

Commentary

Evaluating oral health promotion: need for quality outcome measures

Watt RG, Harnett R, Daly B, Fuller SS, Kay E, Morgan A, Munday P, Nowjack-Raymer R, Treasure ET. Evaluating oral health promotion: need for quality outcome measures. *Community Dent Oral Epidemiol* 2006; 34: 11–7.
© Blackwell Munksgaard, 2006

Abstract – Oral health promotion effectiveness reviews have identified the need to improve the quality of the evaluation of interventions. A project was undertaken to identify and assess the quality of available outcome measures. This paper describes the methodology adopted and highlights the need for further development of oral health promotion outcome measures. Initially a thorough and comprehensive search of both the published and unpublished literature was undertaken to identify potential outcome measures. A set of quality criteria was then developed and used to assess the identified measures. The search identified a total of 1202 outcome measures of which 39% ($n = 466$) were developed for use with schoolchildren. A high proportion of the identified measures were classified as health literacy and healthy lifestyle outcomes, appropriate for the evaluation of oral health education activities. Only 1% ($n = 12$) of measures identified were classified in the healthy public policy category. When reviewed against the quality criteria, 49% ($n = 594$) of the measures were considered satisfactory. The poorest performing measures were those classified as healthy lifestyle and health literacy measures in which only 33% ($n = 72$) and 41% ($n = 240$), respectively, were deemed to be of satisfactory quality. In conclusion, a significant number of oral health promotion evaluation outcome measures have been identified although their quality is highly variable. Very few high-quality outcome measures exist for use in the evaluation of oral health policy and environmental interventions. The lack of appropriate and high-quality outcome measures is hampering the development of oral health promotion.

Richard G. Watt¹, Robert Harnett¹,
Blanaid Daly², Sabrina S. Fuller³,
Elizabeth Kay⁴, Antony Morgan⁵,
Polly Munday², Ruth Nowjack-Raymer⁶
and Elizabeth T. Treasure⁷

¹Department of Epidemiology and Public Health, University College London, London, UK, ²Department of Dental Public Health, Kings College London, London, UK, ³Department of Dental Public Health, Tameside and Glossop Primary Care Trust, Manchester, UK, ⁴Department of Oral Health and Development, University of Manchester, Manchester, UK, ⁵Health Development Agency, London, UK, ⁶National Institute of Dental and Craniofacial Research, Washington DC, USA, ⁷Department of Dental Public Health, University of Wales College of Medicine, Cardiff, UK

Key words: evaluation; oral health promotion; outcome measures

Richard G. Watt, Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London WC1E 6BT, UK
Tel: 00 44 20 7679 1699
e-mail: r.watt@ucl.ac.uk

Submitted 21 September 2004;
accepted 22 June 2005

Oral health promotion seeks to achieve sustainable improvements in oral health and reduce inequalities through action directed at the underlying determinants of oral health (1, 2). An essential component of this process is multidisciplinary action which utilizes a range of complementary strategies. The evaluation of oral health promotion is important in terms of developing effective interventions, disseminating models of good practice, providing feedback to both participants and professionals, ensuring the appropriate use of scarce resources and guaranteeing ethical principles are followed (3). However, evaluation is often a neglected area of practice which faces major

challenges (4). Key difficulties include isolating the effects of complex interventions and measuring change within a practical time frame using appropriate outcomes.

In line with other fields of public health practice, a series of effectiveness reviews have been published which have assessed the value of oral health promotion activities (5–10). One of the consensus criticisms emerging from these reviews was the poor-quality evaluation of many oral health interventions. A major limitation of the evaluations undertaken was that the outcome measures used were of limited value, focused mostly on clinical or behavioural domains and were not comparable. A

key recommendation of the effectiveness reviews was the need to improve the overall standard of oral health promotion evaluation, and in particular, to develop a broader range of quality outcome measures (8, 9).

Within the field of general health promotion, a good deal of discussion and debate has focused upon the most appropriate methods and measures available for the evaluation of interventions (11, 12). Reflecting the nature of contemporary health promotion practice, recognition has been placed upon the need to measure a diverse range of evaluation outcomes (13). Nutbeam (14) has proposed a health promotion evaluation outcome model which provides a theoretical framework for evaluating a range of interventions including policy development, community action and education over an appropriate time scale. The model describes a variety of evaluation measures including *health promotion outcomes* such as policy development, levels of community support and improvements in health knowledge; 'intermediate health outcomes' such as environmental changes and alterations to lifestyle practices such as smoking, and 'health and social outcomes' such as disease markers and quality of life indicators (Fig. 1). Such an evaluation model recognizes the importance of the social determinants of health and the need for interventions to utilize a complementary range of strategies to promote sustainable health improvements and reduce inequalities (15, 16).

A considerable amount of research has been undertaken to develop and test clinical and quality-of-life outcomes (17–19). Very little work has however focused on assessing the quality of health promotion outcomes, although research has been undertaken to develop process and quality assurance indicators (20, 21).

In March 1999, a 2-year study was undertaken which aimed to identify and critically review the quality of outcome measures appropriate for the evaluation of oral health promotion interventions. Based upon the detailed findings of this study, an oral health promotion evaluation toolkit has been published which presents the various outcome measures that were identified and tested (22). This paper aims to describe the methodology employed in the study and to highlight the need for further development of oral health promotion evaluation measures.

Process of searching and reviewing evaluation outcome measures

A variety of population groups may be targeted in oral health promotion. Effectiveness reviews have shown that the majority of oral health promotion interventions target school children (5–10). Increasingly, attention is also being given to addressing the needs of pre-school children and older people (2, 8). Three target population groups were therefore selected as the focus of the search and review:

Health and social outcomes	<u>Morbidity</u> e.g. change in dmft levels	<u>Quality of life, disability</u> e.g. change in no. of episodes of toothache	<u>Equity</u> e.g. reduction in oral health inequalities
Intermediate health outcomes	<u>Healthy lifestyles</u> e.g. change in milk or water consumption at pre-school	<u>Effective dental health services</u> e.g. change in no. of fissure sealant programmes	<u>Healthy environments</u> e.g. change in no. of schools selling healthy snacks
Health promotion outcomes	<u>Health literacy</u> e.g. change in oral health knowledge and skills	<u>Social influence and action</u> e.g. change in public support for water fluoridation	<u>Healthy public policy</u> e.g. change in no. of schools with food policy
Health promotion actions	<u>Education</u> e.g. in-service training for schoolteachers on oral health issues	<u>Facilitation</u> e.g. formation of student schools nutrition action	<u>Advocacy</u> e.g. lobbying for improvements in food labelling

Fig. 1. Oral health promotion evaluation outcome model [modified from Nutbeam (14)].

- parents/carers of pre-school children;
- 12-year-old children;
- older people aged over 65 years.

A description will now be presented of each stage in the search and review process undertaken.

Stage 1: search and development of outcome evaluation measures

A comprehensive and detailed search for published and unpublished outcome evaluation measures was undertaken. First, Medline, Embase, Psychlit and a further 11 clinical and public health electronic databases were searched for any oral health promotion evaluation measures. Key words used in the search included dental health education; oral health promotion; preventive dentistry; effectiveness; evaluation; indicators; outcome measures; oral health; dental diseases; dental caries; oral hygiene; and dental injuries. All records electronically identified were scanned by title, abstract (when available) and/or keywords and the full text of all papers considered potentially relevant were obtained. Evaluation studies that had been published in English, used nonclinical outcome measures appropriate for use in self-complete instruments and were applicable for the three selected target population groups were eligible for inclusion. To capture evaluation studies that were unpublished, the membership of 15 oral health and public health professional organizations in the UK and overseas were also contacted by letter with a request made for any evaluation outcome measures used. Finally, to supplement the studies identified, a search of 30 UK and US national health and oral health surveys was undertaken to identify any other appropriate outcome measures.

The identified outcome measures were then coded, sorted and arranged into a modified version of Nutbeam's theoretical evaluation framework (14). Through this coding and sorting process it became very apparent that, for each of the three target population groups, the majority of the items located were developed for use in assessing oral health literacy and lifestyles, and very few measures had been identified in the healthy public policy and healthy environment dimensions of the framework. To address this problem, a further search was undertaken of the policy development and environmental change literature to identify measures that had been used in other areas of public health practice which could then be modified for use in oral health promotion evaluation.

Identified measures were modified and circulated to the research team for comments and revisions. On completion of the search and development stage, an extensive and varied bank of potential outcome evaluation measures were identified for quality assessment.

Stage 2: formulation of quality criteria

Before a quality assessment of the identified measures could be undertaken, a set of test criteria had to be developed that was deemed appropriate for the purpose of this study. A search of the psychological and research methods literature identified 10 potential criteria on which to judge the quality of outcome evaluation measures (17, 23–25). A consultation process with the research team was then undertaken to determine which of the criteria identified should be considered as core criteria and which should be classified as developmental. (Core criteria were defined as those concerning essential basic criteria, whereas developmental assessed qualities of the measures that could be tested further at a later stage.) Through the consultation, a consensus was reached that the following criteria should be considered core:

- *Content validity*: refers to the adequacy of the measure in assessing comprehensively the domain of interest.
- *Applicability*: assesses the appropriateness of the measure for the purpose of oral health promotion evaluation and for the three target population groups selected.
- *Efficiency*: defined here as the performance of the measure in terms of user friendliness and feasibility of administration.
- *Clarity*: refers to the use of appropriate language and terminology.
- *Sensitivity*: assesses the potential responsiveness of an outcome measure to detect change in a given attribute.

Stage 3: quality review assessment

Based upon the agreed core criteria, a 12-item assessment checklist was developed (see Appendix for a summary checklist). The checklist was initially pilot-tested and members of the research team attended a training session on its use. The training involved the practical application of the checklist with a range of different measures. The training session also provided the team members with opportunities to ask any questions about the assessment procedure. In addition, guidance

notes were prepared to assist the team in the use of the checklist. Four quality assessment exercises were carried out over a 36-week period between August 1999 to February 2000. Each of the outcome evaluation measures was randomly allocated to two members of the research team for assessment. Over the review period, each pair of reviewers assessed 24 sets of measures which contained anything between two and 38 individual items.

Kappa scores were calculated to determine the level of agreement between paired reviewers with a score of ≥ 0.4 being used as an indicator of acceptable agreement (26). Across the four assessment exercises, acceptable agreement was achieved in 76%, 84%, 92% and 91% cases respectively. In cases where an acceptable level of agreement was not reached, a third member of the team arbitrated on points of disagreement. Outcome measures that satisfied the content and applicability criteria and achieved a score of ≥ 8 of a maximum 10 in the other areas of the assessment were judged to be of acceptable quality.

Performance of identified outcome measures

The extensive search of the literature identified a bank of 1202 measures across the three different target population groups (Table 1). Thirty-nine per cent ($n = 466$) of the measures were developed for use with schoolchildren whereas 33% ($n = 394$) and 28% ($n = 342$) were deemed appropriate for use in the evaluation of interventions targeting preschool children and older adults respectively. A wide variation was evident in the numbers of measures identified in the different categories in the theoretical model used. Measures classified within the health literacy category accounted for 48% ($n = 580$) of the total measures identified in the search process. The other two most frequently identified measures were in the healthy lifestyles category with 18% ($n = 217$) and effective dental services with 13% ($n = 153$). Only 1% ($n = 12$) of measures identified were classified in the healthy public policy category.

As regards the quality assessment, 49% ($n = 594$) of the total measures assessed, satisfied the test criteria (Table 1). Only minor differences were evident across the three target groups with 51% ($n = 175$) of the measures appropriate for use with older adults deemed of satisfactory quality, whereas 50% ($n = 198$) and

47% ($n = 221$) of the measures for preschool and school children, respectively, passed the quality assessment. Across the variety of categories of measures tested, a number of differences were observed. The poorest performing measures were those classified as healthy lifestyle, where only 33% ($n = 72$) passed the quality assessment. The other categories which performed poorly were those classified as health literacy and effective dental services, of which only 41% ($n = 240$) and 52% ($n = 80$), respectively, were deemed of satisfactory quality. The best performing measures were in the healthy environments, social influence and action, and healthy public policy categories where over 90% of the measures tested passed the quality criteria.

Methodological challenges

It is widely acknowledged that the quality of oral health promotion evaluation is generally poor and in need of further development (27). There are many reasons for this including limited resources devoted to evaluation and a lack of expertise amongst practitioners. A particular problem identified with the evaluation of oral health promotion is the narrow range and poor quality of outcome measures used and the lack of a theoretical model (8, 9).

A very thorough and comprehensive search identified a large pool of potentially valuable outcome measures. The search initially focused on a very extensive range of electronic databases including clinical, behavioural sciences and health promotion subject areas. In addition, in an attempt to recover unpublished literature, contact was made with a wide selection of individuals working in both general and oral health promotion. Finally, a collection of relevant national health surveys were examined to identify any further potential evaluation outcome measures. The search strategy was both detailed and focused. As a result, a significant number of outcome measures were identified. However, it should be noted that the search was restricted to publications in English, therefore potentially a range of measures may well have been missed if these were published in non-English-speaking publications.

The criteria developed to assess the quality of the identified measures were chosen specifically to be appropriate for the assessment of health promotion outcome measures. Very little research appears to have been conducted in this particular area of

Table 1. Results of quality assessment review for each target population group

Type of measure	Preschool children			12-year-old children			Adults aged 65+ years		
	No. identified	No. meeting criteria	% meeting criteria	No. identified	No. meeting criteria	% meeting criteria	No. identified	No. meeting criteria	% meeting criteria
Morbidity	8	5	63	10	1	10	11	10	91
Quality of life/pain	36	12	33	26	15	58	33	32	97
Healthy lifestyles	113	34	30	63	24	38	41	14	34
Effective dental health services:									
Dental Health Services	48	13	27	24	8	33	17	3	18
Health Visitors	22	19	86	7	7	100	— ^a	— ^a	— ^a
Pharmacists	19	14	74	16	16	100	— ^a	— ^a	— ^a
Subtotal	89	46	52	47	31	66	17	3	18
Healthy environments	25	25	100	17	17	100	43	43	100
Health literacy									
Attitudes	16	11	69	111	35	32	85	25	29
Knowledge	59	33	56	95	39	41	31	15	48
Perceived control	30	15	50	79	41	52	74	26	35
Subtotal	105	59	56	285	115	40	190	66	35
Social influence and action									
Awareness	3	3	100	4	4	100	— ^a	— ^a	— ^a
Opinions	10	10	100	10	10	100	4	4	100
Subtotal	13	13	100	14	14	100	4	4	100
Healthy public policy									
Policy development	3	2	67	2	2	100	2	2	100
Policy implementation	2	2	100	2	2	100	1	1	100
Subtotal	5	4	80	4	4	100	3	3	100
Total	394	198	50%	466	221	47%	342	175	51%

^aNo items identified for this category.

outcome assessment (28–30). Fitzpatrick et al. (17) in a detailed assessment of patient-based outcome measures for use in clinical trials used eight quality criteria: appropriateness, reliability, validity, responsiveness, precision, interpretability, acceptability and feasibility. More recently, an international scientific advisory group recommended a set of quality criteria for assessing health status and quality-of-life measures (18). These criteria included conceptual and measurement model, reliability, validity, responsiveness, interpretability, respondent and administrative burden, alternative modes of administration and cultural and language adaptations. Compared with clinical and quality-of-life outcomes, very little developmental work has been undertaken with oral health promotion outcome measures (27). A set of core criteria which included content validity, applicability, efficiency, clarity and sensitivity were selected by this research team for the initial quality assessment of identified outcome measures. A further set of developmental criteria were also identified which included concurrent validity, test-retest reliability and discriminatory power. These developmental criteria were subsequently used to field test the outcome measures that had passed the initial

quality assessment (22). Overall, a detailed methodology has been developed to search and assess the quality of oral health promotion evaluation measures. However, certain limitations are evident in the approach adopted. The search was restricted to publications in English. It is therefore likely that some high-quality studies published in other languages were missed. Although the criteria selected to review the outcomes was specifically developed for the purpose of assessing oral health promotion evaluation outcomes, other researchers may have selected different quality criteria.

Moving the evaluation agenda forwards

The search identified a significant number of evaluation outcome measures with the highest percentage being designed for use with school-children. This reflects the emphasis that has been traditionally placed on the school as a setting for oral health interventions (1, 2). Although every effort was made to identify a diverse mix of different types of measures, the majority of the items found were measures of health literacy and

healthy lifestyles, applicable for use in the evaluation of educational interventions. This confirms findings from effectiveness reviews which highlighted the limited range of oral health policy and environmental interventions in the oral health promotion literature (5–10). It is also interesting to note that most investigators had developed their own evaluation measures and rarely used those developed by other researchers and practitioners (9). In part, this may explain the large number of measures identified in broadly similar areas. Just under 50% of the identified outcome measures satisfied the quality assessment. The poorest performing measures were in the healthy lifestyle and health literacy categories. This may be due to the lack of a contemporary theoretical basis for many dental health education interventions (8). Within the field of general health promotion, the limited availability of high-quality outcome measures appropriate for use in the evaluation of population and policy interventions has been highlighted (11, 12, 14, 30). It appears that in oral health promotion a similar problem exists.

Conclusion

Evaluation of oral health promotion is a complex and difficult task which has been under-funded and generally neglected. Oral health practitioners have often been given very limited support or training in evaluation methodology. A key element of evaluation is the use of appropriate study designs and outcome measures (12, 14, 27). We have shown that although a significant number of evaluation outcome measures exist, their quality is highly variable and few high-quality measures exist for use in the assessment of policy and environmental change. With the change in emphasis from oral health education towards a broader oral health promotion approach now being widely advocated (31), further research to develop and test outcome measures for use in the evaluation of oral health policy and environmental action, particularly with a focus on tackling inequalities is urgently required.

Acknowledgements

The authors would like to thank all the individuals who assisted in this study. The study was funded by the UK Department of Health, through the Primary Dental Care R & D Programme.

References

1. Schou L, Locker D. Principles of oral health promotion. In: Pine C, editor. Community oral health. Oxford: Wright; 1997. p. 177–87.
2. Sheiham A, Watt R. Oral health promotion and policy. In: Murray J, Nunn J, Steele J, editors. Prevention of oral disease. Oxford: Oxford University Press; 2003. p. 241–58.
3. Blinkhorn A. Evaluation and planning of oral health promotion programmes. In: Schou L, Blinkhorn A, editors. Oral health promotion. Oxford: Oxford University Press; 1993. p. 249–70.
4. Downie R, Fyfe C, Tannahill A. Health promotion: models and values. Oxford: Oxford University Press; 1990.
5. Brown L. Research in dental health education and health promotion: a review of the literature. *Health Educ Q* 1994;21:83–102.
6. Schou L, Locker D. Oral health: a review of the effectiveness of health education and health promotion. Amsterdam: Dutch Centre for Health Promotion and Health Education; 1994.
7. Kay L, Locker D. Is dental health education effective? A systematic review of current evidence. *Community Dent Oral Epidemiol* 1996;24:231–5.
8. Sprod A, Anderson R, Treasure, E. Effective oral health promotion. Literature Review. Cardiff: Health Promotion Wales; 1996.
9. Kay L, Locker D. A systematic review of the effectiveness of health promotion aimed at promoting oral health. London: Health Education Authority; 1998.
10. Department of Human Services. Promoting oral health 2000–2004: strategic directions and framework for action. Melbourne: Health Development Section; 1999.
11. Hepworth J. Evaluation in health outcomes research: linking theories, methodologies and practice in health promotion. *Health Promotion Int* 1997; 12:233–9.
12. World Health Organisation. Evaluation in health promotion. Principles and perspectives. Copenhagen: World Health Organisation; 2001.
13. World Health Organisation. Health promotion evaluation: recommendations to policy makers. Copenhagen: World Health Organisation; 1998.
14. Nutbeam D. Evaluating health promotion – progress, problems and solutions. *Health Promotion Int* 1998; 13:27–44.
15. World Health Organisation. Global strategy for the prevention and control of noncommunicable diseases. Geneva: World Health Organisation; 2000.
16. Smedley B, Syme L. Promoting health. Intervention strategies from social and behavioural research. Washington DC: Institute of Medicine; 2000.
17. Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient based outcome measures for use in clinical trials. *Health Technology Assessment* 1998;2:1–74.
18. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Quality of Life Research* 2002;11:193–205.

19. Slade G. Measuring oral health and quality of life. Chapel Hill, NC: University of North Carolina; 1997.
20. Clark J, Maben J. Health promotion in primary health care nursing: the development of quality indicators. *Health Educ J* 1999;58:99–119.
21. Ader M, Berensson K, Carlsson P, Granath M, Urwitz V. Quality indicators for health promotion programmes. *Health Promotion International* 2001; 16:187–95.
22. Watt RG, Harnett R, Daly B, Fuller SS, Kay E, Morgan A et al. Oral health promotion evaluation toolkit. London: Stephen Hancocks Publishing; 2004.
23. Oppenheim A. Questionnaire design, interviewing and attitude measurement. London: Pinter Publishers Ltd; 1992.
24. Proctor M. Measuring attitudes. In: Gilbert N, editor. *Researching social life*. London: Sage Publications Ltd; 1993.
25. De Vaus D. *Surveys in social research*. London: UCL Press; 1996.
26. Altman D. *Practical statistics for medical research*. London: Chapman Hall; 1991.
27. Watt RG, Fuller SS, Harnett R, Treasure ET, Stillman-Lowe C. Oral health promotion evaluation – time for development. *Community Dent Oral Epidemiol* 2001;29:161–6.
28. Macdonald G. Quality indicators and health promotion effectiveness. *Promotion Educ* 1997;4:10–5.
29. Raphael D, Brown I, Renwick R, Rootman I. *Quality of life indicators and health: current status and emerging conceptions*. Toronto: University of Toronto; 1996.
30. St Leger L. Health promotion indicators. Coming out of the maze with a purpose. *Health Promotion Int* 1999;14:193–5.
31. World Health Organisation. *Oral Health Report 2003*. Geneva: World Health Organisation; 2004.

Appendix

Summary of quality checklist

A – Content validity

Is each item in the set of outcomes relevant to the topic area under investigation?

Does each item in the set of outcome measures conform to the current scientific basis of oral health promotion?

B – Applicability

Is each item in the set of outcome measures applicable to the population group?

Is each item in the set of outcome measures an appropriate measure to evaluate oral health promotion activity?

C – Efficiency

Is each item in the set of outcome measures free from excessive jargon?

Is each item in the set of outcome measures ethically sensitive?

As a whole is the outcome measure too long for either subjects, or health professionals in primary care settings?

D – Clarity

Is each item in the set of outcome measures free of basic grammatical errors?

Is each item in the set of outcome measures free of ambiguity?

Is the arrangement of the sequencing in all of the multiple question items clear and logical?

E – Sensitivity

For each item in the set of outcome measures, are the response categories mutually exclusive?

For each item in the set of outcome measures, are the response categories in the scale exhaustive?

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.