Accuracy of Medical Subject Heading Indexing of Dental Survival Analyses

Danielle M. Layton, BDSc (Hons)(Qld), MSc (Oxon), MDSc (Hons)(Syd)[a]/
Michael Clarke, BA(Oxon), DPhil (Oxon)[b]

Purpose: To assess the Medical Subject Headings (MeSH) indexing of articles that employed time-to-event analyses to report outcomes of dental treatment in patients. ***Materials and Methods:*** Articles published in 2008 in 50 dental journals with the highest impact factors were hand searched to identify articles reporting dental treatment outcomes over time in human subjects with time-to-event statistics (included, n = 95), without time-to-event statistics (active controls, n = 91), and all other articles (passive controls, n = 6,769). The search was systematic (kappa 0.92 for screening, 0.86 for eligibility). Outcome-, statistic- and time-related MeSH were identified, and differences in allocation between groups were analyzed with chi-square and Fischer exact statistics. ***Results:*** The most frequently allocated MeSH for included and active control articles were "dental restoration failure" (77% and 52%, respectively) and "treatment outcome" (54% and 48%, respectively). Outcome MeSH was similar between these groups (86% and 77%, respectively) and significantly greater than passive controls (10%, $P < .001$). Significantly more statistical MeSH were allocated to the included articles than to the active or passive controls (67%, 15%, and 1%, respectively, $P < .001$). Sixty-nine included articles specifically used Kaplan-Meier or life table analyses, but only 42% (n = 29) were indexed as such. Significantly more time-related MeSH were allocated to the included than the active controls (92% and 79%, respectively, $P = .02$), or to the passive controls (22%, $P < .001$). ***Conclusions:*** MeSH allocation within MEDLINE to time-to-event dental articles was inaccurate and inconsistent. Statistical MeSH were omitted from 30% of the included articles and incorrectly allocated to 15% of active controls. Such errors adversely impact search accuracy. *Int J Prosthodont 2014;27:236–244. doi: 10.11607/ijp.3633*

Time-to-event and survival analyses are becoming more common in the dental literature, particularly for the reporting of outcomes of prostheses. However, the concept of "survival" in dentistry differs from that in other medical areas. The concept that an inanimate object, such as a crown, may survive is quite different from the concept that a person with a life-threatening condition may survive. It is hypothesized that this linguistic nuance will affect clinicians' ability to identify articles reporting the survival of prostheses in the dental literature.

To practice evidence-based health care, clinicians, reviewers, and other decision makers must be able to locate relevant evidence. Clinicians generally wish to identify the high-quality synthesized evidence and may seek a systematic review. On the other hand, systematic reviewers wish to identify all evidence relating to a topic to allow analysis and facilitate dissemination. Failing to identify some specific articles is unlikely to greatly hinder individual clinicians, but failing to identify those same articles for a systematic review might have important implications for the results and conclusions, including the introduction of bias. Eventually, that bias will impact individual clinicians when they rely on the aforementioned systematic review to guide clinical practice.

Articles can be identified by hand, electronic full text, and database searches. Hand searching will identify relevant articles, but not in a timely manner. Electronic full-text searches are becoming increasingly available but are not yet possible for a broad range of journals. Multiple bibliographic databases, such as MEDLINE, Embase, and The Cochrane Library index dental journals and facilitate electronic searching using text word or subject heading. This latter search method is commonly used to identify evidence.

A text word search reveals details contained in an article's title and abstract only.[1] Subject headings are assigned to articles based on the complete text and

[a]*Private Practice, Brisbane, Australia.*
[b]*Professor and Director, MRC Hub for Trials Methodology Research, Queens University, Belfast, United Kingdom.*

Correspondence to: *Dr Danielle M Layton, 217 Wickham Terrace, Brisbane, Queensland Australia, 4000.*
Email: laytonpros@dlpros.com.au

are potentially more powerful than keywords. They are embedded in the indexing of databases such as MEDLINE (MeSH, Medical Subject Headings), Embase (Emtree), and parts of The Cochrane Library (MeSH).

Early researchers using Medlars (a precursor to today's MEDLINE) encountered problems identifying relevant dental articles. They found MeSH terms were restrictive and were tailored towards medicine.[2] The MeSH library has since increased; however, it is hypothesized that differences in interpretation of dental vocabulary will continue to impact search yields. Specifically, should the definition of "survival" differ between MeSH indexers and searchers, relevant studies will remain elusive and undermine the provision of evidence-based dentistry.

Assessing the accuracy and consistency of MeSH term assignment is challenging. In 2014,[3] the list of MeSH terms included approximately 27,000 descriptors, more than 200,000 entry terms (those that aid appropriate MeSH identification), and over 200,000 supplementary concept records.[3,4] MeSH are manually allocated to articles; therefore, indexing variation is expected and disparity is not necessarily considered inaccurate. However, omission or misallocation of important terms clearly undermines search performance. Studies evaluating their use found lower consistency associated with subheadings, methodology categories (E: analytic, diagnostic, therapeutic techniques), and those categories whose definitions were less stable (F: psychiatry; N: health care).[5–7] Further errors with methodology MeSH, specifically those for controlled trials, were highlighted in the early years of the Cochrane Collaboration.[8]

This research assesses the allocation of MeSH terms within MEDLINE to articles in the dental literature that employed time-to-event analyses to report outcomes of dental treatment in patients. It is part of a larger research project investigating reporting of time-to-event analyses in dentistry.

Materials and Methods

Search Strategy

The systematic search is divided into four sections: the identification of articles, full-text screening, assessment for eligibility, and selection for inclusion. The flow of the search process is outlined in Fig 1.

The 50 dental journals with the highest impact factors in 2008 were identified from the ISI *Journal Citation Reports* in the Web of Knowledge. In total, 6,955 articles were identified. Articles that reported outcomes of dental treatments in humans over time and may have employed time-to-event analyses were then sought by full-text hand searching of each journal. No electronic search strategy was employed. All articles were published in English.

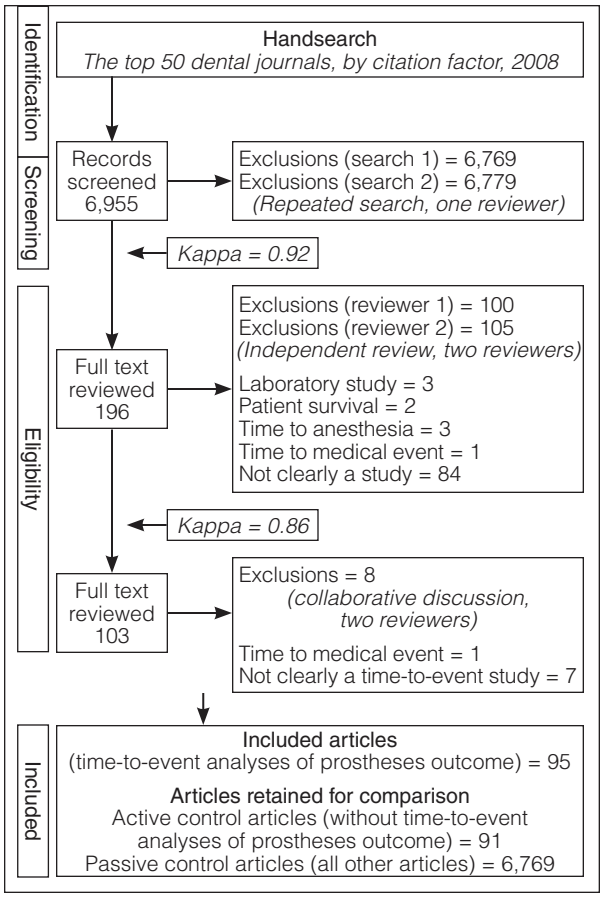

Fig 1 Flow chart of systematic search.

Article identification for full-text screening was accomplished by a single reviewer who completely reviewed the 50 journals twice over 6 months. Article screening, eligibility, and inclusion assessment were accomplished by two independent reviewers. Disagreement was resolved by discussion. The measure of agreement between the two searches and two reviewers (kappa statistic) was determined (Fig 1).

Inclusion criteria were prospective and retrospective studies that employed time-to-event statistics for reporting dental treatment outcomes over time in humans (included articles). Articles that reported such outcomes without using time-to-event statistics were retained for comparison (active controls).

Time-to-event analyses were considered to be those using Kaplan-Meier, life tables, actuarial analyses, and survival functions. Articles using Cox regression, log rank, and hazard ratios, and those reporting outcome as a rate, were retained for further screening to determine whether a time-to-event analysis had also been completed but not clearly reported.

Table 1 The Distribution of Articles Selected From the 50 Journals Studied

Impact Factor	Rank	Journals	Total	Included	Active control	Passive control
1.412	32	Acta Odontologia Scandanavia	59	1	0	58
1.314	36	American Journal of Dentistry	74	3	1	70
1.327	33	American Journal of Orthodontics and Dentofacial Research	313	1	1	311
1.649	24	Archives or Oral Biology	166	0	0	166
1.22	41	Australian Dental Journal	72	1	0	71
1.089	44	British Dental Journal	412	1	2	409
1.327	33	British Journal of Oral and Maxillofacial Surgery	229	0	0	229
2.462	8	Caries Research	60	0	0	60
2.452	9	Clinical Implant Dentistry and Related Research	33	6	3	24
2.92	6	Clinical Oral Implants Research	163	12	19	132
2.233	12	Clinical Oral Investigations	62	1	1	60
0.969	49	Community Dental Health	49	0	0	49
2.418	10	Community Dentistry and Oral Epidemiology	63	0	0	63
2.882	7	Dental Materials	231	2	0	229
1.316	35	Dental Traumatology	173	2	2	169
1.229	39	Dentomaxillofacial Radiology	79	0	1	78
1.024	46	European Journal of Dental Education	55	0	0	55
1.956	19	European Journal of Oral Science	85	0	0	85
0.975	48	European Journal of Orthodontics	94	3	0	91
1.014	47	Gerodontology	38	0	0	38
1.505	28	Implant Dentistry	58	2	2	54
2.223	13	International Endodontic Journal	145	1	0	144
1.978	17	International Journal of Oral and Maxillofacial Implants	130	15	13	102
1.444	31	International Journal of Oral and Maxillofacial Surgery	206	1	3	202
1.141	43	International Journal of Paediatric Dentistry	65	1	1	63
1.702	22	International Journal of Periodontics and Restorative Dentistry	62	1	4	57
1.227	40	International Journal of Prosthodontics	91	11	4	76
1.638	25	Journal of Adhesive Dentistry	59	5	1	53
1.726	21	Journal of the American Dental Association	172	2	1	169
3.549	1	Journal of Clinical Periodontology	66	0	1	65
1.252	38	Journal of Craniomaxillofacial Surgery	151	0	0	151
2	16	Journal of Dentistry	206	0	0	206
1.087	45	Journal of Dental Education	161	2	2	157
3.458	2	Journal of Dental Research	336	2	5	329
2.953	5	Journal of Endodontics	452	2	11	439
1.58	27	Journal of Oral and Maxillofacial Surgery	99	0	0	99
2.144	15	Journal of Oral Pathology and Medicine	121	0	0	121
1.483	30	Journal of Oral Rehabilitation	38	0	0	38
1.263	37	Journal of Orofacial Pain	97	0	0	97
1.966	18	Journal of Periodontal Research	313	9	5	299
2.192	14	Journal of Periodontology	149	3	1	145
1.215	42	Journal of Prosthetic Dentistry	41	0	0	41
0.961	50	Journal of Public Health Dentistry	249	2	0	247
1.683	23	Operative Dentistry	104	1	1	102
1.922	20	Oral Diseases	117	0	0	117
2.336	11	Oral Microbiology and Immunology	80	0	0	80
3.123	3	Oral Oncology	171	0	1	170
1.499	29	Oral Surgery, Oral Medicine, Oral Radiology and Endodontics	440	2	5	433
1.607	26	Orthodontic and Craniofacial Research	30	0	0	30
3.027	4	Periodontology 2000	36	0	0	36
			6,955	95	91	6,769

Table 2 Article Themes of Included and Active Control Articles

	Included		Active control	
	n	%	n	%
1 Implant-related Implant fixtures, temporary anchorage devices, implant prostheses	61	64.21	55	60.44
2 Tooth-related prosthodontics Fixed dental prostheses, crowns, posts, resin bonded prostheses	13	13.68	7	7.69
3 Periodontal related Papilla presence, periodontal indicies	0	0.00	2	2.20
4 Endodontic related Infection resolution, tooth vitality, pulp capping	4	4.21	7	7.69
5 Orthodontic related Bracket/retainer debonding	3	3.16	0	0
6 Tooth-related, other Fillings, tooth survival (ankylosis, resorption, extraction), transplantation, caries	14	14.74	10	10.99
7 Other TMJ fracture healing, orofacial grafts, odontogenic infection resolution	0	0	10	10.99
Total	95	100	91	100

In vitro and animal studies were excluded. Studies where outcomes were not specifically related to dental treatments (such as death, cancer, time to anesthesia) were also excluded. Demographics of included articles are outlined below.

Out of the 6,955 articles published in these 50 journals, full-text screening identified 196 articles (see Fig 1). Of those, 186 and 176 were noted, respectively, during search 1 and search 2 (kappa, 0.92). Both searches were completed by one reviewer over 6 months. Upon further review, 103 different articles were selected for possible inclusion by two independent reviewers (kappa, 0.86), and 95 were subsequently determined by discussion to meet the inclusion criteria. From the initial 196 articles, laboratory (n = 3), patient morbidity (n = 2), time-to-anesthesia (n = 3), and time-to-medical event (n = 2) articles were excluded. Articles reporting a prosthesis outcome over time without using time-to-event statistics were retained as active controls (n = 91). All other articles were retained as passive controls (n = 6,769).

Article Demographics

The distribution of articles from the 50 journals is outlined in Table 1.

The articles in the included group and active controls reported outcomes across a variety of dental treatments. The most common article theme within both the included and active controls related to implant outcomes (64% and 60%, respectively). In total, approximately 80% of the included articles (n = 74) and 70% of the active controls (n = 62) reported outcomes of either implant-related treatment or tooth-supported prostheses. The distribution of the themes of the articles is outlined in Table 2. The themes of the passive control group were not reviewed.

Data Extraction

MeSH allocated by MEDLINE (Ovid) indexers were reviewed for the included (n = 95), active control (n = 91), and passive control (n = 6,769) articles. To be clear, an electronic MeSH search was not used to identify articles. Articles were identified by a hand search, and then MeSH terms that had been allocated to the electronic record by indexers were collated. Those relating to time-to-event statistical techniques, prosthesis outcomes, or the conduct of research over time were identified and extracted manually by one reviewer from the included and active controls. In total, 17 relevant MeSH terms were identified (Table 3), and these had not been determined in advance. These MeSH terms were classified into three groups (Table 4).

Two additional MeSH terms (acturial analysis and survival) may have been applicable to this set of articles, and their definitions have been included with the identified 17 terms (Table 3). However, these two terms had not been allocated to any of the articles.

Following identification of relevant MeSH terms from the included and comparator articles by hand, the search was repeated electronically for the passive controls.

The allocation of MeSH terms was reported as a frequency per article and percentage per article (Table 5).

Table 3 MeSH Definitions from MEDLINE (Ovid)

MeSH term	Definition
Actuarial analysis *(year introduced, 1979)*	The application of probability and statistical methods to calculate the risk of occurrence of any event, such as onset of illness, recurrent disease, hospitalization, disability, or death. It may include calculation of the anticipated money costs of such events and of the premiums necessary to provide for payment of such costs.
Dental prosthesis repair *(year introduced, 1993)*	The process of reuniting or replacing a broken or worn dental prosthesis or its part.
Dental restoration failure *(year introduced, 1995)*	Inability or inadequacy of a dental restoration or prosthesis to perform as expected.
Fatal Outcome *(year introduced, 1994)*	Death resulting from the presence of a disease in an individual, as shown by a single case report or a limited number of patients. This should be differentiated from DEATH, the physiologic cessation of life and from MORTALITY, an epidemiologic or statistical concept.
Follow up studies *(year introduced, 1967)*	Studies in which individuals or populations are followed to assess the outcome of exposures, procedures, or effects of a characteristic, eg, occurrence of disease. A subgroup under longitudinal studies.
Graft survival *(year introduced, 1999)*	The survival of a graft in a host, the factors responsible for the survival, and the changes occurring within the graft during growth in the host.
Kaplan-Meiers Estimate *(year introduced, 2007)* **Kaplan-Meier Estimate** *(amended spelling, 2011)*	A nonparametric method of compiling LIFE TABLES or survival tables. It combines calculated probabilities of survival and estimates to allow for observations occurring beyond a measurement threshold, which are assumed to occur randomly. Time intervals are defined as ending each time an event occurs and are therefore unequal. (From Last, A Dictionary of Epidemiology, 1995)
Life tables *(year introduced, 1990)*	Summarizing techniques used to describe the pattern of mortality and survival in populations. These methods can be applied to the study not only of death, but also of any defined endpoint such as the onset of disease or the occurrence of disease complications.
Longitudinal studies *(year introduced, 1979)*	Studies in which variables relating to an individual or group of individuals are assessed over a period of time.
Proportional hazards model *(year introduced, 1989)*	Statistical models used in survival analysis that assert that the effect of the study factors on the hazard rate in the study population is multiplicative and does not change over time.
Prospective studies *(year introduced, 1967)*	Observation of a population for a sufficient number of persons over a sufficient number of years to generate incidence or mortality rates subsequent to the selection of the study group. A subgroup under Longitudinal studies.
Prosthesis failure *(year introduced, 1999)*	Malfunction of implantation shunts, valves, etc, and prosthesis loosening, migration, and breaking.
Retreatment *(year introduced, 1997)*	The therapy of the same disease in a patient, with the same agent or procedure repeated after initial treatment, or with an additional or alternate measure or follow-up. It does not include therapy that requires more than one administration of a therapeutic agent or regimen. Retreatment is often used with reference to a different modality when the original one was inadequate, harmful, or unsuccessful.
Retrospective studies *(year introduced, 1967)*	Studies used to test etiologic hypotheses in which inferences about an exposure to putative causal factors are derived from data relating to characteristics of persons under study or to events or experiences in their past. The essential feature is that some of the persons under study have the disease or outcome of interest and their characteristics are compared with those of unaffected persons.
Survival *(year introduced, not stated)*	Continuance of life or existence especially under adverse conditions; includes methods and philosophy of survival.
Survival Analysis *(year introduced, 1990)*	A class of statistical procedures for estimating the survival function (function of time, starting with a population 100% well at a given time and providing the percentage of the population still well at later times). The survival analysis is then used for making inferences about the effects of treatments, prognostic factors, exposures, and other covariates on the function.
Survival Rate *(year introduced, 1990)*	The proportion of survivors in a group, eg, of patients, studied and followed over a period, or the proportion of persons in a specified group alive at the beginning of a time interval who survive to the end of the interval. It is often studied using life table methods.
Time factors *(year introduced, 1999)*	Elements of limited time intervals, contributing to particular results or situations.
Treatment outcome (year introduced, 1999)	Evaluation undertaken to assess the results or consequences of management and procedures used in combating disease to determine the efficacy, effectiveness, safety, practicability, etc, of these interventions in individual cases or series.

Statistical Analysis

Differences in the allocation of MeSH terms between the included and active controls were analyzed with Pearson chi-square and Fischer exact tests (where expected cell counts were less than five). Differences between all three groups were analyzed with the Pearson chi-square statistic. Where expected cell counts were less than five, a two by two contingency table was constructed, and Fischer exact test was employed. Significance was set at $P < .05$.

Results

Distribution of MeSH

A total of 530 different MeSH terms were allocated to the included articles (n = 95). Articles had been allocated between 11 and 32 terms, with a mean of 20. In total, 2,031 MeSH were allocated to these 95 articles, with 345 (17%) collated for this research.

There were 611 different MeSH allocated to the active controls (n = 91). Articles had been allocated between 7 and 31 each, with a mean of 20. In total, 1,861 MeSH were allocated to these 91 articles, with 214 (12%) collated for this research.

The MeSH allocated to the 6,769 passive controls were also reviewed, but as there were a large number, the overall distribution and demographics of these terms were not sought.

Comparisons Between Groups

Within both the included articles and active controls, the most frequently allocated MeSH were "dental restoration failure" (77% and 52%, respectively) and "treatment outcome" (54% and 48%, respectively). There was no significant difference between the allocation of outcome MeSH between these groups (86% and 77%, respectively, chi-square 2.75, $P = .10$).

Within the passive controls, "treatment outcome" was the second most frequently used MeSH (7% of articles), while "dental restoration failure" was sixth (2.5% of articles). The allocation of outcome MeSH between these three groups was different, with significantly fewer passive controls being allocated one of these terms (86%, 77%, and 10%, respectively, chi-square 850, $P < .001$).

Significantly more included articles were allocated at least one statistical MeSH compared to the active controls (67% vs 15%, chi-square 51.58, $P < .001$). Specifically, statistical MeSH "Kaplan-Meiers estimate" and "life tables" were not allocated to any active controls, but were assigned to only 21% and 16% of the included articles, respectively. Sixty-nine included articles specifically stated they used either Kaplan-Meier or life table analyses, but only 42% of those (n = 29) were indexed as such.

Table 4 Classification of MeSH Terms

Group	Terms
1 **MeSH: Statistical** MeSH term for a statistical time-to-event technique	Kaplan-Meiers estimate Life tables Proportional hazards model Survival analysis Survival rate
2 **MeSH: Outcome** MeSH term that indicated that an outcome was studied	Dental prosthesis repair Dental restoration failure Fatal outcome Graft survival Prosthesis failure Retreatment Treatment outcome
3 **MeSH: Time** MeSH term that indicated that a study occurred over time	Follow-up studies Longitudinal studies Prospective studies Retrospective studies Time factors

Within the passive controls, 101 statistical MeSH were allocated to 89 of the 6,769 articles. These included 42 patient survival, 5 nonpatient or prosthesis time-to-events, 2 mortality analyses, and 14 review articles. The other 26 articles reported percentage outcomes but did not employ time-to-event analyses.

The allocation of statistical MeSH between these groups was different, with significantly fewer articles from the passive controls being allocated one of these terms (42% and 1%, respectively, Fischer exact test, chi-square 850, $P < .001$).

Significantly more included than active control articles were allocated at least one time-related MeSH (92% and 79%, respectively, chi-square 5.82, $P = .02$). Although these terms indicate that research occurred over time, they do not specifically indicate that time-to-event analyses were used.

Of the 17 MeSH terms, those relating to time were the most frequently allocated for the passive controls. These were assigned to 22% (n = 1,448), with "time factors" the most frequently allocated term (10%). Despite their increased use, they still remained infrequent in comparison with the included and active control articles (92% and 79%, and 22%, respectively, chi-square 414, $P < .001$).

Overall, at least one of the 17 MeSH terms was assigned to 93 included articles (98%, frequency 345), 86 active controls (95%, frequency 214), and 1,872 passive controls (28%, frequency 2,556). There was no significant difference in the allocation of the 17 terms between the included and active control articles

Table 5 Distribution of MeSH Terms Across the Included, Active Control, and Passive Control Articles

Group	Description		Included (n = 95)	Active control (n = 91)	Passive control (n = 6,769)
1	Statistical	Kaplan-Meiers estimate	20	0	25
		Life tables	15	0	1
		Proportional hazards model	6	1	13
		Survival analysis	30	12	42
		Survival rate	1	1	20
		Total frequency*	72	14	101
		Total articles†	67.37%, n = 64	15.39%, n = 14	1.32%, n = 89
		Included vs active controls: chi-square 51.58, $P < .001$‡ Comparison between all groups, chi-square 1262, $P < .001$§			
2	Outcome	Dental prosthesis repair	2	0	12
		Dental restoration failure	73	47	169
		Fatal outcome	0	1	28
		Graft survival	2	3	18
		Prosthesis failure	3	1	6
		Retreatment	4	0	46
		Treatment outcome	51	44	479
		Total frequency*	135	96	758
		Total articles†	86.32%, n = 82	76.92%, n = 70	10.37%, n = 702
		Included vs active controls: chi-square 2.75, $P = .10$‡ Comparison between all groups, chi-square 850, $P < .001$			
3	Time	Follow up studies	49	29	446
		Longitudinal studies	13	4	142
		Prospective studies	23	25	209
		Retrospective studies	37	23	282
		Time factors	16	23	619
		Total frequency*	138	104	1,697
		Total articles†	91.58%, n = 87	79.12%, n = 72	21.39%, n = 1,448
		Included vs active controls: chi-square 5.82, $P = .02$‡ Comparison between all groups, chi-square 414, $P < .001$			
Total	Group 1, 2, or 3	MeSH, as listed above	NA	NA	NA
		Total frequency*	345	214	2,556
		Total articles†	97.89%, n = 93	94.51%, n = 86	27.66%, n = 1,872
		Included vs active controls: chi-square 1.48, $P = .23$‡ Comparison between all groups, chi-square 404, $P < .001$			

* Total frequency of term use. Some articles were allocated more than one term within each group.
† Total number of articles that were allocated at least one term within a given group.
‡ Pearson chi-square test, 2 × 3 contingency table.
§ Fischer's exact test, 2 × 2 contingency table.

(chi-square 1.48, $P = .23$). Overall, significantly fewer passive controls were allocated one of these terms, in comparison with the other groups (98%, 95%, and 28%, respectively, chi-square 404, $P < .001$).

Discussion

This research has identified many challenges that impact identification of published articles reporting "survival" and "time-to-event" outcomes of dental prostheses. It was observed that the allocation of MeSH terms in Medline was inaccurate and inconsistent for the methodology of many of the articles.

These findings are based on assessment of MeSH terms allocated to articles published in 2008 in the 50 dental journals with the highest impact factors. None of these journals are published in languages other than English. The search was completed in a systematic manner by hand. For analysis, all 6,955 articles published in the cohort were allocated to one of three groups: 95 included articles, 91 active controls, and 6,769 passive controls.

The articles within these journals are the "subjects" of this research, and inclusion related to the Institute for Scientific Information (ISI) citation index has preselected articles and edited journals that could be considered "high quality." Misallocation of MeSH may be related to the indexer, but errors could equally be compounded by poor reporting quality. It is likely that articles included in this research are of higher than average quality, with the error rate therefore representing a best-case scenario.

The complexity of the English language confounds word standardization, with fields such as medicine and law choosing to adopt words from Latin, a "dead" language, in an attempt to keep language use and meaning stable. Such stable words, however, are not available for time-to-event outcomes and the associated statistics.

To help clarify definitions, a controlled vocabulary, the Medical Subject Heading system, was introduced in 1960, and these terms are manually assigned to research by indexers who read the full article.[9] These terms provide a standardized vocabulary for searchers to use, increasing their ability to identify and cross reference relevant articles. Their usefulness, however, is directly related to the accuracy of the indexing.

Errors in indexing have been identified in the past, with a particular high-profile error being the misallocation of MeSH to controlled trials, which influenced an extensive program of hand searching for such articles by the Cochrane Collaboration.[8] Also, differences in linguistic vocabulary have been shown to affect the identification of relevant articles in the dental field.[2]

For this study, relevant MeSH terms were identified as individual articles were reviewed, and it is therefore unlikely that pertinent MeSH terms were omitted from the analysis. Terms were classified into three groups: statistical, outcome, and time.

Errors in the allocation of these MeSH terms to the included articles occurred across all three MeSH groups. MEDLINE indexers map MeSH terms to articles by selecting the term from a predetermined list, and therefore spelling errors should not occur.

It is interesting to note that since collation of this data, the spelling of Kaplan-Meiers Estimate has been amended. This term was introduced in 2007, and by 2011 its spelling had changed from Kaplan-Meiers Estimate to Kaplan-Meier Estimate, and MeSH records were updated.

The MeSH term most frequently allocated to included articles related to treatment outcomes: "dental restoration failure." Similar high proportions of both the included and active control articles had been allocated at least one treatment outcome MeSH term (86% and 77%, respectively), but these terms were much less common in the passive controls (10%). It is not surprising that these terms were common for both the included and active control articles because they indicate that the article investigates a treatment outcome, and this is the reason they were selected for this study. However, as both article groups studied an outcome, it was disappointing that such MeSH terms were not allocated to each record.

Of the 17 MeSH terms, those relating to time were the most frequently allocated terms to all three article cohorts. More than 90% of the included articles were allocated at least one such term, with this high frequency reflecting the content of each article. Significantly fewer time-related MeSH terms were allocated to the active controls, but the frequency of nearly 80% remains high. Despite these high frequencies, it is concerning that all the included and active control articles were not allocated at least one MeSH term relating to research over time.

In comparison, significantly fewer of the passive controls were allocated a time-related MeSH term, with an indexing frequency of 22%. Although the frequency is low, inclusion of these terms in a search strategy would result in an additional 1,448 false positive yield, increasing the burden of identification for searchers.

The misallocation of MeSH terms was greatest for statistical themes. First, misallocations related to complete omission of relevant MeSH terms. In total, 67% of included articles were allocated at least one statistical-related MeSH. Therefore, indexers incorrectly omitted vital statistical terms when assigning MeSH to 31 articles, a third of the included cohort.

Secondly, misallocations related to inaccurate use of updated terms. MeSH terms are updated yearly, with an additional survival function term, Kaplan-Meiers estimate, entered in 2007. It is known that at least 42 of the 95 included articles specifically used the Kaplan-Meier survival method, but less than half (16) were assigned the term Kaplan-Meiers estimate. Previous survival function terms of "survival analysis," "survival rate," and "life tables" had been introduced in 1990. Of those 26 articles that were not allocated the updated term, 18 were allocated an alternative, but nearly one-fifth (8 articles) were not allocated any survival MeSH term.

Lastly, misallocation related to incorrect indexing of articles that had not performed survival analyses with such MeSH terms. Among the active controls, 14% were incorrectly allocated the terms "survival analysis" or "survival rate." These false positives represent a moderate percentage, which would lead to a relatively small inaccurate yield that would not be an excessive burden, if identified during an electronic search.

Within the passive controls, allocation of these MeSH terms was uncommon, being allocated to 89 articles (1.3%). These MeSH terms were correctly

allocated to 42 patient survival (cancer) studies, 5 non-patient or prosthesis related time-to-event articles, 2 reporting hazard or mortality rates, and 14 review articles. They were incorrectly allocated to an additional 26 articles, which reported percentage outcomes across laboratory and clinical settings or discussed percentage outcomes in a letter or item of news. These articles did not employ time-to-event analyses, and such MeSH terms should not have been assigned.

It is known that an additional 19 time-to-event articles were present within the passive control group and these should have been allocated a survival MeSH term. This did not occur. Therefore, survival function MeSH terms were also incorrectly omitted from the passive control articles.

Overall, significantly fewer passive controls were allocated at least one statistical MeSH, while they were equally frequent among the included and active control articles. Inclusion of these terms in a search strategy would identify nearly 2,000 false positives, but would correctly identify nearly 98% of articles reporting time-to-event analyses for prostheses outcome.

Error in allocation of MeSH terms among the articles has occurred and contributes to an error rate in identification of relevant articles. It is estimated that the MEDLINE indexing load will increase substantially to over one million articles per year by 2015, a 45% increase from 2007.[6] Therefore, errors may become even more prominent. Automatic indexing strategies including Inductive Logic Programming[5] and Latent Semantic Indexing systems[6] are being trialed, and may reduce error in indexing accuracy and consistency.

However, such errors are multifactorial. They may relate to the misunderstanding of reporting by indexers, poor reporting quality by authors, abbreviation in reporting due to editorial constraints, or human error. It is not yet clear whether one particular factor may have a greater influence on such indexing errors than others. Evidently, the process is not perfect. Additional research is underway with an emphasis on the impact of variation in reporting, and the quality of reporting, on the allocation of MeSH terms.

Conclusions

The allocation of MeSH terms within MEDLINE to articles in the dental literature that used time-to-event analyses to report the outcome of dental treatment in patients was inaccurate and inconsistent. Specifically, statistically related MeSH terms were incorrectly omitted from 30% of the included and incorrectly allocated to 15% of the active control articles. Such errors reduce the accuracy of search strategies and impact the identification of relevant dental outcome articles.

Acknowledgments

The authors reported no conflicts of interest related to this study.

References

1. Lin J. Is searching full text more effective than searching abstracts? BMC Bioinformatics 2009;10:46.
2. Glaser J. Index to dental literature and MEDLINE: A guide to searching the dental literature. Med Ref Serv Q 1984;3:1–16.
3. National Library of Medicine. Fact Sheet: Medical Subject Headings (MeSH®). https://www.nlm.nih.gov/pubs/factsheets/mesh.html. Accessed 13 March 2014.
4. Medical Subject Headings. US National Library of Medicine website. http://www.nlm.nih.gov/mesh/. Accessed 6 April 2010.
5. Neveol A, Shooshan SE, Claveau V. Automatic inference of indexing rules for Medline. BMC Bioinformatics 2008;9 (suppl 11):S11.
6. Neveol A, Shooshan SE, Humphrey SM, Mork JG, Aronson AR. A recent advance in the automatic indexing of the biomedical literature. J Biomed Inform 2009;42:814–823.
7. Funk ME, Reid CA. Indexing consistency in Medline. Bulletin of the Medical Library Association. 1983;71:176–183.
8. Lefebvre C, Clarke M. Identifying randomised trials. In: Egger M, Smith G, Altman D (eds). Systematic Reviews in Health Care Meta-analysis in Context. London: BMJ Publishing Group, 2000:69–86.
9. Coletti MH, Bleich HL. Medical subject headings used to search the biomedical literature. J Am Med Inform Assoc 2001; 8:317–323.