

RESEARCH

Clinical research and diagnostic efficacy studies in the oral and maxillofacial radiology literature: 1996–2005

IH Kim¹, MJ Patel², SL Hirt³ and ML Kantor^{*1,4}

¹Department of Diagnostic Sciences, New Jersey Dental School, University of Medicine and Dentistry of New Jersey, Newark, USA; ²Class of 2010, New Jersey Dental School, University of Medicine and Dentistry of New Jersey, Newark; currently a General Practice Resident at Overlook Hospital, Summit, New Jersey, USA; ³Class of 2004, New Jersey Dental School, University of Medicine and Dentistry of New Jersey, Newark, USA; currently in private practice in Haddon Heights, New Jersey; ⁴Department of Epidemiology, School of Public Health, University of Medicine and Dentistry of New Jersey, Piscataway, USA

Objectives: The aim of this study was to determine the level of evidence that is published in the oral and maxillofacial radiology (OMR) literature.

Methods: OMR papers published in *Dentomaxillofacial Radiology* and *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* between 1996 and 2005 were classified using epidemiological study design and diagnostic efficacy hierarchies. The country of origin and number of authors were noted.

Results: Of the 725 articles, 384 could be classified with the epidemiological study design hierarchy: 155 (40%) case reports/series and 207 (54%) cross-sectional studies. The distribution of study designs was not statistically significant across time (Fisher's exact test, $P = 0.06$) or regions ($P = 0.89$). The diagnostic efficacy hierarchy was applicable to 246 articles: 71 (29%) technical efficacy and 166 (67%) diagnostic accuracy studies. The distribution of efficacy levels was not statistically significant across time ($P = 0.22$) but was significant across regions ($P < 0.01$). Authors from Japan produced 26% of the papers with a mean \pm standard deviation of 5.78 ± 1.98 authors per paper (APP); American authors, 23% (3.78 ± 1.72 APP); and all others, 51% (3.76 ± 1.51 APP).

Conclusion: The OMR literature consisted mostly of case reports/series, cross-sectional, technical efficacy and diagnostic accuracy studies. Such studies do not provide strong evidence for clinical decision making nor do they address the impact of diagnostic imaging on patient care. More studies at the higher end of the study design and efficacy hierarchies are needed in order to make wise choices regarding clinical decisions and resource allocations.

Dentomaxillofacial Radiology (2011) 40, 274–281. doi: 10.1259/dmfr/81879482

Keywords: dental radiography; research design; classification; trends

Introduction

Discovery and development research has provided many new technologies and modalities for patient imaging. However, it is important to assess the utility and efficacy of new methods before they are adopted widely or become the standard of care. In 1991, Samuel Thier, then president of the Institute of Medicine of the US National Academies, said "Practices are built by accretion rather than by assessment. We no longer

know which of the available choices in a given circumstance is best relative to the others. We only know whether something works or does not. That is an unacceptable situation for any of the profession, and particularly unacceptable in health services where costs keep rising. If we cannot determine what is useful and what is appropriate, then it is unlikely that we will make wise choices about what we wish to do."¹ This prescient comment coincided with the beginning of the modern evidence-based practice movement in the early 1990s. Since the mid-1990s there has been a steady growth of evidence-based papers in the dental literature (Figure 1). However, all evidence is not equally

*Correspondence to: Dr Mel L Kantor, UMDNJ New Jersey Dental School, 110 Bergen Street, Room D860, PO Box 1709, Newark, NJ 07101-1709; E-mail: mkantor@umdnj.edu

Received 1 February 2010; revised 21 May 2010; accepted 17 June 2010

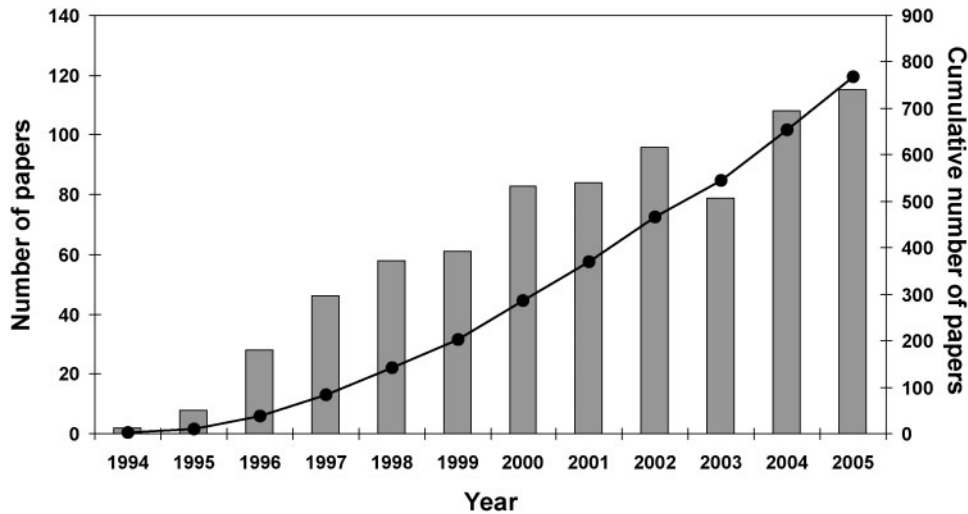


Figure 1 Number of “evidence-based” papers in the dental literature published annually (bars) and cumulatively (line) since 1994. Data were obtained from PubMed (www.pubmed.gov) using the following search strategy: evidence-based (all fields). Search was limited to English-language dental journals

compelling; the strength of the evidence falls along a continuum from weak to strong, from narrow and focused to broad and global. Using the strongest available evidence will optimize patient care.

One way to assess the quality of a clinical study is by classifying its study design. Clinical research designs may be classified by several different schemes,²⁻⁶ including the traditional epidemiological hierarchy where study designs are arranged from less rigorous to more rigorous.^{3,6} Observational studies describe a disease or assess the relation of exposures to outcomes and address the questions of prevalence, natural history, aetiology or risk factors associated with disease. From weakest to strongest, study designs in this group are case reports, case series and cross-sectional, case-control and cohort studies. Higher on the epidemiological hierarchy are the experimental studies. In this group, randomized clinical trials are performed where there is an intervention or prevention to change the course of the disease.

In 1979, Fletcher and Fletcher reviewed a random sample of articles published between 1946 and 1976 in three leading medical journals and found an overall increase in the use of weak research designs.⁷ While clinical trials increased from 13% to 21%, cohort studies decreased from 59% to 34% and cross-sectional studies increased from 24% to 44%. In 1991, a follow-up analysis based on the same three journals revealed an increase in the use of strong research designs.⁸ From 1971 to 1991, clinical trials increased from 17% to 35% and case series decreased from 30% to 4%. No changes were noted in cross-sectional, cohort or case-control studies. Since the Fletcher study in general medicine, similar studies have been conducted in pediatrics, neurosurgery, otolaryngology, obstetrics, nursing, genetics and medical radiology.⁹⁻¹⁵

The epidemiological hierarchy provides a tool for the clinician and researcher to assess the strength of evidence in a given study. However, for imaging studies it does not address the issue of diagnostic efficacy, which is another key element in the clinical

decision-making process. Fryback and Thornbury have aptly defined the concept of diagnostic efficacy with a six-tiered hierarchical model that can be used to classify imaging studies.¹⁶ This model is a global approach where the efficacy of a radiographic image is analysed as part of a larger system and where the aim is to treat patients efficiently. The lower tiers measure image quality and diagnostic accuracy, while the higher tiers measure patient outcome efficacy and societal efficacy. The tiers range from the localized goal of obtaining the most accurate image to the global goal of measuring patient outcome and societal efficacy to effect changes in patient care and health policy. The work of Fryback and Thornbury has been widely cited¹⁷ but rarely applied to assess the overall literature. Rather, the diagnostic efficacy hierarchy has been used as inclusion criteria for selecting articles in systematic reviews.^{18,19} Whereas systematic reviews address a specific topic in radiology, our study applied the efficacy hierarchy to a specific literature in radiology, the oral and maxillofacial radiology literature.

Sound clinical decision making and resource allocations rely on well-founded evidence. The primary objective of our study was to assess the oral and maxillofacial radiology literature for strength of evidence using the epidemiological hierarchy model applied to clinical research and for strength of efficacy using the diagnostic efficacy hierarchy model. Secondary objectives were to assess the regional distribution of the number of published papers and to compare the number of authors per paper across regions.

Materials and methods

The sample for this study came from the two major English-language journals devoted in part or in whole to oral and maxillofacial radiology: *Oral Surgery, Oral*

Medicine, Oral Pathology, Oral Radiology and Endodontology (OOOOE) and *Dentomaxillofacial Radiology* (DMFR). We reviewed all 400 articles in the OMR section of OOOOE published between 1996 and 2005, and the 325 articles in the odd-numbered issues of DMFR published during the same time period. Each article was reviewed independently by two investigators and categorized according to the two classification schemes. Investigators were calibrated prior to the study and when a classification disagreement occurred the papers were reviewed together and a consensus reached.

We classified each clinical research article according to the traditional epidemiological study design hierarchy (Table 1, Study design). We defined clinical research as research in which the objects of study were patients (images), providers (observers) or institutions (environment). The study designs that did not fit the traditional model were classified as "others". This included studies that used phantoms, cadavers or animals; papers that described the development of a new technology or a new application of an existing technology; dosimetry studies; and papers that were comprehensive descriptions of a specific disease or described observed changes over time without any quantitative assessment. We also classified each paper according to the hierarchy of diagnostic efficacy¹⁶ (Table 1, Efficacy level).

In addition to classifying each paper by its study design and diagnostic efficacy level, we recorded the number of authors per paper and the country of origin of the corresponding author. Preliminary analysis revealed that authors from Japan produced 26.5% of the papers; authors from the USA contributed 22.8%, and authors from the remaining 39 countries published the rest (Figure 2). This distribution became the basis for the regional analyses.

SPSS 16.0 (SPSS, Inc., Chicago, IL) was used for data management and statistical analyses. Initial analysis revealed many table cells with low or zero counts. Therefore, data from 2 years were merged, producing 5 bienniums, and SAS 9.1 (SAS Institute, Cary, NC) was used to perform 2-sided Fisher's exact tests (Monte Carlo estimate) to evaluate the distributions of the 2 hierarchies over time and across regions. Analysis of variance (ANOVA) was used to compare the number of authors by region, followed by *post hoc* pairwise comparisons when appropriate (Tukey honestly significant difference (HSD)).

Initially, the data were analysed separately for each journal, which revealed a single minor difference between the journals. Therefore, we report only the results for the combined data except where there was a difference.

Results

Of the 725 articles reviewed, 384 fit the traditional epidemiological study design classification. Of these 384, there were 115 case reports, 40 case series, 207 cross-sectional, 11 case-control, 4 cohort and 7 experimental studies. There was no statistically significant association between study design and biennium (Table 2, Study design; Fisher's exact test, $P = 0.06$) or between study design and region (Table 2, Study design; Fisher's exact test, $P = 0.89$).

A total of 246 of the 725 articles could be classified according to the diagnostic efficacy hierarchy. Only four levels of the hierarchy were represented and a little over two-thirds of the articles were diagnostic accuracy efficacy studies. There was no statistically significant association between efficacy level and biennium (Table 2, Diagnostic efficacy; Fisher's exact test, $P = 0.22$), but

Table 1 Traditional epidemiological study design and diagnostic efficacy hierarchies

Classification	Definition
A. Study design	
Case report	The presentation of one or two new cases
Case series	The presentation of three or more new cases
Cross-sectional	A prevalence study; observations related to one moment in time
Case-control	Subjects are classified on the basis of outcome (<i>e.g.</i> disease or no disease) and previous events (<i>e.g.</i> exposed or non-exposed). It is an attempt to link aetiology with their current disease status
Cohort	Subjects are classified on the basis of exposure and followed through time to see if an outcome develops or not. At least two observational points are chosen. This is an attempt to establish cause and effect
Experimental	Clinical trials with non-random and random allocation. Variables could be manipulated
B. Efficacy level^a	
1. Technical	Technical aspects of radiology equipment; comparison of radiographs based on the technical criteria, such as resolution, modulation, grey-scale range and sharpness
2. Diagnostic accuracy	Yield of normal or abnormal diagnoses in a group of cases; assessed with sensitivity and specificity, predictive values or ROC analysis
3. Diagnostic thinking	Number of cases in which an image judged helpful in making a diagnosis. Clinician changes diagnosis based on pre- and post-test information
4. Therapeutic	Number of times the image was helpful in planning patient management. The number of times planned therapy was changed after test information was obtained
5. Patient outcome	Number of patients improved with test vs without test; morbidity avoided; changes in life expectancy; cost saved with image information
6. Societal efficacy	Cost/benefit from a societal viewpoint

ROC, receiver operating characteristic. ^aModified from Fryback and Thornbury¹⁶

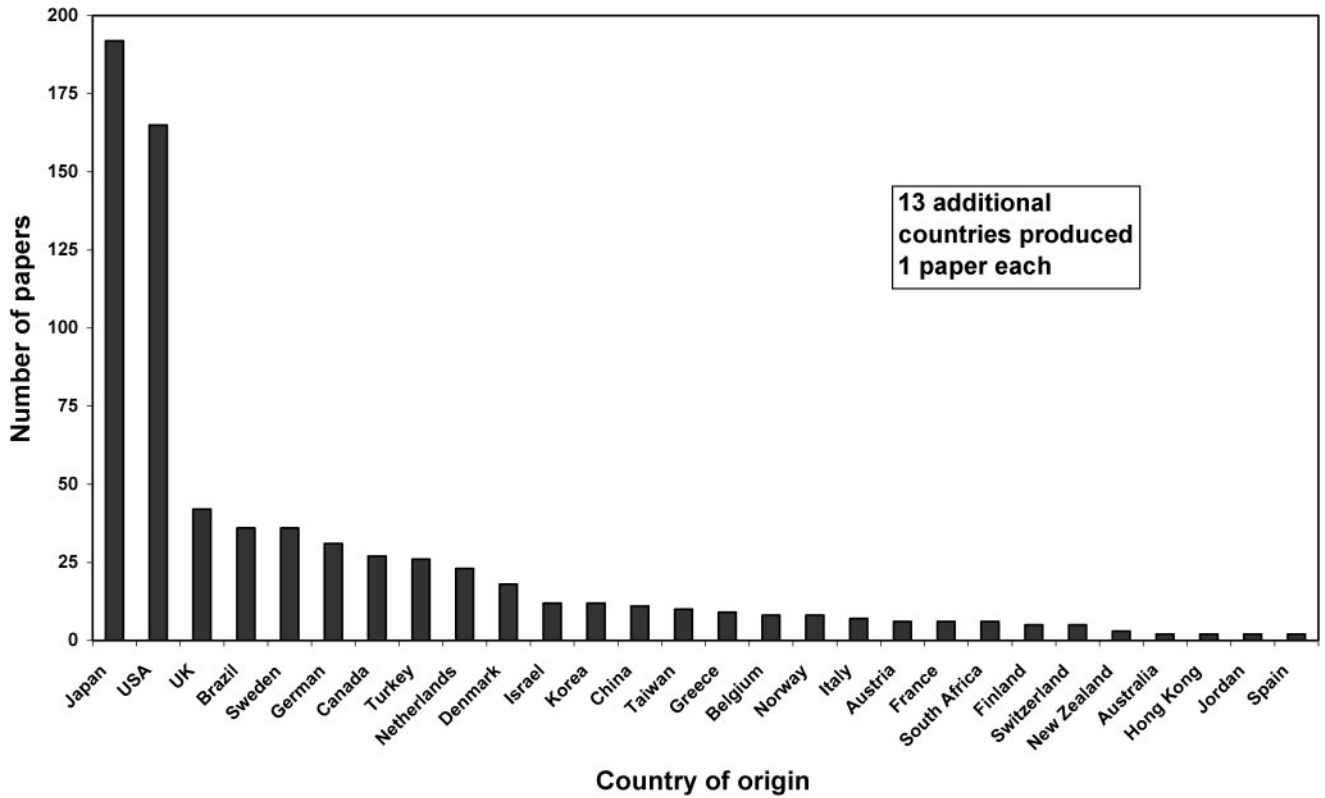


Figure 2 Total number of papers published between 1996 and 2005 by country of origin ($n = 725$)

there was a statistically significant association between efficacy level and region (Table 2, Diagnostic efficacy; Fisher’s exact test, $P < 0.01$). When the data were analysed separately for each journal (results not shown) there was a statistically significant association between efficacy level and region for OOOOE (Fisher’s exact test, $P < 0.01$), but not for DMFR (Fisher’s exact test, $P = 0.26$).

The cross-tabulation of study design with diagnostic efficacy revealed the intersection of these two hierarchies (Table 3). The cross-sectional design was the most

common study design used (65/166 or 39%) in diagnostic accuracy studies. However, among the cross-sectional studies, the diagnostic accuracy efficacy studies were a distant second (65/207 or 31%) to the more common non-efficacy studies (133/207 or 64%).

While the number of papers published from each region fluctuates over the 10-year period, there is a general downward trend for American authors, a flat trend for Japanese authors and an upward trend for authors from other countries (Figure 3). In Japan, the mean number of authors (\pm standard deviation) was

Table 2 Study design and diagnostic efficacy by biennium and by region

Classification	Biennium					Total	Region			Total
	96–97	98–99	00–01	02–03	04–05		Japan	USA	Other	
A. Study design										
Case reports	20	16	26	22	31	115	33	16	66	115
Case series	10	11	3	9	7	40	14	6	20	40
Cross-sectional	27	50	40	48	42	207	63	34	110	207
Case-control	3	0	2	2	4	11	5	2	4	11
Cohort	0	2	2	0	0	4	1	1	2	4
Experimental	1	3	2	1	0	7	1	2	4	7
Total	61	82	75	82	84	384	117	61	206	384
	Fisher’s exact test, $P = 0.06$						Fisher’s exact test, $P = 0.89$			
B. Diagnostic efficacy										
Technical	17	21	18	6	9	71	26	24	21	71
Diagnostic accuracy	35	44	30	35	22	166	30	45	91	166
Diagnostic thinking	0	1	0	2	1	4	1	1	2	4
Therapeutic	1	3	1	0	0	5	0	2	3	5
Total	53	69	49	43	32	246	57	72	117	246
	Fisher’s exact test, $P = 0.22$						Fisher’s exact test, $P < 0.01$			

Table 3 Cross-tabulation of study design vs diagnostic efficacy

Study design	Efficacy					Total
	Technical	Diagnostic accuracy	Diagnostic thinking	Therapeutic	Not applicable	
Case report	0	0	0	0	115	115
Case series	0	1	0	1	38	40
Cross-sectional	2	65	4	3	133	207
Case-control	0	3	0	0	8	11
Cohort	0	0	0	1	3	4
Experimental	1	5	0	0	1	7
Other	68	92	0	0	181	341
Total	71	166	4	5	479	725

5.78 ± 1.98, in the USA it was 3.78 ± 1.72 and elsewhere it was 3.76 ± 1.51 (Table 4). The differences among regions is statistically significant (ANOVA, $F = 99.50$, degrees of freedom (df) = 2722, $P < 0.01$) and *post hoc* pairwise comparisons revealed that the mean number of authors per paper from Japan is significantly greater than the mean number of authors per paper from the USA and from other countries (Tukey HSD, $P < 0.01$). The same results were obtained when the clinical research ($n = 384$) and diagnostic efficacy ($n = 246$) papers were analysed separately (Table 4).

Discussion

During the 10-year study period 1996–2005, there has been no significant change over time in the types of study design used in clinical research in oral and maxillofacial radiology. The study designs were predominantly case reports/series and cross-sectional studies with very few studies in the higher levels of the epidemiological hierarchy.

Overall, case reports/series comprised 21.4% of all 725 oral and maxillofacial radiology papers published in OOOOE and DMFR during the study period, which is consistent with the percentage of case reports (19%)

published in DMFR between 1987 and 2000.²⁰ However, this is substantially less than the 57.8% of oral pathology papers published in OOOOE in 1972 and 1992 that were case reports,²¹ but somewhat greater than the 9.7% of radiology articles published in the *American Journal of Roentgenology (AJR)* and *Radiology* in 1998 and 1999 that were case reports.¹⁵ Direct comparisons for cross-sectional studies cannot be made because of the way the other studies characterized the study designs and reported their results. However, limiting the analysis to the 384 clinical research papers revealed that cross-sectional studies comprised 53.9% of the oral and maxillofacial radiology clinical studies, which is greater than the 22%–44% of clinical research papers published in the general medical literature between 1946 and 1991 that were cross-sectional studies.^{7,8} Cohort and case-control studies combined comprised 2.1% of our total study sample of 725 papers, which is comparable with the 2% found in the paediatric dentistry literature.²² Experimental studies comprised only 1% of our total sample, whereas they comprise 9% of the paediatric dentistry literature²² and 17% of the English-language literature regarding therapy for temporomandibular disorders.²³

Longitudinal studies, especially randomized clinical trials (RCTs), are uncommon in oral and maxillofacial radiology despite the fact that they provide the strongest evidence for assessing risk, establishing causality

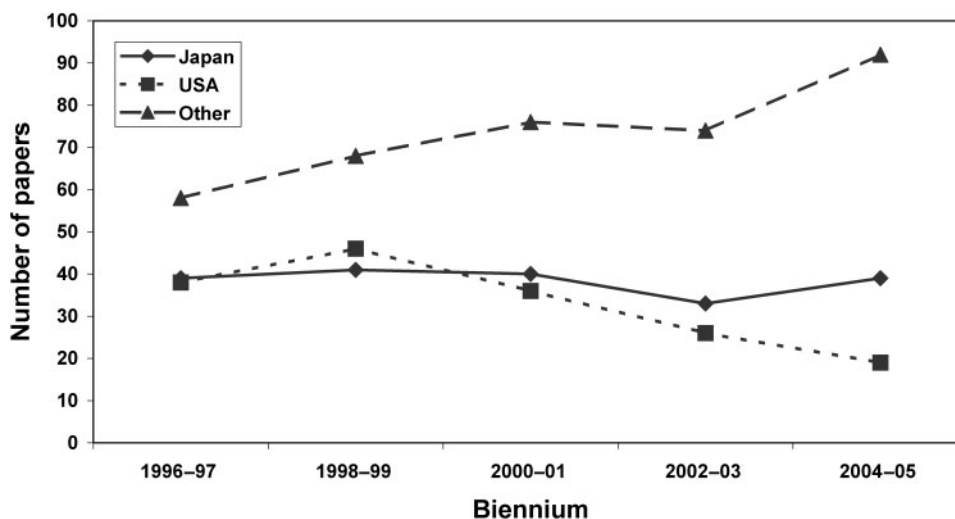
**Figure 3** Number of papers published per biennium by region ($n = 725$)

Table 4 Number of authors per paper by region

	N	Region						F	df	P*
		Japan		USA		Other				
		n	Mean (SD)	n	Mean (SD)	n	Mean (SD)			
All papers	725	192	5.78 (1.98)	165	3.78 (1.72)	368	3.76 (1.51)	99.50	2722	<0.01
Clinical research	384	117	5.97 (1.84)	61	3.67 (1.55)	206	3.72 (1.46)	81.23	2381	<0.01
Diagnostic efficacy	246	57	5.75 (1.92)	72	3.90 (1.73)	117	3.68 (1.31)	34.60	2243	<0.01

df, degrees of freedom; SD, standard deviation. *All *post hoc* pairwise comparisons were statistically significant (Tukey honestly significant difference, $P < 0.01$): Japan > USA, other

and evaluating therapeutic interventions. According to Fryback and Thornbury, RCTs are required to provide a “definitive answer concerning whether a radiographic examination is efficacious with respect to patient outcome[s].”¹⁶ However, others claim that RCTs are not always appropriate or necessary for imaging studies. Because technology changes rapidly and longitudinal studies progress slowly, it is possible, even likely, that the imaging technology will have changed before the study is completed and the results published.²⁴ Indeed, the role of RCTs is at the heart of the controversy regarding CT screening for lung cancer.^{25,26} Some researchers advocate CT screening for lung cancer based on results from their single-arm study that did not include a non-screening control group.²⁷ Other researchers, who conducted similar single-arm studies and obtained similar results, interpret the results differently and urge restraint until an ongoing National Cancer Institute-funded RCT is completed.²⁸

In the general medical literature, clinical research papers published in American journals tended to use stronger study designs than those published in Japanese journals.²⁹ However, in this study of OMR literature, we found no significant difference in the study designs used by authors from different geographic regions or between American and Japanese authors in particular.

During the study period, there has been no significant change over time in the level of efficacy studies in our sample of OMR literature. Technical efficacy (level 1) and diagnostic accuracy efficacy (level 2) comprised 96.3% of the efficacy studies ($n = 246$). Whereas level 1 and 2 studies were conducted with approximately equal frequency in Japan, level 2 studies were nearly twice as common as level 1 studies in the USA and a little more than four times as common elsewhere (“other”), which accounts for the statistically significant difference in diagnostic efficacy among the geographic regions. The reason for this trend is not obvious, but may be related to differences in the introduction of new technology, education of academic oral and maxillofacial radiologists and imaging scientists, and availability of research funding in different parts of the world.

Both level 1 and 2 efficacy studies are relatively simple and inexpensive to conduct and can be completed in a reasonably short time. In the efficacy hierarchy, they provide the foundation necessary for studies at higher levels. However, Fryback and Thornbury describe

patient outcome efficacy (level 5) as “the sine qua non of efficacy from the individual patient’s viewpoint”,¹⁶ of which there were none in the OMR literature. To the best of our knowledge, there are no other studies that have applied the efficacy hierarchy to the OMR or medical radiology literature with which to compare our results. However, studies of published medical radiology articles that fall within a single diagnostic efficacy level suggest a literature that has not improved over time. Receiver operating characteristic (ROC) studies (level 2) published in *Radiology* between 1997 and 2006 were often inadequate for clinical decision making and contained frequently recurring mistakes.³⁰ The quality of diagnostic accuracy studies (level 2) published between 2001 and 2005 was fair to middling throughout the study interval, leading the author to conclude that the “quality of reporting of diagnostic accuracy is substandard, with a great deal of room for improvement”.³¹ While the average number of cost-effectiveness studies (level 6) published each year increased between 1985 and 2005, the average quality remained constant across the 20 years.³²

The cross-tabulation of study design with diagnostic efficacy revealed that most diagnostic efficacy studies were cross-sectional studies. This is a predictable observation given that diagnostic accuracy efficacy studies usually yield sensitivity, specificity and ROC data, which are typically generated from cross-sectional studies. Furthermore, the lack of higher levels of efficacy studies could be predicted from the lack of higher levels of study designs, and conversely higher levels of efficacy (levels 3–6) can also be addressed with decision analytic models.¹⁶ However, in our study the paucity of cohort studies and RCTs was not offset by decision analysis studies. In the absence of cohort studies and RCTs, a sufficient number of well done case–control studies with consistent results may also be sufficient to guide clinical decision making and policy, but these were uncommon in our sample.

Cross-sectional and diagnostic accuracy studies predominate the literature for multiple reasons. Cohort studies and RCTs are difficult to perform, requiring large patient populations, long-term clinical follow-up and sophisticated statistical analyses. They demand greater planning and are much more expensive than case reports/series and cross-sectional studies, and require a cadre of well-trained researchers and a source of adequate funding. In addition, the Fletchers suggest that “...the

approach to clinical research might be influenced by rapid developments in technology, the mobility of both researchers and patients and academic competition (“publish or perish”).⁷

In our sample, Japan and the USA produced nearly half of the OMR papers during the study period, which closely parallels the trend in global biomedical publications during a comparable period (1995–2000).³³ There has been a decrease in the number of papers from the USA and an increase in the number of papers from Japan and others. This is consistent with the trend observed for *Radiology* between 1998 and 2003.³⁴

The present study is unique in its global analysis of the OMR literature. To our knowledge, this is the first time the Fryback and Thornbury model has been applied to this literature. Our study was limited to OOOOE and DMFR, the two major English-language journals for oral and maxillofacial radiology, and therefore may not be representative of the entire OMR literature. Restricting the DMFR sample to papers published in the odd-numbered issues should not bias the results, as there is no systematic difference in the types of papers published in odd- and even-numbered issues (Brooks SL, DMFR editor, personal communication, 10 May

2010). The current study may serve as a baseline for future studies to establish temporal trends, as well as for comparative studies of head and neck articles in the medical radiology literature. In addition, future studies that apply the same methods to a single technology or disease/disorder in the OMR literature may reveal a maturation of evidence that was not manifest in this global assessment of the literature.

In conclusion, the OMR literature consists mostly of case reports/series and cross-sectional and diagnostic accuracy studies. Such studies do not provide strong evidence for clinical decision making nor do they address the impact of diagnostic imaging to patient care. While they are necessary, they are not sufficient to guide evidence-based practice and policy making. More studies at the higher end of the study design and efficacy hierarchies are needed in order to make wise choices regarding clinical decisions and resource allocations.

Acknowledgments

We thank Dr Barbara Greenberg (UMDNJ New Jersey Dental School) for reviewing the manuscript and providing valuable comments and suggestions.

References

- Thier SO. Dental education in the future. *J Dent Educ* 1991; **55**: 353–355.
- Feinstein AR. Clinical biostatistics. XLIV. A survey of the research architecture used for publications in general medicine journals. *Clin Pharmacol Ther* 1978; **24**: 117–125.
- Stephenson JM, Babiker A. Overview of study design in clinical epidemiology. *Sex Transm Infect* 2000; **76**: 244–247.
- Hayes C. Evidence based dentistry: design architecture. *Dent Clin North Am* 2002; **46**: 51–59.
- Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health* 2003; **57**: 527–529.
- Hujoel P. Grading the evidence: the core of EBD. *J Evid Based Dent Pract* 2008; **8**: 116–118.
- Fletcher RH, Fletcher SW. Clinical research in general medical journals: a 30-year perspective. *N Engl J Med* 1979; **301**: 180–183.
- McDermott MM, Lefevre F, Feinglass J, Reifler D, Dolan N, Potts S, et al. Changes in study design, gender issues, and other characteristics of clinical research in three major medical journals from 1971 to 1991. *J Gen Intern Med* 1995; **10**: 13–18.
- Thakur A, Wang EC, Chiu TT, Chen W, Ko CY, Chang JT, et al. Methodology standards associated with quality reporting in clinical studies in pediatric surgery journals. *J Pediatr Surg* 2001; **36**: 1160–1164.
- Gnanalingham KK, Tysome J, Martinez-Canca J, Barazi SA. Quality of clinical studies in neurosurgical journals: signs of improvement over three decades. *J Neurosurg* 2005; **103**: 439–443.
- Rosenfeld RM. Clinical research in otolaryngology journals. *Arch Otolaryngol Head Neck Surg* 1991; **117**: 164–170.
- Dauphinee L, Peipert JF, Phipps M, Weitzen S. Research methodology and analytical techniques used in the journal of *Obstetrics and Gynecology*. *Obstet Gynecol* 2005; **106**: 808–812.
- Jacobsen BS, Meninger JC. The designs and methods of published nursing research: 1956–1983. *Nurs Res* 1985; **34**: 306–312.
- Bogardus ST, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research. *J Am Med Assoc* 1999; **281**: 1919–1924.
- Ozsunar Y, Unsal A, Akdilli A, Karaman C, Huisman T, Sorensen AG. Technology and archives in radiology research: a sampling analysis of articles published in the *AJR* and *Radiology*. *AJR Am J Roentgenol* 2001; **177**: 1281–1284.
- Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991; **11**: 88–94.
- ISI Web of Science [database on the Internet]. Stamford (CT): Thomson Reuters; c2009 [cited 2010 January 10]. Available from: <http://www.isiknowledge.com/>.
- Limchaichana N, Petersson A, Rohlin M. The efficacy of magnetic resonance imaging in the diagnosis of degenerative and inflammatory temporomandibular joint disorders: a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2006; **102**: 521–536.
- Ribeiro-Rotta RF, Lindh C, Rohlin M. Efficacy of clinical methods to assess jawbone tissue prior to and during endosseous dental implant placement: a systematic literature review. *Int J Oral Maxillofac Implants* 2007; **22**: 289–300.
- Benn D, Hirschmann P. An analysis of papers published in the journal *Dentomaxillofacial Radiology*: 1987–2000. *Dentomaxillofac Radiol* 2000; **30** (Suppl 2): S30 (Abstract 18.2).
- Kaugars GE, Riley WT, Grisius TM, Page DG, Svirsky JA. Comparison of articles published in Oral Surgery Oral Medicine Oral Pathology in 1972 and 1992. *Oral Surg Oral Med Oral Pathol* 1994; **78**: 351–353.
- Hashim Nainar SM. Profile of *Journal of Dentistry for Children* and *Pediatric Dentistry* journal articles by evidence topology: thirty-year time trends (1969–1998) and implications. *Pediatr Dent* 2000; **22**: 475–478.
- Antczak-Bouckoms AA. Epidemiology of research for temporomandibular disorders. *J Orofac Pain* 1995; **9**: 226–234.
- Valk PE. Randomized controlled trials are not appropriate for imaging technology evaluation. *J Nucl Med* 2000; **41**: 1125–1126.
- New studies further the debate over CT lung screening. *RSNA News* 2007; **17**: 8.
- Marshall E. A bruising battle over lung scans. *Science* 2008; **320**: 600–603.
- Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med* 2006; **355**: 1763–1771.

28. Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *J Am Med Assoc* 2007; **297**: 953–961.
29. Fukui T, Rahman M, Sekimoto M, Hira K, Maeda K, Morimoto T, et al. Study design, statistical method, and level of evidence in Japanese and American clinical journals. *J Epidemiol* 2002; **12**: 266–270.
30. Shiraishi J, Pesce LL, Metz CE, Doi K. Experimental design and data analysis in receiver operating characteristic studies: lessons learned from reports in *Radiology* from 1996 to 2006. *Radiology* 2009; **253**: 822–830.
31. Wilczynski NL. Quality of reporting of diagnostic accuracy studies: no change since STARD statement publication—before-and-after study. *Radiology* 2008; **248**: 817–823.
32. Otero HJ, Rybicki FJ, Greenberg D, Neumann PJ. Twenty years of cost-effectiveness analysis in medical imaging: are we improving? *Radiology* 2008; **249**: 917–925.
33. Rahman M, Fukui T. Biomedical publication—global profile and trend. *Public Health*. 2003; **117**: 274–280.
34. Scientific abstract and journal manuscript submissions from overseas increase. *RSNA News* 2004; **14**: 9.