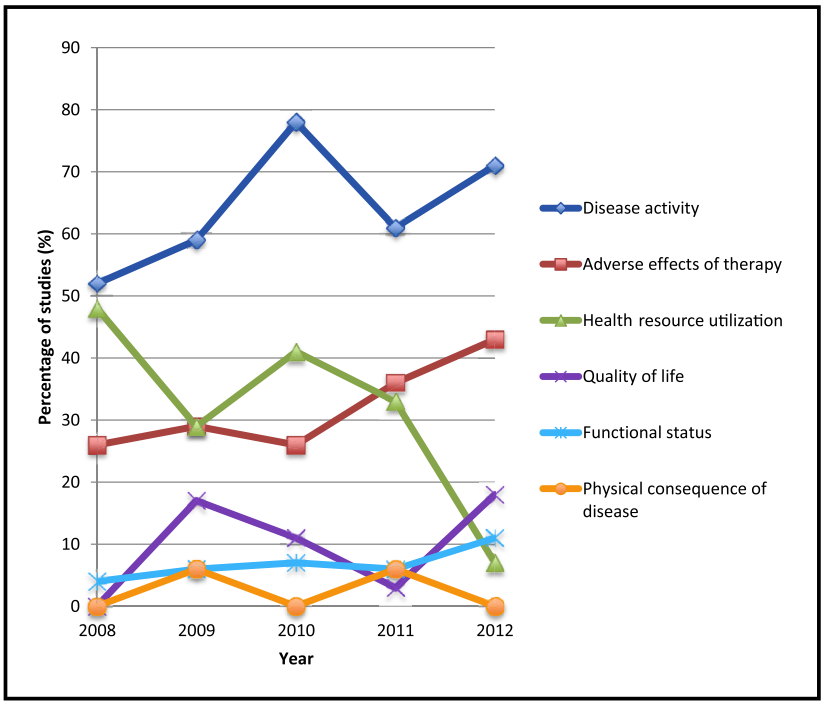

Fig 2. Trends in outcome domain representation across the 5-year period.

already been reported in the literature and have been shown to be valid and effective for use in orthodontic patients.[15,18,19] A recent systematic review, however, by Liu et al,[20] which evaluated evidence from studies on malocclusion and quality of life, identified only 23 studies that used such measures out of 143 eligible articles. This figure is slightly higher but comparable to our findings of 12 studies with quality of life measures. Nevertheless, this trend seems to be gradually increasing in orthodontic research.

The use of standardized tools would also help to overcome issues associated with studies of similar topics employing different outcomes to answer similar questions. In our review, there were some similarities between RCTs in the outcomes used to measure certain domains: for instance, the health resources utilization domain mostly comprising outcomes measuring bracket/band failure and treatment duration. In others, however, it was clear that the outcomes used were very diverse. Moreover, because there is currently no consensus on what should constitute the primary outcome of a study evaluating a particular intervention, selection of these was not homogenous among studies of similar context. This is problematic, especially when attempting to synthesize such heterogeneous data to make evidence-based decisions.

A few initiatives have now been developed in order to identify which outcomes to measure in clinical trials of various conditions. The core outcome measurement in effectiveness trials (COMET) initiative[21] is an example of the extensive work that is already being undertaken in developing universally agreed outcome sets across medical and dental fields, since there is increasing acceptance that standardizing outcomes for measurement in clinical trials is urgently needed.[22] COS development needs to involve appropriate consensus methods and key stakeholders including researchers, clinicians, patients and the public, policymakers, and public health professionals. Proposed methods for COS development were adopted by Harman et al[23] in the development of COS for use in trials of treatment of otitis media with effusion in children with cleft lip and palate and included a 3-stage Delphi process to gather opinions of health care professionals, followed by qualitative interviews with patients and parents, before making decisions during a final consensus meeting. It is clear that investigators need to explore the development of COS for orthodontic research using a similar process.

We clearly defined the inclusion criteria for the articles in this study, and some points require clarification. We based the inclusion criteria on the premise that most orthodontic treatment is carried out on children; hence, studies solely involving adults were excluded. Nevertheless, we did include studies that investigated interventions in both adults and children (eg, those with a

participant age range of 12-33 years). Additionally, in order to ensure that the review was as comprehensive as possible, trials involving orthodontic treatment for children with nonsyndromic clefts were also included. A 5-year time period was chosen so that a significant number of publications relating to orthodontic interventions would be captured. Finally, we only included prospective RCTs because the nature of retrospective studies means that some outcomes that are relevant to patients, such as sociopsychologic measures, could not be collected, and this would have biased our study.

CONCLUSIONS

Orthodontic research has advanced greatly over the years; however, it seems that a substantial amount of current research is epitomized by outcomes used to measure morphologic changes of treatment, whereas those relevant to issues important to patients are largely overlooked. Further research to determine which are the most important and relevant outcomes to measure in trials of orthodontic treatment interventions in children, taking into account the perspectives of key stakeholders including patients and clinicians, is therefore necessary. This should ideally be directed toward the development of an agreed, standardized COS for orthodontic trials to overcome issues relating to heterogeneity and underrepresentation of outcomes used to measure the outcome domains. This would consequently enable transparent interpretation and synthesis of the evidence and, ultimately, produce more meaningful research.

REFERENCES

1. Shaw WC, O'Brien KD, Richmond S, Brook P. Quality control in orthodontics: risk/benefit considerations. Br Dent J 1991;170:33-7.
2. World Health Organization. International classification of functioning, disability and health; 2000. Available at: http://www.who.int/classifications/icf/en/. Accessed May 25, 2013.
3. Dolan TA. Identification of appropriate outcomes for an aging population. Spec Care Dentist 1993;13:35-9.
4. Harrison JE. Clinical trials in orthodontics I: demographic details of clinical trials published in three orthodontic journals between 1989 and 1998. J Orthod 2003;30:25-30.
5. O'Brien K. Orthodontic research: if we know the answers, are we asking the right questions? Semin Orthod 2013;19:137-41.
6. Sinha I, Jones L, Smyth RL, Williamson PR. A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. PLoS Med 2008;5:e96.
7. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet 2014;383(9913):267-76.
8. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine: how to practice and teach EBM. Edinburgh, United Kingdom: Churchill Livingston; 2000.
9. Harrison JE, O'Brien KD, Worthington HV. Orthodontic treatment for prominent upper front teeth in children. Cochrane Database Syst Rev 2007;CD003452.
10. Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. J Health Serv Res Policy 2012;17:1-2.
11. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials 2007;8:38.
12. Sinha IP, Williamson PR, Smyth RL. Outcomes in clinical trials of inhaled corticosteroids for children with asthma are narrowly focussed on short term disease activity. PLoS One 2009;4:e6276.
13. Locker D. Concepts of oral heal, disease, and the quality of life. In: Slade G, editor. Measuring oral health and quality of life. Chapel Hill, NC: University of North Carolina; 1997.
14. Kumar AG, Bansal A. Effectiveness and acceptability of Essix and Begg retainers: a prospective study. Aust Orthod J 2011;27:52-6.
15. Bennett ME, Michaels C, O'Brien K, Weyant R, Phillips C, Dryland K. Measuring beliefs about orthodontic treatment: a questionnaire approach. J Public Health Dent 1997;57:215-23.
16. Trulsson U, Strandmark M, Mohlin B, Berggren U. A qualitative study of teenagers' decisions to undergo orthodontic treatment with fixed appliance. J Orthod 2002;29:197-204.
17. Vig KW, Weyant R, O'Brien K, Bennett E. Developing outcome measures in orthodontics that reflect patient and provider values. Semin Orthod 1999;5:85-95.
18. Mandall NA, Vine S, Hulland R, Worthington HV. The impact of fixed orthodontic appliances on daily life. Community Dent Health 2006;23:69-74.
19. O'Brien K, Wright JL, Conboy F, Macfarlane T, Mandall N. The child perception questionnaire is valid for malocclusions in the United Kingdom. Am J Orthod Dentofacial Orthop 2006;129:536-40.
20. Liu Z, McGrath C, Hägg U. The impact of malocclusion/orthodontic treatment need on the quality of life. A systematic review. Angle Orthod 2009;79:585-91.
21. COMET initiative. Core outcome measures in effectiveness trials; 2011-2013. Available at: www.comet-initiative.org. Accessed June 26, 2013.
22. Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials 2007;8:39.
23. Harman NL, Bruce IA, Callery P, Tierney S, Sharif MO, O'Brien K, et al. MOMENT—Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials 2013;14:70.