The Challenges of Defining Oral Cancer: Analysis of an Ontological Approach

Jose Luis Tapia · Louis J. Goldberg

Received: 4 August 2011 / Accepted: 2 September 2011 / Published online: 14 September 2011
© Springer Science+Business Media, LLC 2011

Abstract An important inconsistency currently exists in the literature on oral cancer. Reviewing this literature, one finds that the term oral cancer is defined and described with great variation. In a search in PubMed, at least 17 different terms were found for titles of papers reporting data on oral cancer. The variability of the terms used for designating anatomic regions and type of malignant neoplasms for reporting oral cancer has hampered the ability of researchers to effectively retrieve information concerning oral cancer. Therefore, it is sometimes extremely difficult to provide meaningful comparisons among various studies of oral cancer. Recently, a new ontological strategy that is rooted in consensus-based controlled vocabularies has been proposed to improve the consistency of data in dental research (Smith et al. in *J Am Dent Assoc* 141:1173–1175, 2010). In this paper, we analyzed the terminology dilemma on oral cancer and explained the current situation. We proposed a possible solution to the dilemma using an ontology-based approach. The advantages for applying this strategy are also discussed.

Keywords Oral cancer · Biomedical ontology · Annotations · Oral cancer ontology

Introduction

It has been recently pointed out in an article in the *Journal of the American Dental Association* (JADA) that the

current explosion of information in biomedical research is creating significant problems for our ability to access relevant data that have been published in the literature. Smith et al. [1] asserted that there is presently “an uncontrolled explosion of different ways of describing information.” The paper goes on to suggest a “strategy to advance the consistency of data in the dental research community” with the goal of solving problems of retrieval and access. Smith et al. [1] suggested the necessity for creating ontologies that are rooted in “consensus-based controlled structural vocabularies” in order that “past, present, and future data” be maintained in an electronic database that permits searches enabling researchers from diverse research communities to readily obtain relevant information from the literature. The goal of the current paper is to explore the use of ontologies in the area of oral cancer.

In the JADA paper, Smith et al. [1] describe the advantages and benefits to building such ontologies. It is suggested that biomedical ontologies should be designed so as to be interoperable with other existing ontologies. The interoperability among different ontologies promotes the search and integration of various types of biomedical data. An ontology can be modified or corrected in accordance with new research findings and can also enable the integration of previously captured data. This integration is possible because an ontology allows the incorporation of traditional terms that have been previously used by the research community.

Terminology Dilemma

An impediment currently exists in the ability of the oral cancer research community to locate relevant information concerning oral cancer in the medical literature. Various

J. L. Tapia (✉) · L. J. Goldberg
Department of Oral Diagnostic Sciences, School of Dental
Medicine, University at Buffalo, 355 Squire Hall, 3435 Main
Street, Buffalo, NY 14214, USA
e-mail: jltapia2@buffalo.edu

uncontrolled, nonconsensus-based terms are presently being used to refer to cancer that occurs in the various regions of the mouth [2]. Attempts to develop a controlled vocabulary for such terms have not been successful [2, 3]. In fact, some authors [2, 4] have concluded that, it is perhaps “too late for a rational solution to this problem”. Although it may be a daunting task to achieve a widely-accepted definition of oral cancer, our obligation as clinicians and researchers, however, is to make a serious attempt to solve this problem by whatever available means. If we let the current situation continue, it will be extremely difficult to provide meaningful comparisons among various studies of oral cancer regarding incidence, mortality rates, and related epidemiologic and clinical data. As Moore et al. [2] state, “given the importance of this disease on a worldwide basis, it would appear to be essential that some form of international consensus be derived regarding the classification of oral cancer.”

As noted above, establishing a consensus on oral cancer terminology is, however, a difficult challenge. This paper has three goals: (1) to present an analysis of the literature on oral cancer in order to delineate the nature of the current terminological inconsistencies; (2) to describe and examine the methods of biomedical ontology, as presented in the Smith et al. [1] paper, to determine their applicability for addressing the oral cancer terminology dilemma; and (3) to discuss the advantages in developing a biomedical ontology in the domain of oral cancer that go beyond providing a remedy for the terminology dilemma.

Analysis of the Literature with Respect to the Terminology Dilemma

A review of the literature indicates that there is no consensus on the definition of the term oral cancer. The definition of the word “cancer” itself is clear enough and well accepted: cancer is a common term used to denote all malignant tumors [5]. However, the definition of oral cancer is problematic. The meaning of the term oral cancer varies considerably in pertinent publications [6]. The use of different terms and definitions for reporting oral cancer has hampered communication between clinicians and researchers.

Lack of Consensus on the Definition of Oral Cancer

The lack of consensus on the definition of oral cancer was well documented by Moore et al. in a letter published in the journal Oral Diseases [2]. They reviewed the oral cancer literature and wrote that “defining oral cancer presents some important challenges to both clinicians and researchers.” From their point of view, the root of the

problem is related to what they called a terminology dilemma. They provide evidence to support this view and conclude by stating that, “as yet there seems no uniformly accepted definition of oral cancer” [2]. Similarly, McCartan [4] describes the oral cancer terminology as “unhelpful and confusing”.

The variety of terms that have been and are currently being used to describe oral cancer is illustrated in Table 1. A PubMed search revealed at least 17 different terms used in titles of papers reporting data on oral cancer.

Lack of Consensus on the Anatomic Sites that are Included under Oral Cancer

The lack of consensus in the definition of oral cancer is reinforced by the variability in the specific anatomic designation of the disease in published reports [24]. Some epidemiological studies focus on cancer that occurs in the area between the vermilion border of the lip and the junction of the soft and hard palate [16]. Some studies add, to the criteria stated above, cancers of the oropharynx [18], nasopharynx, and hypopharynx [20], and others include cancer of the major salivary glands [25, 26]. Still others incorporate cancer of the external (skin) portion of the lip [27]. Furthermore, some papers in the literature even fail to specify which anatomic sites are included [28–30].

Table 1 Terms found in titles of papers involving studies of oral cancer

Terms
Cancer of mouth and pharynx [7] ^a
Cancer of the oral cavity [8]
Cancer of the oral cavity/oropharynx [9]
Cancer of oral cavity and pharynx [10]
Cancer of the tongue and oral cavity and pharynx [11]
Intraoral cancer [12]
Malignant oral tumors [13]
Mouth cancer [14]
Oral cancer [15]
Oral cavity cancer [16]
Oral and pharyngeal cancer [17]
Oral and oropharyngeal cancer [18]
Oral cavity and pharynx cancer [19]
Oral cavity and pharynx-throat cancer [20]
Oral cavity and oropharyngeal cancers [21]
Oral malignant tumours [22]
Tongue and mouth cancer [23]

^a Number after the term refers to a paper in this paper’s reference list

Lack of Consensus in the Designation of the type of Malignant Neoplasms Associated with the Term Oral Cancer

Another issue related to the terminology dilemma is the histological types of malignant neoplasms that are included as oral cancer. It is clear that malignant neoplasms of the mucosal tissue (epithelium) are always considered to be oral cancer. What is not clear is whether malignant neoplasms in tissues adjacent to the oral mucosa should also be designated as oral cancer. Reviewing the literature, one finds that many studies only include squamous cell carcinomas [31]. Others add cancer in adjacent tissues such as salivary glands [19] and connective tissue [13, 15, 25]. In addition, others also include malignant neoplasms of the skin of the lip [27], lymphoid [15, 25], muscle [13] and nerve [19] tissues as oral cancer. There also are studies that include metastases to the oral cavity and melanoma [15] as oral cancer.

Consequences

The variability of the terms used for designating anatomic regions and types of malignant neoplasms for reporting oral cancer has hampered the ability of researchers to effectively retrieve information in the literature. For example, a researcher who wants to retrieve papers containing information concerning the specific location and type of malignant neoplasm found in various mouth regions is confronted with two problems. One problem is that the researcher must know the anatomic definition of oral cancer that is used by each author. Another problem is that the researcher must know what types of malignant neoplasms are included in the paper.

An Ontology-Based Approach to the Oral Cancer Terminology Dilemma

The terminology dilemma has developed because various investigators have selected terms for defining oral cancer with respect to location and type of malignancy without the formation of a consensus concerning the definition of those terms. One way to solve the terminology dilemma requires that the international community reach a consensus upon a standard definition of oral cancer. Attempts have been made in the oral cancer community to define oral cancer, but these attempts have not been successful [2]. It seems that the current practice of using various, loosely defined terms for oral cancer is unlikely to change.

Another strategy to solve the terminology dilemma would be keep the terms currently in use, as illustrated in Table 1, and create consensus-based standard definitions

for each one of them (e.g., intraoral cancer, mouth cancer, oral cavity cancer, oropharyngeal cancer etc.). This would be exceedingly difficult and not very productive even if successful. The reason is because there are at least 17 terms used in the oral cancer literature that we have identified. In our judgment, it would be extremely difficult, if not impossible, to convince experts in the area of oral oncology to convene in order to develop a consensus-based definition for each one of them. It would also not be productive because these variable terms have been developed in the absence of any attempt to establish a controlled vocabulary, which is the very reason for the dilemma we address. The goal is not to solve the dilemma by attempting to define these terms but to move forward with the approach that we are presenting in this paper.

To solve the terminology dilemma, we do not advocate attempting any of the previously mentioned solutions. Rather, we advocate for the development of an oral cancer ontology. The fundamental objective is to allow researchers to obtain specific information they need from the literature concerning oral cancer location and type, regardless of the title of the paper and the definition of oral cancer used by the authors. To this end we believe that the biomedical ontology-based method described by Smith et al. [1] can be employed to achieve this objective.

Smith et al. [1] have stated that ontological strategy is rooted in consensus-based controlled vocabularies. Even though there is no consensus-based definition of the terms used to define oral cancer, we can take advantage of other consensus-based terminologies that are currently being used by researchers in the field. Reviewing the oral cancer literature, it is evident that there are two universally well-accepted terminologies in every paper concerning oral cancer. These terminologies are: (1) the standard anatomic terms used for identifying the locations of malignant neoplasms; and (2) the standard terms used to identify types of malignant neoplasms. Information concerning the locations and types of malignant neoplasms are always included in published reports of oral cancer concerning incidence and prevalence.

Location of Malignant Neoplasms

Standard anatomic terms of location are employed by researchers to report the sites in a paper that will be included in the study. For example, tongue, hard palate, gingiva, or floor of mouth are universally accepted terms. There is currently in existence a hierarchically-structured consensus-based terminology of anatomical terms that is commonly referred to in the oral cancer literature. This terminology is called the International Classification of Diseases for Oncology (ICD-O topography) of possible sites for malignant neoplasms [32]. This ICD-O topography of anatomical sites is accepted as the standard by the World Health

Organization (WHO) [33]. This ICD-O topography for the mouth region is listed under the heading of “LIP, ORAL CAVITY, AND PHARYNX” and includes the following major subclasses:

- Lip
- Base of tongue
- Other and unspecified parts of the tongue
- Gum
- Floor of the mouth
- Palate
- Other and unspecified parts of the mouth
- Parotid gland
- Other and unspecified major salivary glands
- Tonsil
- Oropharynx
- Nasopharynx
- Piriform sinus
- Hypopharynx
- Other and ill-defined sites in lip, oral cavity and pharynx

Under each subclass, there are further subdivisions. For example, under the subclass, “palate,” can be found the following subdivisions:

- Hard palate
- Soft palate
- Uvula
- Overlapping lesions of the palate
- Palate NOS (not otherwise specified)

It is obvious that this classification system is not complete as suggested by the use of such phrases as, “other and unspecified”, “other and ill-defined” and “not otherwise specified”. These phrases indicate a lack of specific information about the location of certain tumors. It is to be expected that in the future new standardized protocols will help to eliminate the designation of particular locations as “ill-defined” or not “otherwise unspecified”. When advances are made in the identification of tumor sites this information can be added to the ontology. In addition, ontology-driven information systems have the capacity to retain legacy terms and designate their relationship to the new terms that have taken their place [1].

We consider the ICD-O site classification system to be a consensus-based, hierarchically-organized controlled vocabulary that can be used as part of an oral cancer ontology. The ICD-O topography can serve as an important component of the oral cancer ontology we are proposing.

Type of Malignant Neoplasm

Many efforts have been made by several groups of pathologists to standardize the terminology used for malignant

neoplasms. These efforts have resulted in widely accepted terminology that is currently used by pathologists in their diagnostic reports of malignant neoplasms. This terminology is included in the World Health Organization Classification of Tumors [34]. Therefore, oral cancer researchers use terms from the WHO classification of tumors to describe the types of malignant neoplasms that are included in their studies. The types of malignant neoplasms that may occur in the mouth region are listed under the heading of “WHO CLASSIFICATION OF TUMOURS OF THE ORAL CAVITY AND OROPHARYNX” and include the following major subclasses:

- Malignant epithelial tumours
- Salivary gland tumours
- Soft tissue tumours
- Haematolymphoid tumours
- Mucosal malignant melanoma
- Secondary tumours

Under each subclass are listed “finer-grained” subdivisions. For example, under the subclass, “malignant epithelial tumors,” can be found:

- Squamous cell carcinoma
 - Verrucous carcinoma
 - Basaloid squamous cell carcinoma
 - Papillary squamous cell carcinoma
 - Spindle cell carcinoma
 - Acantholytic squamous cell carcinoma
 - Adenosquamous carcinoma
 - Carcinoma cuniculatum

We consider the WHO classification of tumors to be a consensus-based, hierarchically-organized controlled vocabulary that can be used as part of the proposed oral cancer ontology. Thus, there are currently in place consensus-based hierarchically-organized terminologies for both location (ICD-O) and type (WHO) of malignant neoplasms. Both of these terminologies will be fundamental elements in the building of an oral cancer ontology. The National Cancer Institute thesaurus (NCIT) is another controlled vocabulary concerning malignant neoplasms [35]. As will be discussed later, links to the NCIT will be possible using the oral cancer ontology.

Implementation of the Ontological Strategy

The next step in the implementation of the ontological strategy is to instantiate the controlled terminologies into a database that can maintain the hierarchical structure of the terminologies. This requires establishing two coding systems. The first is the designation of a computer code to identify each term and the second is the designation of a code that

establishes the hierarchical relationship between the terms. Fortunately, both the ICD-O site terminology and the WHO classification of tumor terminology have already provided identification codes (ICD-O morphology codes) for their terminologies and these codes will be used in the ontology. There are, however, no codes to establish the relationship between terms and these will have to be provided.

There is a well-established and growing science concerning the principles of ontology building, including methodologies for coding relationships between terms [36]. There are software programs for instantiating terminologies such as we have described for ICD-O and WHO in ontological databases [37].

Once the two terminologies in the oral cancer ontology have been placed in an ontological database another coding step must be implemented. The logic of the papers we are interested in makes fundamental correlations between types of cancer and the locations in the mouth region in which they are found. Those relationships must be maintained in the ontology. To do so, relational codes must be established that link every term in the location terminology, with every term in the type of malignant neoplasm terminology.

A semantic query of the ontology might take the following form: give me all the papers in the literature that have studied the occurrence of spindle cell carcinomas in the soft palate and uvula. An answer to that question requires that all the terms in the ontology are cross-linked. Again, the methodology for achieving those cross-linkages has been established in the field of biomedical ontology [36].

Annotation

The next step in the implementation of the “ontological strategy” of Smith et al. [1] requires establishing a relationship between the oral cancer ontology and the literature in oral cancer. This is accomplished through the method of

annotation (tagging). Annotation in the case of the oral cancer ontology requires that location and type of cancer information in each relevant paper be coded into the ontology. Figure 1 illustrates the annotation process of oral cancer literature.

In Fig. 1, column 1 shows examples of terms from the consensus-based terminology for types of malignant neoplasms. The corresponding WHO codes for the types of malignant neoplasms are shown in column 2. Column 3 contains the corresponding ICD-O codes for the consensus-based terminology of a small sample of anatomical terms (column 4). The horizontal and crossed arrows represent the links established by the use of relational codes between the two terminologies in the ontology. The vertical arrows represent the codes used to maintain the hierarchical structure of the terminology in the ontology. Column 5 shows two hypothetical examples of papers in oral cancer literature that are linked to the ontology based on the location and type of oral cancer reported.

For example, paper #1 reports data regarding spindle cell carcinoma of the uvula. Retrieving the information on the location (uvula) and type of malignant neoplasm (spindle cell carcinoma) that are reported on in paper #1 is accomplished by the process of linking the code for paper #1 (0.1) to the appropriate terms in the ontology. The process of linking information in the literature to a computer-based ontology is called annotation. This is represented in Fig. 1 by the placing of the paper #1 code (0.1) under both spindle cell carcinoma and uvula. Thus, paper #1 is coded appropriately into the ontology in both terminologies. Paper #1, along with all related publication information concerning the paper, will be retrieved from the ontology by querying the ontology with the terms “spindle cell carcinoma” (type of malignant neoplasm), “uvula” (location of the oral cancer) or by the combined phrase “spindle cell carcinoma of the uvula”. A second example involves a paper # 2, which includes information about squamous cell carcinoma of the palate. Annotating the paper with the ontology will allow retrieval of the paper

Fig. 1 Oral cancer ontology and annotation of the oral cancer literature

by querying the terms “squamous cell carcinoma”, “hard palate” or the combined phrase, “squamous cell carcinoma of the hard palate”.

After annotating the location and types of malignant neoplasms reported in papers concerning oral cancer, the oral cancer ontology can be used to facilitate oral cancer literature retrieval. An oral cancer ontology will allow researchers to easily access the information that they need. For example, if the investigator is searching for oral cancer that specifically concerns the tongue, the ontology will retrieve all papers that report on cancer of the tongue. If the investigator narrows the search and is interested in papers that report a specific type of malignant neoplasm of the mouth region, such as squamous cell carcinoma, and in a specific location, such as the tongue, the ontology will locate all such papers. The ontology allows investigators to query the literature by using any combination of location and types of malignant neoplasms.

Initially, this oral cancer ontology can be used to annotate existing published literature based on the location and type of malignant neoplasm. The long-term goal, however, is to expand this ontology and integrate new terminologies, as well as modify existing terminologies based on new knowledge that requires the addition of new terms or modifications of old terms [1]. For example, consensus-based terminologies describing factors such as geographic data, risk factors, staging and grading on oral cancer can be added to the oral cancer ontology as they are developed over time.

Additional Advantages of an Oral Cancer Ontology

Biomedical Ontologies

Biomedical ontologies are now commonly used to facilitate searches of the literature in particular domains [38, 39]. The oldest and most well used biomedical ontology is the Gene Ontology (GO). It has been in existence for over 10 years. GO provides a standardized vocabulary for gene product annotations in the areas of molecular function, biological processes, and cellular components [40]. GO has been widely adopted as a vocabulary for annotating functional gene-products data [1]. There are over 11 million annotations relating gene products described in research databases to terms in the GO. Research information regarding approximately 180,000 genes has been tagged using GO [41].

GO has facilitated retrieval of scientific journal articles associated with genes. Today, scientists commonly use GO tools to access past and present scientific literature. This is possible because the GO vocabulary encompasses all terms that have been used by experts to describe gene data. As of

2010, more than 50,000 articles on genes have been annotated through GO and the content is available for computer searches [42].

The proposed oral cancer ontology possesses a similar structure to that of the Gene Ontology. GO contains three controlled vocabularies that describe the functional attributes of gene products [43]. The proposed oral cancer ontology contains two controlled vocabularies that describe the location and type of malignant neoplasm found in the mouth region. The GO terms have unique identifiers (codes) [40]. Location and type malignancy of oral cancer will also have their own unique identifiers. Literature-based-annotation using GO codes is an important way in which biological information about specific gene products can be captured and expressed in a searchable, computable form [44]. In an oral cancer ontology, literature-based-annotation using ICD-O and WHO codes can be the method in which the location and type of malignant neoplasms of the mouth region will be captured and made accessible for computer-based searches.

The Interoperability of Biomedical Ontologies

There are several additional and important benefits in constructing the proposed oral cancer ontology on well-defined ontological principles, such as those used in the GO. An oral cancer ontology so constructed can always be modified to adapt to the discovery of new knowledge or to make changes in accepted terminologies, without compromising information already present in the ontology. High quality biomedical ontologies are already in existence and can be linked to the oral cancer ontology. Table 2 shows a partial list of well-established current biomedical ontologies. An oral cancer ontology will give oral cancer researchers the ability to gain information related to problems in oral cancer from these other biomedical ontologies. The semantic interoperability of these ontologies permits the integration of information from many fields and at many levels of granularity, from the genetic (GO) to the cellular (cell type ontology) to the clinical (as SNOMED-CT and the NCI Thesaurus).

The ICD-O codes concerning the locations of oral cancer can be linked to the Foundational Model of Anatomy (FMA), an ontology with a controlled vocabulary of the human body [46]. As a comprehensive ontology of human anatomy, the FMA includes terms that represent the structural organization of the human body from the macromolecular to macroscopic levels. The types of malignant neoplasms found in the mouth region can be linked through FMA to other comprehensive terminologies related to cancer such as Mouse Pathology [47], and Cancer Research and Management ACGT Master Ontology [48]. Currently, we are developing a standardized vocabulary for the

Table 2 A partial list of ontologies for domains relevant to oral cancer

Title	Domain	Prefix	Website
Gene ontology	Biological process Cellular component Molecular function	GO	http://www.geneontology.org
Foundational model of anatomy	Anatomy	FMA	http://sig.biostr.washington.edu/projects/fm/AboutFM.html
PRotein ontology	Proteins	PR	http://pir.georgetown.edu/pro/pro.shtml
Mouse pathology	Health	MPATH	http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=MPATH
Advancing clinico-genomic trials	Health	ACGT	http://www.ifomis.org/wiki/ACGT_Master_Ontology_(MO)
Ontology of general medical science	Medicine	OGMS	http://code.google.com/p/ogms/
Cell type ontology	Anatomy	CL	https://lists.sourceforge.net/lists/listinfo/obo-cell-type

A more comprehensive list is available at <http://www.obofoundry.org>. [45]

surgical pathology domain and we are planning to link the mouth region cancer terminologies to that vocabulary.

Conclusion

We believe that developing an ontology for oral cancer will be an important step in permitting the field of oral oncology to take advantage of the rapidly developing ontological computer based methodologies. The advantages of the ontology development, as summarized by Smith et al. [1] include:

- Data integration across various biomedical domains and scientific fields
- Researchers create and maintain the high quality of the ontology by a process of peer review
- An ontology is designed to work well with the semantic web
- An ontology provides pretested and well-defined terms that can be used in the development of new databases
- An ontology enables the integration of traditional terms and common nomenclatures such as SNOMED-CT
- An ontology can be designed to contain standard terminologies that allows the data to be easily accessible to other researcher groups as required by governmental institutions such as the National Institute of Health.

How will the development of an ontology for the field of oral cancer be accomplished? This depends upon the degree to which members of the oral cancer community believe in the importance of the creation of ontologies similar to the one we have presented here. It is our view that oral cancer related ontologies can become as indispensable as the GO is to scientists working in the field of genomics. The oral cancer community through its various organizations such as the American Academy of Oral and

Maxillofacial Pathology, International Association of Oral Pathologists, the International Academy of Oral Oncology and North American Society of Head and Neck Pathology should address the issue of the importance of using the ontological methods described here to establish a comprehensive data base that would integrate the past, current, and evolving knowledge of the field. This would permit the oral cancer community to meet the challenge posed by Smith et al. [1] of dealing with the “uncontrolled explosion of different ways of describing information” by creating an electronic “consensus-based controlled structured vocabulary” concerning all aspects of oral oncology that would make the data available for searches and algorithmic processing.

We have proposed a solution to the terminology dilemma involving the common, but ill-defined phrase, “oral cancer”, a dilemma that has prevented effective retrieval of focused oral cancer information from the literature. We have proposed the creation of an oral cancer ontology-based on established ontological principles that captures and integrates site (ICD-O) and type (WHO) of malignant neoplasm information from all papers on oral cancer. The proposed oral cancer ontology will enable researchers to easily access information regarding location and type of mouth region cancer. Furthermore, the oral cancer ontology can be expanded to include, for example, information concerning demographic data, risk factors, staging and grading. In addition, as the field of biomedical ontology expands, our system will help to promote biomedical data integration and semantic interoperability with other existing ontologies, thereby expanding the search capacities of scientists interested in advancing knowledge of cancer of the mouth region.

Acknowledgments The authors wish to express their gratitude to Dr. Mirdza Neiders and Dr. Alfredo Aguirre for their invaluable help in reviewing the manuscript.

References

- Smith B, Goldberg LJ, Ruttenberg A, et al. Ontology and the future of dental research informatics. *J Am Dent Assoc.* 2010; 141:1173–5.
- Moore SR, Pierce AM, Wilson DF. ‘Oral cancer’—the terminology dilemma. *Oral Dis.* 2000;6:191–3.
- Moreno Lopez LA, Esparza Gomez GC. Oral cancer versus oropharyngeal cancer versus head and neck cancer. *Oral Oncol.* 1998;34:576–7.
- McCartan B. ‘Oral cancer’—the terminology dilemma. *Oral Dis.* 2001;7:66.
- Stricker TP, Kumar V. Neoplasias. In: Kumar V, Abbas AK, Fausto N, Mitchell R, editors. *Robbins basic pathology*. 8th ed. Philadelphia: Saunders Elsevier; 2007. p. 174.
- Sankaranarayanan R. Oral cancer in India. An epidemiologic and clinical review. *Oral Surg Oral Med Oral Pathol.* 1990;69: 325–30.
- Tarvainen L, Kyyronen P, Kauppinen T, et al. Cancer of the mouth and pharynx, occupation and exposure to chemical agents in Finland [in 1971–1995]. *Int J Cancer.* 2008;123:653–9.
- Carvalho AL, Singh B, Spiro RH, et al. Cancer of the oral cavity. A comparison between institutions in a developing and a developed nation. *Head Neck.* 2004;26:31–8.
- Merletti F, Boffetta P, Ciccone G, et al. Role of tobacco and alcoholic beverages in the etiology of cancer of the oral cavity/oropharynx in Torino, Italy. *Cancer Res.* 1989;49:4919–24.
- Goldstein BY, Chang SC, Hashibe M, et al. Alcohol consumption and cancers of the oral cavity and pharynx from 1988 to 2009: an update. *Eur J Cancer Prev.* 2010;19:431–65.
- Moller H. Changing incidence of cancer of the tongue, oral cavity, and pharynx in Denmark. *J Oral Pathol Med.* 1989;18:224–9.
- Chandran R, Lalloo R, Myburgh NG, et al. Scientific. Risk of intraoral cancer associated with tobacco and alcohol—a case-control study. *SADJ.* 2005;60:326–8.
- Ostman J, Anneroth G, Gustafsson H, et al. Malignant oral tumours in Sweden 1960–1989—an epidemiological study. *Eur J Cancer B Oral Oncol.* 1995;31B:106–12.
- Moore SR, Johnson NW, Pierce AM, et al. The epidemiology of mouth cancer: a review of global incidence. *Oral Dis.* 2000;6: 65–74.
- Zini A, Czerninski R, Sgan-Cohen HD. Oral cancer over four decades. Epidemiology, trends, histology, and survival by anatomical sites. *J Oral Pathol Med.* 2010;39:299–305.
- de Camargo Cancela M, Voti L, Guerra-Yi M, et al. Oral cavity cancer in developed and in developing countries. Population-based incidence. *Head Neck.* 2010;32:357–67.
- Tarvainen L, Suuronen R, Lindqvist C, et al. Is the incidence of oral and pharyngeal cancer increasing in Finland? An epidemiological study of 17,383 cases in 1953–1999. *Oral Dis.* 2004;10:167–72.
- Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol.* 2009;45:309–16.
- Canto MT, Devesa SS. Oral cavity and pharynx cancer incidence rates in the United States, 1975–1998. *Oral Oncol.* 2002;38:610–7.
- Rodu B, Cole P. Oral cavity and pharynx-throat cancer in the United States, 1973–2003. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2007;104:653–8.
- Gillison ML. Current topics in the epidemiology of oral cavity and oropharyngeal cancers. *Head Neck.* 2007;29:779–92.
- Rojas-Alcayaga G, Brignardello R, Espinoza I, et al. Oral malignant tumours: a report on 728 cases in Chile. *Acta Odontol Latinoam.* 2010;23:136–42.
- Franceschi S, Talamini R, Barra S, et al. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. *Cancer Res.* 1990;50:6502–7.
- Chen J, Katz RV, Krutchkoff DJ, et al. Lip cancer. Incidence trends in Connecticut, 1935–1985. *Cancer.* 1992;70:2025–30.
- Zini A, Czerninski R, Vered Y, et al. Trends of oral and pharyngeal cancer in Israel, by gender, age, ethnic group, and country of origin. 1970–2006. *Community Dent Oral Epidemiol.* 2009;37:547–54.
- Morse DE, Kerr AR. Disparities in oral and pharyngeal cancer incidence, mortality and survival among black and white Americans. *J Am Dent Assoc.* 2006;137:203–12.
- Swango PA. Cancers of the oral cavity and pharynx in the United States: an epidemiologic overview. *J Public Health Dent.* 1996; 56:309–18.
- Kingsley K, O’Malley S, Ditmyer M, et al. Analysis of oral cancer epidemiology in the US reveals state-specific trends: implications for oral cancer prevention. *BMC Public Health.* 2008;8:87.
- Petti S, Scully C. Oral cancer: the association between nation-based alcohol-drinking profiles and oral cancer mortality. *Oral Oncol.* 2005;41:828–34.
- Tramacere I, Negri E, Bagnardi V, et al. A meta-analysis of alcohol drinking and oral and pharyngeal cancers. Part 1: overall results and dose-risk relation. *Oral Oncol.* 2010;46:497–503.
- Brown LM, Check DP, Devesa SS. Oropharyngeal cancer incidence trends. Diminishing racial disparities. *Cancer Causes Control.* 2011;22:753–63.
- International Statistical Classification of Diseases and Related Health Problems. Available from : <http://apps.who.int/classifications/apps/icd/icd10online/>. Accessed 13 June 2011.
- International Classification of diseases (ICD). Available from: <http://www.who.int/classifications/icd/en/>. Accessed 31 July 2011.
- Barnes EL, Evenson JW, Reichart P, Sidransky D, editors. *World Health Organization classification of tumours. Pathology and Genetics of the head and neck tumours*. Lyon, France: IARC Press; 2005. p. 164.
- National Cancer Institute Thesaurus. Available from: <http://ncit.nci.nih.gov/ncitbrowser/>. Accessed 13 June 2011.
- Smith B, Ceusters W, Klagges B, et al. Relations in biomedical ontologies. *Genome Biol.* 2005;6:R46.
- Protégé. Available from: <http://protege.stanford.edu/>. Accessed 3 Aug 2011.
- Rubin DL, Shah NH, Noy NF. Biomedical ontologies: a functional perspective. *Brief Bioinform.* 2008;9:75–90.
- Barrell D, Dimmer E, Huntley RP, et al. The GOA database in 2009—an integrated gene ontology annotation resource. *Nucleic Acids Res.* 2009;37(Database issue):D396–03.
- Huntley RP, Binns D, Dimmer E, et al. QuickGO. A user tutorial for the web-based gene ontology browser. *Database (Oxford).* 2009;2009:2–19.
- Smith B, Ashburner M, Rosse C, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnol.* 2007;25:1251–5.
- Ai J, Smith B, Wong DT. Saliva ontology: an ontology-based framework for a salivaomics knowledge base. *BMC Bioinform.* 2010;11:302.
- Lomax J. Get ready to GO! A biologist’s guide to the gene ontology. *Briefings Bioinform.* 2005;6:298–304.
- Costanzo MC, Park J, Balakrishnan R, Cherry JM, Hong EL. Using computational predictions to improve literature-based gene ontology annotations. A feasibility study. *Database (Oxford).* 2011;2011:1–8.
- The Open Biological and Biomedical Ontologies. Available from: <http://www.obofoundry.org/>. Accessed 3 Aug 2011.
- Rosse C, Mejino JL Jr. A reference ontology for biomedical informatics. The Foundational model of anatomy. *J Biomed Inform.* 2003;36:478–500.

47. Schofield PN, Bard JB, Booth C, et al. Pathbase. A database of mutant mouse pathology. *Nucleic Acids Res.* 2004;32(Database issue).D512–15.
48. Brochhausen M, Spear AD, Cocos C, et al. The ACGT master ontology and its applications—towards an ontology-driven cancer research and management system. *J Biomed Inform.* 2011;44:8–25.