AN INVESTIGATION OF PROGNOSTIC DETERMINANTS AMONG UPPER AERODIGESTIVE TRACT CANCER PATIENTS

PAUL J. ALLISON BDS, FDSRCS(Eng), MSc

FACULTY OF DENTISTRY
McGILL UNIVERSITY
MONTREAL

APRIL 1998

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of PhD.

CONTENTS

ACKNOWLEDGMENTS

First of all, I would like to thank my two co-supervisors, Drs. Jocelyne Feine and David Locker for all the intellectual, emotional and financial support they have given me throughout the four years of my PhD thesis work. I am extremely grateful for their generosity and confidence in taking me on as a student, their enthusiastic guidance of my work and their uncomplaining reading of countless drafts of various protocols, manuscripts and thesis sections.

I would also like to thank Drs. Eduardo Franco and Sharon Wood-Dauphinee for the considerable time they have taken to advise me on the research protocol and subsequent manuscripts relevant to their fields of expertise.

I would also like to express my gratitude to Drs. Martin Black and George Shenouda for their enthusiastic support of my research work, their advice on its conception and data collection and their giving me such ready access to patients under their care.

I would like to thank all the subjects of my study who freely and anonymously donated some of their time to give me data for my research project. And finally, I would like to thank Sophia for her support throughout my PhD.

PREFACE

Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly-duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.

If this option is chosen, **connecting texts that provide logical bridges between the different papers are mandatory.** The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.

The thesis must still conform to all other requirements of the "Guidelines for Thesis Preparation". **The thesis must include:** A Table of Contents, and abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.

Additional material must be provided where appropriate (eg. in appendices) and in sufficient detail to allow clear and precise judgement to be made of the importance and originality of the research reported in the thesis.

In the case of manuscripts co-authored by the candidate and others, **the candidate is required to make explicit statement in the thesis as to who contributed to such work and to what extent.** Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all authors of the co-authored papers.

Declaration of contribution of co-authors of manuscripts contained in this thesis

Manuscripts I (The role of professional diagnostic delays in the prognosis of oral cancer: A review of the literature) and IV (Quality of life: A dynamic construct) are both review articles whose original concept, preparatory work and original preparation was all performed by Paul Allison (the student). The role of the co-authors (David Locker and Jocelyne Feine) for both of these manuscripts was one of review and modification.

Manuscripts II (The role of professional diagnostic delays in the prognosis of upper aerodigestive tract carcinoma), III (Predictors of professional diagnostic delays for upper aerodigestive tract carcinoma) and V (Correlates of quality of life in upper aerodigestive tract cancer patients) all report the findings of statistical analyses of data collected as part of the research reported in this thesis. Again, the original conception of the research protocol, data collection and analysis and preparation of first drafts of these manuscripts was the work of Paul Allison (the student). The role of Dr. Feine in all three of these manuscripts was in supervising the preparation of the research protocol, advising on various statistical analysis strategies and reviewing and modifying the draft texts. In manuscripts II and III, the role of Dr. Franco was in advising on the most appropriate data to collect, suggesting alternative statistical analysis strategies and reviewing and modifying the draft texts. Similarly for manuscript V, Dr. Locker and Dr. Wood-Dauphinee independently advised upon statistical analysis strategies and reviewed and modified the draft text. Finally the role of Dr. Black in manuscripts II and V was to review the draft texts.

ABSTRACT - ENGLISH

The poor prognosis for upper aerodigestive tract (UADT) cancer patients has changed little over the previous 30 years. Recent research attention has subsequently focused on earlier diagnosis and post-treatment health-related quality of life (HRQL). The research reported in this thesis investigates both fields in a sample of UADT cancer patients.

UADT cancer diagnosis is very poorly understood. In view of this, the primary aims of this research were to investigate the association between diagnostic delays and patient prognosis, plus the predictors of such delays. Data were collected from UADT cancer patients diagnosed during an 18 month period at three McGill University Teaching Hospitals. Multivariate logistic regression was used to calculate odds ratios for predictors of late (as opposed to early) stage disease in a sample of 188 UADT cancer patients. Pharyngeal site, professional diagnostic delay >1 month and younger age were significant, independent predictors of increased odds for late stage disease. A further multivariate logistic regression was subsequently performed to calculate odds ratios for professional delays >1 month. The significant independent predictors of increased odds for such a delay were comorbidity present at the time of initial presentation of symptoms, younger age, lower education and non-oral site.

With respect to the investigation of post-therapeutic HRQL, a cross-sectional study design, using the self-complete EORTC QLQ-C30 and H&N37 module instruments as the dependent variable, was employed in a sample of UADT cancer patients. Multivariate regression analysis was used to calculate the relative contribution of clinical and sociodemographic variables in explaining the variance in the global domain of the EORTC QLQ-C30 instrument. Unemployment, female gender, older age, having teeth and more advanced disease stage were significant independent predictors of a worse global evaluation of HRQL, while oral site was a significant predictor of a better global evaluation of HRQL.

Overall, the results of this study suggest that, in a sample of UADT cancer patients, disease site and patient age are independent predictors of diagnostic delays, diagnostic disease stage and post-therapeutic HRQL.

RÉSUMÉ - FRANÇAIS

Le pronostic sévère des cancers des voies aérodigestives supérieures (VADS) ne s'est guère amélioré depuis 30 ans. L'attention des chercheurs s'est donc récement portée sur le diagnostic précoce et la qualité de vie (QV) après le traitement. Les recherches décrites dans cette thèse ont été menées dans ces domaines à partir d'un échantillon de patients atteints de cancer des VADS.

Le diagnostic du cancer des VADS est très mal compris. C'est pourquoi cette recherche visait avant tout à étudier le rapport qui existe entre le délai de diagnostic et le pronostic, ainsi que les paramètres de prédiction de tels délais. Les données ont été recueillies sur 18 mois auprès de patients chez qui on avait diagnostiqué un cancer des VADS dans trois hôpitaux d'enseignement de l'Université McGill. Grâce à une analyse de régression logistique, on a calculé le rapport de cotes associé aux paramètres de prédiction d'un stade avancé (par opposition à précoce) de la maladie dans un échantillon de 188 patients atteints d'un cancer des VADS. Le siège du cancer du pharynx, le délai de diagnostic professionnel >1 mois et l'âge décroissant sont des paramètres de prévision indépendants du risque accru de détection au stade avancé de la maladie. Une autre analyse de régression logistique multiple a permis de calculer le rapport de cotes pour le délai professionnel >1 mois. On a ainsi constaté que la présence de comorbidité à l'apparition des symptômes, l'âge décroissant, le niveau d'instruction moins élevé et le fait que le siège du cancer ne se trouve pas dans la cavité buccale sont des facteurs de prédiction indépendants de risque accru au titre d'un tel décalage.

Pour évaluer la QV, nous avons mené une étude transversale auprès d'un échantillon de patients atteints de VADS. Nous avons à cette fin utilisé les intruments

modulaires QLQ-C30 et H&N37 de l'OERTC comme variables dépendantes. Nous avons calculé au moyen d'une analyse de régression multiple l'effet relatif des variables cliniques et socio-démographiques sur la variance dans le domaine global de l'instrument QLQ C30. Le chômage, le sexe (féminin), l'âge croissant, le fait pour le patient d'avoir conservé ses dents et le stade d'évolution plus avancé sont des paramètres prévisionnels indépendants d'une évaluation globale moins favorable de la QV, tandis que le siège du cancer dans la cavité buccale est un paramètre prévisionnel indépendant d'une évaluation globale plus favorable de la QV.

En générale, les resultats de cette étude incitent à croire que les patients atteints d'un cancer du pharynx courent plus de risques de subir un décalage de diagnostic et d'être diagnostiqués à un stade avancé de la maladie et subissent une baisse QV, tandis que le patients plus âgés courent moins de risques de subir un décalage de diagnostic et d'être diagnostiqués à un stade avancé de la maladie, même s'ils subissent eux aussi une baisse de QV.

oral, pharyngeal and laryngeal cancers (4). Another form of tobacco consumption which is causing increasing concern in Western countries is smokeless tobacco. The strongest evidence linking smokeless tobacco and intra-oral cancer relates to the "snuff-dipping" habits of people in the southern USA (5-7). In Canada, one study has reported several cases of verrucous carcinoma in association with long term tobacco chewing (8). However, data on the consumption of smokeless tobacco in Canada are scant, especially among non-native peoples. Main and Lecavalier reviewed smokeless tobacco and oral disease in relation to the implications for Canadian public health, and they reported that 0.7% of Canadian non-native males over 15 years of age chew tobacco and 0.4% use snuff (9). However, they felt these figures were an underestimate. Studies in the USA and Canada have reported that the use of smokeless tobacco among native American Indians is much more popular. Millar and Rensberg have estimated that 17% of students in the Northwest Territories have tried smokeless tobacco and 9% were regular users (10). Of these people, the vast majority were native American Indians. Among these Indians, and those in the USA, the use of smokeless tobacco appears to be as common among women as it is among men (9). In summary, regarding smokeless tobacco use and cancer, the IARC have concluded that: i) there is *sufficient* evidence that *oral use of snuff* is carcinogenic to humans; ii) there is *limited* evidence that *chewing tobacco* is carcinogenic to humans; and iii) there is *inadequate* evidence that *nasal use of snuff* is carcinogenic to humans (11).

The role of alcohol consumption in the aetiology of UADT cancers is less clear in that it is recognized that ethanol itself is not a carcinogen for the UADT (12), however high levels of alcohol consumption have been consistently associated with an increased risk for UADT cancer (13-17). Most investigators have concluded that tobacco smoking and alcohol consumption combine to increase the risk for UADT cancer in some form of multiplicative relationship (13-19). The exact mechanism by which alcohol is associated with UADT cancers is not clear, however a number have been suggested: i) nutritional deficiencies associated with heavy drinking; ii) the

effects of contaminants and cogeners in alcoholic drinks; iii) the induction of michrosomal enzymes which enhance the metabolic activation of tobacco and other carcinogens; and iv) the capacity of alcohol to act as a solvent facilitating the diffusion of carcinogens through the epithelium (20). In recent years, increasing attention has focussed upon the malnutrition which tends to accompany excessive alcohol consumption (20). Decreased risk for oral and pharyngeal cancers has been shown with increased consumption of vitamins A and C, plus fresh fruit and vegetables (21-23) and more recent work from Brazil (18) and the USA (24) has shown a further protective effect of fresh fruit, not explained by beta-carotene, vitamin C or fibre content. Further recent studies in Italy have shown increased risk for oral and pharyngeal cancer associated with the consumption of rice, maize, pasta, polenta, cheese, eggs and pulses (25,26). However, although the exact nature of this relationship is not clear, it is felt that the increased consumption of these foods is likely to be an indicator of a deficient diet (3).

Other aetiological factors associated with UADT cancers have largely been investigated separately for oral and pharyngeal cancers and laryngeal cancers. With regard to the former group, evidence is beginning to accumulate concerning the link between the intra-oral presence of certain human papillomaviruses and increased risk for oral cancers. However, two recent reviews of the literature concerning the role of these viruses in the aetiology of oral carcinoma have concluded that, while there is some epidemiological and experimental evidence supporting an association, further prospective research is required before an independent, or even a synergistic, aetiological role can be recognized for human papillomaviruses (27,28).

Other risk factors which have been associated with oral cancers include dentition, oral hygiene and mouthwash, and although there is some evidence supporting a link between poorer dental and oral hygiene status and increased risk for oral cancer (29-31), it is recognized that these may be indicators of socio-economic status rather than independent aetiological factors (3).

Finally, research concerning aetiological factors for laryngeal cancer other

than tobacco and alcohol have concentrated largely upon occupational and gender-related risk factors. There is some evidence to support a link between asbestos and laryngeal cancer (32,33), however this is controversial as several well-controlled studies have failed to detect an association (34). Similarly other studies have linked nickel, mustard gas, wood dust and strong acid with laryngeal cancer, but the evidence is inconclusive (34). Male gender has been suggested as a risk factor for a subgroup of laryngeal cancers because of the excessive male:female ratio noted for glottic cancers (35).

In summary, the principal aetiological factors associated with UADT cancers in the developed world are tobacco smoking and alcohol consumption, with a small proportion of their incidence possibly being attributable to certain forms of smokeless tobacco, dietary insufficiency, certain high risk human papillomaviruses and various occupational exposures.

1.3 World, Canadian and Quebec incidence data for UADT cancers

On a worldwide basis, Parkin et al estimated that cancers of the mouth and pharynx collectively accounted for approximately 5.4% of all new cancer cases among men and women throughout the world in 1985 (36). However, quite apart from the fact that laryngeal cancers were not grouped with oral cancers in this analysis, this crude figure masks enormous differences in incidence rates by region in the world, by gender and by age group. Categorising incidence by gender and within developed or developing countries, Parkin et al estimated that, among males, UADT cancers ranked 5th and 3rd in incidence in developed and developing countries respectively, and 12th and 4th among women in developed and developing countries respectively (36). Examining the geographical variation in incidence more specifically, the age-standardized incidence for cancers of the oral cavity (ICD-9 143-5) among males in the Bas-Rhin region of France (13.4/100,000[1]) is 22 times that

[1]Unless otherwise stated, all subsequent incidence figures quoted will be per 100,000 population

among males in the Yamagata region of Japan (0.6); and the age-standardized incidence for cancers of the larynx (ICD-9 161) among males in the Basque region of Spain (20.4) is 12 times that of males in the Khon Kaen region of Thailand (1.7) (37).

These wide variations in incidence are a reflection of differing regional aetiologies. For the purposes of this research, it is therefore most pertinent to examine the incidences of UADT cancers in North America and North-Western Europe. Once again using Parkin et al's crude estimates of cancer incidence for comparison, it is evident that, among the North American, Northern European and Western European regions, the latter has the highest incidence of both oropharyngeal and laryngeal cancers, while Northern Europe has the lowest incidence of both (36). Within Canada itself, among males, the 1997 age-standardized incidence rates predicted for the whole country are 15 for oral cancers (ICD-9 140-9) and 7 for laryngeal cancers (38). Among Canadian females, this figure is only available for oral cancers (1997 predicted incidence: 5) because the figure for laryngeal cancers is so low (38). The latest data for cancer incidence in Canada show that, compared to the other provinces and territories, Quebec appears to have a lower than average incidence of oral cancers (ICD-9 140-9) among men and women but the highest incidence of laryngeal cancer in Canada for both sexes (38). However, these data include lip cancers in the "oral" group (ICD-9 140-9). Other data published by the Canadian Cancer Registry in which lip cancer (ICD-9 140) has been separated from oral cancers (ICD-9 141-9), demonstrate that, among Canadian males, Quebec has a much lower than average incidence for lip cancers (ICD-9 140) but a much higher than average incidence of oral cancers (ICD-9 141-9) (39). However, even within these figures concerning "oral" cancers (ICD-9 141-9) there is a lack of homogeneity. The oral cancer incidence for both genders is highest in the Northern Territories (39); a product of the well-recognized, elevated incidence of both nasopharyngeal and salivary gland cancers in the Inuit population (39). It would appear, therefore, that Quebec has the highest incidence, among Canadian males, of the cigarette and

alcohol-induced UADT carcinomata that are the subject of this thesis.

Apart from the geographical determinants of UADT cancer incidence, the two other most important determinants are gender and age. As has previously been said, these cancers have a higher incidence in males than females. The exact ratios vary, however, according to data for the 1980s, the male to female ratio for oral cancers in North America and North-western Europe is approximately 2-3:1, while the same ratio for laryngeal cancers is approximately 5:1 (36). These data agree with estimates for oral cancer incidence in Canada in which there is a 3:1 male to female ratio (38). With respect to age, as would be expected with a cancer largely associated with chronic exposure to cigarette smoking and excessive alcohol consumption, the highest incidence of UADT cancers is among older people. Among oral cancers (ICD-9 140-9), it has been estimated that 98% of cases arise in people aged 40 years or more (40), while the median age for the diagnosis of laryngeal cancers in Canada has been estimated to be in the 55-64 years for women and 65-74 years for men (41).

Finally, in terms of the relative incidence of UADT compared to other cancers in Canada, oral cancers (ICD-9 140-9) are predicted to comprise 3.1% of new cancers among Canadian males in 1997, making them the 5th most common cancer for this population (38). (These data for women and laryngeal cancers are not given due to the relatively low incidence of UADT cancers among women.)

In summary, therefore, it would seem that the incidence of UADT cancers in Canada is lower than in western Europe (in particular, France and Spain) but higher than in the USA and northern Europe (in particular, Scandinavia and the UK). Within Canada, Quebec males have a higher than average incidence of oral cancers (ICD-9 141-9), and both Quebec males and females have the highest incidence of laryngeal cancers. Furthermore, in Canada and Quebec, UADT cancers have a higher incidence among males and the older population.

1.4 Trends in incidence for UADT cancers

In examining incidence trends for UADT cancers, once again, most of the

literature documents data for oral and laryngeal cancers separately. As far as oral cancers are concerned, their increasing incidence, especially among younger males, has been documented in the USA (42), Scotland (43), Scandinavia (44) and Denmark (45). This evidence is supported by a recent age-cohort analysis of European oral cancer mortality data in which there is an increase in male mortality rate in the majority of countries, especially evident among younger males (35-64 years) (3). In Canada, although no analysis of oral cancer incidence by age groups is available, the age-standardized incidence of these cancers (ICD-9 141-9) over the period 1969-1988 has shown a steady and significant increase from 8.2-10.2 among males, and a slight increase from 3.5-3.9 among females (39). The trend in oral cancer incidence over the same period among Quebec males shows an increase until the period 1979-1983, before falling during the period 1984-1988, while no particular pattern is evident among Quebec's women (39).

With respect to laryngeal cancer in Canada, there was a dramatic increase in incidence documented variously as a 50% increase for males and nearly 100% increase for females over the period 1970-1980 (41) or a 4.3% per year increase for males and 6.8% per year increase among females over the same period (46). These trends in the incidence of laryngeal cancer among Canadian males and females were mirrored in the USA, where the age-standardized incidence for males increased from 5.6 to 9.0, and for females from 0.5 to 1.5, over the period 1947-1984 (47). Similarly, in Quebec, the female incidence increased from 0.9 to 1.9 and the male incidence increased from 6.7 to 10.6 (39). However, over the 10 year period 1985-1994 among Canadian men, laryngeal cancer incidence has been falling by approximately 1.6% per year (38), and after showing a 3.6% per year increase during the period 1982-1989 (48), laryngeal cancer incidence among Canadian women has shown a decrease of 0.7% per year during the period 1985-1994 (38). Thus, it would appear that the incidence of laryngeal cancer among Canadian men has been falling for over a decade, while that for Canadian women has just begun to fall during the past few years.

In summary, due to the lack of a detailed analysis of the trends for UADT cancers in Canada and Quebec, it is difficult to be certain as to the pattern of these diseases in this country. It is evident that there is a general trend for decreasing incidence in these cancers, however, we do not know whether the age-cohort increases reported among younger males in the USA and Europe are present in Canada.

1.5 The survival of UADT cancer patients

Once again, the majority of data concerning the survival rates and mortality following UADT cancers is divided into oral and laryngeal cancers, with the former often including lip cancer which has a much better prognosis than other oral sites. The crude worldwide five-year survival rates for lip and oral cancers are 80% and 30-40% respectively (40). In the USA, the crude oral cancer (ICD-9 140-9) five-year survival rates has remained 55% for whites and 31-36% for blacks over the past 20 years, while the same figure for laryngeal cancer has remained 66-69% for whites and 52-59% for blacks (49). These figures are similar to European data which show no improvements in survival rates for UADT cancers (50,51). Indeed, there is even some evidence, from Switzerland, that survival rates worsened for laryngeal cancer patients between 1974 and 1983 (52). In Quebec, the crude oral cancer (ICD-9 140-9) five-year survival rates during the period 1984-1986 were 52% for females and 40% for males, while the figures for laryngeal cancer during the same period in Quebec were 59% for females and 52% for males (53).

1.6 Summary of UADT cancer epidemiology and justification of research aims

In Canada, UADT cancers are a public health problem, particularly among men, largely caused by cigarette smoking and excessive alcohol consumption. They are a largely preventable health problem, approximately 50% of whose victims will die during the five year period following their diagnosis. This survival rate has remained approximately the same for the past 30 years.

Probably as a reflection of the apparent lack of progress in trying to cure people of their UADT cancer, several major research themes, other than conventional treatments, have recently emerged in the field of UADT cancers. The European School of Oncology Advisory Report to the European Commission for the Europe Against Cancer Programme specifically recognized the lack of progress in improving survival among oral and pharyngeal cancer patients over the past few decades and subsequently recommended research in several areas. Among these areas were the natural history of oral cancers and the investigation of biomarkers, both aimed at promoting earlier diagnosis, plus the investigation of post-treatment quality of life for patients (54). Similarly, in the USA, a National Strategic Planning Conference for the Prevention and Control of Oral and Pharyngeal Cancer recommended a whole series of research strategies aimed at improving prevention of these cancers. Among these strategies, the need for behavioural research integrated with other fields was highlighted (55). In other words, recognising the lack of prognostic improvement for UADT cancer patients brought about by the trial of various treatments, a number of organisations are now promoting a shift of research priorities into screening and early detection strategies as a means to improve prognosis, and upon health status/quality of life assessment following therapy as a means to improve rehabilitation. Indeed, there is a direct link between these fields in that diagnosis earlier in the disease process should result in less aggressive treatment regimes, which should consequently lead to improved post-treatment health status and quality of life for this and other groups of cancer patients. Thus earlier diagnosis should lead to improved patient prognosis in its holistic sense: improved survival rates and improved post-treatment health status and quality of life.

In summary, this introduction gave a broad overview of the current state of knowledge concerning the aetiology, epidemiology and prognosis of UADT cancers (with particular reference to Canada and Quebec). It is evident, however, that only limited progress in reducing the incidence and improving the prognosis for these cancers has been achieved over the past few decades. In view of this situation, the

aim of the research reported in this thesis was to investigate two areas whose understanding could improve the prognosis for UADT cancer patients:

1. An investigation of diagnostic determinants of prognosis among UADT cancer patients; it has been recognized that, with our current state of knowledge, population or even targeted screening programmes for oral cancers cannot be recommended because of the lack of any evidence of effectiveness (54). Nevertheless, it is imperative that every effort is made to improve our understanding of the natural history of UADT cancers and the diagnostic process, so that some form of intervention can be developed which promotes diagnosis of precancerous lesions and early cancer (56). The aim of the first part of the research was, therefore, to investigate the role of diagnostic delays in determining patient prognosis among UADT cancer patients.

2. An investigation of the post-treatment health status outcomes of UADT cancer patients; patient mortality and survival is an extremely important indicator of patient prognosis for all forms of cancer and many other forms of disease. However, it is a crude indicator of outcome which takes no account of the quality of a patient's life following the diagnosis of their cancer. In recent years, with the failure of western medicine to develop treatments which show anything but very marginal improvements in survival for many cancers, increasing attention has turned to the investigation of patient health status and life quality following a cancer diagnosis. UADT cancers are a very good example of this situation. As with cancers at other sites, investigators working in the field of UADT cancers are increasingly documenting the health status and quality of life problems resulting from these cancers and their treatment, although our understanding of this complex issue is still in its infancy. The aim of the second part of the research reported in this thesis was, therefore, to investigate the determinants (or at least correlates) of health status in a sample of UADT cancer patients following their therapy.

The dependent variable in the first section is disease stage, a well-recognized predictor of patient prognosis in many forms of cancer including UADT cancers. Disease stage was carried through into the second phase of the research in which it was hypothesised as one of the principal determinants of health status among post-therapeutic UADT cancer patients. The thread of these two sections of research is therefore, disease stage. Although not specifically tested, it is the broad hypothesis of this research project that diagnostic delays among UADT cancer patients lead to more extensive disease (ie, later stage disease) at diagnosis, and this subsequently results in a worse health status (in addition to worse survival rates) following treatment. In view of these two connected but different lines of investigation, the following text will describe the work in two separate sections devoted to each of the aforementioned research aims, before bringing the findings of all the research together in one final section in which the results of the research as a whole will be summarized and the implications for future work will be discussed.

SECTION 2

AN INVESTIGATION OF DIAGNOSTIC DETERMINANTS OF PROGNOSIS AMONG UADT CANCER PATIENTS

2.1 An introduction to the investigation of the determinants of prognosis among UADT cancer patients

As with all other cancers and diseases, the determinants of patient prognosis in this population can broadly be categorized into three groups:

1. ***Biological variables;***
2. ***Patient-related socio-demographic and behaviour variables;*** and
3. ***Health care professional-related behaviour variables.***

It is arguable whether a fourth category of factors associated with the health care system (HCS) should be added separately, or whether factors associated with the HCS should be integrated with the determinants of patient and health care professional (HCP) behaviour. However, in view of the fact that it is not the aim of this research to compare the effects of differing HCSs on patient prognosis, for the purposes of this discussion and research, factors associated with the HCS will be integrated into the determinants of patient and HCP behaviours. Furthermore, it needs to be recognized that there may be some overlapping in this categorization. For instance, in a certain context, patient age and gender can be considered biological determinants of UADT cancer patient prognosis, while in a different context they can be considered as sociodemographic determinants of patient behaviour.

The vast majority of research investigating the determinants of UADT cancer patients' prognosis has concentrated upon biological variables, with a small amount of work investigating patient behaviour-related variables and very little research examining the role of HCP behaviour-related variables in determining patient prognosis. I will review the literature concerning the determinants of UADT cancer

patients' prognosis, using the categorization outlined, before explaining the specific objectives of the research and reporting the findings.

2.2 Biological determinants of UADT cancer patients' prognosis

The best-recognised biological indicators of UADT cancer prognosis are those included in the TNM cancer staging systems described by the American Joint Committee of Cancer (AJCC) and the Union International Contre le Cancer (UICC). The latest version of the staging system is identical for the two organizations and includes the size of the primary cancerous lesion (the T category), the involvement or not of regional lymph nodes (the N category) and the existence of distant metastatic disease (the M category) (57). The T, N and M categories are defined for each of the UADT (and indeed other anatomic) sites, and then these three indicators are combined to produce a disease stage categorization from I (the least advanced cancer) to IV (the most advanced cancer) for each cancer patient (57). (Table 1 illustrates the stage grouping which is recognized for UADT cancers.)

Table 1. Stage grouping for UADT cancers (taken from ref. 57)

Stage	T category	N category	M category
Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IV	T4	N0	M0
	T4	N1	M0
	Any T	N2	M0
	Any T	N3	M0
	Any T	Any N	M1

NB. Tis: carcinoma in situ

These three biological markers are used because of their significance in the development of a cancer which grows from carcinoma in situ at the primary site, to the point where it spreads typically through the local lymphatic system to regional lymph nodes and then metastasizes to distant anatomical sites. This is well illustrated by the five-year survival rates for oral cancers (ICD-9 140-9) among USA whites during the period 1986-91. The figures for local (any N0, M0 lesion), regional (any N≥1, M0 lesion) and distant (any M1 lesion) cancers were 81%, 43% and 20% five-year survival rates respectively (49). Similar graduations in survival rates were evident for blacks and whites at all cancer sites (49).

Although there are criticisms of the AJCC/UICC staging system, it is a reasonable predictor of survival for UADT cancer patients and is popular with clinicians because of its ease of application. Langdon et al suggested a modification of the TNM system for oral cancers, including anatomical subsite and a pathological categorization of the lesion (58). Initial analysis seemed to suggest that this more complex system was a better predictor of survival (58), however, a subsequent analysis of the same data controlling for patient age and gender and using disease-specific death as the outcome showed that the conventional TNM system is nearly as good a predictor of survival as the more complex system (59). Thus, the simpler system has retained its popularity, especially with clinicians.

Apart from tumour size, regional lymph node involvement and distant metastatic involvement, many other biological predictors of prognosis were investigated by the multi-centre German, Austrian, Swiss Association for Head and Neck Tumours (DOSAK) in a retrospective analysis of data from 802 oral cancer (ICD-9 140-6) patients treated during the period 1952-72 and followed-up until 1978 (60). In a multivariate analysis of 18 pretherapeutic factors, 7 emerged as significant predictors of survival: tumour size; a clinical assessment of tumour infiltration; degree of histological differentiation; tumour site; regional lymph node involvement; distant metastatic involvement; and patient age. Once treatment modalities had been entered into the analysis, the significant predictors of a poor survival were: tumour

size >4 cm; a clinical assessment of tumour infiltration >5 mm; palpable fixed lymph nodes; the presence of distant metastases; and age >70 years (60). A number of publications have subsequently confirmed the findings of the DOSAK study in reporting the prognostic value of some of the variables in that study. The degree of tumour infiltration or tumour thickness has been demonstrated as a predictor of both regional lymph node involvement and overall patient survival among oral cancer patients (61-63), as has histological grading (involving ordinal categorization of degree of keratinization, nuclear polymorphism, pattern of invasion and host response at the deep invasive margins of the tumour) among UADT cancer patients (64-66). The role of tumour site as a predictor of prognosis in laryngeal cancer is well recognised, with glottic cancers having a better survival rate than cancers at other laryngeal subsites (67). Furthermore, as has already been noted, laryngeal cancer subsite is related to gender in that relative to supra- and sub-glottic cancers, glottic cancer is more common in men (35).

Since the 1980s, several avenues of research for prognostic indicators in cancer patients have emerged, all with promising but, as yet, equivocal results. The measurement of cell kinetics using flow cytometry is based on the theory that faster replicating tumours are more aggressive and so a measure of the rate of cell replication in a tumour will be predictive of prognosis (68). This technique has been applied to UADT cancer patients in on-going studies with mixed results. Begg has concluded that, in view of the fact that most of the studies involve small numbers and that few are randomised, controlled trials, we will have to wait for these studies to accrue larger numbers before drawing firm conclusions as to the predictive power of cell kinetics for UADT or any other cancers (69). Another area of research in cancer genetics has been in the attempts to show links between various oncogenes and tumour suppressor genes and tumour behaviour and patient prognosis. The most promising of these is p53, the product of a tumour suppressor gene, whose mutation is one of the most commonly detected abnormalities in human cancer (70). Once again however, although detection of mutations in the p53 gene is relatively easy, so

making the technique popular, the results of studies investigating its' ability to predict patient prognosis are mixed. Thus, the use of p53 expression as a predictor of patient prognosis is of equivocal value (71).

Finally, a completely different avenue has lain in the investigation of the association between tumour angiogenesis (the development of microvessels around the primary tumour) and patient prognosis. This is one of the latest lines in the search for prognostic predictors for cancer patients and the first studies show promising results. A positive association between microvessel proliferation and patient survival has been demonstrated for breast, gastric, prostate and non-small-cell lung carcinoma, plus germinal malignancies of the testes and brain (72,73) and more recently for oral (74) and laryngeal cancers (75).

In summary, the best biological indicators of UADT cancer patient prognosis are tumour site, size, regional and distant metastatic involvement, histological grading and depth of infiltration. Further biological indicators whose ability to predict patient prognosis shows promise include cell kinetics, p53 expression and tumour angiogenesis. These prognostic indicators can be reduced to three fundamental biological determinants of patient prognosis: the primary tumour site; the aggressivity or rate of growth of the tumour; and the degree to which the disease has progressed. These variables relate principally and primarily to the behaviour of the tumour, however, in addition, they include an element of host response.

2.3 Patient-related socio-demographic and behaviour variables as determinants of UADT cancer patients' prognosis

As has already been alluded to, one of the most readily observed variables associated with prognosis is patient gender. In Quebec, the crude oral cancer (ICD-9 140-9) five-year survival rates during the period 1984-1986 were 52% for females and 40% for males, while the figures for laryngeal cancer during the same period in Quebec were 59% for females and 52% for males (53). These figures are borne out by others from elsewhere, for instance the five-year survival rate for males with

UADT cancer in the Vaud Region of Switzerland during the period 1979-83 was 38%, while that for females was 46% (52). It is therefore evident that females with a UADT cancer tend to have a better survival rate than their male peers.

Similarly, readily available data demonstrate differences in survival associated with patient age. Once again, data from Quebec illustrate the differences in 5-year survival by age group. (The figures are shown in Table 2.)

Table 2. Five-year survival rates for oral (ICD-9 140-9) and laryngeal (ICD-9 161) cancers, Quebec, 1984-86

Gender	Age group	Crude 5yr survival rate		Disease-specific 5yr survival rate		% of life expectancy realized	
		oral	larynx	oral	larynx	oral	larynx
Female	0-44yrs	75%	73%	75%	74%	-	-
	45-54yrs	59%	71%	60%	73%	-	-
	55-64yrs	59%	66%	62%	69%	-	-
	65-74yrs	42%	46%	47%	52%	16.9%	25.2%
	75+yrs	32%	30%	50%	47%	17.2%	18.7%
Male	0-44yrs	67%	86%	67%	87%	-	-
	45-54yrs	49%	62%	51%	64%	16.7%	-
	55-64yrs	40%	61%	44%	68%	14.7%	-
	65-74yrs	36%	42%	46%	53%	20.5%	31.3%
	75+yrs	29%	24%	56%	44%	25.4%	30.5%

Data taken from Pelletier G, 1993 (53)

Looking at the crude five-year survival rates for oral and laryngeal cancers, there is a steady decline with increasing age. However, it needs to be recognized that crude five-year survival rates will drop with increasing age in any adult population due to the increased probability, with older age, of dying due to other causes. The disease-specific survival rates are therefore more relevant indicators of prognosis for UADT cancers. Here we see a more complex pattern wherein, among oral cancer patients, survival appears to initially fall and then stabilize or, particularly for males, actually improve with increasing age. These figures for disease-specific survival

among oral cancer patients are supported by the data for the percentage of life-expectancy realized, which increases a small amount for both males and females in old age. Disease-specific survival rates among laryngeal cancer patients however, fall with increasing age. It is evident, therefore, that UADT cancer patient survival varies with age, although the patterns are different with oral and laryngeal cancer patients.

Another well-documented socio-demographic indicator of survival is ethnicity. This is perhaps most easily seen in the USA where cancer registries routinely categorize cancer cases into "white" and "black" Americans. Five-year survival rates over the period 1986-91 for patients in the USA with oral cancers (ICD-9 140-9) were 55% for whites and 33% for blacks (49). This better survival rate for whites when compared to blacks was consistent for cancers at all sites (49). However, it should be recognized that, depending upon the geographical source of the data, ethnicity is likely to be an indicator of socio-economic status. The potential associations between socio-economic status and patient prognosis has seen very limited investigation among UADT cancer patients, however, in relation to the field of breast cancer where ethnic minorities consistently have a poorer prognosis than do whites, Facione has concluded that differences are due to issues of poverty and access rather than ethnicity itself (76). If this is the case, then it would seem highly likely that the association between socio-economic status and patient prognosis is mediated through one or more intermediate variables, most of which will concern some form of patient behaviour, eg, rate and period of consumption of aetiological factors, delay in presentation for diagnosis and response to diagnosis and treatment plan.

The effect of the period of time and the rate at which aetiological factors (principally cigarettes and alcohol) are consumed may be direct or through an intermediate variable like patient delay in presentation for diagnosis. A number of studies have investigated the association between aetiological factor consumption and prognosis for UADT cancer patients. In a univariate analysis of data from a

sample of tongue cancer patients, Johnston and Ballantyne found increased tobacco and alcohol consumption to be associated with higher disease- and non-specific mortality rates and higher incidence of recurrent or new disease (77). Subsequently, stronger evidence from studies using multivariate analyses has confirmed the link. In a sample of oral cancer patients Silverman et al found that those who quit cigarette smoking had a reduced risk for a second primary tumour (78). Browman et al found that UADT cancer patients who continue smoking through their radiotherapeutic treatment had poorer treatment response and survival rates than those who stop (79). Pradier et al found that alcohol consumption was a significant independent predictor of survival in a sample of laryngeal cancer patients (80), and Bundgaard et al found that tobacco consumption was an independent predictor of survival in a sample of oral cancer patients (81). In addition, using disease stage (the TNM system) as an outcome, Elwood and Gallagher found low alcohol consumption to be associated with early stage oral cancer (82), and Kowalski et al found late stage lip cancer to be associated with increased alcohol consumption (83). Only the latter of these studies (83) used a multivariate analysis in which diagnostic delays were included in the model with aetiological factor consumption. However, on balance, it seems likely that tobacco and/or alcohol consumption are directly associated with UADT cancer patient prognosis, but this does not preclude their also being indicators of delayed patient presentation.

The final potential determinant of prognosis among UADT cancer patients in this category of patient sociodemographic and behavioural variables, is delayed presentation for diagnosis. For a detailed discussion of this issue, please see manuscript I (section 2.5) of this thesis.

2.4 Health care professional-related behaviour variables as determinants of UADT cancer patients' prognosis

As previously mentioned, virtually no research has been conducted on this issue in relation to the diagnosis and treatment of UADT cancers. The potential for

HCPs' behaviour to affect patient prognosis basically revolves around their diagnostic and treatment decision-making processes. For a discussion of the role of HCP diagnostic delays in the prognosis of UADT cancer patients, see manuscript I of this thesis. With respect to the role of therapeutic treatment decisions in the prognosis of UADT cancer patients, there has been very little research, but what little evidence there is suggests that the nature of these decisions is largely dictated by specialty training and geography, and is, at best, idiosyncratic and, at worse, prejudiced. A study of a variety of specialists treating UADT cancer in the UK found that physicians often differed on the aim of treatment (eg, palliation or cure) and made subjective value decisions (84). Another, much larger, multinational study focusing on the treatment of glottic laryngeal cancer showed that, apart from the extent of disease, the most significant variables influencing treatment recommendations were the physicians' specialty and location of practice rather than predicted treatment outcome (85). The dilemma concerning the treatment of glottic cancer is radical surgery (ie, removal of the larynx with all its subsequent problems) or radiotherapy with laryngeal preservation. The aim of both treatments is cure. Recognizing the lack of high quality evidence from randomized trials, Stalpers et al had, nevertheless, previously found no difference in survival following surgery or radiotherapy, the inference being that radiotherapy (followed by salvage surgery if necessary) is preferable in that the larynx and voice are preserved (86). It is therefore disturbing to note the predictors of treatment do not include expected outcome. That is, it does not appear that physicians are using the expected outcome of any treatment in the treatment decision process. However, this situation is not found only in relation to UADT cancer patients. Similar variations in practice have been shown concerning breast cancer where the controversy was over breast conservation or not (87,88). Obviously, this documentation of the idiosyncratic or "setting-lead" nature of treatment decisions has not been directly related to patient prognosis, nevertheless it is difficult to believe that treatment decisions of this nature do not affect patient survival and broader health status outcomes.

Even more worrying, however, is documentation which suggests the systematic, prejudicial nature of HCP treatment decisions; especially if these are influencing outcome. The only study to address this issue in relation to UADT cancer patients was a retrospective analysis of the determinants of survival among 4527 oral cancer patients in Brazil (89). Females and non-white patients were more likely than males and whites to receive no treatment, even after controlling for stage, presumably one of the principal determinants of treatment modality. Furthermore, non-whites had increased risk for lip cancer recurrence and death resulting from lip cancer. However, once the analysis was controlled for stage and treatment modality, this racial difference disappeared (89). Further evidence suggesting the racially prejudiced nature of some treatment decisions also exist in the USA (90).

2.5 Manuscript I: "The role of diagnostic delays in the prognosis of oral cancer: A review of the literature"

INTRODUCTION

It is generally believed that cancer mortality can be reduced if lesions are detected, diagnosed and treated at an early stage. This belief is based on our understanding of the temporal progression of tumour growth and on evidence that there is a dose-response relationship between the local, regional or distant spread of cancers and patient survival (eg.1). However, despite the fact that such thinking has driven much of the work on cancer screening and treatment during this century, with the exception of breast cancer (2-6), there remains a lack of unequivocal epidemiological evidence to support a relationship between diagnostic delays and

patient prognosis.

There is now an increased interest in the subject of screening and the promotion of early diagnosis for oral cancers (7,8,9). This may be because changing treatments have failed to improve the survival rates of oral cancer patients over the past 30 years (10), and because oral (if not pharyngeal) precancerous lesions and cancers are, theoretically, relatively easy to detect. However, our current understanding of the interaction of tumour, patient and health care professional behaviours in relation to diagnosis and prognosis of oral cancers is poor. Therefore, the aim of this paper is to review the literature concerning factors affecting the diagnostic process, and hence prognosis, in oral cancer patients, in order to develop a conceptual framework upon which to base future research. To do so, published material concerning this subject for oral and other cancers will be reviewed.

Definitions: Oral Cancers

In the scientific literature, the phrase "oral cancer" is used to describe a multitude of combinations of tumours, benign and malignant, including those of the lip, salivary glands and pharynx, in addition to those of the oral cavity. In the western world, the vast majority of "oral cancers" are malignancies of epidermal tissue lining the upper aerodigestive tract and are caused by smoking tobacco alone or in combination with excessive alcohol consumption or, in the case of lip cancer, sunlight. In view of this, we shall define "oral cancers" as carcinomas of the oral cavity (ICD9: 141, 143-5), pharynx (ICD9: 146-8) and other ill-defined parts of the mouth and pharynx (ICD9: 149). Within the western world, these carcinomas are aetiologically and pathologically very similar. Therefore, it is useful to group them together for epidemiological, public health research of this nature.

Definitions: Diagnostic and Treatment Delays

One of the major problems involved in reviewing the literature concerning delayed cancer diagnoses is that a variety of definitions of delay are used in the

different studies. Delay has most often been categorised as: i) *patient delay* - the period between the patient first noticing a symptom and their first consultation with a health care professional concerning that symptom; and ii) *physician/provider delay* - the period from the patient's first consultation with a health care professional and the definitive pathological diagnosis or the initiation of therapy. Although this categorisation has the benefits of simplicity and applicability, the pejorative nature of the nomenclature can lead to over-simplified interpretation. For example, patients not following physicians' advice could thereby increase "provider delay" (11), or the inaccessibility of provider services could elongate "patient delay" (12).

One of the first attempts to improve the understanding of delayed cancer diagnosis and its determinants came from the work of Safer et al (1979) who successfully tested a model for patient delay which comprised three elements: i) *appraisal delay* - the time a patient takes to appraise a symptom as a sign of illness; ii) *illness delay* - the period between a patient deciding he/she has an illness and deciding to seek medical care; and iii) *utilisation delay* - the period between a patient deciding to seek medical care and actually seeing a health care professional (13). This work has recently been expanded to include the total delay (the period between noticing a symptom and the initiation of cancer therapy) in one model (14). This newer model incorporates appraisal and illness delay as before, then divides utilisation delay into *behavioural delay* (the period between deciding to seek medical care and acting to receive that medical care, eg, making an appointment) and *scheduling delay* (the period between the patient making an appointment and actually seeing a health care professional). Finally this newer model includes *treatment delay* which is defined as the period between the first consultation and the initiation of cancer therapy (14).

Obviously, these aforementioned categorisations and definitions of the various aspects of diagnostic delays reflect the approach and needs of behavioural scientists which may not be the same as clinicians. Nevertheless, even this brief discussion of the definitions of delays illustrates the need for clarity and uniformity.

While recognising the value of the categorisations of Safer et al (13) and Andersen and Cacioppo (14), for the purposes of this review and for the majority of research in this field, we suggest the following definitions:

Patient delay - the period of time between the patient first noticing a symptom and their first consultation with a health care professional concerning that symptom;

Professional delay - the period of time from the patient's first consultation with a health care professional and their first consultation with the treating specialist;

Treatment delay - the period of time between the patients first consultation with the treating specialist and the initiation of ablative or palliative therapy;

Total diagnostic delay - the sum of patient, professional and treatment delays.

Such a classification uses clear and easily defined categories as a starting point for the understanding of the diagnostic process and is a basis upon which to plan and implement interventions to improve that diagnostic process. However, it should be recognised that the aforementioned delays are not mutually exclusive and should not be interpreted pejoratively (ie. patient delays may not necessarily be due to patients etc.). Furthermore, all these delays take place within a health care system which may influence them in different ways and whose influences will differ from system to system. For instance, comparing waiting times for radiotherapeutic cancer treatment in Canada and the USA, Mackillop et al found patients in Canada had to wait significantly longer for therapy than their American peers (15). The latter study obviously describes only treatment delays. However, the organisation of health care services will have effects on all aspects of diagnostic and treatment delays.

Definitions: Prognosis

The most common prognostic outcome variable used in the investigation of delayed diagnoses is disease stage. Cancer staging is now a uniform, internationally recognised form of disease categorisation whose objectives are to: i) aid treatment planning; ii) indicate prognosis; iii) assist in the evaluation of treatment outcomes; iv) facilitate the exchange of information between treatment centres; and v)

contribute to the continuing investigation of cancers (16). Usually evaluated at the time of definitive diagnosis and treatment planning, cancer staging is an indicator of the extent to which the disease has progressed. The staging is based upon the site of the primary tumour, an evaluation of the size of the primary tumour (T category), the extent of regional lymph node involvement (N category) and the involvement or not of distant metastases (M category) (16). This process, outlined by the American Joint Committee on Cancer (AJCC) in agreement with the Union Internationale Contre le Cancer (UICC), concerns cancers at all sites and is designed for use throughout the world. Although it uses the most appropriate variables common to all cancer sites, this general approach lacks some of the detail necessary for within-site comparisons. For instance, the improved prognostic predictive value of a staging system for oral cancers which includes anatomical subsite (eg, tongue and floor of mouth among oral cancers) and histopathological data has been demonstrated (17), and it has been suggested that there is a benefit to assessing tumour growth velocity (18). Hence, while the prognostic predictive power of the TNM staging system is reasonable, it could be improved upon with the addition of certain pathological and anatomical subsite variables (19). However, despite these reports of an improved predictive power with the more detailed site-specific staging system, the cruder non-specific staging systems describing disease as local, regional or distant have been shown to be capable of predicting 5 year survival for oral cancer patients (1).

The other possible outcomes in an investigation of the effect of delayed diagnosis on prognosis are mortality and morbidity. However, appropriate data analysis would require the collection of variables associated with treatment, thereby complicating the investigation.

EVIDENCE FOR THE ROLE OF DIAGNOSTIC DELAYS IN THE PROGNOSIS OF ORAL CANCER

Patient Delays

Early publications concerning patient delays and oral cancer were descriptive.

In 1964, a French study of 904 cases of cancer of the tongue found that the average time for patients to first consult a physician was 4.6 months (20). These results were very similar to the 4.9 month mean patient delay period found in a much smaller study of 34 patients with oral cancer in Denmark (21). In a large series of 869 cases of lip cancer in Denmark, 17% of patients delayed first presentation for 12 months or more (22), while smaller descriptive studies in Scotland and Wales found, respectively, that 12.8% and 58% of patients presented within 1 month of the onset of symptoms (23,24).

One of the first studies which attempted to statistically analyse a possible association between patient delays and prognosis was that of Guggenheimer et al in the USA, who found no such association in a sample of 149 oral and oropharyngeal cancer patients (25). Similarly, a study of 336 lip, oral and oropharyngeal cancer cases in Brazil (26) and another of 167 oral cases in Denmark (27) revealed no association between patient delay and disease stage at diagnosis. No other studies have specifically investigated the association between patient delay and prognosis.

Professional Delays

Again, the earliest studies concerning professional delays in the diagnosis of oral cancers were descriptive. Cooke and Tapper-Jones reported professional delays ranging from 1 day to 7 weeks (24), while Bruun highlighted the possible importance of professional delays when he found the mean professional delay (5.6 months) to be longer than the mean patient delay (21). In their study of diagnostic delays among oral cancer patients, Guggenheimer et al considered professional delay to have occurred if "no treatment or inappropriate treatment had been provided", although this was clarified no further. Using these criteria, they found that professional delay had occurred in 30% of cases (25). The first study to statistically investigate the possibility of an association between professional delay and patient prognosis for oral cancers was that of Kowalski et al in Brazil (26). They found that professional delay greater than 1 month was associated with late stage lip and oral cancers. More

recently, Wildt et al found no association between professional delay and disease stage, although they did find a significant but weak correlation between tumour size and professional delay (27).

THE DETERMINANTS OF DIAGNOSTIC DELAYS

Determinants of Patient Delay

Despite the fact that there is no epidemiological evidence to support an association between patient delay and prognosis, two studies have specifically investigated correlates for patient delay, presumably on the assumption that patient delays do contribute to prognosis. Guggenheimer et al investigated the association between patient delay and age, gender, education and history of alcohol consumption and found no relationship (25). Wildt et al investigated the association between patient delay and age and gender and found no relationship (27). One other study has investigated variables associated with early stage disease at diagnosis in oral cancer and found factors which are possibly related to patient delay and/or tumour behaviour. In a sample of 160 oral cancer patients Elwood and Gallagher found regular dental attendance and low alcohol to be associated with early stage disease and that total diagnostic delay was longer in women than men (28). Unfortunately, however, it is impossible from this work to say whether dental attendance and alcohol consumption are associated with disease stage as indicators of patient, professional or tumour behaviour and we cannot say which aspect(s) of the total diagnostic delay is longer for women. In theory, dental attendance could be associated with patient delay, with those patients more likely to have a lesion noticed early in the disease process by their dentist and/or with patients who drink or smoke less (relative to others diagnosed with an oral cancer) and so have a less aggressive lesion. Similarly, alcohol consumption may be an indicator of dental attendance and/or of tumour aggressivity.

Assuming that patient delay plays a role in determining prognosis for oral cancers, it is evident that we have learned little from the limited evidence available.

A review of the literature concerning the determinants of patient delay in relation to cancers at other sites sheds more light on the subject, but also illustrates the complexity of human behaviour and the lack of a clear picture of the determinants of patient delay. Determinants of patient delay could be categorised as follows:

1. Cognitive Interpretations and Affective Reactions. In their review of delay in the detection of cancer, Antonovsky and Hartman concluded that as many as 75% of cancer patients delay at least one month and 35-50% delay 3 months between first noticing a symptom and consulting a physician (11). They reviewed 22 articles on the subject and observed that low socioeconomic status (SES) and low educational status were the most consistent factors associated with delay and that older age and males were also often similarly associated. They also reported that there seemed a universal belief that ignorance of the significance of certain symptoms was related to delay but that this view of human behaviour ignores the interaction, within an individual, of knowledge about cancer, the various affective orientations and the life context within which this occurs (11). This view of the complexity of the relationship between "cancer knowledge" and patient delay is supported by evidence that delay in seeking medical help appears to be a conscious and deliberate act, rather than a failure to perceive the neoplasm or comprehend its consequences (29). In making this conclusion, Hackett et al observed that patient delays had remained the same during the past 50 years, despite the attempts of the medical profession to educate the public. In their study of 563 patients with cancers in various sites they found that discovery of the cancer during a routine physical examination was the best predictor of reduced patient delay (29). Although this finding has important implications for the promotion of regular check-ups with a health care professional (at least among certain populations), it is important to realise that the incidental discovery of a cancer during routine examination implies that the patient may not have recognised the symptom and so, by definition, cannot be considered to be "patient delay". The second most important predictor of reduced delay found in this study was worry.

However, the authors discussed the complexity of this variable and its interaction with other affective factors such as fear, anxiety, denial and fatalism. They suggested that worrying about the symptom itself or worrying about cancer while simultaneously recognising that the latter is curable predicts a reduced delay. On the other hand, individuals who worry about their health in general or about cancer in general are more likely to prolong the period before presentation to a health care professional (29).

Other affective factors suggested to be associated with patient delay include fear, denial and perception of social responsibilities. Denial in the form of patients' reluctance to use the word cancer has been found to be associated with patient delay (29,30). However, evidence for the association between delay and fear or anxiety is equivocal or conflicting (31). The issue of social responsibility in relation to delay has hardly been investigated. Nevertheless, the finding of Safer et al that having a recent, competing problem was associated with delay (13), plus the qualitative finding of Dignan et al that women with cervical cancer would look after their child before seeking care for their symptoms (32) suggest that this may be an important issue.

Research like this led to the work of Safer and colleagues in their development of a model of delay for cancer diagnosis. When testing the model, they found that: i) having a recent, competing problem or life change, reading about the symptom and older age correlated significantly with longer total delay; ii) the presence of pain or bleeding correlated with shorter total delay; iii) negative imagery concerning treatment significantly increased illness delay; and iv) patient concern about treatment cost and pain and belief about the curability of the symptom correlated signficantly with utilisation delay (13).

2. Sociodemographics. The evidence of an association between sociodemographic variables and patient delay among the different cancer sites is equivocal. In their review of the literature on delayed diagnoses and cancer, Antonovsky and Hartman

made the generalised conclusion that older, less educated and poorer people are more likely to delay presentation of a cancer (11). However, a more recent study among patients with cancers at various sites found no association between any sociodemographic variables and patient delay (33). Looking at more specific groups, Marshall and Funch found that women with rectal cancer were more likely to delay than their male peers, but that there was no difference for colonic cancer (34). More recently, a study of the sociodemographic determinants of patient delay among people with lung, breast and colorectal cancers, revealed that, with the exception of older age being predictive of increased patient delay for colorectal cancers only, no other sociodemographic variables were implicated (35). In her review of diagnostic delays in breast cancer, Facione concluded that there is no relationship between age and patient delay. The author also commented that, although many studies have reported that black women are more likely to be diagnosed with late stage disease and had worse survival, they failed to control for the fact that racial minorities are over-represented in lower SES groups (31). She concluded that poverty, rather than race, was probably a stronger predictor of patient delay, although this may be related to the issue of access to services in that minority populations may delay help-seeking based on perceptions of how they will be treated by health care providers (31).

3. Previous Health-Related Experiences. The evidence concerning the relationship between patient delay and past health-related experiences is not clear. Hackett et al found a trend between people saying that cancer ran in their family and delayed presentation (29) and Gould-Martin et al found that among a sample of women with breast cancer, those with a previous history of a benign lump delayed longer than those with no such history (36). In addition, if one includes previous health-related practices in this category, it has been postulated that people who attend health services regularly for screening or who perform self-examination (eg for breast lumps) are less likely to delay. In this respect, overall indicators of health practices, such as having a family physician (4,36) or receiving regular medical check-ups

(4,29) have been shown to be associated with earlier presentation. However, in their reviews of delayed diagnoses for breast cancer, both Facione (31) and Caplan and Helzlsouer (37) concluded that the relationship between breast self-examination or attendance for screening and reduced patient delay is not proven.

4. Symptomology. With regard to symptomology, the evidence is again unclear. Once more, the majority of research relating to this subject has been done with breast cancer patients. As Facione pointed out in her review, a painless breast lump is the most common presenting symptom for breast cancer (31). However, Mor et al found that although those with breast cancer were more likely to attribute their symptoms to cancer, this did not lead to reduced patient delay (35). Furthermore, among several variables hypothesised to affect patient delay, Hackett et al found that pain was not a good predictor of reduced delay, even though it was stated as the reason for medical consultation in 33% of the sample (29).

5. Access. The issue of the influence of the accessibility of the health care system for cancer diagnosis has seen very little research. Using the definition of access outlined by Penchansky and Thomas (12) (which includes availability, accessibility, accommodation, affordability and acceptability of health care services as dimensions of the overall access construct), illustrates the complexity of the subject. The issue of the doctor patient relationship (acceptability) has been addressed in a few studies of breast cancer, showing that the perceived closeness of the doctor/patient relationship (38) and the gender of the doctor (39) influence patient delay. Furthermore, there is an increasing body of evidence suggesting reduced utilisation of health care services by black Americans (40,41), even among Medicare beneficiaries (42,43).

Determinants of Professional Delay

Evidence concerning misdiagnosis or mistreatment of patients and the type

of health care professional was documented first in research on determinants of professional delays for oral cancers. Shafer reported the rate of mismanagement of oral cancers referred over a 24-year period to one American institution as being 14.8% (44), and a study in the UK found that 33% of oral cancer patients had initially been misdiagnosed (23). Another study in the UK a decade later showed that, while there was little difference in the period of professional delay among referring physicians and dentists, the former requested urgent consultations 63% of the time and the latter group did so only 17% of the time (45). This, they suggested, was a reflection of a higher level of suspicion among physicians. In a study of 373 head and neck cancer patients, Amsel found patterns of referral were related to "anatomical provinces" with dentists referring a higher proportion of oral cancers and physicians referring patients with cancers at virtually all other head and neck sites (46). In addition, he found that patients with late stage disease were more likely to be referred from physicians. He suggested that this was because these patients waited until the disease was more serious and then presented to their physician (46). More recently, Dimitroulis et al have described a similar pattern of referrals among oral cancer patients, with dentists being the predominant referral source, mismanagement being present in 33% of cases and physicians referring predominantly late stage disease patients (47).

Unfortunately these studies were descriptive and only three studies have statistically investigated correlates of professional delay. In the UK, a study of 96 cases of oral cancer demonstrated that family physicians were significantly less likely to delay referral and more likely to make the correct diagnosis than dentists (48). However, delayed referral was defined as beyond 2 days, apparently in view of the fact that 2 days is the recommended time for referral of such cases in the UK (no reference for this recommendation was given). However, there was no statistical difference between the proportion of dentists and physicians delaying 3 weeks (48). In Denmark, Wildt et al found longer professional delay to be correlated with small tumour size and female and older patients (27). However, it needs to be recognised

that these correlates were the product of a series of univariate analyses and, although the correlations were statistically significant, they were weak. Furthermore, the use of a Spearman's rank correlation test with a nominal variable such as gender is inappropriate. Finally, in a retrospective study of 543 oral and oropharyngeal cancer patients in Israel, Gorsky and Dayan found that physicians are significantly more likely to refer patients with late stage disease (49).

Once again, the research related specifically to oral cancers yields little concrete evidence concerning the determinants of professional delays. The literature relating to other cancers provides only a few indications of potential determinants as there has been very little research concerning professional delays in the diagnosis of cancers. In a closely related subject, among a sample of 4527 oral cancer patients treated over a 28-year period in one Brazilian hospital, Franco et al (50) found that black and white patients within the same disease stage strata were treated differently and that once treatment modality was controlled for, blacks and whites had the same disease recurrence and mortality rates. These data strongly suggest that treatment decisions affecting prognosis are related to race (50). This evidence is linked to the previously mentioned work on race and access to health care services, implying that not only are racial minorities less inclined to use medical services, but that when they do, decisions concerning their treatment have a racist element (51).

METHODOLOGICAL PROBLEMS WITH THE AVAILABLE RESEARCH

It is evident that there is very little evidence supporting the existence of an association between delayed diagnosis and stage among oral cancer patients. However, it is extremely difficult to reject the hypothesis that diagnostic delays contribute to worse prognoses, in view of our understanding of the temporal progression of tumour growth and based on evidence confirming that patients diagnosed with cancers that have progressed further have a worse prognosis than patients with cancers diagnosed early in the disease progression. The problem, therefore, almost certainly lies with methodological problems in the available

research:

i) The majority of the studies are descriptive. While it is perfectly reasonable to perform such studies to describe a perceived problem, some of the authors of these reports assume (and others suggest) that there is a link between diagnostic delays and patient prognosis. However no statistical analyses to support their assumption/suggestions are carried out;

ii) Of those studies that statistically test possible associations, several use sample sizes with insufficient power reducing the possibility of finding associations even though they may exist (ie, there remains the possibility that they are committing a type II error). A factor which contributes to the need for increased sample sizes in such studies is the use of categorical data. Unfortunately, by their very nature, the majority of data collected in studies of this nature can only be categoric and even those data which, theoretically, could be collected as continuous variables (eg the length of the various delays) are probably better collected categorically to improve validity (see below for a discussion of data validity);

iii) In the majority of analytical studies, some important variables have not been collected and/or have not been controlled for in the analyses. Although it is not clear from those studies exactly why this is the case, the probable causes are insufficient power for the collection of sufficient numbers of variables and/or the lack of a theoretical model upon which to base the research. In fact, it appears that one of the most important variables to consider is tumour growth rate. To explain not having found an association between diagnostic delay and patient prognosis, Kaufman et al (52) and Evans et al (18) suggested that patients with fast growing tumours present quickly, but still have a poor prognosis, while those with a slow growing tumour have a good prognosis despite delaying presentation for a long period. Since no uniformly recognised direct measure of tumour growth rate exists, various indicators

could be used. While it is not the goal of this review to discuss the measurement of tumour growth rate, possibilities include tumour thickness (53, 54, 55), histological malignancy grading (56,57), the assessment of cell kinetics through flow cytometry (58,59), p53 protein expression (60) and tumour angiogenesis (61);

iv) There remains a question about the validity of the data which, in many studies, were collected retrospectively. Some of the data (eg, patients' recall of length of time between first symptoms and consultation with a physician) can, by definition, only be collected retrospectively. Nevertheless, the collection of retrospective data within in a prospective study design is preferable to a study completely dependent upon old records or on patients' distant recall. In their review of the literature, Antonovsky and Hartman observed that no studies had even attempted to address the problem of the reliability and validity of their data. They suggested that the only possible solution to the problem of potential distortion in recall data would be to carry out a prospective cohort study in which subjects and their symptoms were followed. However, they pointed out that an enormous sample size would be required and that it would be ethically impossible to allow subjects to behave "normally" once symptoms were recorded (11). The comments in that review, published over 20 years ago, remain pertinent today;

v) A major obstacle to trying to make comparisons of results from the available research is the fact that several studies have failed to define their diagnostic delay periods, while those which have defined them often use differing definitions;

vi) Finally, the geographic locations of the studies must be considered, as the health systems and cultures of different countries could contribute to the variability in patient and professional delays.

SUMMARY AND CONCEPTUAL FRAMEWORK

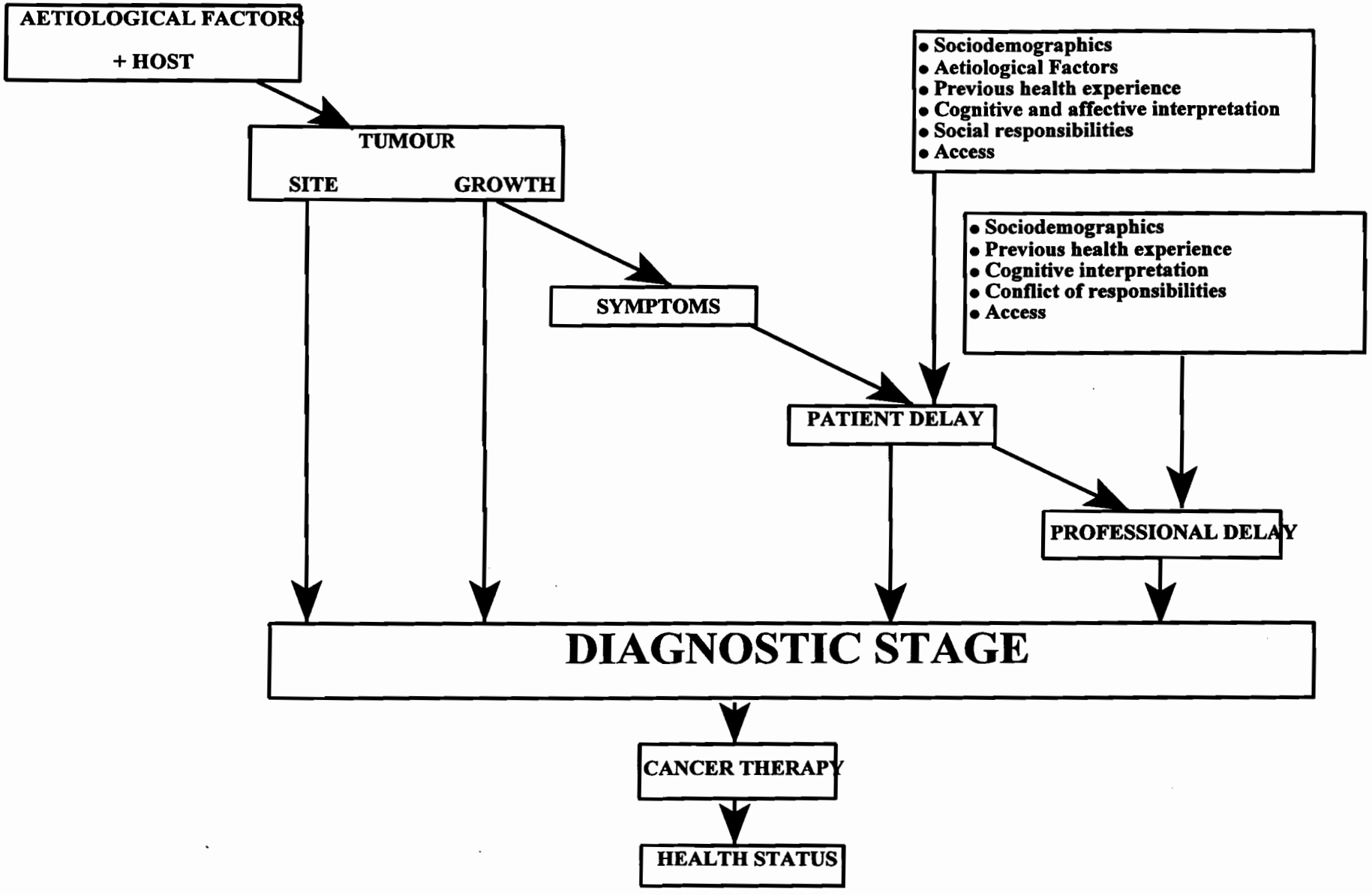
The available evidence concerning the role of diagnostic delays in the prognosis of oral cancer provides no clear picture of the situation, and the research concerning such delays related to cancers at other sites only serves to widen the field of potentially associated variables without further clarifying the situation. However, if we are to improve the prognosis of oral cancer patients through earlier diagnosis, it is essential that the diagnostic process is further investigated. As a synthesis of the reviewed data concerning the diagnostic process involved with oral cancers, we propose a conceptual framework based on three key actors; the tumour, the patient and the health care professional (see Figure 1). In addition to facilitating improved comprehension of the diagnostic process, such a categorisation of the principal actors enables the appropriate targeting of strategies to improve that process. However, it is important to recognise that this diagnostic process takes place within a health care system which will almost certainly influence patient and professional behaviours.

Tumour Factors

The available evidence suggests that the tumour variables with a role in the diagnostic process are site, growth velocity and symptomology. Tumour site may be important in that certain sites are more visible or noticeable than others and/or tumours at certain sites may have a greater or lesser capacity for rapid growth (eg, through lymphatic drainage). With respect to the latter point, it is important to note that although many people assume oropharyngeal cancers grow faster than their oral cavity counterparts, based on an analysis of the DOSAK data of Platz et al (53), Howaldt et al could find no evidence to support such a hypothesis (62).

Tumour growth velocity is an important variable because the evidence suggests that rapidly growing tumours are more likely to have a poor prognosis than slow growing tumours. Finally, tumour symptomology may be an important variable in the diagnostic process through its influence on both patient and professional behaviours.

Figure 1. The diagnostic process for oral cancers: A conceptual framework



Patient Factors

Once the patient has noticed a symptom(s), he/she needs to recognise it as a sign of illness and then act upon that recognition as described in the models of delay (13,14). Certain patient variables will play a role in the recognition of those symptoms and the decision as to how and when to act on them. As with most forms of human behaviour, the list of potential determinants and indicators of patient delay is daunting, probably involving complex interactions. However, these factors could broadly be categorised as follows:

i) Sociodemographics - age, gender, race, education, employment, income, etc.

ii) Rate and volume of aetiological factor consumption - alcohol, cigarettes, etc.

iii) Previous health experiences - previous medical/dental screening practices, previous personal and family medical history

iv) Cognitive interpretation of symptoms - ignorance/knowledge of potential implications of the symptoms, subsequent treatment and prospects for survival

v) Affective interpretation of symptoms - fear of symptoms, cancer and/or treatment, denial etc.

vi) Conflict of responsibilities- family, job etc.

vii) Access.

It is important to recognise that, in addition to influencing patient delay, some of these variables (such as patient age or rate and volume of aetiological factor consumption) could influence tumour behaviour. For instance, Elwood and Gallagher found that low alcohol consumption was associated with early stage oral cancer (28), Kowalski et al found that increased alcohol consumption was a predictor for late stage lip cancer (26) and Bundgaard et al found tobacco consumption to be predictive of oral cancer 5-year survival (63) and to have highly significant correlations with a number of histological markers of tumour behaviour (64). The latter study strongly suggests a direct association between aetiological factor consumption and tumour behaviour, while the results of the other studies could reflect a direct association and/or an association through the intermediate variable of patient delay.

Professional Factors

Finally, once the patient has consulted a health care professional, the potential for professional and/or system delays affecting diagnostic stage becomes a possibility. As previously mentioned, there are very few investigations of the determinants of these factors. However, it seems perfectly reasonable to categorise the potential determinants of professional delay in the same way as those of patient delay:

i) Sociodemographics - age, gender, race;

ii) Previous health experiences - type of health care professional, relevant medical/dental education, previous screening practices, previous professional experience with respect to oral cancers and cancers in general, previous personal and family medical history;

iii) Rate and volume of aetiological factor consumption - alcohol, cigarettes, etc.

iv) Cognitive interpretation of symptoms - ignorance/knowledge of potential implications of the symptoms leading to correct or mismanagement of the patient;

v) Conflict of responsibilities - patient comorbidity;

vi) Access - distance from and availability of relevant specialist services.

Because of the near-total absence of research into the determinants of professional delays in cancer diagnosis, the above list is almost entirely hypothetical. Nevertheless, examination of the determinants of patient and professional behaviour using such a theoretical framework facilitates clarity. Furthermore, as mentioned previously, it is theoretically important to recognise the role of all potential variables further up the aetiological pathway. For instance, the sociodemographic determinants of professional delay include those of the patients' and of the professionals'.

In conclusion, despite the lack of epidemiological evidence available to support the hypothesis, it seems highly likely that diagnostic delays have a role in determining patient prognosis for oral cancer. This review has demonstrated the weak

nature of the available evidence on the subject, and a theoretical framework upon which to base future research has been proposed. While perhaps oversimplifying the diagnostic process, the categorisation of the factors involved into tumour, patient and professional variables, nevertheless clarifies a process which is currently very poorly understood and lays out a clear aetiological framework upon which to base future research. Furthermore, such a categorisation enables clarity in the description of targets for future interventions aimed at improving aspects of the diagnostic process among oral cancers. And finally, such a conceptual framework also facilitates comparative work among other cancers which are aetiologically and epidemiologically very similar to the oral cancers discussed in this paper.

References

1. Parker SL, Tong T, Bolden S and Wingo PA (1996) *Cancer statistics 1996.* CA Cancer J. Clin 46(1):5-28

2. Facione NC (1993): *Delay versus help seeking for breast cancer symptoms: a critical review of the literature on patient and provider delay.* Soc. Sci. Med. 36(12):1521-34

3. Robinson E, Mohilever J, Zidan J and Sapir D (1984): *Delay in diagnosis of cancer. Possible effects on stage of disease and survival.* Cancer 54:1454-60

4. Samet JM, Hunt WC, and Goodwin JS (1990): *Determinants of cancer stage. A population-based study of elderly New Mexicans.* Cancer 66:1302-7

5. Porta M, Gallen M, Malats N and Planas J (1991): *Influence of diagnostic delay upon cancer survival: an analysis of 5 tumour sites.* J. Epidemiol. Comm. Hlth 45(3):225-30

6. Maguire A' Porta M, Malats N, Gallen M, Pinol JL and Fernandez E (1994): *Cancer survivaland the duration of symptoms. An analysis of possible forms of the risk function.* Eur. J. Cancer 30A(6):785-92.

7. UK Working Group on Screening for Oral Cancer and Precancer (1993): *Screening for cancer: Should oral screening be introduced?* Community Dental Health 10(Suppl 1):1-3

8. Boyle P, Macfarlane GJ, Blot WJ, Chiesa F, Lefebvre JL, Mano Azul A, de Vries N and Scully C (1995): *European School of Oncology Advisory Report to the European Commission for the Europe Against Cancer Programme: Oral carcinogenesis in Europe.* Oral Oncol. Eur. J. Cancer 31B(2):75-85

9. Horowitz AM, Goodman HS, Yellowitz JA and Nourjah PA(1996): The need for health promotion in oral cancer prevention and early detection. J. Public Health Dent. 56(6):319-30

10. Binnie WH and Rankin KV: *Epidemiology of oral cancer* In Wright BA, Wright JM and Binnie WH eds. *Oral Cancer: Clinical and Pathological Considerations.* Boca Raton, FL: CRC Press, 1988

11. Antonovsky A and Hartman H (1974): *Delay in the detection of cancer: A review of the literature.* Health Education Monographs 2(2):98-128

12. Penchansky R and Thomas JW (1981): *The concept of access. Definition and*

relationship to consumer satisfaction. Med. Care 19:127-40

13. Safer MA, Tharps QJ, Kackson TC and Leventhal H (1979): *Determinants of three stages of delay in seeking care at a medical clinic.* Med. Care 17:11-29

14. Andersen BL and Cacioppo JT (1995): *Delay in seeking a cancer diagnosis: Delay stages and psychophysiological comparison processes* British Journal of Social Psychology 34:33-52

15. Mackillop WJ, Zhou Y, and Quirt CF (1995): *A comparison of delays in the treatment of cancer with radiation in Canada and the United States.* Int. J. Radiation Oncology Biol. Phys. 32(2):531-9

16. American Joint Committee on Cancer: *Manual for Staging of Cancer.* 4th Ed'n. J.B.Lippincott Co., Philadelphia, 1992

17. Rapidis AD, Langdon JD, Patel MF and Harvey PW (1977): *STNMP a new system for the clinico-pathological classification and identification of intra-oral carcinomata.* Cancer 39:204-9

18. Evans SJW, Langdon JD, Rapidis AD and Johnson NW (1982): *Prognostic significance of STNMP and velocity of tumour growth in oral cancer* Cancer 49:773-6

19. Langdon JD: *Classification and staging. Mouth cancer and jaw tumours.* In Langdon JD and Henk JM (Eds) *Malignant tumours of the mouth, jaws and salivary glands.* 2nd Ed'n Edward Arnold, London, 1995

20. Flamant R, Hayem M, Lazar P and Denoix P (1964): *Cancer of the tongue; a study of 904 cases.* Cancer 17:377-85

21. Bruun JP (1976): *Time lapse by diagnosis of oral cancer.* Oral Surg. Oral Med. Oral Path. 42(2):139-49

22. Jorgensen K, Elbrond O and Anderson AP (1973): *Carcinoma of the lip, a series of 869 cases* Acta Radiol. 12:177-90

23. Pogrel MA (1974): *The dentist and oral cancer in North East Scotland.* Br. Dent. J. 137:15-20

24. Cooke BED and Tapper-Jones L (1977): *Recognition of oral cancer - cases of delay.* Br. Dent. J. 142:96-8

25. Guggenheimer J, Verbin RS, Johnson JT, Horkowitz CA and Myers EN (1989):

Factors delaying the diagnosis of oral and oropharyngeal cancers. Cancer 64:932-5

26. Kowalski KP, Franco EL, Torloni H, Fava AS, Sobrinho JdeA, Ramos G, Oliveira BV and Curado MP (1994): *Lateness of diagnosis of oral and oropharyngeal cancer: factors related to the tumour, the patient and health professionals.* Oral Oncol. Eur. J. Cancer 30B(3):167-73

27. Wildt J, Bundgaard T and Bentzen SM (1995): *Delay in the diagnosis of oral squamous cell carcinoma.* Clin. Otolaryngol. 20:21-5

28. Elwood JM and Gallagher RP (1985): *Factors in the early diagnosis of cancer of the oral cavity.* Can Med. Assoc. J. 133:651-6

29. Hackett TP, Cassem H and Raker JW (1973): *Patient delay in cancer.* N.E.J.M. 289(1):14-20

30. Worden JW and Weisman AD (1975): *Psychosocial components of lagtime in cancer diagnosis.* J. Psychosomatic Research 19:69

31. Facione NC (1993): *Delay versus help seeking for breast cancer symptoms: a critical review of the literature on patient and provider delay.* Soc. Sci. Med. 36(12):1521-34

32. Dignan M, Michielutte R, Sharp P, Bahnson J, Young L and Beal P (1990): *The role of focus groups in health education for cervical cancer among minority women.* J. Community Health 15:369

33. Wool MS: *Understanding denial in cancer patients.* In Goldberg RJ (Ed) *Psychiatric aspects of cancer* Basil Karger 1988 p37-53

34. Marshall JR and Funch DP (1987): *Gender and illness behavior among colorectal cancer patients.* Women and Health 11(3-4):67-82

35. Mor V, Masterson-Allen S, Goldberg R, Guadagnoli E and Wool MS (1990): *Pre-diagnostic symptom recognition and help seeking among cancer patients.* J. Comm. Health 15(4):253-66

36. Gould-Martin D, Paganini-Hill A, Casagrande C, Mack T and Ross RK (1982): *Behavioural and biological determinants of surgical stage of breast cancer.* Preventive Medicine 11:429-40

37. Caplan LS and Helzlsouer KJ (1992/93): *Delay in breast cancer: a review of the literature*. Public Health Review 20:187-214

38. Henderson JG (1965): *Denial and repression as factors in the delay of patients with cancer presenting themselves to the physician.* Annals of the New York Academy of Sciences 125:856-64

39. Williams EM, Baum M and Hughes LE (1986): *Delay in presentation of women with breast cancer disease.* Clinical Oncology 2:327-331

40. Broman CL (1987): *Race differences in professional help seeking* Am. J Community Psychol. 15:473-89

41. Blendon RJ, Aiken LH, Freeman HE and Corey CR (1989): *Access to medical care for black and white Americans. A matter of continuing concern.* JAMA 261:278-81

42. Gornick ME, Eggers PW, Reilly TW Mentnech RM, Fitterman LK, Kucken LW and Vladeck BC (1996): *Effects of race and income on mortality and use of services among medicare beneficiaries* N.E.J.M. 335(11):791-9

43. Peterson ED, Wright SM, Daley J and Thibault GE (1994): *Racial variation in cardiac procedure use and survival following acute myocardial infarction in the Department of Veterans Affairs.* JAMA 271:1175-80

44. Shafer WG (1975): *Initial mismanagement and delay in diagnosis of oral cancer.* JADA 90:1262-64

45. Scully C, Malmos D, Levers BGH, Porter SR and Prime SS (1986): *Sources and patterns of referrals of oral cancer: role of general practitioners.* BMJ 293:599-601

46. Amsel Z, Strawitz JG and Engstrom PF (1983): *The dentist as a referral source of first episode head and neck cancer patients.* JADA 106(2):195-7

47. Dimitroulis G, Reade P and Wiesenfeld D (1992): *Referral patterns of patients with oral squamous cell carcinoma, Australia.* Oral Oncol. Eur. J. Cancer 28B(1):23-7

48. Schnetler JFC (1992): *Oral cancer diagnosis and delays in referral.* Br. J. Oral Maxillofacial Surgery 30:210-3

49. Gorsky M and Dayan D (1995): *Referral delay in diagnosis of oro/oropharyngeal cancer in Israel.* Oral Oncol. Eur. J. Cancer 31B(3):166-8

50. Franco EL, Dibb LL, Pinto DS, Lombardo V and Contesini H (1993): *Race and gender influences on the survival of patients with mouth cancer.* J. Clin. Epidemiol. 46(1):37-46

51. Geiger HJ (1996): *Race and health care - an American dilemma?* N.E.J.M. 335(11):815-6

52. Kaufman S, Grabau JC and Lore JH (1980): *Symptomology in head and neck cancer; a quantitative review of 385 cases.* Am. J. Public Health 70:520-2

53. Platz H, Fries R and Hudec M (1986): *Prognoses of oral cavity carcinomas.* Carl Hanser, Munich.

54. Spiro RH, Huvos AG, Wong GY, Spiro JD, Gnecco CA and Strong EW (1986): *Predictive value of tumour thickness in squamous cell carcinoma confined to the tongue and floor of the mouth.* Am. J. Surg. 152:345-50

55. Mohit-Tabatabal MA, Sobel HJ, Rush BF and Mashberg A (1986): *Relation of thickness of floor of mouth stage I and II cancers to regional metastasis.* Am. J. Surg. 152:351-3

56. Anneroth G, Batsakis J and Luna M (1987): *Review of the literature and a recommended system of malignancy grading in oral squamous cell carcinomas.* Scand J. Dent. Res. 95:229-49

57. Byrne M, Koppang HS, Lilleng R and Kjaerheim A (1992): *Malignancy grading of the deep invasive margins of oral squamous cell carcinomas has high prognostic value.* J. Pathol. 166:375-81

58. Forster G, Cooke TG, Cooke LD, Stanton G, Bowie G and Stell PM (1992): *Tumour growth rates in squamous cell carcinoma of the head and neck measured by in vivo bromodeoxyuridine incorporation and flow cytometry.* Br. J. Cancer 65:698-702

59. Begg AC (1995): *The clinical status of Tpot as a predictor? Or why no tempest in the Tpot!* Int. J. Radiation Oncology Biol. Phys. 32(5):1539-41

60. Jaros E, Lunec J, Perry RH, Kelly PJ and Pearson ADJ (1993): *P53 protein overexpression identifies a group of central primitive neuroectodermal tumours with poor prognosis.* Br. J. Cancer 66:373-85

61. Penfold CN, Partridge M, Rojas R and Langdon JD (1996): *The role of angiogenesis in the spread of oral squamous cell carcinoma.* Br. J. Oral maxillofacial Surg. 34(1):37-41

62. Howaldt H-P, Frenz M and Pitz H (1992): *Proposal for a modified T-classification for oral cancer.* J. Cranio-Maxillo-Facial Surgery 22:96-101

63. Bundgaard T, Bentzen SM and Wildt J (1994): *The prognostic effect of tobacco and alcohol consumption in intr-oral squamous cell carcinoma.* Oral Oncol. Eur. J Cancer 30 B(5):323-8

64. Bundgaard T, Bentzen SM and Sogaard H (1995): *Histological differentiation of oral squamous cell cancer in relation to tobacco smoking.* Oral Oncol. Eur. J Cancer 31B(2):118-21

NB. End of manuscript I

2.6 Summary of literature review on the prognostic determinants in UADT cancer patients

Using the categorization of determinants previously described, it would seem that tumour site, size, histological grading, depth of infiltration and degree of regional and distant spread are currently the best biological determinants. However, work concerning cell kinetics, tumour genetics and tumour angiogenesis has produced promising prognostic indicators for UADT cancer patients.

Among patient sociodemographical determinants, age, gender and race influence survival, although the exact reasons for this are not clear. The fact that older patients have a worse survival rate than younger patients is not UADT cancer-specific and there is some evidence to suggest that older UADT cancer patients live a larger proportion of their life-expectancy than do younger patients. Similarly with gender, it is clear that women almost uniformly have a better survival rate than do men. Once again, this is not UADT cancer-specific and within UADT cancers, it is not clear why this is the case. Finally racial determinants would seem to be related to prognosis through other socio-economic determinants such as poverty, education and problems of access etc., although there may be some forms of behaviour specific to certain ethnic groups.

With respect to patient behavioural determinants, the rate of consumption of aetiological factors would seem to be an important prognostic determinant, as would

delayed presentation for diagnosis. This delay may be related to any number of factors but is most likely to be associated with patient sociodemographics, previous health experiences, cognitive and affective reaction to symptoms and issues of access.

Finally, HCP determinants of prognosis revolve around the issues of professional diagnostic delays and treatment decisions. It would seem that HCP diagnostic delays do occur and that they may influence patient prognosis. Furthermore, the limited evidence concerning treatment decisions raises concerns over their apparent lack of consideration of the expected outcome. This is probably a product of treatment decisions being, to a large extent, determined by the specialist-type and geographic location involved, and the fact that the evidence concerning treatment outcomes for this group of patients is poor.

2.7 Diagnostic determinants of prognosis: project aim

In view of the position stated in section 1.6 of the introduction and the relative paucity of conclusive evidence concerning the role of diagnostic delays in the determination of patient prognosis, the aim of this section of the research project was to investigate the hypothesis that, in a sample of UADT cancer patients, controlling for cancer site, patient and professional diagnostic delays are independently associated with patient prognosis.

2.8 Methodology

Patients diagnosed with oral, pharyngeal and laryngeal carcinomas (ICD-9 141, 143-9 and 161) during the 18 month period beginning July 1st, 1995 were included in this study. They were all recruited from one of three McGill University hospitals in Montreal. The study was approved by the human ethics committees of the relevant hospitals and all subjects read and signed an informed consent before participating in the study. Data were collected from subjects in a 20-30 minute interview using a standardised questionnaire (see Appendix 1) eliciting information

on socio-economic and demographic variables, plus information concerning the cancer development and symptomology, health care professionals consulted and the period of time taken for each stage in the diagnostic process. Interviews were performed between subjects' initial consultation with the treating specialist and the initiation of their cancer therapy. A random sample of 30 subjects' responses concerning professional delay were validated by confirming them with their primary family physician or dentist and hospital charts.

2.8.1 Methodology - variables

The dependent variable for this section of the project was the dichotomous outcome, disease stage; early or late. Clinically, disease stage (as described in section 2.2) is most often used as a four-category ordinal scale (stages I-IV). However, it is often collapsed to a two-category scale in epidemiological research of this nature using the involvement of regional lymphadenopathy as the cut-off point between stages II and III or early and late stage disease (ie. stages I and II, where the disease is still local, are early stage disease, while stages III and IV, involving regional and/or distant metastatic disease, are late stage disease).

The majority of independent variables were nominal (patient gender, cohabitation status, education status, comorbidity status, dental status, presenting symptomology and disease site), while three independent variables were ordinal (patient, professional and total delay) and one variable was continuous (patient age). For the purposes of the statistical analyses, patient age was converted into a dichotomous variable (<65 years v ≥ 65 years) centred around the sample mean age (64.6 years).

Patient delay was considered to be the time between subjects' initially noticing a symptom and their first presentation to a health care professional (HCP) concerning that symptom. Professional delay was considered to be the period of time between the initial consultation with a HCP and the first consultation with the treating specialist and total delay was considered to be the sum of patient and

professional delays. So as to improve the validity of subjects' recall of patient and professional delays, the unit used for these time periods was months.

For the purposes of the statistical analyses, patient and professional delays were then categorised into <1 month, 1-3 months and >3 months, with the former having a further, separate category of "no symptom" for those patients who noticed no symptom and whose cancer was discovered by accident while consulting an HCP for a different problem. It can be argued that the most appropriate way to categorize the delay variables would have been to use the mean/median delay periods or some arbitrarily decided percentile within the range. However, the former would have resulted in dichotomous variables and so less information, while the latter would have been an arbitrary decision. An *a priori* decision was made to categorize the delay variables as described with a view to the use of the results in the future. It was felt that a delay of <1 month could justifiably be stated to be as quick as possible; that a delay of 1-3 months may leave some room for improvement; and that delays >3 months were unacceptable if they affected patient prognosis. The idea being that if this categorization of the delay variables yielded an association between diagnostic delays and patient prognosis in this and later research, they could be used as key periods in the promotion of early diagnosis for this group.

Clinico-pathological data were collected from subjects' hospital charts. Disease stage was classified according to the American Joint Committee on Cancer (57), although for the purposes of statistical analysis, subjects were classified with early (T1 or T2 and N0) or late (T3, T4 or N>0) stage disease.

2.8.2 Methodology - statistical analyses

Following descriptive statistics, a series of univariate analyses were performed to determine the odds ratios (OR) plus their 95% confidence intervals for late vs. early stage disease in relation to all independent variables. Subsequently, a multiple logistic regression was used to formulate a multivariate model containing variables with an independent association with disease stage. The model originally

contained all independent variables and those variables which explained the least amount of variation were successively removed from the model until it contained only those variables with a significant independent association with disease stage.

2.9 Manuscript II: "The role of professional diagnostic delays in the prognosis of upper aerodigestive tract carcinoma"

This manuscript is reprinted from Oral Oncology (in press) Allison PJ, Franco E, Black M and Feine JS: *The role of professional diagnostic delays in the prognosis of upper aerodigestive tract carcinoma*, Copyright 1998, with kind permission from Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington 0X5 1GB, UK.

Abstract

Despite the belief that cancer mortality can be reduced if lesions are detected, diagnosed and treated at an early stage, only one study, among a number concerning cancers of the upper aerodigestive tract (UADT), has found any relationship between such delays and prognosis for this population of cancer patients. The aim of this study was, therefore, to investigate the relationship between patient and professional diagnostic delays and patient prognosis in a group of UADT cancer patients. **Method:** Patients diagnosed with squamous cell carcinoma of oral cavity sites (ICD-9 141, 143-5) oro-, naso- and hypopharynx (ICD-9 146-8) and larynx (ICD-9 161) were included in the study. Stepwise multiple logistic regression was used to calculate the odds ratio (OR) of late vs. early stage disease for selected study variables. **Results:** The sample comprised 188 subjects. Multivariate analysis found that having a pharyngeal cancer (OR: 9.26; 95%CI: 4.02-21.32; p: 0.0001) a professional delay >1 month (OR: 2.28; 95%CI: 1.13-4.64; p: 0.022) and age <65 years (OR: 2.22; 95%CI: 1.11-4.54; p: 0.024) were predictive of late stage disease. A dose-response relationship between professional delay and OR for late stage disease for the whole sample (p for trend 0.03) and among those with oral cancer (p

for trend 0.0001) was found. **Conclusions:** The results of this study suggest that, among patients with an UADT cancer, professional delays >1 month are contributing to an increased risk for being diagnosed with late stage disease.

Key words: diagnosis, head and neck cancer, prognosis.

Introduction

It is generally believed that cancer mortality can be reduced if lesions are detected, diagnosed and treated at an early stage. This belief is based on our understanding of the temporal progression of tumour growth and on evidence that there is a dose-response relationship between the local, regional or distant spread of cancers and patient survival (1). Although a number of studies have been published documenting so-called patient- and/or professional-related delays in the diagnosis of upper aerodigestive tract (UADT) cancers, the relationship between diagnostic delays and patient prognosis in this population remains unclear. This results, in part, because the majority of published material is purely descriptive (2,3,4,5,6,7,8). Other studies have involved small samples (9,10), while some of those with larger sample sizes did not reveal significant associations (11,12). Another reason for our poor understanding of the relationship between diagnostic delays and prognosis is the confounding effect of the rate of tumour growth. Evans et al (1982) found that longer diagnostic delay was associated with improved survival, and hypothesised that a fast-growing tumour (with a poor prognosis) promotes rapid diagnosis, while a slow-growing tumour leads to longer delays (13).

Despite these problems, several studies have contributed to our understanding of the causes of late diagnoses and the association between late diagnoses and prognosis among UADT cancer patients. Elwood and Gallagher studied a group of 160 oral cancer patients in Canada. They found that regular dental care and low alcohol consumption were associated with early stage disease and that the delay between first symptom and histological diagnosis was longer among women (14). In

a retrospective investigation of the influence of race and gender on the survival of 4527 oral cancer patients diagnosed and treated over a 28 year period in one Brazilian hospital, Franco et al found that race and gender were strong predictors of diagnostic stage and treatment modality. However, once stage and treatment were controlled for, only gender remained as a predictor of recurrent disease or survival, with females having an improved prognosis for both outcomes (15). A prospective study of 336 lip, oral cavity and oropharyngeal cancer patients in Brazil found that male gender and less visible tumour sites were associated with advanced stage. However, controlling for cancer site, it was found that late stage lip cancers were associated with painful ulcers, alcoholism and professional delay, and that late stage oral cancers were associated with infiltrative lesions, dysphagia and professional delay (16). A retrospective study of 543 cases of lip, oral and oropharyngeal cancer cases in Israel found that patients with lip cancers had a longer diagnostic period than those with cancers at other sites, but that lip cancers were associated with early stage disease (17). And finally, in Denmark, a study of 167 patients with oral cancer demonstrated no association between patient or professional delay and stage, but found that small tumour size, women and older aged patients were associated with increased professional delay (18).

The available evidence suggests that tumour growth rate is a confounding factor in investigations of the association between diagnostic delays and prognosis. Based on the assumption that cancer growth patterns can be controlled for, at least partially, by classifying tumours by site into oral cavity, pharyngeal and laryngeal cancers, the aim of this study was to test the hypothesis that, controlling for tumour site, patient, professional and/or total diagnostic delays are associated with disease stage at diagnosis among a sample of UADT cancer patients.

Methodology

Patients diagnosed with squamous cell carcinoma of oral cavity sites (ICD-9 141, 143-5) oro-, naso- and hypopharynx (ICD-9 146-8) and larynx (ICD-9 161)

were included in the study. All subjects came from one of three McGill University hospitals in Montreal and were diagnosed during an 18 month period beginning July 1, 1995. This study was approved by the Human Ethics Committees of the relevant hospitals and all subjects read and signed an informed consent. Data were collected from subjects in a 20-30 minute interview using a standardised questionnaire eliciting information on socio-economic and demographic variables, plus information concerning the cancer development and symptomology, health care professionals consulted and the period of time taken for each stage in the diagnostic process. Interviews were performed between subjects' initial consultation with the treating specialist and the initiation of their cancer therapy. A random sample of 30 subjects' responses concerning professional delay were validated by checking them against their primary family physician or dentist and hospital charts.

Patient delay was considered to be the time between subjects' initially noticing a symptom and their first presentation to a health care professional (HCP) concerning that symptom. Professional delay was considered to be the period of time between the initial consultation with a HCP and the first consultation with the treating specialist and total delay was considered to be the sum of patient and professional delays. So as to improve the validity of subjects' recall of patient and professional delays, the unit used for these time periods was months.

Clinico-pathological data were collected from subjects' hospital charts. Disease stage was classified according to the American Joint Committee on Cancer (19), although for the purposes of statistical analysis, subjects were classified with early (T1 or T2 and N0) or late (T3, T4 or N>0) stage disease. Stepwise multiple logistic regression was used to calculate the odds ratio (OR) of late vs. early stage disease for selected study variables. All possible variables were entered into the original model.

Results

One hundred and eighty-eight subjects were included in the sample. Subjects'

age ranged from 34-91 years, with a mean of 64.6 years. Seventy one percent (134) of subjects were male. The distribution of cancer site and stage in the sample is shown in Table 1. The distribution by site showed a slight predominance of oral tumours (40.4%), while the distribution by stage showed slightly increased numbers of stages I and IV, with half the sample (50.5%) diagnosed with late stage disease. Within sites, the large majority of subjects with a pharyngeal cancer had late stage disease (80.7% with stages III or IV), while the majority of those with oral or laryngeal cancers had early stage disease (65.8% of oral and 58.2% of laryngeal cases with stages I or II).

The distribution of patient, professional and total delays is shown in Table 2. Patient delay for 36.7% of the sample was 1 month or less, with ranges of a few days to 18 months (median 2 mths). Almost 65% of patients were delayed by 1 month or less with their primary HCP. This professional delay ranged from a few days to 12 months (median 1 mth). Total patient and professional delay ranged from 1 week to 20 months (median 4 mths). Twenty five of the subjects (13.3%) did not notice any symptoms and their cancer was detected coincidentally when they were consulting an HCP for some other reason. This group were treated as a separate category in the analysis of patient delays. Seventy seven percent (144 subjects) of the sample presented initially to a family physician and 16.5% (31 subjects) consulted a dentist. Of the remaining subjects (6.9%) 2 were already seeing an ear nose and throat specialist concerning a previously diagnosed problem, and the remainder were consulting a specialist in a field other than otorhinolaryngology for pre-existing morbidity. Of the 31 subjects who first consulted a dentist, 27 (87.1%) had oral cancers and the remainder had pharyngeal cancers. Thus, 35.5% of oral cancer subjects, 7.0% of pharyngeal cancer subjects and none of the laryngeal cancer subjects first presented to a dentist.

Table 3 provides the crude ORs for late stage disease by selected demographic and social variables, with respective frequencies. These data suggest that patients under the age of 65yrs have a significantly increased risk of being

diagnosed with late stage disease when compared with those 65yrs or older. Furthermore, the risk of late stage disease appears to be increased among those who live alone, although the significance of this is marginal. Gender and education were not associated with disease stage.

Table 4 gives the crude OR estimates for selected health status and tumour variables in association with early and late stage disease. The general health status indicators of comorbidity and dental status at the time of diagnosis were not associated with disease stage. However, those subjects who had a mucosal lesion or voice change as their presenting symptom had a significantly reduced risk of being diagnosed with late stage disease when compared with those subjects presenting with a swelling. Furthermore, subjects with a pharyngeal cancer had odds of being diagnosed with late stage disease 8 times those of subjects with oral cancer.

The crude ORs for late stage disease associated with the different forms of diagnostic delay are shown in Table 5. No association was found between increased patient delay and risk of late stage disease. However, there was a pattern of increased odds for late stage disease with increased professional delay, with these odds being 3 times greater among those subjects delayed more that 3 months compared to those with less than 1 month's professional delay (p for trend 0.03). In addition, there was a significantly increased risk for late stage disease among all categories of total delay above 1 month. There was a pattern of increasing risk with increasing total delay with the exception of the >12 months category whose risk was approximately the same as the 1-3 and 4-6 month categories and of borderline significance. With regard to the first HCP consulted, those subjects who first consulted a dentist, rather than a family physician, had a reduced risk of late stage disease of borderline significance.

Table 6 shows the OR estimates for late stage with patient and professional delays, controlling for site. Again there is no clear pattern for laryngeal cancer. However, among oral cancer cases, there is a marked increase in odds for late stage with 1-3 months professional delay and a further increase in odds with >3 months professional delay, when compared with those odds for a delay <1 month (p for trend

0.0001). Conversely, with pharyngeal cancer cases, there is a strong tendency for increased patient delay to increase the odds for late stage disease. Professional delays appear to have little effect.

Finally, stepwise multiple logistic regression was used to build a model of variables which could best explain the risk of late stage cancer (see Table 7). This model contains tumour site, professional delay and age as independent explanatory variables. It demonstrates that: i) pharyngeal cancers have 9 times the odds of oral or laryngeal cancers for late stage disease; ii) professional delay >1 month has approximately twice the odds for late stage of professional delay <1 month; and that iii) older age (≥65yrs) patients have approximately half the odds for late stage cancer of those <65 yrs.

Discussion

The aim of this study was to investigate the hypothesised association between diagnostic delays and disease stage at diagnosis in a sample of UADT cancer patients using primary tumour site as a proxy for tumour behaviour and/or rate of growth. We found that having a pharyngeal cancer is very strongly associated with late stage disease, and that professional delays are also associated with late stage. Furthermore, our finding of a dose-response relationship for professional delay and late stage suggests that professional delay is causing an increased risk of late stage disease and so worsening patient prognosis, especially among those with oral cancers.

The findings of our study corroborate and extend those of Kowalski et al (16). Like the latter study, we found that patient delay did not predict stage, but that professional delay greater than 1 month predicted an increased risk of late stage disease. The definitions of patient and professional delay used in the two studies were the same. However, for the purposes of their statistical analyses, Kowalski et al defined the presence of patient delay as being in excess of the sample median. The finding in both studies that professional delay greater than a month is predictive of significantly increased risk of late stage disease, could be explained by the reasoning

that many patients are presenting to HCPs already with late stage disease and are being delayed further due to the obscurity of their symptoms. However, the finding in our study of a dose-response relationship between late stage and professional delay and between late stage and total delay, but not between late stage and patient delay, suggests that the professional delay is causing late stage disease. Furthermore, it would appear from Table 6 that a large part of the increased risk for late stage associated with professional delay relates to oral cancers. Despite the small numbers in this site-controlled analysis, it would appear that a large majority of patients with pharyngeal cancer are presenting to an HCP already with late stage disease and that patient delay is increasing their risk of late stage prior to presentation. It is not clear whether this situation is due to genuine patient delay (ie, patients are genuinely delaying presentation to an HCP, perhaps due to the obscurity of symptoms), or whether the onset of symptoms is relatively late in pharyngeal cancers and the length of patient delay is the same as for other sites. The vast majority of oral cancer patients are, however, presenting to an HCP with early stage disease and professional delay is leading to significant proportions of these becoming late stage cases. These data suggest that our controlling for site as a proxy for tumour behaviour, at least in relation to oral and pharyngeal cancers, has some validity and that it is sufficient to demonstrate the effect of professional delays but not patient delays. We suggest that more directly related markers of tumour behaviour are required to prevent the confounding of the effect of patient delays. The fact that we found no particular pattern for laryngeal cancers throughout the analyses suggests that our categorisation of all laryngeal cancers together is invalid. This is supported by reports of different survival rates for supraglottic, glottic and subglottic cancers (20,21). Finally, the observation of a slight reduction in risk for late stage between total delay of 6-12 months and >12 months is consistent with the idea that people with slow growing (and so low risk for late stage) tumours delay longer and have a lower risk of late stage disease than those presenting quickly. Nevertheless, the data from our sample suggest that people with a total delay >12 months still have an increased risk of late

stage disease compared with those whose total delay was <1 month.

Our finding that pharyngeal cancer patients are at much greater risk of being diagnosed with late stage disease than patients with either laryngeal or oral cancers is similar to that of Platz et al, (22). There are two possible explanations for this. Firstly, pharyngeal cancers grow and metastasize more rapidly than oral and laryngeal cancers; and/or secondly, the diagnostic delays associated with pharyngeal cancers are longer. With regard to the former, Howadlt et al (23), using the same data as Platz et al (22), concluded that there is no evidence that the biological behaviour of oral and oropharyngeal cancers is different. Nevertheless, hypopharyngeal cancers may grow more rapidly and/or they may present symptoms much later in the disease process.

The fact that, in the univariate analysis, patients presenting with a painless mucosal lesion or vocal changes had reduced risk of late stage cancer when compared with those presenting with some form of swelling is as one would expect. A painless mucosal lesion is suggestive of an early carcinoma and it is well recognised that many vocal cord cancers present early with vocal changes (24), while the presence of some form of swelling is suggestive of a large primary tumour or regional metastasis to the lymph nodes, both of which indicate late stage disease.

The finding, in the univariate analysis, that patients presenting to a dentist had a somewhat reduced risk of early stage cancer can probably be explained by two phenomena. Firstly, it needs to be recognised that 81.7% of patients presenting to a dentist had oral cancers and that the risk of late stage disease at that site is much less than that of pharyngeal cancers. Secondly, several studies have reported finding an association between physicians rather than dentists as the referral source when patients have late stage disease (17,25,26). The point being that the HCP first consulted is the dependent variable in this particular relationship. The fact that the type of primary HCP first consulted no longer remained a significant predictor of disease stage in the multiple regression analysis would tend to support the idea that this variable is related more to disease site than stage.

Our finding that age is a significant predictor of disease stage is consistent with Kowalski et al (16) who found that, in a univariate analysis, those patients 65 years and older had a marginally reduced risk of late stage disease, although this variable did not remain in their multifactorial model. The possibility that older patients have a reduced risk of late stage disease could be a reflection of the fact that the cancer of someone diagnosed at 70 years of age is highly likely to have grown more slowly than an individual whose cancer was diagnosed at 50 years of age. Thus, age may be inversely correlated with tumour growth rate for UADT cancers.

In making these inferences, it should be recognised that this was a relatively small sample size and the sample could only said to be representative of those patients diagnosed and treated at the 3 McGill University teaching hospitals concerned. Furthermore, the referral patterns to the 3 hospitals may have been different, but the sample size from each hospital did was not conducive to including this as an independent variable in the analysis. However, in Canada the majority of UADT cancer treatment is performed in a few tertiary referral centres (the case for all of the hospitals in this study) throughout the country, with many patients travelling large distances for treatment. Therefore, this sample is probably representative of patients referred to such centres in Canada, although there may be inter-provincial differences. The other potential problem is the validity of the crucial delay data which is largely dependent upon subject recall. The fact that our random validity check for professional delay demonstrated 100% accuracy suggests that these data are valid. Although one could argue that, by inference, the patient delay data is also accurate, recall of the date that a chronic symptom first appears may be less easy to remember than the date of a specific appointment with an HCP. It is conceivable that misclassification of the patient delay information may have led to a decreased statistical association between this variable and late stage.

In summary, the findings of our study suggest that, among patients with a UADT cancer, those with a pharyngeal cancer, those with a professional delay greater than 1 month and those diagnosed at a younger age (<65 years) are at

increased risk of being diagnosed with late stage disease, and consequently, of having a poorer prognosis. This study has also demonstrated a dose-response relationship between the odds of having late stage disease and professional delay. This suggests that professional delays by primary health care professionals may have a critical role in the aetiology of late stage UADT cancers, especially oral cancers. We are still far from fully understanding the inter-relationships between tumour, patient and professional behaviours, and further research in this field could provide useful information to promote the earlier diagnosis of UADT cancers.

Acknowledgements

We would like to thank Dr. George Shenouda for his help with this research project, plus the Fonds de la recherche en sante du Quebec and the Canadian Medical Research Council for providing financial support.

Table 1. Distribution of tumour site and stage.

	Early stage disease		Late stage disease		
Site	**Stage I**	**Stage II**	**Stage III**	**Stage IV**	**Total (%)**
Mouth (%)	30 (39.5)	20 (26.3)	13 (17.1)	13 (17.1)	76 (40.4)
Pharynx (%)	2 (3.5)	9 (15.8)	18 (31.6)	28 (49.1)	57 (30.3)
Larynx (%)	23 (41.8)	9 (16.4)	11 (20.0)	12 (21.8)	55 (29.3)
Total (%)	55 (29.3)	38 (20.2)	42 (22.3)	53 (28.2)	188

Table 2. Distribution of patient, professional and total delays.

Variable	**Category**	**N (%)**
Patient delay	**<1 month**	69 (36.7%)
	1-3 months	47 (25.0%)
	>3 months	47 (25.0%)
	"no symptom"	25 (13.3%)
Professional delay	**<1 month**	122 (64.9%)
	1-3 months	38 (20.2%)
	>3 months	28 (14.9%)
Total delay	**<1 month**	37 (19.7%)
	1-3 months	41 (21.8%)
	4-6 months	54 (28.7%)
	>6 months	46 (29.8%)

NB. "no symptom" refers to a group of patients whose cancer was first noticed by an HCP.

Table 3. Crude OR estimates for late stage UAT cancer according to selected demographic and social variables.

Variable	Category	Early/late stage disease	OR	95% CI
Age	<65yrs:	37/53	1.91	1.07-3.41
	≥65yrs:	56/42	1.0	Ref.
Gender	male:	66/68	1.0	Ref.
	female:	27/27	0.97	0.52-1.82
Cohabitation	living with family:	80/72	1.0	Ref.
	living alone:	13/23	1.97	0.93-4.17
Educational level	secondary education or less:	60/60	1.0	Ref.
	more than secondary education:	33/35	1.06	0.58-1.92

OR: odds ratio. CI: confidence interval.

Table 4. Crude OR estimates for late stage UAT cancer according to selected health status and tumour variables.

Variable	Category	Early/late stage disease	OR	95% CI
Comorbidity status	comorbidity absent:	63/64	1.0	Ref.
	comorbidity present:	30/31	1.01	0.55-1.86
Dental status	1 or more teeth:	54/59	1.0	Ref.
	edentate:	39/36	0.84	0.47-1.51
Presenting symptomology	swelling:	25/44	1.0	Ref.
	pain:	25/29	0.66	0.32-1.36
	mucosal lesion:	15/6	0.23	0.08-0.67
	voice change:	28/16	0.32	0.15-0.7
Disease site	oral:	50/26	1.0	Ref.
	pharynx:	11/46	8.04	3.57-18.09
	larynx:	32/23	1.38	0.67-2.82

Table 5. Crude OR estimates for late stage UAT cancer according to various forms of delay.

Variable	Category	Early/late stage disease	OR	95% CI
Patient delay	<1 month	37/32	1.0	Ref.
	1-3 months	20/27	1.57	0.74-3.31
	>3 months	21/26	1.44	0.68-3.03
	"no symptom"	15/10	0.78	0.31-1.98
Professional delay	<1 month	68/54	1.0	Ref.
	1-3 months	17/21	1.56	0.75-3.27
	>3 months	8/20	3.16*	1.29-7.73
Total delay	<1 month	26/11	1.0	Ref.
	1-3 months	19/22	2.76	1.08-7.03
	4-6 months	25/29	2.76	1.14-6.69
	7-12 months	12/20	3.98	1.46-10.87
	>12 months	11/13	2.81	0.97-8.18
Type of HCP first consulted	family physician	66/78	1.0	Ref.
	dentist	20/11	0.47	0.21-1.05
	other specialist	7/6	0.73	0.23-2.28
Tumour first noticed by	patient	78/85	1.0	Ref.
	HCP	15/10	0.61	0.26-1.44

"no symptom": those patients whose cancer was first noticed by an HCP

* p value for trend 0.03

Table 6. OR estimates for late stage according to patient and professional delay, controlling for cancer site.

Site	Delay	Early/late stage disease	OR	95% CI
Mouth	patient delay			
	<1 month	18/6	1.0	Ref.
	1-3 months	12/7	1.76	0.47-6.54
	>3 months	11/6	1.67	0.43-6.49
	"no symptoms"	9/7	2.36	0.61-9.13
	professional delay			
	<1 month	42/12	1.0	Ref.
	1-3 months	5/6	4.4	1.07-15.96
	>3 months	3/8	9.2*	2.11-40.17
Pharynx	patient delay			
	<1 month	6/21	1.0	Ref.
	1-3 months	1/16	4.57	0.5-41.86
	>3 months	0/8	∞	
	"no symptoms"	3/2	0.19	0.03-1.41
	professional delay			
	<1 month	6/28	1.0	Ref.
	1-3 months	2/8	0.86	0.14-5.11
	>3 months	2/11	1.18	0.21-6.76
Larynx	patient delay			
	<1 month	12/7	1.0	Ref.
	1-3 months	8/3	0.66	0.13-3.34
	>3 months	10/11	1.9	0.54-1.91
	"no symptoms"	3/1	0.57	0.05-6.59
	professional delay			
	<1 month	23/11	1.0	Ref.
	1-3 months	8/9	2.35	0.71-7.75
	>3 months	2/2	2.08	0.26-16.76

* p value for trend 0.0001

Table 7. Multifactorial model for the OR estimates for late stage UAT cancer.

Variable	Category	OR	95% CI	p
Site	pharynx	9.26	4.02-21.32	0.0001
Professional delay	>1 month	2.28	1.13-4.64	0.022
Age	≥65 years	0.45	0.22-0.91	0.024

References

1. Parker SL, Tong T, Bolden S and Wingo PA (1996) *Cancer statistics 1996.* CA Cancer J. Clin 46(1):5-28

2. Flamant R, Hayem M, Lazar P and Denoix P (1964): *Cancer of the tongue; a study of 904 cases.* Cancer 17:377-85

3. Jorgensen K, Elbrond O and Anderson AP (1973): *Carcinoma of the lip, a series of 869 cases* Acta Radiol. 12:177-90

4. Pogrel MA (1974): *The dentist and oral cancer in North East Scotland.* Br. Dent. J. 137:15-20

5. Shafer WG (1975): *Initial mismanagement and delay in diagnosis of oral cancer.* JADA 90:1262-64

6. Cooke BED and Tapper-Jones L (1977): *Recognition of oral cancer - cases of delay.* Br. Dent. J. 142:96-8

7. Scully C, Malmos D, Levers BGH, Porter SR and Prime SS (1986): *Sources and patterns of referrals of oral cancer: role of general practitioners.* BMJ 293:599-601

8. Dimitroulis G, Reade P and Wiesenfeld D (1992): *Referral patterns of patients with oral squamous cell carcinoma, Australia.* Oral Oncol. Eur. J. Cancer 28B(1):23-7

9. Bruun JP (1976): *Time lapse by diagnosis of oral cancer.* Oral Surg. Oral Med. Oral Path. 42(2):139-49

10. Jovanovic A, Kostense PJ, Schulten EAJM, Snow GB and van der Waal I (1992): *Delay in diagnosis of oral squamous cell carcinoma; a report from The Netherlands.* Oral Oncol. Eur. J. Cancer 28B(1):37-8

11. Kaufman S, Grabau JC and Lore JH (1980): *Symptomology in head and neck cancer; a quantitative review of 385 cases.* Am. J. Public Health 70:520-2

12. Guggenheimer J, Verbin RS, Johnson JT, Horkowitz CA and Myers EN (1989): *Factors delaying the diagnosis of oral and oropharyngeal cancers.* Cancer 64:932-5

13. Evans SJW, Langdon JD, Rapidis AD and Johnson NW (1982): *Prognostic significance of STNMP and velocity of tumour growth in oral cancer* Cancer 49:773-6

14. Elwood JM and Gallagher RP (1985): *Factors in the early diagnosis of cancer of the oral cavity.* Can Med. Assoc. J. 133:651-6

15. Franco EL, Dib LL, Pinto DS, Lombardo V and Contesini H (1993): *Race and gender influences on the survival of patients with mouth cancer.* J. Clin Epidemiol. 46(1)37-46

16. Kowalski LP, Franco EL, Torloni H, Fava AS, Sobrinho JdeA, Ramos G, Oliveira BV and Curado MP (1994): *Lateness of diagnosis of oral and oropharyngeal cancer: factors related to the tumour, the patient and health professionals.* Oral Oncol. Eur. J. Cancer 30B(3):167-73

17. Gorsky M and Dayan D (1995): *Referral delay in diagnosis of oro/oropharyngeal cancer in Israel.* Oral Oncol. Eur. J. Cancer 31B(3):166-8

18. Wildt J, Bundgaard T and Bentzen SM (1995): *Delay in the diagnosis of oral squamous cell carcinoma.* Clin. Otolaryngol. 20:21-25

19. American Joint Committee on Cancer: *Manual for Staging of Cancer.* 4th Ed'n. J.B.Lippincott Co., Philadelphia, 1992

20. Austin DF: *Larynx.* In Schottenfeld D and Fraumeni JF (eds): *Cancer epidemiology and prevention.* WB Saunders Co, Philadelphia 1982

21. Kowalski LP, Franco Le, de Andrade Sobrinho J, Oliviera BV and Pontes PL (1991): *Prognostic factors in laryngeal cancer patients submittes to surgical treatment.* J. Surg. Oncol. 48(2):87-95

22. Platz H, Fries R and Hudec M (1986): *Prognoses of oral cavity carcinomas.* Carl Hanser, Munich.

23. Howaldt H-P, Frenz M and Pitz H (1992): *Proposal for a modified T-classification for oral cancer.* J. Cranio-Maxillo-Facial Surgery 22:96-101

24. Benner SE, Lippman SM and Hong WK: *Cancer of the head and neck.* In Greenwald P, Kramer BS and Weed DL (eds): *Cancer Prevention and Control.* Marcel Dekker Inc., New York, 1995.

25. Amsel Z, Strawitz JG and Engstrom PF (1983): *The dentist as a referral source of first episode head and neck cancer patients.* JADA 106(2):195-7

26. Schnetler JFC (1992): *Oral cancer diagnosis and delays in referral.* Br. J. Oral Maxillofacial Surg. 30:210-13

NB. End of manuscript II

2.10 The predictors of professional diagnostic delays: methodology

In view of the finding that professional delays greater than 1 month were a significant predictor of increased odds for late stage disease (as reported and discussed in manuscript II), a further analysis was performed to investigate those variables independently associated with a professional diagnostic delay >1 month. Thus, in this analysis, the dependent variable was professional diagnostic delays, which was subsequently converted to a dichotomous variable of delay <1 month and delay >1 month. For this analysis, the majority of independent variables were nominal (patient gender, cohabitation status, education status, comorbidity status, dental status, presenting symptomology and disease site), the ordinal variable patient delay was converted to a dichotomous one (<1 month v >1 month) and the one remaining continuous variable (patient age) was converted into a dichotomous variable (<65 years v ≥65 years) centred around the sample mean age (64.6 years). The statistical analysis strategy used to formulate a multivariate explicative model for professional diagnostic delay >1 month was the same as that described in section 2.8.2.

2.11 Manuscript III: "Predictors of professional diagnostic delays for upper aerodigestive tract carcinoma"

This manuscript is reprinted from Oral Oncology (in press) Allison PJ, Franco E and Feine JS: *Predictors of professional diagnostic delays for upper aerodigestive tract carcinoma,*

Abstract

Despite the belief that cancer mortality can be reduced if lesions are detected, diagnosed and treated at an early stage, the predictors of diagnostic delays for upper aerodigestive tract (UADT) cancers have been the subject of little research. This study was aimed to investigate the role of selected variables as predictors of professional diagnostic delays in a sample of UADT cancer patients. Patients diagnosed with squamous cell carcinoma of oral cavity sites (ICD-9 141, 143-5) oro-, naso- and hypopharynx (ICD-9 146-8) and larynx (ICD-9 161) were included in the study. Multiple logistic regression was used to calculate the odds ratio (OR) for professional delay >1 month vs. professional delay <1 month for selected study variables. The sample comprised 188 subjects. Multivariate analysis found that the presence of comorbidity at the time of presentation of UADT symptoms (OR:2.84; 95%CI:1.35-5.98), age ≥65yrs (OR:0.31; 95%CI:0.15-0.64), higher education (OR:0.45; 95%CI:0.22-0.93) and cancer at an oral cavity site (OR:0.31; 95%CI:0.15-0.64) were the explanatory variables for professional delay. This study suggests that, among UADT cancer patients, the presence of comorbidity at the time of presentation increases the odds for a professional delay >1mth, while older age, higher education and oral cancer reduce the odds.

Key words: diagnosis, head and neck cancer, prognosis

Introduction

Based upon our understanding of the temporal growth and development of cancers, it is generally believed that the earlier their diagnosis and treatment, the

better the prognosis for the patient. With this in mind, there have been a number of publications describing and investigating the role of diagnostic delays in relation to cancers of the upper aerodigestive tract (UADT). Early studies concentrated more upon patients as the source of diagnostic delays (1,2); ie, after having noticed some form of symptom(s) patients delayed presentation to a health care professional (HCP). However, it was soon recognised that primary HCPs (principally family physicians and dentists) were also an important source of diagnostic delays in that they prolonged the period between the patients' first presentation to them and the definitive diagnosis with the treating specialist. Professional misdiagnoses and delays have been documented in several countries (3-9). However, such descriptive studies shed little light on the roles of patient and professional delays in determining patient prognosis, and few analytic studies have been published with reference to this subject and UADT cancers.

Indeed, among UADT cancer patients, no studies have found an association between increased patient delays and worse prognosis, while only one study (10) has found an association between professional delays and worse prognosis. It has been suggested that the reason for this is the confounding effect of tumour growth. Evans et al (11) have suggested that patients with fast growing tumours were being diagnosed quickly but with late stage disease, while those with slow growing tumours had long diagnostic delays but were diagnosed with early stage disease.

Despite the equivocal nature of the evidence concerning an association between diagnostic delays and UADT cancer patient prognosis, a number of studies have investigated factors associated with professional diagnostic delays in this group. In the UK, a study of 96 cases of oral cancer demonstrated that family physicians were significantly less likely to delay referral and more likely to make the correct diagnosis than dentists (12). In a retrospective study of 543 cases of oral and oropharyngeal cancer cases in Israel, it was found that diagnostic delay and early disease stage are associated with lip cancers (13). Finally, in Denmark, a study of 167 patients with oral cancer demonstrated no association between patient or professional

delay and stage, but correlations were found between increased professional delay and small tumour size, women and older aged patients (14).

In summary, although the picture is unclear, it would seem that professional diagnostic delays associated with UADT cancers do occur. Furthermore, the evidence that exists suggests that tumour variables, such as site and size, patient variables, such as age, and professional variables such type of health care professional, are probably important factors associated with professional diagnostic delays for this group of patients. In light of this, a study was undertaken to investigate variables associated with professional delays in a sample of UADT cancer patients.

Methodology

Patients diagnosed with squamous cell carcinoma of the oral cavity and contents (ICD-9 141, 143-5) oro-, naso- and hypopharynx (ICD-9 146-8) and larynx (ICD-9 161) were included in the study. All subjects were diagnosed during an 18 month period from July 1st 1995 in one of three McGill University teaching hospitals in Montreal. Data were collected from subjects in a 20-30 minute interview using a standardised questionnaire eliciting information on socio-demographic and health status variables, plus information concerning the cancer development and symptomology, health care professionals consulted and the period of time taken for each stage in the diagnostic process. Interviews were performed between subjects' initial consultation with the treating specialist and the initiation of their cancer therapy. A random sample of 30 subjects' responses concerning professional delay were validated by checking them against their primary family physician or dentist and hospital charts.

Patient delay was considered to be the time between subjects' initially noticing a symptom and their first presentation to an HCP concerning that symptom. Professional delay was considered to be the period of time between the initial consultation with an HCP and the first consultation with the treating specialist, and total delay was considered to be the sum of patient and professional delays. Multiple

logistic regression was used to calculate the odds ratio (OR) of a professional delay >1 month Vs. a professional delay <1 month for selected study variables.

Results

One hundred and eighty-eight subjects were included in the sample. Subjects' age ranged from 34-91 years, with a mean of 64.6 years. One hundred and thirty four subjects (71%) were male. The distribution by tumour site showed a slight predominance of oral tumours (40.4%), with pharyngeal (30.3%) and laryngeal (29.3%) tumours occupying approximately equal proportions of the remainder of the sample (Table 1). Professional delays of less than 1 month comprised the majority for each tumour site. However, this proportion was larger among oral cancer cases.

The distribution of professional delays by the type of health care professional consulted is shown in Table 2. In the sample as a whole, approximately 35% of patients were delayed by 1 month or more with their primary HCP. This professional delay ranged from a few days to 12 months (median 1 mth). One hundred and forty four subjects (77%) presented initially to a family physician and 31 subjects (16.5%) initially consulted a dentist. Of the remaining subjects (6.9%) 2 were already seeing an ear nose and throat specialist concerning a previously diagnosed problem, and the remainder were consulting a specialist in a field other than otorhinolaryngology for pre-existing morbidity. Of the 31 subjects who first consulted a dentist, 27 (87.1%) had oral cancers and the remainder had pharyngeal cancers. Thus, 35.5% of oral cancer subjects, 7.0% of pharyngeal cancer subjects and none of the laryngeal cancer subjects first presented to a dentist.

The crude estimates for the OR for a professional delay >1 month, according to selected sociodemographic variables, are demonstrated in Table 3. Of these variables, age exerts the strongest influence on delay, with a significant increase in the odds (OR 2.75) for a professional delay >1 month among those younger than 65 years compared to those aged 65 years or more. There was a non-significant tendency towards a reduced odds for professional delay among those with higher education.

However, no association was found between professional delay and gender or cohabitation.

Table 4 shows crude estimates for the odds of professional delay according to selected patients' health status variables. There were no significant associations between these health status variables and professional delay. However, there was a tendency for concurrent comorbidity to increase the odds for professional delay and for oral cancers and painless mucosal lesions to be associated with reduced odds for professional delay.

Crude estimates for the OR for professional delay associated with previous patient delay and the type of HCP consulted are given in Table 5. There was no association between patient delay and professional delay. However, there was a tendency for reduced odds for professional delay if the patient consulted a dentist with their symptoms and a significant reduction in the odds for professional delay if the first HCP consulted regarding tumour symptoms was a specialist.

Finally, multiple logistic regression was used to build a model for independent variables which could best explain the risk for a professional delay of >1 month (see Table 6). This model contains four variables significantly associated with delay. The presence of concurrent comorbidity when the patient was consulting the HCP concerning their tumour symptoms doubled the odds for having a professional delay compared to patients with no such comorbidity. Patients aged 65 years or older and those with an oral cancer had approximately one third the odds for a professional delay compared to patients under 65 years of age and to patients with cancers at other sites, respectively. In addition, patients with higher education had approximately half the odds for a professional delay compared to those with minimum education or less.

Discussion

The aim of this study was to investigate selected tumour, patient and health care professional variables as predictors of professional delay in a sample of UADT

cancer patients. We have found that having some form of comorbidity present at the time of presentation to an HCP increases the odds for professional delay by that HCP. We have also found that older age (≥65 years), higher education and having an oral (as opposed to pharyngeal or laryngeal) cancer are predictive of reduced odds for professional delay.

There are few published analytic studies concerning the role of diagnostic delays in the prognosis of UADT cancers and even fewer with professional delay as the dependent variable. Our finding that patients with concurrent comorbidity at the time of presenting their UADT cancer symptoms to an HCP had nearly 3 times the odds for a professional delay has not been previously mentioned in the UADT literature. If this finding is confirmed by future work, it has important implications. It is perhaps understandable that an HCP already monitoring a patient with a previously diagnosed illness may prioritize the stabilisation of that illness over the diagnosis of the cause of relatively obscure and apparently innocuous UADT symptoms. However, the majority of patients diagnosed with a UADT cancer are elderly people who have smoked for a long period and often been drinking excessive quantities of alcohol. Such people are highly likely to have a collection of chronic illnesses, and it should be a practice of primary HCPs to closely screen those with one chronic illness diagnosis for other chronic illnesses of similar aetiology.

With regard to the association between tumour site and professional delay, the findings in our study agree with those of Kowalski et al (10) and Wildt et al (14), although their samples respectively included lip, oral and oropharyngeal cancers (10), and oral cancers only (14). Although the former study did not investigate the predictors of professional delay as a dependent variable, they found that "less visible tumours" predicted late stage and that professional delay >1 month predicted late stage lip and oral cancers (10). We have found that oral (more visible than pharyngeal and laryngeal) cancers are associated with reduced odds (OR: 0.31) for professional delay compared with other sites. Similarly, Wildt et al (1995) found that small (and so less visible) oral tumours increased professional delay (14).

With respect to sociodemographic variables, our findings differ from those of several other studies. Wildt et al (1995) found a significant positive correlation between older age and increased professional delay. However, their sample included only oral cancer cases; they used univariate analyses only and their correlation, although significant, was weak ($r=0.19$, $p=0.02$) (14). We found a significant association between older age and reduced odds for professional delay in the univariate analysis, which remained as a predictive factor in the multivariate analysis. We found no association between gender and diagnostic delays. This differs from other investigators who found professional delay (14) and total diagnostic delay (15) to be longer for women. However, apart from the sample differences, the correlation found by Wildt et al (1995) was, once again, significant but weak and was calculated using Spearman's rank correlation test which is inappropriate for a nominal variable such as gender (14). Elwood and Gallagher's (1985) analysis was between gender and total, rather than professional, delay (15). Finally, with respect to sociodemographic variables, we have found that patients with higher education have reduced odds for professional delay. This finding is consistent with the idea that people with a higher education feel more comfortable with HCPs, and may be more likely to request referral if they feel it appropriate, than would those with a minimal education (16).

One of the variables which we did not find predictive of professional delay on multivariate analysis was the type of health care professional responsible for the delay. In a sample of oral cancer cases, Schnetler (1992) found physicians significantly less likely to delay referral to a specialist than dentists when delay was defined as beyond 2 days (12). However, he found no statistical difference between dentists and physicians with a professional delay of 3 weeks, a delay much closer to the 1 month period used in our study (12). In the present study (which included pharyngeal and laryngeal cancers), the univariate analysis showed a substantial reduction in the odds for delay with dentists and a significant reduction in the odds for delay with specialists. However, the type of referring HCP did not remain in the

multiple regression model. The tendency for reduced odds for a delay among dentists can be explained by the fact that, in this sample, 81.7% of the cases they referred were oral cancers which itself is associated with reduced odds for delay. There remains the possibility that dentists are less likely to delay patients with an oral cancer longer than a month. However, although a site-controlled analysis showed a tendency for reduced odds for delayed referral among oral cancers, our sample was insufficient to demonstrate a significant reduction in likelihood of delay among dentists. As for the fact that the odds for delay with a specialist were significantly reduced in the univariate analysis, one would expect that most specialists would immediately refer a medical problem outside the realm of their expertise to the appropriate colleague.

While making these inferences, certain limitations of the study should be recognised. The sample size was relatively small and could only said to be representative of those patients diagnosed and treated at the 3 participating McGill University hospitals. Nevertheless, in Canada the majority of UADT cancer treatment is performed in a few tertiary referral centres, therefore this sample is probably representative of patients referred to such centres in Canada, although there may be inter-provincial differences. Another potential problem is the validity of the crucial delay data which is largely dependent upon subject recall. However, the fact that our random validity check for professional delay demonstrated 100% accuracy suggests that these data are valid. Furthermore, one could argue that, by inference, the patient delay data is also accurate. However, recall of the date that a chronic symptom first appears may be less easy to remember than the date of a specific appointment with an HCP.

In summary, the findings of this study suggest that, in a sample of UADT cancer patients, the presence of comorbidity at the time of presentation of tumour symptoms to an HCP increase the odds for professional delay, while older age, higher education and oral cancers reduce the odds for professional delay. While these findings require confirmation through investigations elsewhere and with larger

samples, they give important indications of possible targets for interventions to reduce diagnostic delays among UADT patients and others with epidemiologically and aetiologically similar cancers.

Acknowledgements

We would like to thank Dr. Martin Black for assistance in patient accrual. This research project was supported by the Fonds de la recherche en santé du Québec and the Medical Research Council of Canada.

Table 1. Distribution of professional delays by tumour site.

Site	Delay <1 mth	Delay 1-3 mths	Delay >3 mths	Total (%)§
Oral (%)*	54 (71.0%)*	11 (14.5%)*	11 (14.5%)*	76 (40.4%)§
Pharyngeal (%)*	34 (59.6%)*	10 (17.5%)*	13 (22.8%)*	57 (30.3%)§
Laryngeal (%)*	34 (61.8%)*	17 (30.9%)*	4 (7.3%)*	55 (29.3%)§
Total (%)*	122 (64.9%)*	38 (20.2%)*	28 (14.9%)*	188

* Percentages of oral, pharyngeal and laryngeal cancer cases, plus the whole sample with delays of <1 mth, 1-3 mths and >3 mths respectively.
§ Percentages of the total sample with oral, pharyngeal and laryngeal cancers

Table 2. Distribution of professional delays by type of health care professional (HCP) first consulted

HCP	Delay <1 mth	Delay 1-3 mths	Delay >3 mths	Total (%)§
Family physician (%)*	88 (61.1%)*	32 (22.2%)*	24 (16.7%)*	144 (76.6%)§
Dentist (%)*	23 (74.2%)*	4 (12.9%)*	4 (12.9%)*	31 (16.5%)§
Specialist (%)*	11 (84.6%)*	2 (15.4%)*	0	13 (6.9%)§
Total (%)*	122 (64.9%)*	38 (20.2%)*	28 (14.9%)*	188

* Percentages of patients first consulting a family physician, dentist or specialist, plus the total sample with professional delays of <1 mth, 1-3 mths and >3 mths respectively

§ Percentages of the total sample first consulting a family physician, dentist or specialist

Table 3. Crude OR estimates for professional delay of >1 month, according to selected sociodemographic variables related to patients

Variable	Category	No delay /delay	OR	95% CI
Age	<65 yrs	48/42	2.75	1.48-5.11
	≥65 yrs	74/24	1.0	Ref.
Gender	male	88/46	1.0	Ref.
	female	34/20	1.13	0.59-2.18
Cohabitation	living with family	98/54	1.0	Ref.
	living alone	24/12	0.91	0.42-1.96
Education	minimum education	73/47	1.0	Ref.
	post-secondary education	49/19	0.61	0.32-1.16

OR: odds ratio. CI: confidence interval

Table 4. Crude OR estimates for professional delay of >1 month, according to selected health status variables related to patients

Variable	Category	No delay/delay	OR	95% CI
Comorbidity	absent	87/40	1.0	Ref.
	present	35/26	1.61	0.86-3.02
Dentition	dentate	71/42	1.0	Ref.
	edentate	51/24	0.80	0.43-1.48
Symptomology	swelling	44/25	1.0	Ref.
	pain	34/20	1.04	0.51-2.18
	mucosal lesion	17/4	0.42	0.13-1.39
	voice change	27/17	1.11	0.51-2.42
Tumour site	oral	54/22	1.0	Ref.
	pharyngeal	34/23	1.67	0.81-3.52
	laryngeal	34/21	1.51	0.72-3.15

Table 5. Crude OR estimates for professional delay of >1 month, according to selected consultation process variables

Variable	Category	No delay/delay	OR	95% CI
Patient delay	<1 mth	44/25	1.0	Ref.
	1-3 mths	29/18	1.09	0.51-2.34
	>3 mths	30/17	1.0	0.46-2.16
	"no symptoms"	18/7	0.68	0.26-1.79
HCP consulted	family physician	88/56	1.0	Ref.
	dentist	23/8	0.55	0.23-1.31
	specialist	11/2	0.28	0.04-0.84

NB. "no symptoms" refers to those subjects who noticed no symptoms

HCP: health care professional

Table 6. Multifactorial model for the OR estimates for a professional delay >1 month

Variable	Categories contrasted	OR	95% CI
Comorbidity	*present* v absent	2.84	1.35-5.98
Age	*≥65* v <65yrs	0.31	0.15-0.64
Education	*post secondary* v minimum or less	0.45	0.22-0.93
Tumour site	*oral* v pharyngeal and laryngeal	0.31	0.15-0.64

References

1. Flamant R, Hayem M, Lazar P and Denoix P (1964): *Cancer of the tongue; a study of 904 cases*. Cancer 17:377-85

2. Jorgensen K, Elbrond O and Anderson AP (1973): *Carcinoma of the lip, a series of 869 cases* Acta Radiol. 12:177-90

3. Pogrel MA (1974): *The dentist and oral cancer in North East Scotland.* Br. Dent. J. 137:15-20

4. Bruun JP (1976): *Time lapse by diagnosis of oral cancer.* Oral Surg. Oral Med. Oral Path. 42(2):139-49

5. Shafer WG (1975): *Initial mismanagement and delay in diagnosis of oral cancer.* JADA 90:1262-64

6. Cooke BED and Tapper-Jones L (1977): *Recognition of oral cancer - cases of delay.* Br. Dent. J. 142:96-8

7. Scully C, Malmos D, Levers BGH, Porter SR and Prime SS (1986): *Sources and patterns of referrals of oral cancer: role of general practitioners.* BMJ 293:599-601

8. Dimitroulis G, Reade P and Wiesenfeld D (1992): *Referral patterns of patients with oral squamous cell carcinoma, Australia.* Oral Oncol. Eur. J. Cancer 28B(1):23-7

9. Jovanovic A, Kostense PJ, Schulten EAJM, Snow GB and van der Waal I (1992): *Delay in diagnosis of oral squamous cell carcinoma; a report from The Netherlands.* Oral Oncol. Eur. J. Cancer 28B(1):37-8

10. Kowalski LP, Franco EL, Torloni H, Fava AS, Sobrinho JdeA, Ramos G, Oliveira BV and Curado MP (1994): *Lateness of diagnosis of oral and oropharyngeal cancer: factors related to the tumour, the patient and health professionals.* Oral Oncol. Eur. J. Cancer 30B(3):167-73

11. Evans SJW, Langdon JD, Rapidis AD and Johnson NW (1982): *Prognostic significance of STNMP and velocity of tumour growth in oral cancer* Cancer 49:773-6

12. Schnetler JFC (1992): *Oral cancer diagnosis and delays in referral.* Br. J. Oral Maxillofacial Surg. 30:210-13

13. Gorsky M and Dayan D (1995): *Referral delay in diagnosis of oro/oropharyngeal cancer in Israel.* Oral Oncol. Eur. J. Cancer 31B(3):166-8

14. Wildt J, Bundgaard T and Bentzen SM (1995): *Delay in the diagnosis of oral squamous cell carcinoma.* Clin. Otolaryngol. 20:21-25

15. Elwood JM and Gallagher RP (1985): *Factors in the early diagnosis of cancer of the oral cavity.* Can Med. Assoc. J. 133:651-6

16. Antonovsky A and Hartman H (1974): *Delay in the detection of cancer: A review of the literature.* Hlth Educ Monographs 2(2):127-40

NB. End of manuscript III

2.12 Predictors of patient delays

The findings of the analysis to evaluate predictors of late vs. early stage disease (reported in manuscript II) were that professional delays are associated with increased odds for late stage disease, while patient delays were not. Although this latter finding agrees with work from other groups, it should be recognized that this

may be a genuine feature of the diagnostic process for this group of cancer patients or it could be a false negative finding. The latter is a strong possibility for four principal reasons:

i) The validity of the delay data is questionable due to its dependence on recall. It could be argued that the fact that the validity check for professional delay data (see section 2.8) suggested good validity for this variable infers good validity for the patient delay variable. However, it is probably more difficult to recall the commencement of the patient delay period as defined by the moment of onset of an insidious symptom, which for many patients was painless, than to recall a consultation with an HCP. Although no evidence exists to support or contradict this assumption, it seems a strong possibility that patient delay data has poorer validity than professional delay data.

ii) The categorization of patient delays used was inappropriate. Although the a priori categorization of delays seems to be have been appropriate for professional delays, it may not be so for patient delays. It may be that the relationship between disease development and patient delays is different to that with professional delays. Perhaps disease development is not uniform (as is the case with diseases other than UADT cancers) and the effect of professional delays is stronger because a number of patients are presenting at a critical moment in the disease process. Once again, this is all hypothetical, but demonstrates our lack of understanding of the progression of UADT cancers.

iii) The relationship between patient delay and disease stage is still confounded by tumour growth rate. As previously mentioned, no effort to directly measure tumour growth rate was made in this study, although this variable was hypothetically controlled for by categorising subjects into anatomical subsites, assuming that there is some form of uniformity of growth rate within those subsites. This assumption

appears to be at least partially supported by the findings of this research. However, it may be the case that the relationship between patient delays and disease stage is too subtle to be demonstrated with such a relatively crude control for growth rate.

iv) The sample size may not be large enough to demonstrate any effect. The data suggest a slight (statistically insignificant) increase in odds for late stage disease with a patient delay >1 month (see Table 5, manuscript II) and the possibility of a fair increase in odds for late stage disease among pharyngeal cancer patients who delayed >1 month (see Table 6, manuscript II). Both of these findings may have been confirmed as statistically significant with a larger sample size.

In view of these observations, and the subsequent doubt in the validity of the finding of no association between patient delay and disease stage, a third analysis was performed to evaluate those variables which best explained patient delay >1 month. With the exception of the patient delay variable itself, the independent variables used in this analysis were the same as those used in the evaluation of predictors of professional delay. Similarly, the analytic strategy to formulate a multivariate model best explaining the variation in patient delays was the same as that used for manuscripts II and III (see section 2.8.2).

The result of this multivariate logistic regression analysis was that none of the independent variables had a statistically significant association with the patient delay variable. Once again, while it is quite possible that none of these variables are associated with patient delays, it seems more likely that this finding is a product of invalid patient delay data, or at least categorization thereof.

2.13 Summary of the findings of the investigation into diagnostic delays among UADT cancer patients

Notwithstanding the problems of sample size and representativeness and concerns about the validity of the crucial delay data as discussed in manuscripts II

and III, the findings of this research confirm and extend previous work in this field. In terms of tumour biology, among the variables collected in this project, tumour site seems to have an extremely important impact upon the odds of being diagnosed with early or late stage disease and upon the odds of having a professional delay >1 month. Results from this research suggest that, compared to patients with oral or laryngeal cancer, those with a pharyngeal cancer have approximately 9 times the odds for late stage disease. However, it is not clear from these results why this is the case. Is it because i) pharyngeal cancers tend to be more aggressive; ii) pharyngeal cancers are detected later in the disease process; or iii) a combination of the two? Data from manuscripts II and III suggest that patients with an oral cancer are less likely to have a professional delay >1 month, but that when they do this significantly increases the odds for late stage disease. Thus, compared to patients with a pharyngeal cancer, those with an oral cancer independently have reduced odds for late stage disease and professional diagnostic delay. Although it is impossible to draw any firm conclusions from the data in this research, they suggest that patients with pharyngeal cancer (compared to those with oral cancer) have increased odds for professional delay, and then (independent of professional delay) increased odds for late stage disease. This would suggest that there is a problem with the detection of the disease by both patients and professionals, such that by the time it is treated, it has already progressed to a late stage. The subsequent question would then be: is this because (again, compared to oral cancer symptoms) i) pharyngeal cancer symptoms present later in the disease process; or ii) pharyngeal cancer symptoms are more benign or less evident for patients and professionals alike, thereby provoking less reaction by both? These very tentative observations agree with those of Howaldt et al who have concluded that, contrary to the popular assumption, there is no evidence to support the hypothesis that oropharyngeal cancers grow/spread faster than do oral cancers (91).

With respect to patients with laryngeal cancer, it is very difficult to draw any conclusions other than that, compared to those with pharyngeal cancers, they

evidently have reduced risk for late stage disease. The problems of the validity of categorizing all forms of laryngeal cancer together have been discussed (see manuscript II), nevertheless, the sample size in this project was not sufficient to permit further categorization of tumour site. Conclusions concerning the nature of diagnostic problems for laryngeal cancers will therefore have to await further research with an appropriate sample.

The next important predictor for patient prognosis is age. The finding in the first analysis that older patients had reduced risk for late stage disease, independent of professional delays, and in the subsequent analysis that older patients had reduced risk for professional delay, confirms the problems associated with diagnosing UADT cancer in younger patients. The first finding suggests that age is a biological variable associated with the speed of development of the cancer. That is, among patients in whom the cancer is diagnosed at an older age, the cancer is slower growing, providing more time to diagnose and treat the disease at an early stage in its progression, thus permitting a better prognosis. The finding in the second analysis, that age is associated with professional delays, suggests that it also acts through sociodemographic and/or behavioural factors. Although it is impossible to know exactly why older patients have a reduced risk for professional delay, the most plausible explanation would be that HCPs have a higher degree of suspicion of symptoms with older patients than with their younger peers.

The third predictor of late stage disease suggested by this work was professional delays >1 month. The validity of this observation was reinforced by the additional finding of a significant trend for increased odds for late stage disease with increased professional delay. Assuming this is a true positive finding, its value lies in it being the only variable suitable as a target for an intervention aimed at improving the diagnostic process for this group of cancer patients. The second analysis of the predictors of professional delay provides further information which could be of use in such an intervention. The findings suggest that primary HCPs should be more suspicious of symptoms in younger patients, be more vigilant with

less (visibly) evident symptoms, be more ready to listen to the complaints of poorly educated patients who may be less able to express themselves, and be aware of the possibility of comorbidity among a group who, due to their risk factor consumption, are highly likely to have more than one chronic health problem.

In conclusion, the findings of this research suggest that tumour site, patient age and professional behaviour affect patient prognosis in UADT cancer patients, but that the finding of no association between patient delays and patient prognosis is probably a product of invalid data and insufficient study power.

SECTION 3

AN INVESTIGATION OF THE POST-TREATMENT HEALTH STATUS OUTCOMES OF UPPER AERODIGESTIVE TRACT CANCER PATIENTS

3.1 An introduction to the investigation of health status in UADT cancer patients

This section of the research concerns an investigation of those factors associated with UADT cancer patients' health status following their treatment. It will firstly include a discussion of the definition of health status and subsequently discuss issues concerned with validly measuring health status, part of which will be in the form of a manuscript (manuscript IV). This will then be followed by a review of the literature concerning the health status of UADT cancer patients specifically, and then a report of the research methodology and findings, part of which will again be in the form of a manuscript (manuscript V).

3.2 Defining health status

The starting point for most discussions of the definition of health status is that of the World Health Organization (WHO) who, in 1948, defined health as "a state of complete, mental, physical and social well-being, and not merely the absence of disease or infirmity" (92). The problem with this definition is that, while it may be a succinct "dictionary" definition of health, it is extremely difficult to operationalize for the purposes of measurement. This difficulty in conceptualizing health is perhaps the major constraint on the development and usefulness of health status indicators (93). The majority of workers in the field have subsequently decided to concentrate upon what could be called the negative side of the health concept, because the positive aspect is too extensive and too vague (94). Boorse has argued that positive health is a concept which involves maximizing ones full potential, the achievement

of which can be assessed on an individual, species or limitless level. Ideal positive health at an individual level is judged in relation to that individual's abilities and potential. Ideal positive health at a human species level is judged according to the maximum potential of the species in all fields of life, and the limitless view of ideal positive health sets no parameters (95). All of these concepts of the standards by which to assess positive health have their problems in terms of measurement. The individual level would require the definition of positive health for each person and make valid between-subject comparisons extremely difficult. The species level would be irrelevant for or unacceptable to a large proportion of people and the limitless interpretation cannot be operationalized because one cannot deviate from an undefinable state. Thus the focus of "health status" evaluation has been on disease, the medical, biological view of ill-health; sickness, the subjective, patient view of ill-health; and illness behaviour, those behaviours associated with ill-health (96). In other words, unable to conceptualize health, most work in the field of health status measurement has concentrated upon ill-health as a measurable construct.

In this tradition, one of the best-recognized frameworks which begins to conceptualize ill-health is that developed by Wood and recognized by the WHO in its International Classification of Impairments, Disabilities and Handicaps (97):

Impairment - any loss or abnormality of psychological, physiological or anatomical structure or function

Disability - any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being

Handicap - a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfilment of a role that is normal (depending upon age, sex, and social and cultural factors) for that individual.

One of the principal virtues of this framework is that it provides a clear distinction for the consequences of ill-health at an organ/tissue/body level, at an individual level and at a societal level. It was, however, intended to be a framework

for the evaluation of ill-health rather than an all-embracing classification. And indeed, several alternative and broader frameworks have subsequently been developed. Patrick and Elinson (98) described 5 indicators of health status: mortality, morbidity, discomfort, dissatisfaction and disability. Mortality rates, life expectancy and other measures associated with death are essential health indicators, however they have limitations, especially in Western countries where life expectancy is long and the main problem is the assessment of chronic ill-health. Morbidity data is based upon the clinical assessment and diagnosis of a pathological condition (disease). The prevalence and incidence of morbidity is then recorded by the likes of cancer registries, infectious disease registries and hospital discharge summaries. These data, again, provide valuable information but are limited by the pathological definitions of disease and are often subject to bias and error. Patrick and Elinson defined discomfort as self-reported feelings of pain, aches, anxiety, tiredness or sadness (98). These feelings may be part of, or precursor to a diagnosable pathological condition or identifiable injury, however this is often not the case. Dissatisfaction was defined either as an individual's lack of satisfaction over the state of their health, or a lack of satisfaction over the medical and/or social services available to meet that individual's perceptions of their needs (98). Finally they discussed disability using the WHO classification of impairment, disability and handicap (98).

Yet another conceptualization of a framework for health status evaluation was developed by Ware who classified seven dimensions (99):

Physical - the capacity of an individual to perform activities of daily living such as eating, bathing and dressing;

Mental - the psychological and emotional status of an individual;

Social well-being - this involves the concepts of social contacts and social resources. The former relates to observing, for example, how often an individual meets friends or family, or practices a hobby, while the latter relates to the individual's feelings as to the value of those and other non-observable contacts;

Role functioning - relates to an individual's ability to perform their usual role

activities such as school, occupation, housework;

General health perceptions - this relates to an individual's perception of their overall health and well-being. It often involves a simple categorisation such as "excellent" or "poor";

Symptoms - these can be both physical, such as pain and swelling, and psychological symptoms, such as anxiety and fear. The psychological symptoms are differentiated from the mental dimension in that the former are manifestations of an underlying psychological condition;

Physiological status - these are a heterogeneous collection of assessments which are disease-specific.

It is evident that while a number of dimensions in each of these frameworks are immediately recognizable in terms of their relationship with health status (eg, death, morbidity, impairment, physical and mental status), others (eg, discomfort, dissatisfaction and social well-being) expand the concept of health status towards something more akin to the broader concept of quality of life (QOL) (100). Torrance has stated that there are many factors which contribute to or detract from an individual's QOL, and health is but one, albeit important, of these factors (101). However, many of the workers in the field of the development of "health status" measures have adopted phrases such as "Quality of Life", "Health-related Quality of Life" and "Well-being" as the goal of measurement in terms of health care outcomes and needs assessment. Unfortunately, while the names of the instruments have changed, their content has remained broadly similar (102). As Hunt recently observed, indicators of "QOL" have ranged from the purely physiological, through functional capacity, to complex series of questionnaires on social activities and psychological problems (103).

Thus, while most people in the field of health status evaluation would agree that it is a multidimensional construct, there is very little agreement evident on which dimensions should be included in a health status instrument. Moreover, Wilson and

Kaplan have argued that "the conceptual models that underlie the measurement of health status, as they are currently operationalized, do not meet the important need of clinicians to understand causation and mechanism, without which rational and effective therapy is difficult" (104). In an attempt to clarify the inter-relationships between health status indicators, health-related quality of life (HRQL) and QOL, Wilson and Cleary have described a conceptual model which links biological and physiological variables, symptoms, functional and health status, HRQL and QOL (105). This model shows how measures of health can be thought of as existing on a continuum of increasing biological, psychological and social complexity and attempts to demonstrate how individual and environmental factors affect these measures (105). This model has a theoretical hierarchy similar to, and an extension of, the WHO classification of impairments, disabilities and handicaps.

In summary, while it is evident that there has been some development in our understanding of the concept of health over the past few decades, there remains little agreement upon how it should be defined in terms that facilitate measurement. In view of this situation, and the fact that it is not the aim of this research to develop a measure of health status for use among UADT cancer patients, I will take the pragmatic approach for this research in using the most appropriate and valid instrument available. For the purposes of this research project, I will define health status as a subjective, multidimensional construct and review the relevant literature to justify the use an instrument which measures such a construct validly for UADT cancer patients.

3.3 Health status measurement issues: subjectivity and construct dynamism

It is evident from the previous discussion of the definition of health status that it includes both disease (ie, medically or biologically defined abnormality) and sickness or illness (ie, patient-based reports of their experiences and perceptions) indicators. If it is the intention of the investigator to evaluate any element of health status beyond the biological indicators, then since these include a large element of

individual patient perception, it is essential that such evaluations are patient-based (106,107). This has been emphasized with the demonstration that clinicians' and patients' evaluation of patients' health status is different, and that the latter's evaluations are more reliable (108). This subsequently implies that the subject of the evaluation should be asked to make that evaluation him/herself. The most popular means by which this is achieved is through a self-administered questionnaire, although occasionally interview-based questionnaires, and in special cases where the subject is, for various cognitive reasons (eg, young child, mentally handicapped individual), unable to respond directly, observer-based questionnaires are used.

For a further discussion of issues concerning the subjectivity and dynamism of the "health status" construct, see manuscript IV.

3.4 Manuscript IV: "Quality of life: A dynamic construct"

This manuscript is reprinted from Social Science and Medicine, Volume 45, Number 2, Allison PJ, Locker D and Feine JS: *Quality of life: A dynamic construct,* pages 221-230, Copyright 1997, with kind permission from Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington 0X5 1GB, UK.

Abstract

The principle of Einstein's theory of special relativity is that an observer of an apparently moving body cannot be sure if the body really has moved, if he/she has moved or if both events have occurred. Although Einstein was discussing physical events, a similar hypothesis may apply to quality of life. When using quality of life instruments, one presumes that the point of reference (the observer in Einstein's terms) does not move, ie. that an individual's attitude towards a particular construct will remain stable. Otherwise, changes in response to particular variables cannot be interpreted. However, attitudes are not constant: they vary with time and experience and are modified by such psychological phenomena as adaptation, coping,

expectancy, optimism, self-control and self-concept. For example, eating problems may be extremely important at one point in a person's life. However, when oral discomfort has been diagnosed as cancer and treated with surgery or radiation, the same individual may "objectively" demonstrate more problems when eating, but report them as less because they have now become relatively unimportant. Furthermore, paradoxical reports that some groups of ill individuals rate their quality of life higher than do "healthy" persons raise similar questions concerning between-group point of reference differences. Investigators in the fields of organisational management, education and psychology have developed techniques such as "Then ratings", saliency indicators and individualised questionnaires in attempts to quantify within-subject variability and between-group differences pertaining to point of reference. We suggest that similar methods may help us to measure change in the impact of the different items of quality of life instruments. In this paper, we will describe the theories of change associated with quality of life measurement. In addition, we will present evidence suggesting that the point of reference does change, the reasons for this and possible solutions to the problem.

Key words: adaptation, measurement, quality of life, response-shift bias.

INTRODUCTION

During the past 15 years, there has been a dramatic increase in literature concerning quality of life and health-related quality of life (QOL) . A great deal of this research has concentrated on defining the content of the large array of QOL instruments. However, there has been relatively little investigation into the more subtle, but nonetheless important, complexities of the nature of the QOL construct.

One such complexity is the possibility that QOL is a dynamic phenomena. Investigators have implicitly recognised that there are between-subject differences in determining instrument content. However, the possibility of within-subject QOL

construct dynamism (i.e. an individual changing the standards by which he/she assesses his/her QOL) and its' subsequent effects upon valid QOL measurement have largely been ignored.

Evidence suggesting the dynamic nature of the QOL construct exists in several forms. As early as the 1960s and 1970s, Campbell(1976) found that while "objective" social indicators among the population of the the USA improved during the period 1957-72, the number of people reporting themselves to be very happy declined, and that this was especially true amongst the most affluent. Brickman et al(1978) found that a group of lottery winners (of $50,000 - $1,000,000) were not significantly happier than a control group, and that a group of people rendered paraplegic through accidents were not as unhappy as could be expected through "objective" observation. In light of the findings of this latter study, Brickman et al commented that there is also evidence that the inhabitants of poorer cities, regions or countries are not less happy than their more affluent counterparts(Brickman et al, 1978). As subjective assessment has moved into the field of health care, so similar examples have come to light in the medical literature. Cassileth et al(1982) found the psychological well-being of a group of melanoma patients to be significantly better than that of a group of patients with other dermatological disorders, and slightly (but insignificantly) better than a group of the general population. Similarly, the life satisfaction of a group of elderly and middle-aged spinal cord injury patients was reported as only slightly worse than that of controls(Decker and Schultz, 1985), and the life satisfaction among cancer survivors has been reported as better than that of controls(Irwin et al, 1982). Furthermore, Evans(1991) has reported that, in a series of studies of transplant recipients, haemodialysis patients and others, patients level of happiness, satisfaction or life quality often exceeded that of a healthy population. All of these studies are examples of research in which subjective indicators are at odds with objective indicators or observer expectations. They are not necessarily representative of all QOL research. However, they highlight the possibility of

problems inherent in the measurement of QOL. The observation that subjective and objective measures of a certain phenomenon can produce apparently paradoxical results, is partially explained by the fact that these measures are not assessing exactly the same thing. However, the difference between observer expectations and the results of subjective health measurement is more difficult to explain.

This latter issue has been raised in the literature, but only to a limited degree. de Haes and van Kippenberg(1985), in their review of the QOL of cancer patients, suggested that adaptation, downward social comparison and the different contributions that affective and cognitive components make towards quality of life assessments, might be influential in these paradoxical observations. The issue was also discussed by Breetvelt and van Dam(1991) under the guise of cancer patients apparently underreporting their problems. They categorised explanations for "underreporting" into the "defence mechanism", "crisis" and "judgement theories", and concluded that a form of testing called "Then rating" is the most appropriate solution. Most recently, de Haes et al(1992) have further tested their theory that differences in the relative contribution of affective and cognitive components can explain the lack of difference in QOL between apparently ill and healthy people. Similarly, Hyland(1992) has concluded that there are two ways in which people judge their health, QOL or other aspects of their life: "problems (of health)" and "evaluations (of health)". The former involves assessment of the more objective aspects being measured, while the latter entails an evaluative, emotional approach. Hyland et al(1993) reported that this theoretical division is valid, by demonstrating that the problems subscale is more sensitive than the evaluations one.

Finally, the possibility that subjects may simply be misrepresenting their true feelings has to be recognised, however this form of validity problem is another subject all together and will not be covered in this review. The aim of this paper is to present theoretical explanations for QOL construct dynamism, suggest possible solutions and

make recommendations for the design of QOL instruments.

THE THEORY OF A DYNAMIC CONSTRUCT

Practically speaking, the problem of dynamic constructs can best be understood by examining the measurement of within-subject change and between-subject difference. The situation has been most clearly explained by Golembiewski et al(1976) in their article on the measurement of change and persistence in human affairs, especially in relation to self-reports (i.e. subjective assessment). They characterise three types of change; alpha, beta and gamma.

"Alpha change involves a variation in the level of some existential state, given a constantly calibrated measuring instrument related to a constant conceptual domain"(Golembiewski et al,1976).

Alpha change is the conventional conception of change in which one assumes that the construct under assessment is stable. Examples of this sort of assessment in health care could be that of blood haemaglobin or blood pressure. It is this form of change that the vast majority of studies attempt to measure.

"Beta change involves a variation in the level of some existential state, complicated by the fact that some intervals of the measurement continuum associated with a constant conceptual domain have been recalibrated"(Golembiewski et al,1976).

Beta change can be illustrated with reference to the assessment of pain. A patient given a self-report instrument to assess the levels of their pain 1) at the moment of diagnosis, 2) immediately after treatment and 3) two months after treatment may give the following indications:

How much pain do you have at the moment?

no pain at all		rating 3	rating 1	rating 2	The worse pain imaginable

If only alpha change is assumed, one would conclude that the pain became worse immediately after the treatment but two months later had improved to such an extent that it was less than before the treatment. If, however, one acknowledges the possibility of beta change another possible explanation becomes tenable. Between times 1 and 2 the patient's terms of reference for "the worst pain imaginable" could have changed in light of her[1] experience with the treatment. In this case, the difference in ratings of pain at times 1 and 2 would not be valid.

"Gamma change involves a redefinition or reconceptualisation of some domain, a major change the perspective or frame of reference within which phenomena are perceived and classified, in what is taken to be relevant in some slice of reality"(Golembiewski et al,1976).

Gamma change adds a further level of complexity to the interpretation of measured change and is of great relevance in the measurement of multidimensional constructs like QOL. For example, when asked what features of his life make the most important contributions to its quality, a "healthy" individual may reply; "my social and family life, my work, my health and my financial circumstances". However, one year later, the same individual, after being diagnosed and treated for laryngeal cancer, may reply to the same question; "my ability to speak and eat, my family and social life and my spiritual happiness". The terms of reference (and so the standard) of this individual's personal assessment of his QOL have completely changed. It is important to realise that there are two dimensions to this change in standard. Not only have the domains contributing to that individual's evaluation of his own quality of life

---

1

Note: to avoid the cumbersome use of "his/her" throughout the text, we are distributing references to each gender equally.

changed, but also the relative importance of those domains have changed. Gamma change is the most profound form of change that can confuse the evaluation of alpha change and is, therefore, the most important to evaluate or to control for when measuring change in QOL.

Having established the theoretical dynamism of QOL constructs, the next stage is to explore the causes of this phenomenon. What factors make human subjectivity dynamic?

THE POSSIBLE CAUSES OF QOL CONSTRUCT DYNAMISM

The generation of well-being

Deiner(1984) has described two perspectives concerning the generation of an individual's feeling of well-being; the so-called "bottom-up" and "top-down" approaches. The former is based upon the theory that well-being (and other subjective overall evaluations of an individual's life) is a product of life's experiences. This suggests that an individual who has more negative experiences will have a worse feeling of well-being than another individual whose life experiences have been mainly positive. Alternatively however, one could theorise that the generation of well-being works from the top downwards. This means that an individual's feeling of well-being will affect the way she feels about the different parts of life. An individual who feels positive overall, will feel good about her experiences. If this is the case, then the determinants of well-being will lie within the person(Heyink, 1993). In this respect, Guttman and Levy(1982) considered well-being an attitude, and Ormel(1983) described the lack of well-being as a personality characteristic.

This "top-down" perspective is supported by research in which happiness is not dependent upon people's experience of positive and negative events(Costa and McCrae, 1980) and on evidence that demonstrates the maintenance of well-being

over time and through differing circumstances. Palmore and Kivett(1977) concluded that past life satisfaction was the best predictor of future life satisfaction in a longitudinal study of middle-aged and elderly persons. Kreitler et al(1993) found that the life satisfaction of a group of cancer patients was not significantly worse than that of a group of orthopaedic patients and healthy people. However, the cancer patients considered more and different domains in making this assessment than the two other groups. The phenomenon of downward social comparison, in which people often compare themselves with those worse off in order to retain their level of well-being(Diener, 1984, Bunk et al, 1990), is a good example of the top-down theory. Downward comparison tends to occur among people in extremely threatening positions where no effective action can be taken to alleviate the situation(Wills, 1981). These changes in priority and number of domains and changes in groups of social comparison are good examples of gamma change.

If the "top-down" perspective to the generation of well-being is the case, then it could cause a dynamic subjective standard. If an individual's external environment is changing, yet his evaluation of his well-being remains fairly constant, then he must be altering his internal standards to accomodate the change. If this is the case, what psychological processes are occurring that contribute to these changing internal criteria?

Adaptation

Heyink(1993) has described adaptation as an intrapsychic process in which past, present and future situations and circumstances are given such cognitive and emotional meaning that an acceptable level of well-being is achieved. Events and situations can affect well-being, sometimes seriously. However, in due course, adaptive processes are usually put into action to achieve a level of well-being "belonging to the person". Heyink classified adaptation processes into three distinct categories: shifting intrapsychic criteria, cognitive reconstruction and future-time

perception. In so doing, he recognised the problem of altered internal standards in measuring subjective well-being.

One of the key conceptualisations of adaptation has been Helson's adaptation theory, which states that an individual's adaptation level is a product of all stimuli, past and present, and their effects upon the attribute under consideration(Helson, 1964). The adaptation level changes constantly as new stimuli are experienced, and all judgements are made relative to this adaptation level. In the short term, the effect of extreme stimuli is reduced through contrast, in which the meaning and importance of other domains of an individual's life are altered. In the long term, the affective component of a strong stimulus is lessened through habituation. In other words, adaptation is an individual's attempt to normalise stimuli. The effect of any extreme stimuli, whether good or bad, will, to one extent or another, be negated through the normalising effect of the adaptation level. Based on this theory, Chamberlain and Zika(1992) have developed a model of well-being in relation to adaptation in which well-being is stable in the absence of situational change, sensitive to change when it occurs and adaptive to the occurrence of change. However a model of adaptation is described, it is evident that it involves individuals shifting internal standards and so creating the possibility of beta and gamma changes.

Coping

The concept of coping is very similar to that of adaptation. However, coping emphasises stressors and the individual's attempts to deal with them. Lazarus and Folkman(1984) have defined coping as "constantly changing cognitive and behavioural efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person". They categorise two fundamental forms of coping according to its function: *problem-focused coping* processes directed at managing or altering the problem causing the distress, and *emotion-focused coping* processes directed at regulating the emotional response to

the problem. The former coping strategies are more likely to be mobilised if a stress is appraised as manageable and amenable to change. These include information-seeking, aid-seeking and direct action. On the other hand, emotion-focused coping strategies will tend to be used when a stressor is appraised as being beyond control. These strategies include trying to see humour in the situation, avoidance behaviour and detachment(Folkman and Lazarus, 1980). The coping strategies used by an individual in a given situation are determined by several factors:

1. Cognitive appraisal. Lazarus and Folkman describe this as "the process of categorising an encounter, and its various facets, with respect to its significance for well-being"(Lazarus and Folkman, 1984). This comprises two evaluative processes; "Am I in trouble or being benefited, now or in the future, and in what way?" (primary appraisal) and "What, if anything, can be done about it?" (secondary appraisal). Primary appraisal will, in turn, result in one of three categorical responses; irrelevant, benign-positive and/or stressful. The first results from an encounter judged to have no effect upon the individual and the second from something judged to be beneficial. The stressful appraisal can result from the individual having already experienced harm or losses or from the perception of a threat or a challenge. These primary appraisals are not necessarily mutually exclusive. Secondary appraisal is a complex evaluation of what is at stake, available coping options, what each option is likely to achieve and the likelihood of that individual being able to apply that coping strategy effectively.

2. An individual's coping resources. This involves many factors such as health and energy, beliefs, commitments, problem-solving skills, social skills, social support and material resources.

3. Mitigating factors. These can be internal, personal factors (eg. beliefs and cultural values), environmental factors which compete for the same resources, or excessive

threat rendering coping skills inoperable.

To complicate matters further, coping efforts are viewed as unstable over time or between problems. Coping is a multidimensional construct that varies independently in several dimensions(Englert et al, 1994). Therefore this psychological process is also bound to contribute to the dynamism of an individual's evaluation of her QOL, creating beta and gamma changes.

Affect/cognition and problem/evaluation

Affect refers to the individual's emotional response to experience, and cognition refers to the rational appraisal of the experience(de Haes et al, 1992). Andrews and McKennell(1980) and McKennell and Andrews(1980) have demonstrated the difference between affective and cognitive appraisal in relation to self-reported well-being. The importance of affect and cognition to QOL construct dynamism is that the cognitive component of QOL assessment is believed to be less sensitive to change than the affective component(McKennell, 1978). In other words, an assessment of satisfaction (cognition) with a certain situation is more likely to remain fairly constant, while affective (emotional) response to the same situation may fluctuate enormously. This difference in sensitivity to change between affect and cognition is supported by research from Headey et al(1984) and de Haes et al(1992 and 1987). The results of their investigations suggest that cognitive appraisal corresponds with the "top-down" generation of QOL and affective appraisal corresponds with the "bottom-up" approach.

A similar categorisation of items that implicates construct dynamism is that of "problems of health" and "evaluation of health"(Hyland, 1992). "Problems" relate to an individual's cognitive knowledge of his health status and corresponds well with the functional subscales in QOL instruments. "Evaluations" relate to the individual's personal appraisal of her health status, involving emotions and personal

significance(Hyland et al, 1994a). However, in a longitudinal clinical trial, "problems" were found to be sensitive and "evaluations" insensitive to change(Hyland et al, 1994). "Evaluations" are further categorised into negative emotional evaluation and positive emotional evaluation. These have been demonstrated to be weakly correlated, but independent phenomena(Hyland et al, 1994a). Negative and positive evaluations are related to personality traits, as well as environmental factors, and are therefore likely to demonstrate different levels of sensitivity within and between subjects.

Research on the cognitive/affective and problem/evaluation categorisation of QOL instrument items appears to be contradictory in that de Haes et al(1987) assert that affective items are more sensitive to change, while Hyland et al(1994) maintain that problem items are more sensitive to change. This may be partially explained by the fact that de Haes et al were referring to short term change, while Hyland et al were referring to long term change. However, it is evident that whatever model of affective and cognitive evaluation is used, there are differences which are liable to create beta changes when assessments are made.

Uncertainty

In her uncertainty in illness theory, Mishel(1988) defines uncertainty as an individual's inability to determine the meaning of illness related events. The theory posits that such uncertainty triggers two processes, inference and illusion, that determine whether the value of the uncertainty is positive (opportunity) or negative (danger). In the inference process, the evaluation of uncertainty is based upon related situations. The illusion process allows an individual to construct a generally positive outlook(Mishel, 1988). In situations where an individual feels helpless to affect the outcome or when conditions progressively worsen, the existence of uncertainty is paramount in the creation of illusion to maintain hope. In the absence of hope, life is senseless and accompanied by a lack of motivation to achieve anything, often

resulting in seriously depressive illness. Although there is only limited empirical evidence to support this theory and its relationship to adaptation and quality of life(Padilla et al 1992), it is highly probable that individuals illuding themselves, for whatever reasons, are likely to evaluate QOL using changing standards, thereby creating the possibility of both beta and gamma changes in addition to alpha changes..

Self-control

Cybernetic or control theory is a general approach to the understanding of self-regulating systems whose breadth of application is ubiquitous (whether engineering, mathematics, economics or medicine)(Carver and Scheier, 1982). More specifically, the principles can be applied to the way in which people maintain their QOL. Control theory is based upon a hierarchical negative feedback loop. The outcome of any action by an individual will be assessed in relation to a standard (of that individual's making), and that action may subsequently be adjusted to achieve the desired outcome. Within this regulatory system there is a hierarchical structure of reference values whose aim is to control behaviour at levels ranging from the molecular and cellular, to the cultural and philosophical. The standard for each feedback loop originates from the output of that loop immediately superior to it in the hierarchy(Powers, 1973). Powers(1973) postulated a hierarchy in which the "system concept" and "principle" levels are the highest levels of control. Carver and Scheier(1982) further developed this model suggesting that people normally function at the "programme" level. For example, you go to the bank to take out some money (programme) to return to a friend from whom you borrowed it (principle) because you believe, as an honest, responsible individual in a society, that you should return the money to your benefactor (system concept). Carver and Scheier discussed control theory in relation to health and illness behaviour (eg., an individual has a headache, so she takes an aspirin). However it is important to recognise the implications of control theory for QOL measurement: an individual's standards for the assessment

of his QOL could move up and down the aforementioned hierarchical feedback system, thereby causing gamma change. For example, an individual may smoke because it makes her feel good, but, following severe ill-health due to the smoking, decide to quit because she feels the responsibility of staying at work to support her family.

Self-concept

Self-concept has been defined as "the sum total of all that a person feels about himself/herself"(Schain, 1980). It can be subdivided into four compartments: the body self (physical function and body image); the interpersonal self (psychosocial and sexual interaction); the achievement self (job/role function); and the identification self (spiritual and ethical beliefs)(Foltz, 1987). These domains are evidently very similar to those of a well-being/ QOL/ health construct, and changes in self-concept are bound to influence the measurement of such a construct. Indeed, self-concept and quality of life are interdependent(Foltz, 1987).

As with QOL, self-concept has often been treated as an outcome indicator. However, Curbow et al(1990) point out that it should also be considered as a predictor and moderator of health care interventions. Although there is some debate, the dominant view is that some aspects of self-concept are susceptible to change, while others remain constant(Curbrow et al, 1990). Taylor and Brown(1988) have suggested that people respond to chronic negative feedback by downgrading the importance of the relevant domain. Similarly, McCrae and Costa(1988) have argued that people will downgrade the importance of a domain about which they are unable to feel good. This psychological behaviour in relation to self-concept is essentially the same as adaptation and downward social comparison and as such is a potential source, particularly, of gamma change.

Expectancy, optimism and self-efficacy

According to psychological theories of motivation, people's actions are greatly affected by their beliefs about the probable outcomes of those actions(Bandura, 1977). People who see desired outcomes as attainable will continue to exert effort to achieve those outcomes, even if the process is difficult. However, when an outcome becomes unattainable, people disengage themselves from their pursuit. Outcome expectancies are therefore a major determinant in continued striving versus giving up(Scheier and Carver, 1987). In his self-efficacy theory, Bandura(1982) differentiates between self-efficacy expectancies and outcome expectancies. The latter refers to an individual's belief that a certain action will result in a particular outcome, while the self-efficacy expectancy refers to that individual's belief that he has the ability to perform that action and so achieve the resulting outcome. Scheier and Carver(1987) have added external circumstances to the equation producing the decision to continue or to give up. In particular, they have postulated that optimism can affect expectancies, behaviour and health. They described optimism as a stable personality characteristic with a predictive ability for coping strategies and health(Scheier and Carver, 1985) and have concluded that optimists have a higher level of well-being(Scheier and Carver, 1987), recover more quickly and return to work more quickly following coronary artery bypass surgery(Scheier et al, 1989) and are more likely to use problem-focused coping strategies than are pessimists. Furthermore, they have concluded that optimists are more likely to use acceptance/resignation as a coping strategy when they perceive the situation as uncontrollable. In light of the finding that denial of chronic or terminal illness is associated with poor long term coping(Suls and Fletcher, 1985), optimists would seem to have an advantage in uncontrollable as well as controllable situations. In other words, an optimist is more adaptable(Bandura, 1977). Perhaps one of the most powerful examples of the effect of expectancies upon subjective health is the placebo effect, in which a patient's expectancies of treatment benefits become realised in their subjectivity. Since the goal of many health care interventions is to improve the

subjective health of an individual with chronic problems, then a strong placebo effect can provide a very real benefit(Wennberg et al, 1993).

In summary, it seems that the optimistic/pessimistic nature of individuals is an important determinant of their expectancies and, therefore, will affect the evaluation of their QOL. These expectancies will certainly differ between individuals and may alter in nature within the same individual under extreme circumstances, for example, severe ill-health. Again, these phenomena are potential causes of both beta and gamma changes.

THE POSSIBLE CAUSES OF QOL CONSTRUCT DYNAMISM: A SUMMARY

The previous discussion has highlighted some of the most important of the possible causes of QOL construct dynamism in order to indicate the plausibility of QOL construct dynamism due to well-known psychological phenomena. If such adjustments of standards can occur when conventional QOL instruments are used to assess change, then the possibility arises that we may actually be measuring a combination of QOL change and ability to adapt and/or cope, plus various other psychological phenomena (ie. we may be measuring a combination of alpha, beta and gamma change rather than the absolute, alpha change desired). This raises questions concerning the validity of within-subject comparison of QOL data.

THE SUGGESTED SOLUTIONS

Cronbach and Furby(1970) discussed the issue of the measurement of change 25 years ago , concluding that investigators should frame their questions in ways that do not involve the investigation of change. They advised that comparisons be made using only between-group post-intervention scores. However, Golembiewski et al(1976) have subsequently argued that the recommendations of Cronbach and Furby are inappropriate, as they did not recognise all the possible dimensions of change.

Then ratings

Howard et al(1979) proposed retrospective pretests, or "Then Ratings" as a solution to the problem of beta and gamma changes. They noted that conventional assessment of the effect of an intervention involving pre- and post-intervention measurements is no longer valid when beta and gamma changes have occurred. They suggested an additional post-intervention test (the "then rating") in which the individual is asked to score the instrument in reference to how she *now* perceives herself to have been *before* the intervention. This theory is based on the assumption that the individual will use the same criteria for the conventional post-test and the "then rating". Any difference between the pre-test and the "then rating" is presumed to be due to beta change. The difference between the "then rating" and post-test is then accepted as true alpha change. This theory was discussed further by Terborg et al(1980) who reviewed evidence concerning the validity of the "then rating" theory. They found that in 5/11 studies comparing the pre/post and "then"/post forms of analysis, the latter yielded a drastically different set of conclusions regarding the effectiveness of the interventions concerned. They also found, in another 5 studies comparing the results of pre/post and "then"/post analyses, that the latter analyses correlated better with objective indicators of the same phenomena. They concluded that in no study was the conventional pre/post analysis superior to a "then"/post analysis. However, the validity of "then" ratings is entirely dependent upon the accuracy of an individual's memory regarding his previous situation. Outcomes research in pain(Feine et al, 1989, and Erskine et al, 1990) and prosthodontics(De Grandmont et al, 1991) suggests that memory of chronic conditions is highly inaccurate. Therefore, the validity of "then ratings" is questionable. The other problem with "then ratings" is that they use the same questionnaire for all the tests thereby only permitting evaluation of beta change; the possibility of gamma change is ignored.

Individualised questionnaires

Other approaches to solving the problem of dynamic standards involve the use of

individualised instruments. These include individualised items and/or individualised weights for each of the items. The MACTAR Questionnaire was designed to assess the effectiveness of therapy in patients with rheumatoid arthritis and, at the time, incorporated several novel components: i)individualised assessment of disabilities, as well as generalised questions; ii)a value component in which activities most affected by arthritis are identified; and iii)a focus on change rather than absolute values(Tugwell et al, 1983). Similarly, Guyatt et al(1987) designed a questionnaire to measure the fuctional status of patients with cardiorespiratory disease. This instrument incorporated some general questions and some individual ones. The authors concluded that, although the individualised questions create problems for between-subject comparison, they are extremely responsive and could be an excellent measure of outcome for within-subject trial designs.

A different approach to the individualising of questionnaires was taken by McGee et al(1991) whose Schedule for the Evaluation of Individual Quality of Life (SEIQoL) was aimed at determining individual domains of QOL, rather than actual items. For each individual, five areas of life considered important in assessing QOL were elicited through a semi-structured interview. The subject then indicated her current status for each of the five domains and an overall evaluation of QOL using visual analogue scales. This technique was found to be valid and reliable among healthy individuals, although less so among idividuals with irritable bowel syndrome. Although this approach is strong in terms of individualisation, it is very time-consuming and not specific enough to be of practical use in most large scale health care assessments.

A compromise approach to individualisation in which subjects can indicate weighting of importance of domains or items, while not actually chosing the questions themselves, has been used by a few others. Ferrans and Powers(1985) developed a quality of life instrument in which subjects are able to indicate the unique importance

of each domain, and Ferrans(1990) later refined a quality of life instrument in which subjects could indicate the importance for each item.

Assessment of premorbid characteristics

Finally, a number of investigators have suggested the tactic of evaluating individual characteristics before assessing QOL. Discussing QOL with special reference to patients undergoing coronary artery bypass surgery, Cohen(1982) suggested that patients' "life plans" in terms of their goals and hopes need to be taken into account when assessing their QOL. A very similar idea was promoted by Calman(1984) who defined an individual's quality of life as the gap between her expectations and her reality. Furthermore, Goodinson and Singleton(1989) have concluded that information relating to an individual's QOL cannot be abstracted in isolation from coping strategies, past experiences of illness and other variables. The principle common to these suggestions is that the characteristics of each individual need to be obtained prior to any subjective assessment or subsequent intervention. Otherwise, the data will be compromised due to dispositional characteristics. Cella and Tulsky(1993) have argued that such a perspective is required to compare different treatments, even though no measurement technique has yet successfully bridged the problem of separating premorbid characteristics from disease and treatment morbidity. However, a phenomenological perspective would be that the distinction between premorbid characteristics and disease/treatment effects is irrelevant and that the subjective assessment is valid for its own sake(Cella and Tulsky, 1993). Furthermore, it is evident from this review that there are a number of behavioural and psychological factors (such as coping, adaptation, expectation and optimism) that contribute to an individual's system of dynamic standards. Routine assessment of all these factors with QOL measurement would be impractical, even if all such factors and their contributions to QOL were fully understood. In addition, it seems that some of these behavioural and psychological characteristics are person-specific, some are situation-specific and many are a combination of the two. Finally, as Hughes(1985)

has pointed out, depressive illness can be caused by cancer and its therapy, but it may have been present before the diagnosis. Indeed it may even have been an aetiological factor. This raises the question: are some aspects of QOL outcome confounding or aetiological factors? Indeed, is QOL itself an outcome or an aetiological factor?

Implications for study design and data analysis

The implication of construct dynamism for research in which QOL is an outcome measure, is that one cannot be sure whether the apparent change is alpha (ie. the absolute change due to the intervention) or beta or gamma (ie. "change" due to altered standards). This, subsequently, raises the whole question of the appropriateness of QOL as an outcome measure for clinical or other health care intervention trials. That is not to say that QOL instruments should not be used, rather, that they should be used in conjunction with other variables. A health care intervention is never approved as a result of only one trial in which it is seen to be of benefit, but after several trials using various different outcome indicators. QOL instruments can be one of those indicators, contributing valuable qualitative data to the overall information. As with all forms of data, investigators need to be aware of the limitations of QOL assessment.

In the light of the discussion in this review, one particular form of trial design for which QOL outcomes may be inappropriate would be the within-subject crossover trial. This design requires a washout period following the first intervention so that the subject's baseline status is the same before the next intervention. If QOL is used as a primary outcome, this should include psychological parameters(Spilker, 1991). However, it would seem highly possible that an individual could adjust their standards (through, for example, adaptation or change of expectancies) following the first intervention, thus invalidating subsequent QOL assessment.

Another important implication of QOL construct dynamism is the manner in which

the data are treated. In view of the observation that different aspects of a QOL instrument may demonstrate differing levels of sensitivity, the most valid comparison of QOL data will involve like items or subscales, rather than a comparison of an overall evaluation/score. In addition, the most valid comparisons will be those of the within-subject pre- to post-intervention change.

SUMMARY AND CONCLUSION

In order to assess the effectiveness and efficiency of health care, valid and appropriate measurement instruments need to be developed. QOL instruments can make an important contribution to such health care evaluation. However, the problems inherent in measuring QOL need to be recognised and dealt with either through altered instrument design, or through appropriate interpretation of the data. Although the abstract nature and between-subject variability of QOL constructs are dealt with through the development of "consensus instruments", the latter do not control for the within-subject dynamism of the QOL constructs.

This review presents evidence supporting the presence of QOL construct dynamism and provides a theoretical explanation of how construct dynamism can affect the measurement of change. The causes of construct dynamism lie in such psychological phenomena as adaptation, coping, self-control, uncertainty, self-concept, expectations, optimism and in the differing sensitivities of the affective and cognitive components of an instrument. These factors have been highlighted because they are the phenomena best understood in relation to QOL.

The possible solutions to the problem of measuring dynamic constructs are categorised into three groups: "then" ratings, individualised questionnaires and determination of pre-intervention characteristics. The problems with "then" ratings are that their validity can be confounded by inaccurate memory and that by using the same questionnaire for post-test and "then" rating, one is measuring only beta and not

gamma change. The determination of preintervention characteristics is potentially too complex and time-consuming to be practically useful, even if one is in a position to know exactly which characteristics to assess. Individualised questionnaires with saliency indicators would, therefore, appear to be the most valid solution presently available. They facilitate the evaluation of gamma as well as alpha and beta change. Totally individualised questionnaires can be time-consuming and difficult to develop, and will make valid between-subject comparison more difficult. However, a form of partially individualised questionnaire is possible. Part of the questionnaire could be generic and the other part individual. Subjects could chose from a list of alternatives, or fit personally appropriate questions into a template of domains. Saliency indicators can also be added to each item, thereby providing some degree of individualisation to a "consensus instrument". It does need to be pointed out however, that this suggested use of some form of saliency indicator has very little evidence to support its validity and that research needs to be carried out to assess whether this is indeed a valid solution.

Furthermore, having decided what measurement approach to use, potential difficulties in the interpretation of results using (partially) individualised QOL instruments need to be recognised. The ability to directly compare between-subject and within-subject data will vary, depending upon the degree of individualisation. When investigating the effectiveness of an intervention, one can assume a large degree of non-compatability of the absolute data. Therefore, the most appropriate data for comparison may be measures of within-subject pre- and post-intervention change. In addition, due to the varying sensitivities of the different aspects of the instrument, the most valid comparisons are those that involve like items and subscales, rather than overall evaluations or scores.

QOL instruments contribute valuable information to the overall evaluation of many health care interventions. However, the problems inherent in measuring a dynamic

construct need to be recognised and dealt with to ensure their validity. The most valuable information current QOL instruments can contribute towards the evaluation of the effectiveness of a health care intervention is through comparison of within-subject, like-item/subscale change. Further research is required to investigate beta and gamma change as it relates to health care interventions and the appropriateness of (partially) individualised instruments and various forms of saliency indicator in relation to subjective health measurement.

REFERENCES

Andrews FM and McKennell AC. (1980): *Measures of self-reported well-being: their affectice, cognitive and other components.* Social Indicators Research 8:127-55

Bandura A. (1977): *Self-efficacy: Toward a unifying theory of behaviour change.* Psychology Review 84:191-215

Bandura A. (1982): *Self-efficacy mechanism in human agency.* American Psychologist 37:122 47

Breetvelt IS. and van Dam FSAM. (1991): *Underreporting by cancer patients: the case of response shift.* Social Science and Medicine 32(9):981-7

Brickman P., Coates D. and Janoff-Bulman R. (1978): *Lottery winners and accident victims: Is happiness relative?* Journal of Personality and Social Psychology 36(8):917-27

Bunk B., Collins R., Taylor S., van Yperen N. and Dakof G. (1990): *The affective consequences of social comparison: either direction has its ups and downs.* Journal of Personality and Social Psychology 59:1238-49

Calman KC. (1984): *Quality of life in cancer patients - an hypothesis.* Journal of Medical Ethics 10:124-27

Campbell A. (1976): *Subjective measures of well-being.* American Psychologist 31:117-24

Carver CS. and Scheier MF. (1982): *Control theory: A useful conceptual framework for personality-social, clinical and health psychology.* Psychological Bulletin 92(1):111-35

Cassileth BR., Lusk EJ. and Tenaglia ANA. (1982): *Psychological comparison of patients with malignant melanomas and other dermatological diseases.* American Academy of Dermatology 7:742-6

Cella DF. and Tulsky DS. (1993): *Quality of life in cancer: Definition, purpose and method of measurement.* Cancer Investigation 11(3):327-36

Chamberlain K. and Zika S (1992): *Stability and change in subjective well-being over short time-periods.* Social Indicators Resarch 26:101-17

Cohen C. (1982): *On the quality of life: some philosophical reflections.* Circulation 66(suppl III):29-33

Costa PT. and McCrae RR. (1980): *Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people.* Journal of Personality and Social Psychology 38:668-78

Cronbach LJ. and Furby L. (1970): *How we should measure "change" - or should we?* Psychological Bulletin 74(1):68-80

Curbow B., Somerfield M., Legro M. and Sonnega J. (1990): *Self-concept and cancer in adults: theoretical and methodological issues.* Social Science and Medicine 31(2):115-28

Decker SD. and Schultz R. (1985): *Correlates of life satisfaction and depression in middle-aged and elderly spinal cord injured patients.* American Journal of Occupational Therapy 39:740-5

De Grandmont P, Lamarche C, Donahue WB, Boudrias P, Feine JS and Lund JP (1991): *Evaluation of patient's satisfaction in prosthodontics.* Journal of Dental Research 70:559(abstract)

Diener E. 1984): *Subjective well-being.* Psychological Bulletin 95:542-75

Englert JS., Ahrens B., Gebhardt R., Kleifoth M., Saupe R., Steiglitz R-D. and Unnewehr S. (1994): *Implications of the concepts of coping and quality of life for criteria of course and outcome.* Pharmacopsychiatrica 27(suppl):34-6

Erskine A., Morley S. and Pearce S. (1990): *Memory for pain: a review.* Pain 41:255-65

Evans RW. (1991): *Quality of life.* Lancet 338:636

Feine JS., Dao TTT., Lavigne GJ. and Lund JP. (1989): *Is treatment success due to faulty memory of chronic pain?* Journal of Dental Research 68:1018

Ferrans CE. and Powers MJ. (1985): *Quality of life index: development and psychometric properties.* Advances in Nursing Science 8(1):15-24

Ferrans C. (1990): *Development of a quality of life index for patients with cancer.* Oncology Nursing Forum 17:15-19

Folkman S. and Lazarus RS. (1980): *An analysis of coping in a middle-aged community sample.* Journal of Health and Social Behaviour 21:219-39

Foltz AT. (1987): *The influence of cancer on self-concept and life quality.* Seminars in Oncology Nursing 3(4):303-12

Golembiewski RT., Billingsley K. and Yeager S. (1976): *Measuring change and persistence in human affairs: types of change generated by OD designs.* Journal of Applied Behavioural Science 12:133-57

Goodinson SM. and Singleton J. (1989): *Quality of life: a critical review of current concepts, measures and their clinical implications.* International Journal of Nursing Studies 26(4):327-41

Guttman L. and Levy S. (1982): *On the definition and varieties of attitude and well-being.* Social Indicators Research 10:159-74

Guyatt G. Walter S. and Norman G. (1987): *Measuring change over time: assessing the usefulness of evaluative instruments.* Journal of Chronic Disease 40(2):171-8

de Haes JCJM. and van Kippenberg FCE. (1985): *The quality of life of cancer patients: a review of the literature.* Social Science and Medicine 20(8):809-17

de Haes JCJM, Pennink BJW and Welvaart K (1987): *The distinction between affect and cognition.* Social Indicators Research 19:367-78

de Haes JC., de Duiter JH., Tempelaar R. and Pennink BJ (1992): *The distinction between affect affect and cognition in the quality of life of cancer patients - sensitivity and stability.* Quality of Life Research 1(5): 315-22

Headey B, Holmstrom E and Wearing A (1984): *The impact of life events and changes in domain satisfactions on well-being.* Social Indicators Research 15:203-27

Helson H. *Adaptation Level Theory.* Harper and Row, New York, 1964

Heyink J. (1993): *Adaptation and well-being.* Psychological Reports 73:1331-42

Howard GS., Ralph KM., Gulanick NA., Maxwell SE., Nance D. and Gerber SL. (1979): *Internal invalidity in pretest-posttest self-reporting evaluations and a reevaluation of retrospective pretests.* Applied Psychological Measurement 3:1-23

Hughes JE. (1985): *Depressive illness and lung cancer: depression before illness.* European Journal of Surgical Oncology 11:15-20

Hyland ME. (1992): *A reformation of quality of life for medical science.* Quality of Life Research 1:267-72

Hyland ME., Kenyon CAP. and Jacobs PA. (1994): *Sensitivity of quality of life domains and constructs to longitudinal change in a clinical trial comparing salmeterol with placebo in asthmatics.*Quality of Life Research 3:121-6

Hyland ME., Bott J., Singh S. and Kenyon CAP. (1994): *Domains, constructs and the development of the breathing problems questionnaire.*Quality of Life Research 3:245-56

Irwin PH., Allen G., Kramer S. and Danoff B. (1982): *Quality of life after radiation therapy: A study of 309 cancer survivors.* Social Indicators Research 10:187-210

Kreitler S., Samario C., Rapoport Y., Kreitler H. and Algor R. (1993): *Life satisfaction and health in cancer patients, orthopaedic patients and healthy individuals.* Social Science and Medicine 36(4):547-56

Lazarus RS. and Folkman S. *Stress, appraisal and coping.* Springer Publishing, New York, 1984

McCrae RR. and Costa PT. (1988): *Age, personality and spontaneous self-concept.* Journal of Gerontology and Social Science 43:S177-85

McGee HM., O'Boyle CA., Hickey A., O'Malley K. and Joyce CRB. (1991): *Assessing quality of life of the individual: The SEIQoL with a healthy a gastroenterology unit population.* Psychological Medicine 21:749-59

McKennell AC. (1978): *Cognition and affect in perceptions of well-being.* Social Indicators Research 5:389-426

McKennell AC and Andrews FM. (1980): *Models of cognition and affect in perceptions of well-being.* Social Indicators Research 8:257-98

Mischel MH (1988): *Uncertainty in illness.* Journal of Nursing Scholarship 20(4)225-32

Ormel J. (1983): *Neurotisism and well-being inventories: measuring traits or states?* Psychological Medicine. 13: 165-76

Padilla G., Mischel MH. and Grant MM. (1992): *Uncertainty, appraisal and quality of life.* Quality of Life Research 1:155-65

Palmore ED. and Kivett V. (1977): *Change in life satisfaction: A longitudinal study of persons aged 46-70.* Journal of Gerontolgy 32(3):311-6

Powers WT. (1973): *Feedback, beyond behaviourism.* Science 179:351-56

Schain W. (1980): *Sexual functioning, self-esteem and cancer care.* Front Radiation Therapy Oncology 14:12-29

Scheier MF. and Carver CS. (1985): *Optimism, coping and health: assessment and implications of generalised outcome expectancies.* Health Psychology 4(3):219-47

Scheier MF. and Carver CS. (1987): *Dispositional optimism and physical well-being: The influence of generalised outcome expectancies on health.* Journal of Personality 55(2):169-210

Scheier MF. et al (1989): *Dispositional optimism and recovery from coronary artery bypass surgery: The beneficial effects upon physical and psychological well-being.* Journal of Personality and Social Psychology 57(6):1024-40

Spilker B: *Guide to Clincal Trials.* Raven Press Ltd, New York. 1991 p29.

Suls J. and Fletcher B. (1985): *The relative efficacy of avoidant and non-avoidant coping strategies: a meta analysis.* Health Psychologist 4:249-88

Taylor SE. and Brown JD. (1988): *Illusion and well-being: a social psychological perspective on mental health.* Psychological Bulletin 103:193-210

Treborg JR., Howard GS. and Maxwell SE. (1980): *Evaluating planned organisational change: a method for assessing alpha, beta and gamma change.* Academic Management Review 5(1):109-21

Tugwell PX. et al. (1983): *The ability of the MACTAR questionnaire to detect sensitivity to change in rheumatoid arthritis.* Clinical Research 31:239A

Wennberg JE., Barry MJ., Fowler FJ. and Mulley A. (1993): *Outcomes research, PORTS and health care reform.* Annals of New York Academy of Science 703:52-62

Wills TA. (1981): *Downward comparison principles in social psychology.* Psychological Bulletin 90:245-71

NB. End of manuscript IV

3.5 Health status measurement issues: validity and reliability

Broadly speaking, measurement validity refers to the extent to which an instrument measures what it purports to (109). Traditionally, the assessment of an instrument's validity has been through evaluation of its content, construct and criterion validity (109):

Construct validity - the extent to which the measurement corresponds to theoretical concepts (constructs) concerning the phenomenon under study. For example, a measure of health status should reflect the differences between an individual with a broken leg and another individual with no leg problems;

Content validity - the extent to which the measurement incorporates the domain of the phenomenon under study. For example, a measure of health status should embrace physical, psychological, social and role dimensions;

Criterion validity - the extent to which the measurement correlates with an external criterion of the same phenomenon. This in turn can be subdivided into *concurrent criterion validity* (correlation with another criterion used at the same time), for example, the relevant parts of a measure of health status should correlate with measures of physical, psychological and social functions; and *predictive criterion validity* (the ability of the test measure to predict the external criterion), for example, a measure of health status could be expected to predict the period of time off work for a group of patients.

The reliability of an instrument refers to the stability of a measure when it is repeated under identical conditions (109). Integral to the concept of

measurement reliability is the idea that a certain proportion of the variation in any measure is due to measurement error (110). Some error is inevitable. The question is how much is acceptable for the circumstance under which the instrument will be used? For example, one would expect that the proportion of error permitted in the measurement of body temperature would be considerably less than that permitted in the assessment of outside temperature. Obviously this is a value judgement to be made in relation to the particular purpose and target group of each instrument. The proportion of measurement variation due to error can be assessed in several different ways according to the instrument (110):

Intra-observer variation - the variation due to one "measurer" evaluating the same phenomenon more than once;

Inter-observer variation - the variation due to more than one "measurers" evaluating the same phenomenon;

Test-retest variation - the variation due to the same constant phenomenon being measured more than once with a short time interval between measurements;

Internal consistency - the variation in scores of different items or dimensions in an instrument measuring a multi-dimensional construct.

When considering measurement validity and reliability, in addition to the specific domain the instrument purports to measure (eg, QOL, functional status, psychological status, pain etc.) it is also important to remember the target group and construct of the instrument. Some instruments are designed to be used among general populations (eg, Sickness Impact Profile and Nottingham Health Profile), while others are designed to be used in relation to specific diseases (eg, Arthritis Impact Measurement Scale, European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire and the Hospital Anxiety and Depression Scale), others are designed to assess problems at specific anatomic sites (eg, the Oral Health Impact Profile) and others are designed to assess specific aspects of health status (eg, McGill Pain Questionnaire and the Index of Independence in Activities of Daily Living).

Another aspect of an instrument's validity, which needs to be considered, is its application. Kirshner and Guyatt have categorized instruments as discriminative, predictive and evaluative (111). The former is used to discriminate between groups according to a certain phenomenon where no "gold standard" exists (eg, scales which distinguish between those with psychological disorders and the general population, or between those with functional disabilities and the general population). A predictive index is used to classify individuals into a group of predefined categories where a gold standard is available. These are typically cheap and easy tests used in screening where a positively identified individual will then be subjected to the gold standard (eg breast palpation leading to biopsy). Finally, an evaluative index is used to evaluate the magnitude of change in a certain individual or group in relation to a certain phenomenon (typically a treatment). In the case of evaluative instruments, an important aspect of their reliability/validity is their sensitivity to change (112). In conclusion therefore, it is extremely important to be aware of the target population and the purpose of the instrument, because validity testing of that instrument in one group and for one purpose is not necessarily relevant to a different population and purpose.

3.6 Health status measurement in UADT cancer patients

In 1992, Gotay and Moore published a review of what they called "QOL" assessment in UADT cancer patients covering the period 1980-90 and found that no studies had defined QOL. They also reported that most studies did not investigate QOL in a comprehensive fashion, rather they used a miscellaneous collection of instruments and questionnaires covering one or two of the domains discussed in section 3.2 or only parts of such domains (113). Furthermore, no studies reported on the validity of the instruments they had used, although one study had used a previously validated cancer-specific (not UADT cancer-specific) QOL scale, and others had used previously validated measures of speech, hand-grip strength, performance status, depression and psychological well-being (113). Some of these

phenomena may be relevant to health status measurement in UADT cancer patients (and some may not), however, none of the instruments cover all relevant dimensions.

Since the 1980s, the assessment of health status, QOL and HRQL and the development of appropriate instruments has expanded dramatically in all fields of health care, including that of UADT cancers. I shall now review the instruments used to evaluate health status, QOL or HRQL among UADT cancer patients in studies published since 1990.

3.6.1 The Performance Status Scale - Head and Neck (PSS-HN): One of the earliest health status evaluations specific to UADT cancer patients was that developed by List et al (114). This is a "performance status scale" (the PSS-HN) which is clinician-rated and covers the domains of understandability of speech, normalcy of diet and eating in public (114). Each of the domains is rated on a categorical scale from 0-100, with higher scores indicating better performance (114). Although not made explicitly clear, the construct upon which this instrument was developed appears to have been what the developers believed were three principal domains of disability suffered by UADT cancer patients following treatment. Furthermore, it appears to have been based upon the principles used in the development of the Karnofsky performance scale, a clinician-rated instrument used extensively in the field of cancer care and aimed at evaluating cancer patients' mobility and ability to maintain employment, live at home and care for themselves (115). Its developers wanted the PSS-HN to be a simple, practical assessment tool which was acceptable to patients and their carers and that required no formal training for administration (114). There was, however, no discussion of whether a patient- or clinician-rated instrument would be more appropriate. Psychometric testing of the instrument demonstrated that it had reasonable inter-rater reliability; an internal consistency which suggested some correlation between the domains but also that each domain was contributing some unique information; a criterion validity tested against the Karnofsky performance scale which again suggested some correlation

between the two instruments but that the PSS-HN contributed some unique information; and that it was able to discriminate between UADT and breast cancer patients, and within the UADT group between those treated with different forms of surgery (114). They did not test the sensitivity to change or the test-retest reliability of the instrument. As a clinician-rated measure of performance specific to UADT cancer patients and within the limits of its conceptual framework, this testing of the instrument suggests that the PSS-HN has reasonable psychometric properties as a discriminative instrument. However, in terms of a measure of health status its validity is not good in that its content is too limited and it is not rated by the patients themselves.

Nevertheless, the instrument has been used in several studies and has contributed some information for treatment decisions concerning this group of patients. In an observational study of 36 base of tongue cancer patients, Harrison et al found that, controlling for disease stage, those treated with radiotherapy alone were consistently evaluated with better performance status in all three domains (understandability of speech, normalcy of diet and eating in public) than their peers treated with surgery alone (116). And in a similar observational study of 29 disease-free, long term survivors of oral and oropharyngeal carcinoma all treated with surgery and post-operative radiotherapy, the same group of investigators found that stage was predictive of performance status in all three of the PSS-HN domains, and that cancer site was predictive for understandability of speech and normalcy of diet (117). Another observational study of 49 patients treated with external beam radiotherapy for base of tongue cancer, (using descriptive statistics only) found that the eating in public and normalcy of diet domain scores tended to decrease with worsening disease stage (118). It is worth noting, however, that these aforementioned studies (116-118) using the *clinician-rated* PSS-HN did so in a *patient-rated* manner. In their original development of the instrument, List et al did examine the inter-rater reliability comparing "trained" and "untrained" personnel and found that the agreement for two of the domains (normalcy of diet and eating in public) was good, while that for the

third (understandability of speech) was poor (114). However, they did not examine the results of any patient-rated evaluations. It is difficult to predict the affects on the aforementioned studies of using the instrument in such an erroneous fashion, however, this fact must be considered when making any inferences from them.

Other than these studies, the group who developed the PSS-HN instrument have also used it in a prospective, non-randomised, observational investigation of the recovery of 3 groups of laryngeal cancer patients over a 6-month, post-therapeutic period (119). The 21 subjects were divided equally into total laryngectomy, hemilaryngectomy and radiotherapy only groups (7 subjects in each). They found that the total laryngectomy group took longer to recover performance status than the hemilaryngectomy group but that the performance status was not significantly affected at any time during the post-therapeutic period in the radiotherapy only group (119). This suggests that the instrument has limited sensitivity to changes as one would expect there to be some form of treatment-related morbidity following radiotherapy. Furthermore, as the authors observe, their findings were probably strongly influenced by disease stage, which was the principal factor determining which treatment group each patient entered (119). We therefore do not know whether the observed effects are due to disease stage or treatments. Furthermore, as the original testing of the instrument did not involve investigating its sensitivity to change or test-retest reliability, we cannot be sure what proportion of the observed change is true and what proportion is erroneous.

3.6.2 The Functional Assessment of Cancer Therapy-General (FACT-G) and Head and Neck Scale (FACT-H&N): The same group that developed the PSS-HN have also used the latter in an assessment of the validity of an UADT cancer-specific HRQL measure; the FACT-H&N (120). The FACT-H&N is patient-rated HRQL measure consisting of a 28-item core questionnaire (the FACT-G), designed to be used among patients with any form of cancer, which can be complemented by site and/or treatment-specific subscales (121), of which the 9-item "head and neck"

subscale is one. The 28-item core comprises five domains: physical well-being; emotional well-being; social/family well-being; functional well-being; and relationship with doctor; all of which contain items that are scored on a 5-category Likert scale (0-4). The scores from the items in each domain are combined to produce a domain rating, with higher scores indicating a better HRQL (121). Testing the FACT-G and FACT-H&N instruments in a sample of 151 UADT cancer patients, List et al found that the domains and overall instruments had reasonable internal consistency as measured by Cronbach's α (range: α=0.59-89), with the emotional and social domains having the worst internal consistency (α=0.59 for both) (120). When testing discriminant validity, they found that, with the exception of the emotional domain, the FACT-G and FACT-H&N subscale were able to distinguish subjects by Karnofsky performance score, and that with the exception of the emotional and social domains, the FACT-G and FACT-H&N subscale were able to distinguish subjects by treatment status (120). They further tested construct validity of the FACT-G and FACT-H&N subscale by analysing the correlation between their domain scores and those of PSS-HN domains. Consistent with their hypotheses, the best correlations were between the PSS-HN normalcy of diet and eating in public domains and the FACT-H&N subscale (120). They concluded that the FACT-H&N was a reliable and valid instrument for use among UADT cancer patients (120), although they did not test intra-rater or test-retest reliability. In the original development of the FACT-G, the test-retest reliability was evaluated among 70 patients with mixed cancer diagnoses with 3-7 days between evaluations. The reliability was found to be good (range: r=0.82-0.92) (121). However, there does not appear to have been any such testing of the FACT-H&N subscale. This is especially important in light of the fact that the same group used the FACT-H&N in the previously discussed, prospective investigation of health status among laryngeal cancer patients (119). They found no significant associations between treatment type or post-therapeutic time and FACT-H&N domain scores, raising the question: Are there no differences with time, is the instrument unreliable or is it not sensitive enough to detect changes in this group?

The development of the FACT-G other than the investigation of its test-retest reliability was an extensive process with many stages (121). Item generation originally involved semi-structured interviews with lung, breast and colonic cancer patients, plus oncology specialists and then review by independent judges and comparison of items generated for the 3 disease sites to eradicate overlap. Further piloting of the items among different lung, breast and colonic cancer patients enabled reduction of the questionnaire to 38 items. This version of the questionnaire was then extensively tested among 545 patients with a variety of cancers. The subsequent data were tested for their fit to the hypothesised underlying concept of HRQL and subjected to factor analysis. The result of this stage was the elimination of 10 items and the production of 6 significant factors able to explain 51% of the total variation. Criterion and discriminant validity of the subsequent 28-item questionnaire was then tested and finally test-retest reliability and sensitivity to change were also tested, all demonstrating that the FACT-G is a valid HRQL instrument for patients with a variety of cancers (121). The FACT-G therefore appears to be a cancer HRQL instrument with fairly extensive development and validity testing demonstrating that it has reasonable psychometric properties. However, as pointed out by the authors, the FACT-G was designed to be complemented by disease and/or treatment specific subscales of which the FACT-H&N is one.

Further testing of the validity of the FACT-H&N has subsequently been performed by comparing the scores of that instrument with those of the PSS-HN and another UADT cancer-specific HRQL measure, the University of Washington Quality of Life Questionnaire for Head and Neck Cancer Patients (UW-QOL) (122).In a study sample of 50 UADT cancer patients, D'Antonio et al found good correlations between the FACT-H&N subscale and UW-QOL and PSS-HN scores (122). And in one of the few studies to investigate the role of multiple variables in predicting HRQL in UADT cancer patients, using data from the same 50 subjects, the same group of workers found FACT-H&N scores to be associated with disease site

(laryngeal site better HRQL rating than oral or pharyngeal) and FACT-G scores to be associated with marital status (married patients better HRQL rating than others) (123). The FACT-H&N therefore appears to fulfill some aspects of validity and reliability testing (in particular, criterion and discriminant validity and internal consistency), however, it is noteworthy that no reference to the generation of the subscale items has been made and no investigation of its test-retest reliability has been reported, therefore making its sensitivity to change questionable.

3.6.3 The University of Washington Quality of Life Questionnaire for Head and Neck Cancer Patients (UW-QOL): With respect to the other UADT cancer-specific HRQL instrument (the UW-QOL) used in the aforementioned studies (122,123), this is a patient-rated questionnaire with 9 items covering 8 domains: pain; disfigurement; activity; recreation/entertainment; employment; eating (encompassing swallowing and chewing); speech; and shoulder disability (124). Subjects mark a phrase best describing their situation with respect to each of the domains. The phrases have scores ranging from 0-100, with higher scores meaning better HRQL, and the scores are summed to give an overall HRQL rating (124). The instrument was tested among 75 UADT cancer patients for acceptability, intra-rater reliability, criterion and construct validity and sensitivity to change. It was demonstrated to be more acceptable among subjects (ie they found it more relevant and shorter to complete) than the Sickness Impact Profile (SIP - a measure of the impact of ill-health problems containing 136 weighted items in 14 domains and designed to be used in a general population (125)); to have good intra-rater reliability; to have good criterion validity when comparing UW-QOL scores with those of the SIP and Karnofsky scales; to show a tendency for decreasing HRQL score with increasing disease stage, as would be expected; and to be more responsive to change than both the SIP and Karnofsky scales when evaluated prior to, immediately following and 3 months following cancer therapy (124). At first glance, the UW-QOL therefore seems to have reasonable psychometric properties. However, its major drawback is the fact that

there is no explanation of any underlying construct or the subsequent item generation and categorisation of domains. There is subsequently no testing of its internal consistency or demonstration that these domains are appropriate for measuring HRQL in UADT cancer patients, although the decreasing HRQL score with increasing disease stage suggests some construct validity. In other words, this instrument may have good construct validity and internal consistency, but this has not been demonstrated. This may explain why it has rarely been used in other research projects.

3.6.4 The European Organization for Research and Treatment of Cancer quality of life core questionnaire (EORTC QLQ-C30) and head and neck module (EORTC QLQ-H&N37): The EORTC QLQ-C30 and EORTC QLQ-H&N37 were developed along similar lines to the FACT-G and FACT-H&N. The EORTC chose this modular approach to simultaneously facilitate generalizability in allowing cross-study comparisons, plus a level of specificity adequate for addressing research questions relevant to particular cancer groups (126). Although the construct of the EORTC QLQ-C30 is similar to that of the FACT-G in that it is designed to be a cancer specific, patient-rated, multidimensional HRQL instrument, it has the added element of being cross-culturally equivalent (127). It being developed by a European organization, the developers wanted it to be applicable and equivalent in several of the principal European languages. Item generation for the core and modular instruments was based upon a hypothesised multidimensional construct and involved a literature review, interviews with patients and relevant health care professionals plus pilot testing (128). The subsequent draft questionnaire was then tested among 346 lung cancer patients in 13 countries, using factor analysis to test for domains, plus tests for internal consistency, discriminant validity and sensitivity to change. This investigation of the psychometric properties of the instrument supported the hypothesised domain structure of the instrument which comprised 9 multi-item domains (physical, role, cognitive, emotional, social, fatigue, pain, nausea and

vomiting and global health and QOL domains), plus 6 single items (dyspnea, appetite loss, sleep disturbance, constipation, diarrhoea and financial problems) (127). The majority of items have a 4-category Likert response scale, while the two global health and QOL items have 7-category Likert response scales. The testing of internal consistency showed that while the domains were correlated to the underlying HRQL construct, they were additionally all giving some unique information (127). With respect to discriminant validity, only the emotional domain showed significantly different scores by disease stage, however, the majority of domains were able to distinguish between different groups by pretreatment weight loss and on treatment toxicity scores (127). Finally, the analysis of the instrument's sensitivity to change found that some of the domains (physical, role, fatigue, nausea and vomiting and global) correlated well with the changes as measured by a performance status indicator (127).

In a simultaneous study, the EORTC QLQ-C30 (the core instrument) was also tested in a sample of 126 UADT cancer patients in Norway (129). Again the instrument was tested for internal consistency, and criterion, discriminant and construct validity. Internal consistency was found to be satisfactory for all except the cognitive domain (α=0.28), criterion validity for those domains expected to correlate with the General Health Questionnaire used in this study was good and those domains hypothesised as being able to discriminate between different patient groups did so (129). It is evident, therefore, that while the EORTC QLQ-C30 has a relatively sound theoretical construct (accepting the arguments detailed in section 3.2 confirming the lack of consensus on a definition of health status, QOL or HRQL) in which the multidimensional and subjective nature of HRQL is central and it has undergone fairly extensive psychometric testing, no attempt to assess the intra-rater or test-retest reliability of the instrument has been made, raising questions concerning its sensitivity. However, the EORTC group have argued that demonstrating the instrument's internal consistency (or internal reliability) is another means of showing reliability in that patients have scored all domains in the instrument in a correlated

manner (127). They accept that one would ideally investigate test-retest reliability separately but that this presents too many problems in calculating the most appropriate time interval when HRQL would not have changed and patients would not remember their initial responses (127). Furthermore, the fact that the EORTC QLQ-C30 appears to be sensitive to change when compared with a performance status indicator, indirectly suggests its test-retest reliability. A separate observation concerning the instrument is that its design is complex with several multi- and single item domains which should be treated separately rather than combined to give an overall score. This may be a more valid representation of the HRQL construct than the more simple instruments, however, it produces a series of outcome scores rather than one overall score (other than that of the global health and QOL domain), thereby rendering interpretation more complex and necessitating more complex analytic techniques. It does, however, have the advantage of having been simultaneously validated in several languages.

With respect to the development of the accompanying modular instrument for UADT cancer patients (EORTC QLQ-H&N37), this process started with the testing of a 14-item questionnaire in the UK (130) and in Norway (129). Once again, the items were generated in a procedure identical to that used for the core instrument, and in pilot testing, the questionnaire proved acceptable to patients and produced scores in the expected direction with patients with recurrent disease having consistently, significantly worse scores than patients with laryngectomies (130). The questions in the H&N module also proved to be consistently able to discriminate between patients at different stages of treatment (129). Following these pilot studies of the draft H&N module, it underwent further development including the use of a literature search, interviews with relevant patients and specialists, further pilot testing and comparison of the content with the core instrument to ensure its complementary nature (131). Once again, this was all done simultaneously in several European languages to ensure cross-cultural equivalency. As yet, the psychometric properties of the EORTC QLQ-H&N37 are still under investigation and no published data

concerning the instruments validity or reliability have been published.

Nevertheless, there are some reports of data obtained using the core instrument and draft versions of the H&N module among UADT cancer patients. In an investigation of the long term follow-up of a sample of UADT cancer patients randomized to receive different radiotherapy regimes, 245 patients were mailed the core and a 19-item H&N module instruments, 7-11 years following the completion of therapy (132). Seven of the core domains demonstrated a consistently better level of functioning and lower level of symptoms for one of the regimes, while 3 items in the H&N module (regarding problems swallowing and nose breathing) showed lower levels of problems for the same regime. However, when the effect of concomitant surgical treatment was controlled for in a multivariate regression equation, none of the domains tested were significantly associated with any clinical variables (132). And in another investigation of the psychological status of the same sample, low cognitive and social function and high levels of pain, in addition to disease stage were found to be significant independent predictors of psychological distress (133). In a cross-sectional study of 50 patients with a UADT cancer 1-6 years following treatment, using descriptive statistics only, the same group of workers reported that this patient group had similar ratings on the EORTC core and on a 12-item H&N module, compared to those 7-11 years following treatment (134). In an investigation of the feasibility of administering the core and H&N module instruments on a long term basis among 50 UADT cancer patients, Hammerlid et al found that the module had greater sensitivity to change than the core instrument over a 1-year post-treatment follow-up period (135). They also found that the greatest level of functional and symptomological problems in this group was during the period 2-3 months following therapy (135). In a subsequent prospective, non-randomized investigation of the merits of external beam radiotherapy (ER) vs. external beam radiotherapy and brachytherapy (EBR) among a sample of 105 oral and pharyngeal cancer patients, the same group of researchers reported finding that, 12 months following therapy, the ER group reported lower levels of pain but the EBR group reported lower levels of

xerostomia (136). These more recent studies would appear to validate the need for a disease-specific module to compliment the core instrument and that a number of items in that module concerning pain, xerostomia and swallowing difficulties would appear to be particularly pertinent when trying to differentiate different treatment modalities in UADT cancer patients.

To summarize the evidence concerning the psychometric properties of the EORTC QLQ-C30 and H&N module, the core appears to have good construct, content, criterion and discriminant validity and internal consistency, and although it has not been tested for intra-rater or test-retest reliability specifically, it appears to be sensitive to changes in the predicted manner. The H&N module has not yet been fully validated, however, draft items appear to have construct validity and are consistently sensitive to changes in the expected directions. The EORTC instruments do have the additional benefit of having being simultaneously developed in several languages to facilitate cross-cultural comparisons.

3.6.5 The Functional living Index - Cancer (FLIC): Another measure of HRQL designed to be used among cancer patients (although not UADT cancer patients specifically), which was one of the earliest of such measures to be developed, is the FLIC. It consists of 22 items concerning physical, psychological and family/social well-being, each of which have 7-category Likert scale response (137). The items are designed such that responses can be summed to produce an overall score. Its development involved an extensive literature review, interviews with relevant health care professionals and extensive pilot testing among cancer patients. The instrument showed scores according to treatment status largely as would be expected, suggesting its ability to discriminate, and the different domains of the instrument correlated as expected with various criterion instruments, suggesting good criterion validity (137).

The FLIC has, however, only been used to investigate the HRQL of UADT cancer patients on a few occasions. Using a series of univariate analyses, a cross-sectional study of post-treatment HRQL in a sample of 135 oral cancer patients

found tumour size, defect location, type of soft tissue reconstruction and the resection of mandibular bone to be significantly associated with FLIC score (138). However, following multivariate logistic regression to determine the role of surgical variables in predicting FLIC score, the only variable which remained significant was the need for mandibular resection as part of the ablative treatment, which was associated with a worse HRQL (138). In view of the FLIC's lack of a UADT cancer specific module, the researchers in this study designed their own brief module and tested associations between it and FLIC scores. They found that patients with dysphagia, food reflux, decreased appetite and persistent pain had significantly worse FLIC scores (138). In a prospective study of 85 oral cancer patients over a period of 12 months post-therapeutically, the same group of researchers confirmed their findings in the aforementioned retrospective study. In a series of bivariate analyses including FLIC score, time following treatment and various surgical parameters, they found that the defect location and type of reconstruction were associated with FLIC score. They also found that dysphagia, food reflux and limited diet were associated with FLIC score and that overall FLIC score improved significantly with time, but that the improvement in the physical domain was most marked (139). Finally, a cross-sectional study among 45 UADT cancer patients investigated the role of social support in predicting HRQL in this group. Using the FLIC as the dependent outcome, they found that satisfaction with family physician support, stage of cancer, site of cancer and patient gender were significant predictors of HRQL (140). In summary, the FLIC has good face, content, construct and discriminant validity and good internal consistency and appears to have good sensitivity. However, it does not have a UADT cancer specific module necessary to explain the more subtle differences in health status variables evident within such a group.

3.6.6 Miscellaneous instruments used to measure health status among UADT cancer patients: Other than these relatively well-recognised and well-used measures of the different aspects of health status in UADT cancer patients (the PSS-HN,

FACT-G and FACT-H&N, the EORTC QLQ-C30 and EORTC QLQ-H&N, the FLIC and the UW-QOL), since the previously mentioned review by Gotay and Moore covering the period 1980-90 (113), a number of other studies investigating various aspect of health status in this group, using alternative instruments, have been published. Rathmell et al developed their own UADT cancer-specific HRQL questionnaire with 19 items covering elements of symptomology, function, daily activities, psychological and social well-being, and tested it in a cross-sectional survey of 96 UADT cancer patients (141). Although the instrument has some face validity, there was no justification for its construct, no testing of its psychometric properties, and no statistical analysis of the findings was performed. It is therefore, very difficult to draw any conclusions concerning this questionnaire. Browman et al developed a head and neck radiotherapy questionnaire (HNRQ) designed to evaluate aspects of QOL affected in patients undergoing radiotherapy for an UADT cancer (ie, a disease and treatment-specific instrument) (142). The instrument was developed through literature review and interviews with relevant specialists and patients and consists of 22 items with a dichotomous response, divided into 6 domains (skin, throat, oral stomatitis, digestion, energy, psychosocial). For those responding "yes" to any of the items, there is a supplementary question concerning "how troublesome" that particular problem is. They tested the instrument's construct, criterion and discriminant validity, plus its sensitivity to change in an on-going, randomized clinical trial comparing the use of a radioenhancing drug against a placebo during radiotherapy for UADT cancer (142). They found that the HNRQ had good criterion validity in that the relevant domains of the instrument correlated well with other appropriate measures of toxicity, that it was able to discriminate between the active and placebo patients in the trial, that it was sensitive to change over time and that it behaved in the hypothesised manner, thereby supporting its construct validity (142). Although internal consistency or factor analysis was not investigated to support the domains the developers created, there is indirect evidence to support this through the criterion and construct validity. It therefore appears to be an instrument with

reasonable psychometric properties but with application limited to one disease- and treatment-specific situation.

Another cross-sectional study of clinician- and patient-rated functional status and HRQL among 43 oral cancer patients used two untested questionnaires designed specifically for that study (143). The clinician-rated questionnaire covered issues of dental status, oral hygiene, oral mucosal status, tongue and mandibular mobility, speech intelligibility and cosmetic results, while the patient-rated questionnaire contained 25 items covering diet, ability to eat, eating out, oral symptomology, problems with speech daily activity and overall HRQL (143). Using descriptive statistics only, they reported that their findings suggested that patients with stage I tumours treated with surgery only had a better HRQL (143). Once again, although there are aspects of the patient-rated questionnaire with face validity, its content validity for a HRQL questionnaire appears poor and the questionnaire has been subject to no psychometric testing.

Another study investigated the functional status and coping of a sample of 42 oral and pharyngeal cancer patients using the SIP and the sense of coherence scale (SOC) (144). They found that the overall SIP score and some of its domains varied with time since treatment and extent of surgery, and that the SIP scores of the sample were consistently worse than those of a reference group with no diagnosed health problem. However, the SOC varied within the UADT cancer patient sample only by gender (females coping worse), although the SOC and the scores for some of the SIP domains (emotional behaviour, social interaction, sleep/rest, psychosocial and overall score) demonstrated significant correlations at 12 months following therapy (144). Although both the instruments employed in this study had been validated prior to their use here, neither were designed to assess health status in UADT cancer patients. In this study, the SIP particularly demonstrated some useful information, however, it is difficult to comment on its relative validity compared to an appropriate disease-specific instrument other than to say that the latter is likely to be more sensitive and more able to discriminate between groups.

In an investigation of some aspects of the HRQL differences among laryngeal cancer patients treated with total, near-total and partial laryngectomy, DeSanto et al used the Psychosocial Adjustment to Illness Scale (PAIS) plus a disease-specific questionnaire that the investigators designed themselves in a cross-sectional study of 172 patients (145). The PAIS consists of 46 items, each with a 4-category Likert scale, grouped into 7 domains (health attitudes, work/school, relationship with spouse, sexuality, family relationship other than spouse, hobbies and activities and psychologic (146). According to this instrument, there was virtually no difference between total and near-total laryngectomees, with the exception of the work domain wherein the near-total group reported a worse evaluation. However, the PAIS scores in all domains were significantly better in the partial laryngectomy group than the other two groups (145). With respect to information gathered in this study using the disease-specific questionnaire, very limited information concerning its content is given and only descriptive statistics are used to explain any differences between the treatment groups. It is, therefore very difficult to comment on the merits of this questionnaire. With respect to the merits of the PAIS among UADT cancer patients, it appears to contribute some information in this group, however, its limited use prevents comment on its properties relative to disease-specific instruments.

Finally, another UADT cancer-specific, patient-rated, functional status instrument, containing 16 items in domains concerning symptomology, function, appearance, fatigue and overall HRQL, was developed by Baker (147). Each item has a 5-category Likert response scale, and items are summed to give an overall score with a higher score indicating a better HRQL. Once again, the items were generated from a combination of literature review, and patient interviews and the subsequent questionnaire was assessed for discriminant, construct and criterion validity, plus internal consistency. In a sample of 172 UADT cancer patients and 31 patients with gynaecological or breast cancers, the instrument was able to discriminate between the UADT and non-UADT cancer patients; it was associated with disease stage and

extent of surgery; it showed good correlations with several other criterion instruments; and had a Cronbach's α of 0.88 (147). This instrument therefore appears to have reasonable psychometric properties for a measure of functional status, however, it lacks some of the domains appropriate for the broader concepts of health status or HRQL.

3.6.7 A summary of health status measurement in UADT cancer patients: A number of instruments have been used, some of which are *disease-specific* (the PSS-HN, FACT-G and FACT-H&N, the EORTC QLQ-C30 and EORTC QLQ-H&N, the UW-QOL, the HNRQ and the Baker measure of functional status); while others are only *cancer-specific* (the FLIC); and others have been developed for use in *general populations* (the SIP, the PAIS and the SOC); and some investigators have used untested questionnaires. While it is very difficult to make comparisons concerning the relative merits of these instruments, if the goal is to understand the determinants of health status within UADT cancer patients as a whole, then patient-rated, disease-specific instruments covering appropriate domains are the most appropriate. Making these qualifications reduces the list of possible instruments to the FACT-G and FACT-H&N or the EORTC QLQ-C30 and EORTC QLQ-H&N. Both core instruments have undergone relatively extensive validation and been demonstrated as having reasonable psychometric properties among UADT cancer patients. However, both modular instruments have, as yet, undergone only limited validation, although this work suggests that each will be a psychometrically sound instrument. Broadly speaking, therefore, the psychometric status of these instruments is equivalent, however, several differences in their operationalization do exist. The EORTC instrument (including core and module) is more complex, containing a larger number of multi-item and single-item domains covering a broader HRQL construct than the FACT. Both instruments are designed to produce separate domain scores which may be analysed on their own. The FACT can produce an overall rating by summing the domain scores, while the EORTC produces an overall rating through

a separate "global rating of health and QOL" domain. The final difference is that the EORTC instrument has been simultaneously developed and psychometrically tested in several languages, while the FACT has been developed and validated in English only.

3.7 Determinants of health status in UADT cancer patients: project aim

The aim of this aspect of the research was to investigate the determinants (or at least correlates) of health status in a sample of UADT cancer patients following their therapy. In particular, it was the aim of this study to evaluate the relative contribution of clinical, sociodemographic and patient-rated variables in the variance of global health ratings in this population.

3.8 Methodology

Patients diagnosed with and treated for oral, pharyngeal and laryngeal carcinomas (ICD-9 141, 143-9 and 161) were included in this study. They were all recruited from one of three McGill University hospitals in Montreal and the data were collected during the 18 month period July 1st 1995 to December 31st 1996. The study was approved by the human ethics committees of the relevant hospitals and all subjects read and signed an informed consent before participating in the study. Data were collected from subjects in a 20-30 minute interview using a standardised questionnaire (see Appendix 2) eliciting information on socio-economic, demographic and clinical variables, plus the EORTC instruments. Interviews were performed on a convenience sample of patients while they were attending for routine post-therapeutic follow-up consultations.

3.8.1 Methodology - variables

In view of the fact that this research was located in Montreal, Quebec, (a Canadian province which officially recognises two languages: English and French), the EORTC QLQ-C30 and QLQ-H&N37 instruments, simultaneously

developed in, among other languages, English and French, were used to evaluate health status. As described in section 3.6.4, the EORTC QLQ-C30 comprises 9 multi-item domains (physical, role, cognitive, emotional, social, fatigue, pain, nausea and vomiting and global health and QOL domains), plus 6 single items (dyspnea, appetite loss, sleep disturbance, constipation, diarrhoea and financial problems). The majority of items have a 4-category Likert response scale, while the two global health and QOL items have 7-category Likert response scales (127). The instrument is designed in such a way that scores for items within multi-item domains are averaged to give a domain score and all domain scores are treated separately. This means that rather than the instrument providing an overall rating for health status, it effectively provides a collection of evaluations covering those domains included in the HRQL construct. The only rating which could be conceived as an overall rating is the "global health and QOL rating" domain. For the purposes of this research project, this global rating domain is considered as the dependent variable, and all other domains are considered as independent variables. Among functional domains, a higher score reflects better function, while among symptomological domains a higher score reflects a higher level of symptoms, and a higher score in the global domain reflects a better health and HRQL rating (127).

With respect to the module EORTC QLQ-H&N, although this instrument has undergone some preliminary testing, it remains to be fully validated; indeed data collected in this study are contributing to that process. It contains 35 items covering symptomological, functional, psychosocial and sexual domains, plus several items covering medication and weight loss. However, a factorial analysis of the questionnaire has not yet been published and none of the items has been assigned to a specific domain (see Appendix 3 for both EORTC instruments). In view of the incomplete validation of the H&N module, the outcomes represented by the core instrument were used for the primary analysis of predictors of health status.

The majority of independent variables were nominal (patient gender, cohabitation status, education status, comorbidity status, dental status, disease site,

treatment modality and the presence or not of recurrent or new UADT malignant disease), one was ordinal (disease stage) and two were continuous (patient age and time since the completion of treatment). Comorbidity status was defined as the presence or not of any diagnosed comorbidity at the time of health status evaluation; dental status was dichotomized into those with one or more teeth and those edentulous at the time of evaluation; disease site was categorized as oral (ICD-9 141 and 143-5), pharyngeal (ICD-9 146-9) and laryngeal (ICD-9 161) cancer; treatment modality was categorized as surgery only, radiotherapy only or combination therapy; and time since treatment was evaluated in months since the completion of patients' last treatment (for those with recurrent or new disease, this was the time since the completion of the treatment for that disease). Disease stage was classified according to the American Joint Committee on Cancer (57) as shown in Table 1, section 2.2. Separate analyses were performed with overall stage used as the independent variable to describe the extent of the disease at diagnosis, or with the T and N categories entered as separate independent variables used to describe the same variable.

3.8.2 Methodology - statistical analyses

Following descriptive statistics, a series of univariate analyses were performed to evaluate the associations between all independent variables and health status ratings. The univariate associations between nominal variables and the dependent outcomes were tested with a student's t-test, that between the ordinal variable and dependent outcomes was tested using Spearman's rank correlation coefficient, and those associations between the continuous independent variables and dependent outcomes were tested using Pearson's correlation coefficient. These analyses demonstrated those variables with significant univariate associations with health status ratings, thereby giving an indication of those variables most likely to be part of a multivariate model explaining the variation in health status.

In view of the direction of the potential associations between different independent variables and the global HRQL rating, separate multivariate models

were constructed. That is, the relationship between some of the independent variables and global HRQL can only be in one direction (eg. in the relationship between age or gender and global HRQL, only the latter could be the dependent variable), while among other independent variables (especially other subject-rated variables within the EORTC instrument), the direction could be in either or both directions (ie. although it is often assumed that the domains which make up HRQL instruments, such as physical, emotional, social etc., contribute to an individual's evaluation of their overall HRQL, it is conceivable that HRQL is not the dependent variable but the determining variable in these relationships.). If this is the case, mixing subject-rated (potentially bidirectional) variables with non-subject-rated (unidirectional) variables in the same multivariate model will result in the latter being impossible to interpret. Thus, two multivariate models were evaluated, one with subject-rated (potentially bidirectional) "independent" variables and the other with non-subject-rated (unidirectional) independent variables. The models originally contained all independent variables and those variables which explained the least amount of variation were successively removed from the model until it contained only those variables with a significant independent association with the outcome variable. The results of these analyses and further discussion of the potential direction of the possible relationships in this study are reported in manuscript V.

3.9 Manuscript V: "Correlates of quality of life in upper aerodigestive tract cancer patients" (submitted manuscript)

Abstract

Reflecting a limited understanding of the definition and determinants of health-related quality of life (HRQL), the majority of research in this field has concentrated upon the effect of disease- and treatment-related variables. That work specifically investigating HRQL among upper aerodigestive tract (UADT) cancer patients is no exception to this observation. Treating subject-rated and non-subject-

rated variables separately, the aim of this study was to investigate predictors of global HRQL rating in a sample of UADT cancer patients, concentrating upon the relative importance of sociodemographic and clinical variables. A cross-sectional study design was used with a sample of 188 UADT cancer patients. Global HRQL was assessed using the EORTC QLQ-C30 instrument, global domain (global QL). Other study variables were collected by subject interview and chart review. Two multivariate regression models were independently developed, containing respectively subject-rated and non-subject-rated variables. In the model containing subject-rated predictors of global QL, emotional, breathing, physical, financial, pain and appetite problems were significant predictors ($F = 14.6$; $p<0.0001$; $r^2 = 0.54$). Among non-subject-rated sociodemographic and clinical variables tested, unemployment, older age, female gender, being dentate and a more advanced disease stage predicted worse global QL rating, while oral, as opposed to pharyngeal or laryngeal, cancer predicted a better global QL rating ($F = 5.1$; $p<0.0001$; $r^2 = 0.21$). In the latter model, a greater proportion of the variance was explained by sociodemographic variables than by clinical variables.

Key words: head and neck cancer, quality of life, predictors

Introduction

Over the past 10-15 years, there has been an enormous increase in the volume of publications concerning health-related quality of life (HRQL). Unfortunately, this rush to measure HRQL has proceeded with relatively little discussion of its definition as it relates to health care (1). This has inevitably led to a situation in which our understanding of the complex interaction of factors which determine HRQL is poor. As Wilson and Kaplan have argued, "the conceptual models that underlie the measurement of health status, as they are currently operationalized, do not meet the important need of clinicians to understand causation and mechanism, without which rational and effective therapy is difficult" (2). With

the exception of randomized clinical trials, where all variables are theoretically distributed evenly between groups, it is important to have an understanding of the determinants of HRQL if they are to be used as outcome/dependent variables, as is increasingly the case. Unfortunately, the well-intentioned and increasing use of HRQL instruments in health care-related situations seems to be a product of the assumption, among health care providers, that the principal determinants of HRQL are disease- and treatment-related variables. While this is certainly the case for some HRQL domains (eg. particularly the symptomological and functional ones), there is not a lot of evidence to support this assumption for the psychosocial domains or overall evaluation of HRQL. Furthermore, the diversity of domains employed in HRQL instruments suggests that disease- and treatment-related variables will explain only a small proportion of the variation in this outcome. For example, the HRQL instrument employed in this study (the EORTC QLQ-C30) contains domains concerning physical, role, cognitive, emotional and social functioning, various forms of symptomology, finances and an overall evaluation of health and HRQL (3).

Responding to this lack of clarity concerning the determinants of HRQL, Wilson and Cleary have described a conceptual model linking clinical variables and HRQL, while incorporating individual and environmental variables (4). An important aspect of this model is that it recognizes the theoretically bidirectional nature of the inter-relationship between some of the measures within the framework (4). For instance, pain may lead to a worse emotional or psychological evaluation, or psychological problems may cause a worse evaluation of pain. In relation to this, there is some evidence to support HRQL being an attitude (5) or a personality characteristic (6), and consideration should be given to what Deiner has described as the "bottom-up" and "top-down" perspectives concerning the generation of well-being (7). The "bottom-up" theory is based on the more conventional idea that well-being/HRQL is the product of a person's circumstances, while the "top-down" approach suggests that well-being/HRQL is a more permanent phenomenon affecting the way people see the components of their lives.

In order to understand the correlates and determinants of HRQL, it is important to recognise those relationships which may or may not be bidirectional. At one extreme, while it is conceivable that HRQL is associated with age and gender, it is inconceivable that HRQL determines age or gender. If, however, one accepts the possibility that HRQL rating is actually an indicator of a personality trait, it is certainly conceivable that the relationship between HRQL and such variables as physical, role, emotional and social functioning could be in either or both directions (ie, we cannot be sure which are dependent and independent variables). The potential for bidirectionality in this situation is even stronger given the fact that these variables are all subject-rated. In view of this situation, when investigating the correlates/determinants of HRQL, it is important to distinguish between subject- (eg, psychological and symptomological) and non-subject-rated (eg, education and clinical status) variables.

With this current lack of understanding concerning the determinants and predictors of HRQL in mind, we were interested in studying HRQL among upper aerodigestive tract (UADT) cancer patients. In particular, we wanted to investigate the relative importance of sociodemographic and clinical variables as predictors of HRQL and to investigate those (subject-rated) aspects of HRQL (eg, physical, role, psychological and financial domains) which correlate most strongly with overall HRQL rating. In their review of studies of HRQL among UADT cancer patients, Gotay and Moore (1992) stated that, other than a few observational studies documenting the differences in various aspects of HRQL according to disease site or treatment modality, the determinants of HRQL in this population had not been systematically explored using validated instruments (8). Since that publication, the majority of studies investigating HRQL in this group of patients have been aimed at comparing the impacts of two or more treatments (eg.9-15), and have therefore concentrated upon the effects of treatment- and disease-related variables. Only a few have explored the role of non-clinical variables in relation to HRQL (16-18). The result is that, while we may have some idea as to which disease sites and treatment

modalities have an impact upon HRQL in this group, we have very little understanding of how wider psychosocial and demographic variables influence the HRQL of UADT cancer patients. The aim of our study was, therefore, to investigate the predictors of overall HRQL rating in a sample of upper aerodigestive tract (UADT) cancer patients, separately analyzing subject-rated and non-subject-rated (clinical and sociodemographic) variables. Although we recognize that it is important to look at predictors of all aspects of HRQL, this paper will report on predictors of overall HRQL rating (as measured by the "global" domain of the EORTC QLQ-C30) only, so as to concentrate upon the issue of the relative importance of clinical and sociodemographic variables.

Methodology

Subject selection

Subjects were recruited from the head and neck oncology clinic at the Jewish General Hospital, Montreal, Canada from July 1st 1995 to December 31st 1996. All patients having been diagnosed and treated for a squamous cell carcinoma of oral cavity sites (ICD9 141, 143-5) pharyngeal sites (ICD9 146-9) and laryngeal sites (ICD9 161) were eligible. All subjects read and signed a consent form prior to participating in the study.

HRQL measurement

HRQL data were collected using the EORTC QLQ-C30 (version 3), an instrument developed for use among patients diagnosed with any form of cancer (3) and validated among patients with head and neck cancers (19). The process of development and psychometric testing for this instrument involved simultaneous translation and validation in several (principally European) languages (20), an especially important feature for use in a city such as Montreal where two languages (French and English) are officially recognised and spoken. Data were gathered using the French and English versions of the EORTC QLQ-C30. The instrument is a self-complete questionnaire involving 30 questions in 15 separate domains. Twenty eight

of the questions have a 4-category Likert scale and two have a 7-category Likert scale. The questionnaire is divided into 5 multi-item functional scales (physical, role, cognitive, emotional and social), 3 multi-item symptom scales (fatigue, nausea and vomiting and pain) and 1 multi-item global HRQL scale, plus 5 single-item symptom measures and an item assessing perceived financial impact of the cancer and its treatment (5). The dependent variable in the study was the global domain (items 29 and 30).

Classification of sociodemographic and clinical data

These data were gathered through a combination of subject interviews and chart review. Tumour site was categorised as oral cavity (ICD9 141, 143-5), pharyngeal (ICD9 146-9) or laryngeal (ICD9 161). Disease stage was categorised into stages 1-4 according to each tumour site, as described by the American Joint Committee on Cancer (1992) (21). Treatment was categorised as surgery alone, radiotherapy alone and a combination of surgery and radiotherapy. Additional clinical data were collected concerning the presence or not of comorbidity at the time of HRQL evaluation, the period of time between the end of treatment (including recurrences) and HRQL evaluation, recurrent or new UADT cancer and subjects' dental status (edentulous v partially or fully dentate). Sociodemographic data concerning age, gender and living arrangements (living alone v living with others), education (high school education or less v college education or more) and employment (unemployed v employed or retired) status were also collected.

Statistical analyses

The dependent variable in this study was the global domain of the EORTC QLQ-C30 (global QL). This figure is calculated for each subject using the mean of two 1-7 Likert scale scores, which is then algorithmically converted to a 1-100 scale (with a higher score meaning a better HRQL evaluation) (3). The association between the global QL and study variables was assessed first through uni- and bivariate analyses, and then multivariate analyses. The univariate analyses testing the association between dichotomous variables and global QL score was performed

using a student's t-test. The univariate analyses testing the association between continuous variables and global QL score were performed using Pearson's correlation coefficient, and those testing associations between ordinal variables and global QL were performed using Spearman's rank correlation coefficient. In line with the introductory discussion, for the purposes of the multivariate analyses, the subject-rated functional, psychosocial and symptomological variables which make up the EORTC QLQ-C30 were tested as correlates of global QL separate from other clinical and sociodemographic study variables. The two multivariate models were created using general linear modeling with a backwards elimination technique; ie, the procedure was started with a model containing all possible variables, and the latter were eliminated one-by-one until the model contained only independent variables.

Results

The sample included 188 subjects with a mean age of 64.6 years and 71.3% of which were males. The full sample characteristics are summarized in Table 1. The variation in global QL score according to selected sociodemographic variables is shown in Table 2. The univariate analysis of the relationship between global QL score and age reveals no correlation between these variables (Pearson's correlation coefficient $r = -0.03$, $p = 0.65$). There were small, insignificant differences in global QL score by gender, cohabitation status and education, with males, those living with someone and the higher educated subjects having slightly higher global QL scores than their comparison groups. There was, however, a highly significant difference in global QL score in the work status category, with the unemployed having a significantly lower score than those with a job or retired.

The variation in global QL score according to selected clinical variables is demonstrated in Table 3. The majority of these clinical variables (including the presence or not of comorbidity, dental status, treatment modality and the presence or not of recurrent/new disease) had only small, insignificant effects on the mean global QL score. Among these clinical variables, the only one to

demonstrate any effect was disease site, in which there was a difference in mean global QL score of borderline significance at the 0.05 level between oral (the best mean score of the 3 categories) and pharyngeal (the worst mean score) cancers. The correlation between disease stage and global QL is small but significant ($r = -0.21$, $p = 0.005$). The univariate analysis of the variation in global QL score with time since treatment revealed a non-significant correlation coefficient of $r = 0.08$ ($p = 0.28$).

Table 4 demonstrates the inter-relationships between the EORTC QLQ-C30 domains. Global QL scores had statistically significant correlations with all of the EORTC QLQ-C30 domains and items, with correlation coefficients ranging from $r = 0.54 - 0.2$. The physical, fatigue and emotional domains had the strongest correlations with global QL score, while the weakest correlations involved the constipation and diarrhea single-item measures.

Following this series of uni- and bivariate analyses, two multivariate analyses were performed to elucidate the models which best explain the variation in global QL score in this sample. The first model included clinical and sociodemographic variables as independent variables, and the second model included all subject-rated domains and items within the EORTC QLQ-C30 instrument, other than the global QL domain itself. The results of the first analysis are shown in Table 5. This demonstrates that being unemployed (as opposed to employed or retired), being older, being female, being fully or partially dentate and being diagnosed with a later stage cancer were predictive of a worse global QL score, and that having an oral, as opposed to pharyngeal or laryngeal cancer was predictive of a better global QL score. This model, containing 6 variables, had an r^2 of 0.21.

The results of the second multivariate analysis are demonstrated in Table 6. It is evident that increased emotional, breathing, physical and financial problems, plus increasing pain and a lack of appetite were significant predictors of a worse global QL score. The complete model contained 7 of the 14 possible domains and items tested, and had an r^2 of 0.54.

Discussion

The aim of this study was to investigate the role of selected clinical, sociodemographic and subject-rated variables as predictors of global QL rating in a sample of UADT cancer patients. Treating the subject-rated, physical, functional, psychosocial and symptomological variables within the EORTC QLQ-C30 questionnaire separate from the non-subject-rated sociodemographic and clinical variables (as discussed in the introduction), we have constructed two multivariate models that explain some of the variation in global QL scores. Of the latter group of variables, unemployment, older age, female gender, being dentate and having more advanced stage disease were associated with a worse global QL rating, while oral, as opposed to pharyngeal or laryngeal, cancer was associated with a better global QL rating. Among the EORTC QLQ-C30 subject-rated domains and items tested, emotional, breathing, physical and financial problems, in addition to worse pain and a lack of appetite were associated with a worse global QL rating.

Sociodemographic variables

It is interesting to note that the three strongest predictors of global QL rating were non-clinical factors; unemployment, age and gender (see Table 5). Among other investigators examining potential non-clinical predictors of HRQL and similar outcome variables in a population of UADT cancer patients, Bjordal et al found that a higher level of education was related to greater life satisfaction (16), Long et al found that married patients, those living with somebody else and non-church goers had higher HRQL ratings (as measured by the FACT-G) (17) and Mathieson et al found that support from a family physician and male gender were predictive of a better HRQL rating (as measured by the FLIC) (18). The only finding common to our study and one of the latter group is, therefore, that male gender predicts a better overall HRQL rating. If this finding that males tend to report better HRQL than females is true for this and other patient groups, it lends support to the theory that HRQL rating is (at least partially) an indicator of a more permanent personality characteristic (7).

With the exception of the inclusion of employment status in our study, the sociodemographic variables in the aforementioned studies (16,17,18) were very similar to those in ours, thus highlighting the disparity in findings between the four studies. This could be explained by the fact that i) four different instruments were used to evaluate HRQL/life satisfaction; ii) two of the studies had a small sample size (17,18); and iii) the studies were performed in 3 different countries (Norway, USA and Canada). Nevertheless, it is clear that further work is required to clarify the situation. Of itself, the finding in our study that unemployment is associated with a lower global QL score is not surprising, while the age and gender associations will need confirmation and further exploration before being accepted as predictors of global QL in this group.

Clinical variables

With respect to the clinical variables tested, the findings of Bjordal et al, that pharyngeal site predicted lower life satisfaction (16), and Mathieson et al, that laryngeal site predicted lower HRQL ratings (18) were in indirect agreement with our finding that oral site was predictive of a better global QL rating. Contrary to our findings, Long et al found that site was not predictive of HRQL rating, even though they used a general HRQL instrument similar to the EORTC QLQ-C30 (the FACT-G). However, they found that laryngeal cancer patients scored better than other UADT cancer patients on a head and neck cancer specific module (17).

Another important clinical variable that one would expect to be associated with HRQL rating is disease stage. In agreement with our results, Mathieson et al found more severe disease stage to be predictive of worse HRQL (18). However, Bjordal et al found no such association (16) and Long et al did not test it (17). Closely related to disease stage and site is treatment modality, which was not associated with global QL in our study. However, we recognise that one of the reasons why treatment modality did not remain within our multivariate model may have been due to the crudeness of our categorisation of treatments. Having said that, cancer site and stage, both of which lie higher up the theoretical aetiological pathway

whose outcome is patient HRQL, remained within our multivariate model. It may be, therefore, that the interaction between these three variables is such that treatment modality is forced out of the multivariate model.

Looking at the effect of treatment modality on HRQL in other studies, the evidence is not clear. Using a series of univariate analyses in a cross-sectional study of post-treatment HRQL among 135 oral cancer patients, Schliephake et al found tumour size, defect location, type of soft tissue reconstruction and the resection of mandibular bone to be significantly associated with FLIC score (12). Following multivariate logistic regression to determine the role of surgical variables in predicting FLIC score, the only variable which remained significant was the need for mandibular resection as part of the ablative treatment, which was associated with a worse HRQL (12). Using multivariate regression to predictors of physical health in a sample of 204 UADT cancer patients, Bjordal et al found that those receiving pre-operative radiotherapy reported a worse physical health (16). However, in a small sample of laryngeal cancer patients treated with three different approaches, List et al could find no differences in overall HRQL rating between treatment modality or with time (14). And finally, comparing HRQL ratings using the EORTC instrument, in a sample of oral and pharyngeal cancer patients treated with external beam radiotherapy or external beam radiotherapy plus brachytherapy, Hammerlid et al could find no differences between the two groups over a 12 month period (15).

With regard to other clinical variables, no other study has assessed the effect of dental status on HRQL rating. On initial examination, our finding that having one or more teeth predicted a worse global QL rating may run counter to the expectation that most people would want to retain as much of their anatomy (including their teeth) as possible, and that loss thereof would tend to diminish HRQL. However, this reduction in global QL rating associated with having teeth is probably a reflection of the well-recognised oral problems (eg, radiation caries) irradiated dentate UADT cancer patients suffer.

Among those studies using the EORTC instrument to compare the

impact of various treatment- or disease-related variables on HRQL in samples of UADT cancer patients, Jones et al found that patients with recurrent disease rated their HRQL significantly worse than other treatment groups (9). Neither the presence or absence of recurrent or new UADT disease, nor the presence or absence of comorbidity had a significant effect upon global QL rating in our study. Comparing the HRQL ratings of two groups of UADT cancer patients 7-11 years following their treatment with conventional or hyperfractionated radiotherapy, Bjordal et al found that the latter group had a significantly better global QL rating than those treated with conventional radiotherapy (10). Comparing the mean global QL ratings of two samples of UADT cancer patients 1-6 years and 7-11 years following therapy, Bjordal et al reported no difference (11). Similarly, neither our study nor those of List et al (14), Long et al (17) and Mathieson et al (18) found time to be an explanatory factor for HRQL. This runs contrary to intuition and clinical experience which says that, at least among non-terminal cases, patients' symptoms and physical and functional problems improve with time after treatment. Herein lies a good example of the problems of making generalisations about population groups using an instrument which measures a dynamic and individual phenomenon (22). That is, HRQL instruments may demonstrate within-subject changes over time, but are less likely to show within-group changes over time because of the range of baseline values and the different ways in which individuals perceive change. It is interesting to note that, using the Functional Living Index-Cancer (FLIC) instrument in a sample of oral cancer patients, Schliephake et al (13) demonstrated a significant improvement in HRQL during the 6-12 month period following cancer therapy. Their finding time since treatment to be related to HRQL rating, while other studies have not, could be explained by several factors: i) different HRQL rating technique - while the contents of the FLIC and EORTC may be similar, the overall HRQL rating from the former is a product of all the domains and items within the instrument (23), while the same rating in the EORTC comes from a specific domain separate from all others (3); ii) the study of Schliephake et al included a more uniform group of patients - oral

cancer only; and iii) their finding of a relationship was the product of a univariate analysis.

Taking the sociodemographic and clinical variables together, it is interesting to note that in the univariate analyses in our study, age, gender and dental status were not significantly related to global QL rating, although they were in the multivariate model. Age was probably confounded by stage in that older patients are more likely to be diagnosed with early stage disease (24), gender may have been confounded by employment status in that no women described themselves as unemployed and dental status could have been confounded by disease stage in that those with later stage cancers will have received more radical treatment and are subsequently more likely to have lost all their teeth.

Looking at the non-subject-rated model (containing sociodemographic and clinical variables only) as a whole, our finding that this model explained 21% of the variation in global QL score differs considerably with the only other studies to use multivariate analysis in this way. Bjordal et al found that cancer site and education explained only 9% of the variation in life satisfaction, and that cancer site, education and treatment modality explained only 7% of the variation in physical health score in their sample of 204 UADT cancer patients. They concluded that "clinical and sociodemographic variables were poor predictors of patients' responses" (16). Matheison et al, however, found that family physician support, disease stage, disease site and gender explained 68% of the variation in their sample of 44 UADT cancer patients (18).

Finally, having discussed the likelihood that clinical and sociodemographic variables are associated with HRQL in UADT cancer patients, we also need to recognise the possible directions of these associations. In this respect, the associations between global QL rating and age and gender, can only be in one direction (with global QL dependent) and the associations between global QL rating and unemployment, dental status, disease site and stage seem highly likely to be unidirectional (with global QL rating dependent).

Subject-rated variables

With respect to the relationship between the subject-rated EORTC QLQ-C30 variables and overall global QL rating, the findings of the correlation matrix analysis (Table 4) are consistent with those of Aaronson et al in a sample of lung cancer patients (3) and Bjordal and Kaasa with a sample of head and neck cancer patients (19). Our multivariate model suggests those subject-rated domains which have a "statistically independent" association with global QL rating (namely, emotional, breathing, physical, financial, pain and appetite), and hence those components of the HRQL construct making the more important contribution to its rating. However, the dynamic inter-relationships between the components of an HRQL instrument have to be recognised, so we cannot automatically assume, for instance, that emotional problems cause a reduction in global QL rating; the relationship could be in the opposite direction. Furthermore, in the univariate analysis, the role, fatigue and physical domains had good correlations with global QL, but only the latter remained in the model. Looking at the inter-relationships between the physical, fatigue and role domains, we find correlation coefficients of $r = 0.53\text{-}0.69$, suggesting good correlations. Clearly the fatigue and role domains are an important element of any HRQL instrument but they do not remain in our multivariate model because of their interaction with the physical domain. Finally, looking at the multivariate model as a whole, it predictably explained more of the variance (54%) in global QL score than the non-subject-rated variables model. However, these figures demonstrate that there is a large proportion of the variance in HRQL which remains to be explained.

In making these observations concerning this study's findings, the limitations of the study design need to be recognized. Most importantly, it was a cross-sectional study, meaning that we cannot draw any conclusions about the direction of the associations found in the study. In addition, the crude nature of some of the sociodemographic and clinical data and the lack of understanding of the

determinants of HRQL described in the introductory section further limit inferences from the study findings. Nevertheless, the findings of this study suggest that three sociodemographic variables (age, gender and employment status), and three clinical variables (disease site, disease stage and dental status) are independently associated with overall HRQL rating in this sample of Canadian UADT cancer patients; the three sociodemographic variables explaining the larger proportion of the variation. Of the three clinical variables, disease site and stage are, at least indirectly, in agreement with the few previous studies making this sort of analysis among UADT cancer patients, while the association between dental status and global QL is a new finding. If dental status is confirmed as a predictor of HRQL in this group, this has important implications for their rehabilitation therapy. Finally, we have also found six subject-rated variables with a significant, association with overall HRQL rating. That is not to say that the other variables comprising the EORTC QLQ-C30 instrument are not important, but it begins to suggest those with a more direct association with global QL rating, at least among UADT cancer patients. It is, however, important that further work with this population is performed to confirm or otherwise the findings of this study and improve our understanding of those factors which determine a patient's evaluation of his/her HRQL.

Table 1. Sample characteristics

Variable	Category	N (%)	Mean (range)
Age			64.6 (34-91) years
Gender	male	134(71.3%)	
	female	54(28.7%)	
Living arrangements	living with someone	152(80.9%)	
	living alone	36(19.1%)	
Education	high school or less	120(67.6%)	
	>high school	68(36.2%)	
Work status	retired	121(64.4%)	
	employed	47(25.0%)	
	unemployed	20(10.6%)	
Comorbidity	absent	127(67.6%)	
	present	61(32.4%)	
Cancer site	oral	76(40.4%)	
	pharyngeal	56(29.8%)	
	laryngeal	56(29.8%)	
Disease stage	I	57(30.3%)	
	II	38(20.2%)	
	III	40(21.3%)	
	IV	53(28.2%)	
Treatment	radiotherapy alone	62(33.0%)	
	surgery alone	48(25.5%)	
	combined therapy*	78(41.5%)	
Recurrent/new disease	absent	163(86.7%)	
	present	25(13.3%)	
Dental status	(partially) dentate	112(59.6%)	
	edentate	76(40.4%)	
Time between last treatment and the HRQL evaluation			22.2 (1-168) months

*Combined therapy - patient has received both surgery and radiotherapy

Table 2. Variation in global QL score according to selected sociodemographic variables

Variable	Category	Mean global QL score	95% CI
Age*	(continuous variable)	r = -0.03	p = 0.65
Gender	male	64.5	60.0-69.0
	female	60.5	54.3-66.7
Cohabitation	living with someone	64.2	60.2-68.2
	living alone	60.2	51.0-69.3
Education	high school or less	62.5	58.0-67.0
	>high school	65.0	58.8-71.2
Work status	retired	63.5	57.0-69.3
	employed	67.9	63.8-71.8
	unemployed**	43.8	33.7-53.8

* Pearson's correlation coefficient (r, with p value) for the relationship between age and HRQL score.

**Significant difference (using student's T test at the p=0.05 level) in global QL score between unemployed subjects v. retired or employed subjects.

Table 3. Variation in global QL score according to selected clinical variables

Variable	Category	Mean global QL score	95% CI
Comorbidity	absent	63.8	59.3-68.3
	present	62.5	56.7-68.3
Dental status	dentate	61.5	57.2-66.3
	edentate	65.0	61.0-71.3
Disease site	oral	68.2	62.6-73.7
	pharyngeal*	56.5	50.3-62.6
	laryngeal	63.7	56.5-70.8
Treatment	radiotherapy alone	59.8	53.3-66.3
	surgery alone	68.2	60.7-75.7
	combined therapy	63.2	58.0-68.3
Recurrent/ new disease	absent	63.5	59.5-67.5
	present	62.7	52.8-72.5
Disease stage**	(ordinal variable)	$r = -0.21$	$p < 0.01$
Time since last treatment***	(continuous variable)	$r = 0.08$	$p = 0.28$

*Borderline significance (using student's T test) at the p=0.05 level for the difference in global QL score between oral and pharyngeal cancers.

**Spearman's rank correlation coefficient (r, with p value) for the relationship between disease stage and global QL score.

***Pearson's correlation coefficient (r, with p value) for the relationship between time since last treatment and global QL score.

Table 4. Inter-correlation between EORTC QLQ-C30 multi-item domains

	Physical	Role	Emotional	Cognitive	Social	Fatigue	NV	Pain
Global	-0.51	-0.47	-0.54	-0.43	-0.32	-0.54	-0.33	-0.48
Physical		0.53	0.27	0.35	0.28	0.69	0.26	0.36
Role			0.41	0.32	0.52	0.54	0.35	0.46
Emot'l				0.47	0.49	0.44	0.3	0.41
Cog've					0.28	0.35	0.25	0.35
Social						0.4	0.46	0.34
Fatigue							0.4	0.49
NV								0.41

NB. i) All figures quoted are Pearson's correlation coefficient, r;

ii) Values for Pearson correlation coefficients between global QL score and single-item measures are not shown and are as follows: global QL/breathing r = -0.43; global QL/difficulty sleeping r = -0.36; global QL/appetite r = -0.45; global QL/constipation r = -0.2; global QL/diarrhea r = -0.21; and global QL/finances r = -0.32;

iii) All correlation coefficients quoted in the table and in point ii) were significant at the $p<0.01$ level.

Table 5. Multivariate model best explaining the variation in global QL score with clinical and sociodemographic factors as independent variables

Variable	Parameter estimate	Standard error	p
Unemployed	1.9	0.38	0.0001
Age	-0.04	0.01	0.0003
Female gender	-0.57	0.23	0.0170
Dentate	-0.57	0.23	0.0147
Stage	-0.21	0.1	0.0294
Oral cancer site	0.47	0.24	0.0483
Constant	8.22		

NB. F value for model = 5.117; $p<0.0001$; $r^2 = 0.21$

Table 6. Multivariate model best explaining the variation in global QL score using EORTC QLQ-C30 domains as independent variables

Variable	Parameter estimate	Standard error	p
Emotional problems	-0.54	0.14	0.0003
Breathing problems	-0.35	0.11	0.0017
Physical problems	-0.54	0.23	0.0194
Financial problems	-0.24	0.1	0.0211
Pain	-0.3	0.13	0.0269
Appetite problems	-0.24	0.11	0.0298
Social problems	0.23	0.13	0.0850
Constant	8.45		

NB. F value for model = 14.615; $p<0.0001$; $r^2 = 0.54$

References

1. Hunt SM (1997): *The problem of quality of life.* Qual Life Res 6:205-12

2. Wilson IB and Kaplan S (1995): *Clinical practice and patients' health status: How are the two related?* Med Care 33(1):AS209-14

3. Aaronson NK et al (1993): *The European Organization for Research and Treatment of Cancer QLQ-C30: A quality of life instrument for use in international clinical trials in oncology.* JNCI 85:365-76

4. Wilson IB and Cleary PD (1995): *Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes.* JAMA 273(1):59-65

5. Guttman and Levy (1982): *On the definition and varieties of attitude and well-being.* Soc Ind Res 10:159-74

6. Ormel J (1983): *Neurotism and well-being inventories: measuring traits or states?* Psychological Medicine 13:165-76

7. Deiner E (1984): *Subjective well-being.* Psychological Bulletin 95:542-75

8. Gotay CC and Moore TD (1992): *Assessing quality of life in head and neck cancer.* Qual Life Res 1:5-17

9. Jones E, Lund VJ, Howard DJ, Greenberg MP and McCarthy M (1992): *Quality of life of patients treated surgically for head and neck cancer.* J Laryngol Otol 106:238-42

10. Bjordal K, Kaasa S and Mastekaasa A (1994): *Quality of life in patients treated for head and neck cancer: A follow-up study 7-11 years after radiotherapy.* Int J Radiation Oncology Biol Phys 28(4):847-56

11. Bjordal K, Freng A, Thorvik J and Kaasa S (1995): *Patient self-reported and clinician-rated quality of life in head and neck cancer patients: a cross-sectional study.* Oral Oncol, Eur J Cancer 31B(4):235-41

12. Schliephake H, Neukam FW, Schmelzeisen R, Varoga B and Schneller H (1995): *Long-term quality of life after ablative intraoral tumour surgery.* Journal of Cranio Maxillo-facial Surgery 23:243-9

13. Schliephake H, Ruffert K and Schneller H (1996): *Prospective study of the quality of life of cancer patients after intraoral tumour surgery.* J Oral Maxillofac

Surg 54:664-9

14. List MA, Ritter-Sterr CA, Baker TM, Colangelo LA, Matz G, Paoloski BR and Logemann JA (1996): *Longitudinal assessment of quality of life in laryngeal cancer patients.* Head & Neck 18:1-10

15. Hammerlid E, Mercke C, Sullivan M and Westin T (1997): *A prospective quality of life study of patients with oral or pharyngeal carcinoma treated with external beam irradiation with or without brachytherapy.* Oral Oncology 33(3):189-96

16. Bjordal K, Mastekaasa A and Kaasa S (1995): *Self-reported satisfaction with life and physical health in long-term cancer survivors and a matched control group.* Oral Oncol, Eur J Cancer 31B(5):340-5

17. Long SA, D'Antonio LL, Robinson EB, Zimmerman G, Petti G and Chonkich G (1996): *Factors related to quality of life and functional status in 50 patients with head and neck cancer.* Laryngoscope 106:1084-8

18. Matheison CM, Logan-Smith LL, Phillips J, Macphee M and Attia EL (1996): *Caring for head and neck oncology patients. Does social support lead to better quality of life?* Can Fam Physician 42:1712-20

19. Bjordal K and Kaasa S (1992): *Psychometric validation of the EORTC core quality of life questionnaire, 30-item version and a diagnosis-specific module for head and neck cancer patients.* Acta Oncologica 31(3):311-21

20. Aaronson NK, et al (1992): *International quality of life assessment (IQOLA) project.* Qual Life Res 1:349-51

21. American Joint Committee on Cancer: *Manual for Staging of Cancer.* 4th Ed'n. J.B.Lippincott Co., Philadelphia, 1992

22. Allison PJ, Locker D and Feine JS (1997): *Quality of Life: A dynamic construct.* Soc Sci Med 45(2):221-30

23. Schipper H, Clinch J, McMurray A and Levitt M (1984): *Measuring the quality of life of cancer patients: The functional living index-cancer: development and validation.* J Clin Oncol 2:472-83

24. Allison PJ, Franco EL, Black M and Feine JS (1998): *The role of professional diagnostic delays in the prognosis of upper aerodigestive tract cancer.* Oral Oncology In press

NB. End of manuscript

3.10 Summary of the findings of the investigation of health status among UADT cancer patients

It is evident that, while the evaluation of health status in health care settings has changed enormously over the previous 20 years, there are manifestly major disagreements over what domains are most appropriate to measure, how to standardize a dynamic phenomenon and how to interpret the findings of health status evaluations. The disagreement over which domains to assess is somewhat reduced when the discussion concerns the evaluation of health status in a patient population with a specific disease (in the case of this research, UADT cancer). However, even in a more homogenous population such as this, it is evident from the review in section 3.5 that inter-instrument content differences do exist. Once again, the evidence of section 3.5 suggests that while two instruments emerge as the most appropriate and better validated ones for UADT cancer patients (the FACT and EORTC instruments), both have their problems and neither has a fully validated H&N module. The choice of the EORTC instrument in this research was due to the fact that it had already been validated in French, in addition to English, thus making it more appropriate for research in a patient population where the two languages are spoken.

In addition to the equivocal nature of health status evaluation, the findings of this section of the research need to be considered in light of other methodological problems with the study. Firstly, the study design was cross-sectional in nature and so no causal inferences may be made concerning any of the associations found as a result of the statistical analyses. Secondly, the crude nature of some of the independent variables limits further any inferences drawn from finding, or indeed not finding, associations between them and health status ratings. For instance, one would expect treatment modality to be associated with health status rating. However, the crude nature of its categorization (surgery only, radiotherapy only and combination therapy), combined with the fact that it is intimately related to disease stage and, to a lesser a extent, disease site, both of which were significant predictors of health

status, were probably important factors contributing to the finding of no such association in this study. Similarly, the finding of a significant independent association between dental status and health status, while interesting and plausible, needs to be both confirmed and further explored for an explanation. Do dentate people rate their health status worse than their edentulous peers because of the well-recognized post-radiotherapeutic dental problems UADT cancer patients suffer? Or do edentulous people rate their health status better than dentate peers because they have already lost a substantial element of their oral anatomy (all their teeth) and are consequently perhaps better able to adapt to the post-therapeutic problems suffered by this group? The data in this study do not differentiate between those who were edentulous prior to their cancer treatment, those who became edentulous as a direct result of therapy (ie remaining teeth removed as part of ablative surgery) and those who became edentulous as an indirect result of ablative therapy (ie, remaining teeth removed following radiotherapy-related caries, or removed to facilitate prosthetic rehabilitation). The data also do not indicate the period of time people had been edentulous. Thus, this research has an interesting and plausible finding, however, the reason is unclear. Similar criticisms of the nature of the cohabitation, education and comorbidity status variables need to be recognised.

Having admitted those problems with the research methodology, all the findings were plausible and some were confirmations of those in previous studies. Although it is not immediately evident why, the finding that women had a significantly worse health status rating than men was a confirmation of the finding in one previous study. The finding that older age predicts a worse health status could be explained by older patients being less able to adapt to post-therapeutic problems, and the finding that unemployed people had a worse health status rating is not surprising. Similarly one would expect (indeed, given that it is designed as a prognostic indicator, hope) disease stage to be a predictor of health status and it is not surprising that disease site is associated with this variable. Future work should concentrate on more specific associations between various sociodemographic,

clinical and disease-related variables and the various aspects (domains) of health status.

SECTION 4

AN INVESTIGATION OF PROGNOSTIC DETERMINANTS AMONG UPPER AERODIGESTIVE TRACT CANCER PATIENTS: SUMMARY, CONCLUSIONS AND IMPLICATIONS FOR FUTURE WORK

4.1 Summary of thesis research

The central theme of this research has been an investigation of the influence of diagnostic delays upon disease stage and the subsequent role of disease stage as a predictor of post-therapeutic health status in a sample of UADT cancer patients. The association between diagnostic disease stage and patient survival is well recognised for UADT and other cancers. However, the association between disease stage and post-therapeutic health status is less clear. The idea of the research was, therefore, to investigate whether diagnostic delays lead to patients being diagnosed with a later disease stage, and whether the latter subsequently predicts poorer post-therapeutic health status.

The findings suggest that health care professional delays make a significant contribution to increasing the risk for late stage disease at diagnosis, and that the cancer site and patient age also independently affect the risk for late stage disease, with oral cancer patients and older patients having reduced risk for late stage disease. The finding that patient delays were not associated with disease stage at diagnosis was in common with previous studies. However, it is the view of the author that this finding is due to the invalid nature of patient delay data, the insufficient study power and the fact that rate of tumour growth could not be controlled for sufficiently with the available data.

The significance of these findings is important. Evidently, interventions designed to reduce professional delays need to be developed and evaluated, while recognizing the fact that for primary health care professionals, UADT cancers are

rare and often present with apparently innocuous signs or symptoms. The fact that older age patients are at reduced risk for late stage disease could be explained by two phenomena: i) genetically - patients in whom the disease onset is at a younger age have it in a more aggressive form than those in whom the disease onset is at an older age; or ii) rate of consumption of aetiological factors - patients in whom disease onset is at a younger age may have consumed cigarettes and/or alcohol at a higher rate and may have consumed them for a longer period than those in whom disease onset is at an older age. It could also be related to both phenomena. Unfortunately, the research reported in this thesis cannot answer this question because data concerning the consumption of aetiological factors and patients' genetic make-up were not collected. It also needs to be remembered that cancer is a disease with a long latent period and so defining "disease onset" is notoriously difficult. Finally, the finding that, compared to their pharyngeal and laryngeal cancer peers, patients with an oral cancer had reduced odds for being diagnosed with late stage disease could also be related to two phenomena: i) oral cancers have an intrinsic growth or development rate slower than pharyngeal and laryngeal cancers; or ii) because of their anatomic site, oral cancers tend to be detected earlier within the disease process than pharyngeal or laryngeal cancers. Once again, the answer may also be a combination of the two possibilities. The findings from the analysis of predictors of professional delays (in which oral cancers had reduced risk for such delays) tend to suggest that the second possibility is true. Nevertheless, this does not negate the involvement of tumour growth rate.

The analysis of predictors of health care professional delays suggest that comorbidity present at the time of presentation of UADT cancer symptoms increases the risk of such delay, while older age, higher education and oral cancer site reduce the risk. Once again, although these findings need to be confirmed, they are of significant importance in the planning of interventions to reduce diagnostic delays. In the large majority of cases, patients diagnosed with a UADT cancer are or have been long term smokers and/or alcohol drinkers, are of 50 years or older and come from lower socio-economic groups. As such, they are at high risk for many forms of

chronic ill-health which may be manifested in the form of a cancer, heart disease, diabetes, chronic pulmonary disease and many other diseases. It is somewhat troubling therefore, that such people are at risk for increased professional diagnostic delays. Similarly, the finding that people with a higher education are at reduced risk for professional delays is not surprising, but once again highlights the need for health care professionals to communicate better with people from lower social classes, a group who tend to make up the majority of their patients. The finding concerning age is, again, not surprising in that health care professionals understandably tend to have a higher index of suspicion of finding disease with the very young and the old.

The finding that disease stage is associated with post-therapeutic health status supports the original hypothesis that (professional) diagnostic delays increase the risk for late stage disease which subsequently leads to a worse health status. In that it is recognised as a good predictor of mortality, it is simultaneously unsurprising and reassuring (in terms of instrument validation) that disease stage predicts health status in this study. Of the other variables found to be associated with health status, that with the greatest potential for an intervention designed to improve health status in UADT cancer patients is dental status. Once this finding is confirmed, further research is required to explain this association and subsequently to develop diagnostic and prognostic models for the dental treatment of this group. Among the other variables found to be associated with health status in this study, unemployment is well recognised in many spheres of society to be undesirable and much work beyond the scope of health care is being performed to reduce unemployment. The influences on health status, of gender and age, reported in this thesis need first to be confirmed and then, if they prove to be definite, further investigated for an explanation. By their nature, however, they are likely to be predictors of psychosocial perceptions and behaviour which impact upon health status, and so more complex as the target of an intervention to improve health status. Finally, the finding that disease site is associated with health status in this study, is probably a reflection of this variable being closely related to disease stage and treatment modality. Oral cancers

tend to be diagnosed at an early stage and often require relatively minor surgery with subsequently minor impacts upon health status. Pharyngeal (and to a lesser extent laryngeal) cancer patients are much more likely to be diagnosed with late stage disease, which leads to more extensive therapy involving a combination of surgery and radiotherapy. The problem with disease site as a predictor of health status is that there is little to be done to change it, in that it is probably largely associated with the aetiology and genetic predisposition specific to each patient and, as such, is more appropriate for interventions aimed at disease prevention (although it would seem to result in a better prognosis, there seems little point in designing an intervention to promote UADT cancers to arise in the mouth rather than the pharynx). Having said that, it would be very useful to have a better understanding of why oral cancer patients rate their health status better than pharyngeal and laryngeal cancer patients.

4.2 Methodological limits of the research

With respect to that section of the project investigating the role of diagnostic delays in determining disease stage, the principal limitations concern the validity of the delay data and the size and representativeness of the sample. The evidence points to the professional delay data being valid, however, the validity of the patient delay data is more questionable. This observation is confirmed by the finding of no association between the latter and disease stage where one would expect to find one. As previously argued, it may be that the patient delay data is valid but that it is confounded by rate of tumour growth which was assessed only indirectly and hypothetically through disease site. Further research, using more direct methods to assess tumour growth rate, is required before this situation is resolved. A second issue of validity is raised by the categorization of disease site into oral, pharyngeal and laryngeal cancers. The first two categorizations appear to be valid while that of laryngeal cancers appears to be problematic. Future work, with a larger sample, should categorize laryngeal cancer into subglottic, glottic and supraglottic cancers. Another variable in which one could suspect problems of validity was that of

symptomology. This variable was not associated with either disease stage or professional delays, when one would expect it to be implicated somehow. The variable was categorized into four symptoms which were probably not specific enough to produce the expected association. For instance, the category "swelling" could have been divided into "intra-oral/pharyngeal" and "extra-oral/neck".

Finally in this section of the study, the issue of sample representativeness is important. The sample may be said to be representative of those patients referred to the relevant hospitals, however it is difficult to know whether referral patterns and service availability are similar to other hospitals in Quebec and Canada, let alone elsewhere in the world. However, in defense of the sample representativeness, most UADT cancer patients in Canada are treated in tertiary referral centres similar to the ones in the study and the findings were similar to those of a study in Brazil, a country with a totally different social and health care structure. Once again, this question can only be answered through further research in other regions of the country and elsewhere in the world.

With respect to the investigation of the predictors of health status, the principal limits to the methodology were the validity of the dependent variable(s) and the understanding of the determinants of health status which currently exists, plus the cross-sectional nature of the study design and the nature of some of the independent variables. The crude nature of some of the independent variables (eg, treatment modality, comorbidity, etc.) has already been discussed. The cross-sectional nature of the study means that no causal inferences can been drawn from the study findings. However, the main problems lie in the limited understanding of health status as it relates to this and other patient groups currently exhibited in the field of health care research. As discussed in sections 3.2-3.5 and manuscripts IV and V, there is only limited consensus on the definition of health status, QOL or HRQL and how to measure these phenomena, and the validity of those instruments which have been developed is consequently equivocal. Furthermore, the evident complexity of these phenomena as outcome variables leads to difficulties in the interpretation of results.

Having made these observations, the findings reported in this thesis are all plausible, although further confirmatory and explanatory work will be necessary.

4.3 Conclusions

Accepting the previously mentioned caveats concerning the research methodology and the fact that some of the findings need confirming in studies with larger sample sizes and involving different regions, the findings of the research reported in this thesis suggest that among UADT cancer patients:

1. Professional diagnostic delays of more than one month result in increased odds for late stage disease (OR: 2.28; 95%CI: 1.13-4.64) and later stage disease subsequently results in a worse post-therapeutic health status;
2. Patients with a cancer in the mouth have reduced odds for professional diagnostic delays greater than 1 month (OR: 0.31; 95%CI: 0.15-0.64), plus reduced odds for late stage disease (odds for late stage disease among pharyngeal compared to oral and laryngeal cancer patients OR: 9.29; 95%CI: 4.02-21.32) and a significantly higher health status rating than patients diagnosed with a pharyngeal or laryngeal cancer (all of these associations being independent);
3. Older age patients have reduced odds for professional diagnostic delays greater than 1 month (OR: 0.31; 95%CI: 0.15-0.64) and for late stage disease (OR: 0.45; 95%CI: 0.22-0.91) but have a significantly lower health status rating than younger patients;
4. Patients with higher education have reduced odds (OR: 0.45; 95%CI: 0.22-0.93) and those with comorbidity present at the time of presenting their symptoms have increased odds (OR: 2.84; 95%CI: 1.35-5.98) for professional diagnostic delays greater than 1 month; and
5. Patients who are female, those who are unemployed and those with one or more teeth have a lower health status rating than males, the employed or retired and edentate patients.

4.4 Original contributions of the work within this thesis

The original contributions of the research reported in this thesis to the scientific knowledge within the field of oncology, and that of UADT sites in particular, are both theoretical and practical in nature:

1. Manuscript I demonstrates the paucity of work concerning diagnostic delays in this population, both in terms of quantity and quality, and suggests a theoretical framework upon which future research in the field can be based;
2. In reporting the findings of the analysis of predictors of diagnostic disease stage, manuscript II largely confirms those of the one previous study using multivariate techniques but lends weight to the evidence concerning the effect of professional delays by finding a significant dose-response relationship;
3. The analysis of the predictors of professional diagnostic delays specifically, rather than diagnostic delays in general, reported in manuscript III is an original contribution which has findings with potential use in the planning of interventions aimed at reducing such delays; and
4. The findings reported in manuscript V use an original approach to analysis which attempts to clarify and categorize predictors of self-rated health status in this population, and which suggest the role of one variable (dental status) which has not previously been investigated and which could be used in the planning of rehabilitation in this population.

4.5 Implications for future research

The major implications of the work reported in this thesis for future research concern further investigation of the diagnostic process and of those factors which affect rehabilitation in UADT cancer patients. With respect to the former, a detailed study of the interaction of tumour, patient and professional behaviours and how they influence patient prognosis is required. Data collected should include all those discussed in manuscript I. This could lead to a more profound understanding of the diagnostic process for this group of patients and the generation of information useful

in the design and planning of interventions aimed at improving that process among both UADT cancer patients and other groups with similar chronic diseases.

With respect to improved rehabilitation for UADT cancer patients, the research reported in this thesis has highlighted the potential role of dental status as a determinant of health status. The potential for radiation-induced caries and related oral health problems for this group of patients are well documented. But, if confirmed, the research reported in this thesis suggests that "preventive" tooth extraction should be contemplated more often. Heretofore, there has been very little systematic examination of the determinants of tooth loss and its sequelae or of the decision-making process over pretherapeutic extractions in UADT cancer patients. Such work could lead to improved rehabilitation, and subsequently improved health status, for UADT cancer patients.

REFERENCES

1. *WHO international statistical classification of diseases, injuries and causes of death based on the recommendations of the 9th Revision Conference, 1975, vols I and II.* WHO, 1977

2. Berg JW: *Morphologic classification of human cancer.* In Schottenfeld D and Fraumeni JF (eds): *Cancer epidemiology and prevention.* WB Saunders, Philadelphia, 1982

3. La Vecchia C, Tavani A, Franceschi S, Levi F, Corrao G and Negri E (1997): *Epidemiology and prevention of oral cancer.* Oral Oncology 33(5):302-12

4. IARC monographs. *Tobacco smoking. IARC monographs for the evaluation of carcinogenic risks to humans*. Vol 37 IARC. Lyon 1986.

5. Winn DM, et al (1981): *Snuff-dipping and oral cancer among women in southern USA.* NEJM 304:745-749

6. Rosenfeld L and Callaway J (1963): *Snuff-dipper's cancer.* Am. J. Surgery 106:840-844

7. Wilkins SA and Volger WR (1957): *Cancer of the gingivae.* Surg. Gyncol. Obst. 105:145-152

8. Sorger K and Myrden JA (1960): *Verrucous carcinoma in the buccal mucosa in tobacco chewers.* Can. Med. Assoc. J. 83 1413-1417.

9. Main JHP. and Lecavalier DR (1988): *Smokeless tobacco and oral disease. A review.* Can. Dent. Assoc. J. 54(8):586-591.

10. Millar W.J. and van Rensberg S. (1984): *Use of chewing tobacco and snuff in the Northwest Territories 1982.* Chronic Disease Canada 4:54-56

11. IARC Monograph. *Tobacco habits other than smoking; betel quid and areca nut chewing; and some related nitrosamines.* IARC Vol 37 Lyon 1985.

12. Doll R and Peto R (1981): *The causes of cancer: Quantitative estimates of avoidable risks of cancer in the United States today.* JNCI 66:1191-1308

13. Rothman KJ and Keller AZ (1972): *The effect of point exposure to alcohol and tobacco on the risk of cancer of the mouth and pharynx.* J Chronic Diseases 25:711-16

14. Blot WJ, McLaughlin JK and Winn DM (1988): *Smoking and drinking in relation to oral and pharyngeal cancer.* Cancer Research 48:3282-87

15. Tuyns AJ, Estève J, Raymond L et al (1988): *Cancer of the larynx/hypopharynx, tobacco and alcohol: IARC International case-control study in Turin and Varese (Italy), Zaragoza and Navarra (Spain), Geneva (Switzerland) and Calvados (France).* Int. J. Cancer 41:483-91

16. Franceschi S, Talamini R, Barra S et al (1990): *Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx and esophogus in Northern Italy.* Cancer Research 50:6502-07

17. Elwood JM, Pearson JCG, Skippen DH and Jackson SM (1984): *Alcohol, smoking, social and occupational factors in the aetiology of cancer of the oral cavity, pharynx and larynx.* Int. J. Cancer 34:603-12

18. Franco EL, Kowalski LP, Olivieira BV et al (1989): *Risk factors for oral cancer in Brazil: a case-control study.* Int. J. Cancer 43:992-1000

19. Rothman KJ, Cann CI and Fried MP (1989): *Carcinogenicity of dark liquor.* Am. J. Public Health 79:1516-20

20. Marshall JR and Boyle P (1996): *Nutrition and oral cancer.* Cancer Causes Control 7:101-11

21. Rossing MA, Vaughan TL and McKnight B (1989): *Diet and pharyngeal cancer.* Int. J. Cancer 44:593-97

22. Marshall J, Graham S, Mettlin C, Shedd D and Swanson M (1982): *Diet in the epidemiology of oral cancer.* Nutrition and Cancer 3:145-49

23. Winn DM, Zeigler RG, Pickle LW, Gridley G, Blot WJ and Hoover RN (1984): *Diet in the etiology of oral and pharyngeal cancer among women in the Southern United States.* Cancer Research 44:1216-22

24. McLaughlin JK, Gridley G, Block G et al (1988): *Dietary factors in oral and pharyngeal cancer.* JNCI 80:1237-43

25. Franceschi S, Bidoli E, Baron AE and La Vecchia C (1990): *Maize and risk of cancers of the oral cavity, pharynx and esophohus in North-Eastern Italy.* JNCI 82:1407-11

26. Franceschi S, Bidoli E, Baron AE et al (1991): *Nutrition and cancer of the oral*

cavity and pharynx in North-East Italy. Int. J. Cancer 47:20-25

27. Franceschi S, Munoz N, Bosch XF, Snijders PJF and Walboomers JMM (1996): *Human papillomavirus and cancers of the upper aerodigestive tract: a review of epidemiological and experimental evidence.* Cancer Epidemiology Biomarkers Prevention 5:567-75

28. Sugerman PB and Shillitoe EJ (1997): *The high risk human papillomaviruses and oral cancer: evidence for and against a causal relationship.* Oral Diseases 3(3):130-147

29. Bundgaard T, Wildt J, Frydenberg M, Elbrønd O and Nielsen JE (1995): *Case-control study squamous cell cancer of the oral cavity in Denmark.* Cancer Causes Control 6:57-67

30. Marshal JR, Graham S, Haughey PB et al (1992): *Smoking, alcohol, dentition and diet in the epidemiology of oral cancer.* Oral Oncology, Eur. J. Cancer 28B:9-15

31. Zheng T, Boyle P, Hu H et al (1990): *Dentition, oral hygiene and risk or oral cancer: a case-control-study in Beijing, People's Republic of China.* Cancer Causes Control 1:235-41

32. Stell PM and McGill T (1973): *Asbestos and cancer of the head and neck.* Lancet i:678

33. Newhouse ML and Berry G (1973): *Asbestos and laryngeal carcinoma.* Lancet ii:615

34. Benner SE, Lippman SM and Hong WK: *Cancer of the head and neck.* In Greenwald P, Kramer BS and Weed DL (eds): *Cancer prevention and control.* Marcel Dekker Inc., New York, 1995

35. Yang PC, Thomas DB, Daling JR and Davis S (1989): *Differences in the sex ratio of laryngeal cancer incidence rates by anatomic subsites.* J. Clin. Epidemiol. 42(8):755-758

36. Parkin DM, Pisani P and Ferlay J (1993): *Estimates of the worldwide incidence of eighteen major cancers in 1985.* Int. J. Cancer 54:594-606

37. Parkin DM, Muir CS, Whelan SL et al: *Cancer incidence in five continents, Vol VI.* IARC Scientific Publications no 120, Lyon, France, 1992

38. National Cancer Institute of Canada: *Canadian Cancer Statistics 1997.* Toronto,

Canada, 1997

39. Canadian Council of Cancer Registries, Health and Welfare Canada, Statistics Canada: *The making of the Canadian Cancer Registry: Cancer Incidence in Canada and its Regions, 1969 to 1988.* Ottawa 1993

40. Johnson NW: *The global view of the epidemiology of oral cancer.* In Johnson NW (ed): *Risk Markers for Oral Diseases, Volume 2. Oral Cancer. Detection of patients and lesions at risk.* Cambridge University Press, Cambridge, 1991

41. Ayiomamatis A. (1989): *The epidemiology of malignant neoplasia of the larynx in Canada: 1939-1984.* Clinical Otolaryngology 14:349-355

42. Davis S and Severson RK (1987): *Increasing incidence of cancer of the tongue in the United States among young adults.* Lancet ii:910-11

43. McFarlane GJ, Boyle P and Scully C (1987): *Rising mortality from cancer of the tongue in young Scottish males.* Lancet ii:912

44. Hakulinen T, Andersen AA, Malker B, Pukkala E, Schou G and Tulinius H (1986): *Trends in cancer incidence in the Nordic countries.* Acta Pathologica Scandanavica (suppl 188):94

45. Moller H (1989): *Changing incidence of cancer of the tongue, oral cavity and pharynx in Denmark.* J. Oral Pathology Medicine 18:224-29

46. Hill GB (1985): *Cancer in Canada: an epidemiological perspective.* Can. Family Physician 31:1219-1228

47. Devassa SS, et al (1987): *Cancer incidence and mortality trends among whites in the United States 1947-1984.* JNCI 79:701-770.

48. National Cancer Institute of Canada: *Canadian Cancer Statistics 1994.* Toronto, Canada, 1994

49. Parker SL, Tong T, Bolden S and Wingo PA (1996): *Cancer statistics, 1996.* CA Cancer J Clin 65:5-27

50. La Vecchia C, Lucchini F, Negri E, Boyle P and Levi F (1993): *Trends in cancer mortality, 1955-1989: Asia, Africa and Oceania.* Eur J Cancer 29A:2168-2211

51. La Vecchia C, Lucchini F, Negri E, Boyle P, Maisonneuve P and Levi F (1992): *Trends of cancer mortality in Europe, 1955-1989: I, Digestive sites.* Eur J Cancer

28(1):132-235

52. Franceschi S, Levi F and La Vecchia C (19992): *Decline in 5-year survival rates for cancer of head and neck.* Lancet 340:47

53. Pelletier G: *La survie reliée au cancer. Étude des nouveau cas déclarés au Québec au cours des années 1984,1985 et 1986.* Planification-Évaluation santé services sociaux, 1993

54. Boyle P, Macfarlane GJ, Blot WJ, Chiesa F, Lefebvre JL, Mano Azul A, de Vries N and Scully C (1995): *European School of Oncology Advisory Report to the European Commission for the Europe Against Cancer Programme: Oral carcinogenesis in Europe.* Oral Oncology, Eur. J. Cancer 31B(2):75-85

55. *Proceedings of the National Strategic Planning Conference for the Prevention and Control of Oral and Pharyngeal Cancer, August 7-9, 1996.* American Dental Association, Centers for Disease Control and Prevention and the National Institute of Dental Research/National Institutes of Health

56. Horowitz AM, Goodman HS, Yellowitz JA and Nourjah PA (1996): *The need for health promotion in oral cancer prevention and early detection.* J. Public Health Dentistry 56(6):319-30

57. American Joint Committee on Cancer: *Manual for Staging of Cancer.* 4th Ed'n. J.B.Lippincott Co., Philadelphia, 1992

58. Langdon JD, Rapidis AD, Harvey PW and Patel MF (1977): *STNMP - a new classification for oral cancer.* Br. J. Oral Surgery 15:49-54

59. Evans SJW, Langdon JD, Rapidis AD and Johnson NW (1982): *Prognostic significance of STNMP and velocity of tumour growth in oral cancer.* Cancer 49:773-76

60. Platz H, Fries R and Hudec M (1986): *Prognoses of oral cavity carcinomas.* Carl Hanser, Munich.

61. Spiro RH, Huvos AG, Wong GY, Spiro JD, Gnecco CA and Strong EW (1986): *Predictive value of tumour thickness in squamous cell carcinoma confined to the tongue and floor of the mouth.* Am. J. Surgery 152:345-50

62. Mohit-Tabatabal MA, Sobel HJ, Rush BF and Mashberg A (1986): *Relation of thickness of floor of mouth stage I and II cancers to regional metastasis.* Am. J. Surgery 152:351-3

63. Urist MM, O'Brien CJ, Soong SJ, Visscher DW and Maddox A (1987): *Squamous cell carcinoma of the buccal mucosa: analysis of prognostic factors.* Am. J. Surgery 154:411-14

64. Byrne M, Koppang HS, Lilleng R and Kjaerheim A (1992): *Malignancy grading of the deep invasive margins of oral squamous cell carcinomas has high prognostic value.* J. Pathology 166:375-81

65. Byrne M, Jenssen N and Boysen M (1995): *Histological grading in the deep invasive front of T1 and T2 glottic sqamous cell carcinomas has high prognostic value.* Virchows Arch. 427:277-81

66. Anneroth G, Batsakis J and Luna M (1987): *Review of the literature and a recommended system of malignancy grading in oral squamous cell carcinomas.* Scand. J. Dental Research 95:229-49

67. Austin D.F. *Larynx.* In Schottenfeld D and Fraumeni JF (eds): *Cancer epidemiology and prevention.* WB Saunders Co, Philadelphia 1982.

68. Begg AC, McNally NJ, Shrive DC and Karcher HA (1985): *A method to measure the duration of DNA synthesis and the potential doubling time from a single sample.* Cytometry 6:620-26

69. Begg AC (1995): *The clinical status of Tpot as a predictor? Or why no tempest in the Tpot!* Int. J. Radiation Oncology Biology Physics 32(5):1539-41

70. Lee JM and Bernstein A (1995): *Apoptosis, cancer and the p53 tumour suppressor gene.* Cancer Metastasis Reviews 14:149-161

71. Skuse GR (1996): *The search for prognostic genetic indicators of cancer recurrence.* Int. J. Radiation Oncology Biology Physics 34(2):1509-10

72. Maeda K, Chung YS Takatsuka S et al (1995): *Tumour angiogenesis and tumour proliferation as prognostic indicators in gastric carcinoma.* Br. J. Cancer 72:319-23

73. Horvak RE, Leek R Klenk N et al (1992): *Angiogenesis, assessed by platelet/endothelial cell adhesion-molecule antibodies, as indicator of node metastases and survival in breast cancer.* Lancet 340:1120-4

74. Penfold CN, Partridge M, Rojas R and Langdon JD (1996): *The role of angiogenesis in the spread of oral squamous cell carcinoma.* Br. J. Oral maxillofacial Surg. 34(1):37-41

75. Beatrice F, Giusti U, Bisciari T et al (1996): *Microvessel count is predictive of patient's survival in laryngeal squamous cell carcinoma.* Int. J. Cancer 69:426-7

76. Facione NC (1993): *Delay versus help seeking for breast cancer symptoms: a critical review of the literature on patient and provider delay.* Social Science Medicine 36(12):1521-34

77. Johnston WD and Ballantyne AJ (1977): *Prognostic effect of tobacco and alcohol use in patients with oral tongue cancer.* Am. J. Surgery 134:444-47

78. Silverman S, Greenspan D and Gorsky M (1983): *Tobacco usage in patients with head and neck carcinomas: a follow-up study on habit changes and second primary oral/oropharyngeal cancers.* JADA 106:33-35

79. Browman GP, Wong G, Hodson I et al (1993): *Influence of cigarette smoking on the efficacy of radiation therapy in head and neck cancer.* NEJM 328:159-63

80. Pradier R, Gonzalez A, Matos E et al (1993): *Prognostic factors in laryngeal carcinoma.* Cancer 71:2472-76

81. Bundgaard T, Bentzen SM and Wildt J (1994): *The prognostic effect of tobacco and alcohol consumption in intr-oral squamous cell carcinoma.* Oral Oncol. Eur. J Cancer 30 B(5):323-8

82. Elwood JM and Gallagher RP (1985): *Factors in the early diagnosis of cancer of the oral cavity.* Can Med. Assoc. J. 133:651-6

83. Kowalski KP, Franco EL, Torloni H, Fava AS, Sobrinho JdeA, Ramos G, Oliveira BV and Curado MP (1994): *Lateness of diagnosis of oral and oropharyngeal cancer: factors related to the tumour, the patient and health professionals.* Oral Oncology, Eur. J. Cancer 30B(3):167-73

84. Maher EJ and Jeffries AF (1990): *Decision making in advanced cancer of the head and neck: variation in the views of medical specialists.* J. Royal Society Medicine 83:356-59

85. O'Sullivan B, Mackillop W, Gilbert R, Gaze M,. Lundgren J, Atkinson C, Wynne C and Fu H (1994): *Controversies in the management of laryngeal cancer: results of an international survey of patterns of care.* Radiotherapy Oncology 31:23-32

86. Stalpers LJ, Verbeek AL and van Daal WAJ (1987): *Results of radiotherapy and surgery for glottic cancer.* Cancer Treatment Review 14(2):1131-41

87. Farrow DC, Hunt WC and Samet JM (1992): *Geographic variation in the treatment of localized breast cancer.* NEJM 326:1097-1102

88. Nattinger AB, Gottlieb MS, Veum J, Yahnke D and Goodwin JS (1992): *Geographic variation in the use of breast-conserving treatment for breast cancer.* NEJM 326:1102-07

89. Franco EL, Dibb LL, Pinto DS, Lombardo V and Contesini H (1993): *Race and gender influences on the survival of patients with mouth cancer.* J. Clinical Epidemiology 46(1):37-46

90. Geiger HJ (1996): *Race and health care - an American dilemma?* NEJM 335(11):815-6

91. Howaldt H-P, Frenz M and Pitz H (1992): *Proposal for a modified T-classification for oral cancer.* J. Cranio-Maxillo-Facial Surgery 22:96-101

92. *The first 10 years of WHO.* WHO, Geneva, 1958

93. Goldsmith MB (1972): *The status of health status indicators.* Health Services Report 87:213

94. Elinson J. *Introduction.* In Elinson J and Seigman AE (eds) *Sociomedical health indicators.* Boywood, New York, 1979

95. Boorse C (1977): *Health as a theoretical concept.* Philosophy of Science 44:542-73

96. Mechanic D: *Symptoms, illness behaviour and help-seeking.* Prodist, New York, 1982

97. *International classification of impairments, disabilities and handicaps.* WHO, Geneva, 1986

98. Patrick D and Elinson J: *Methods of social research.* In Aiken LH and Patrick D (eds): *Applications of social science to clinical medicine and health policy.* Rutgers University Press, New Brunswick, 1979

99. Ware JE: *The assessment of health status.* In Aiken LH and Patrick D (eds): *Applications of social science to clinical medicine and health policy.* Rutgers University Press, New Brunswick, 1979

100. Locker D: *An introduction to behavioural science and dentistry.* Routledge,

London, 1989

101. Torrance GW (1987): *Utility approach to measuring health related quality of life.* J. Chronic Disease 40(6):593-600

102. Wilkin D, Hallam L and Doggett M-A: *Measures of need and outcome for primary health care.* Oxford University Press, 1992

103. Hunt SM (1997): *The problem of quality of life.* Quality of Life Research 6205-212

104. Wilson IB and Kaplan S (1995): *Clinical practice and patients' health status: How are the two related?* Medical Care 33(1):AS209-14

105. Wilson IB and Cleary PD (1995): *Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes.* JAMA 273(1):59-65

106. Cella DF (1992): *Quality of life: the concept.* J. Palliative Care 8(3):8-13

107. Cella DF and Tulsky DS (1993): *Quality of life in cancer: definition, purpose and method of measurement.* Cancer Investigations 11(3):327-36

108. Slevin ML, Plant H, Lynch D, Drinkwater J and Gregory WM (1988): *Who should measure the quality of life, the doctor or the patient?* Br. J. Cancer 57:109-12

109. Last JM: *A dictionary of epidemiology.* Third Edition. Oxford University Press, 1995

110. Streiner DL and Norman GR: *Health measurement scales. A practical guide to their development and use.* Second Edition. Oxford University Press, 1995

111. Kirshner B and Guyatt G (1985): *A methodological framework for assessing health indices.* J. Chronic Disease 38(1):27-36

112. Guyatt G, Walter S and Norman G (1987): *Measuring change over time: assessing the usefulness of evaluative instruments.* J. Chronic Disease 40(2):171-78

113. Gotay CC and Moore TD (1992): *Assessing quality of life in head and neck cancer.* Quality of Life Research 1:5-17

114. List MA, Ritter-Sterr CA and Lansky SB (1990): *A performance status scale for head and neck cancer patients.* Cancer 66:564-69

115. Karnofsky DA and Burchenal JH: *The clinical evaluation of chemotherapeutics in cancer.* In MacLeod CM (ed): *Evaluation of chemotherapeutic agents.* Columbia Press, New York, 1949

116. Harrison LB, Zelefsky MJ, Armstrong JG, Carper E, Gaynor JJ and Sessions RB (1994): *Performance status after treatment for squamous cell cancer of the base of tongue - a comparison of primary radiation therapy versus primary surgery.* Int. J. Radiation Oncology Biology Physics 30(4):953-57

117. Zelefsky MJ, Gaynor JJ, Kraus D, Strong EW, Shah JP and Harrison LB (1996): *Long-term subjective outcome of surgery plus postoperative radiotherapy for advanced stage oral cavity and oropharyngeal carcinoma.* Am. J. Surgery 171:258-62

118. Moore GJ Parsons JT and Mendenhall WM (1996): *Quality of life outcomes after primary radiotherapy for squamous cell carcinoma of the base of tongue.* Int. J. Radiation Oncology Biology Physics 36(2):351-54

119. List MA, Ritter-Sterr CA, Baker TM, Colangelo LA, Matz G, Pauloski BR and Logemann JA (1996): *Longitudinal assessments of quality of life in laryngeal cancer patients.* Head and Neck 18:1-10

120. List MA, D'Antonio LL, Cella DF, Siston A, Mumby P, Haraf D and Vokes E (1996): *The performance status scale for head and neck cancer patients and the functional assessment of cancer therapy - head and neck scale.* Cancer 77:2294-301

121. Cella DF, Tulsky DS, Gray G et al (1993): *The functional assessment of cancer therapy scale: Development and validation of the general measure.* J. Clinical Oncology 11:570-79

122. D'Antonio LL, Zimmerman GJ, Cella DF and Long SA (1996): *Quality of life and functional status measures in patients with head and neck cancer.* Arch. Otolaryngology Head and Neck Surgery 122:482-87

123. Long SA, D'Antonio LL, Robinson EB, Zimmerman G, Petti G and Chonkich G (1996): *Factors related to quality of life and functional status in 50 patients with head and neck cancer.* Laryngoscope 106:1084-88

124. Hassan SJ and Weymuller EA (1993): *Assessment of quality of life in head and neck cancer patients.* Head and Neck 15:485-96

125. Bergner M and Bobbit RA (1981): *The sickness impact profile: development and final revision of a health status measure.* Medical Care 19:787-805

126. Aaronson NK, Bullinger M and Ahmedzai S (1988): *A modular approach to quality of life assessment in cancer clinical trials.* Recent Results Cancer Research III:231-49

127. Aaronson NK, Ahmedzai S, Bullinger M et al for the European Organization for Research and Treatment of Cancer Study Group on Quality of Life (1993): *The EORTC QLQ-C30: A quality of life instrument for use in international clinical trials in oncology.* JNCI 85:365-76

128. Sprangers MAG, Cull A, Bjordal K, Groenvold M and Aaronson NK for the EORTC Study Group on Quality of Life (1993): *The EORTC approach to quality of life assessment: Guidelines for developing questionnaire modules.* Quality of Life Research 2(4):287-95

129. Bjordal K and Kaasa S (1992): *Psychometric validation of the EORTC core quality of life questionnaire, 30-item version and a diagnosis-specific module for head and neck cancer patients.* Acta Oncologica 31(3):311-21

130. Jones E, Lund VJ, Howard DJ, Greenberg MP and McCarthy M (1992): *Quality of life of patients treated surgically for head and neck cancer.* J. Laryngology Otology 106:238-42

131. Bjordal K, Ahlner-Elmqvist M, Tollesson E, Jensen AB, Razavi D, Maher EJ and Kaasa S for the EORTC Study Group on Quality of Life (1994): *Development of a EORTC questionnaire module to be used in quality of life assessments in head and neck cancer patients.* Acta Oncologica 33(8):879-85

132. Bjordal K, Kaasa S and Mastekassa A (1994): *Quality of life in patients treated for head and neck cancer: A follow-up study 7-11 years after radiotherapy.* Int. J. Radiation Oncology Biology Physics 28(4):847-56

133. Bjordal K and Kaasa S (1995): *Psychological distress in head and neck cancer patients 7-11 years after curative treatment.* Br. J. Cancer 71(3):592-97

134. Bjordal K, Freng A, Thorvik J and Kaasa S (1995): *Patient self-reported and clinician-rated quality of life in head and neck cancer patients: a cross-sectional study.* Oral Oncology, Eur. J. Cancer 31B(4):235-41

135. Hammerlid E, Bjordal K Ahlner-Elmqvist M et al (1996): *A prospective longitudinal quality of life study of patients with head and neck cancer.* Otolaryngology Head and Neck Surgery 116:666-73

136. Hammerlid E, Mercke C, Sullivan M and Westin T (1997): *A prospective*

quality of life study of patients with oral or pharyngeal carcinoma treated with external beam irradiation with or without brachytherapy. Oral Oncology 33(3):189-96

137. Schipper H, Clinch J, McMurray A and Levitt M (1984): *Measuring the quality of life of cancer patients: The functional living index - cancer: Development and validation.* J. Clinical Oncology 2(5):472-83

138. Schliephake H, Neukam FW, Schmelzeisen R, Varoga B and Schneller H (1995): *Long-term quality of life after ablative intra-oral surgery.* J. Cranio Maxillo-Facial Surgery 23:243-49

139. Schliephake H, Ruffert K and Schneller H (1996): *Prospective study of the quality of life of cancer patients after intra-oral tumour surgery.* J. Oral Maxillo-Facial Surgery 54:664-69

140. Mathieson CM, Logan-Smith LL, Phillips J MacPhee M and Attia EL (1996): *Caring for head and neck oncology patients. Does social support lead to better quality of life?* Can. Family Physician 42:1712-20

141. Rathmell AJ, Ash DV, Howes M and Nicholls J (1991): *Assessing quality of life in patients treated for advanced head and neck cancer.* Clinical Oncology 3:10-16

142. Browman GP, Levine MN, Hodson DI, Sathya J, Russell R, Skingley P, Cripps C, Eapen L and Girard A (1993): *A head and neck radiotherapy questionnaire: A morbidity/quality of life instrument for clinical trials of radiation therapy in locally advanced head and neck cancer.* J. Clinical Oncology 11:863-72

143. Bundgaard T, Tandrup O and Elbrønd O (1993): *A functional evaluation of patients treated for oral cancer.* Int. J. Oral Maxillofacial Surgery 22:28-34

144. Langius A, Björvell H and Lind MG (1994): *Functional status and coping in patients with oral and pharyngeal cancer before and after surgery.* Head and Neck 16:559-68

145. DeSanto LW, Olsen KD, Perry WC, Rohe DE and Keith RL (1995): *Quality of life after surgical treatment of cancer of the larynx.* Ann. Otology Rhinology Laryngology 104:763-69

146. Derogatis LR (1986): *The psychosocial adjustment to illness scale (PAIS).* J. Psychometric Research 30:77-91

147. Baker C (1995): *A functional status scale for measuring quality of life outcomes*

in head and neck cancer patients. Cancer Nursing 18(6):452-57

APPENDIX 1: Questionnaire for investigation of diagnostic delays.

1. Today's date _____
2. Subject initials _____
3. Code number _____
4. Date of birth _____
5. First language

   French(1) English(2) Other(3) _____
6. Gender

   Male(1) Female(2) _____
7. Marital status

   Married(1) Cohabiting(2) Single(3) Separated(4) Divorced(5) Widowed(6) _____
8. Cohabitation status

   Alone(1) Living with spouse/family(2) Living with others(3) _____
9. Highest level of education

   Less than compulsory school(1) Compulsory school(2) College(3) University(4)_____
10. Employment status

    Full time employment(1) Part time employment(2) Housewife/homemaker(3) Student(4)

    Retired(5) Unemployed(6) _____
11. What was the first sign/symptom you noticed?

    Pain(1) Swelling(2) Painless ulcer/lesion(3) Voice change(4) Other(5) _____
12. Who did you first consult concerning this sign/symptom?

    Family physician(1) Dentist(2) Specialist(3) _____
13. How long was the period between your first noticing the sign/symptom and consulting somebody about it? ___wks
14. How long was the period between this first consultation and your first consultation in this department? ___wks
15. Were you being treated for any other medical problem at the time you presented this sign/symptom? Yes(1) No(2) _____

16. Do you have any of your own teeth?

Yes(1) No(2) _____

17. Disease site

oral(1) pharyngeal(2) laryngeal(3) _____

18. T stage of tumour

T1(1) T2(2) T3(3) T(4) _____

19. N stage of cancer

N0(1) N1(2) N2(3) N3(4) _____

20. M stage of cancer

M0(1) M1(2) _____

21. Overall disease stage

I(1) II(2) III(3) IV(4) _____

APPENDIX 2: Questionnaire for investigation of health status.

1. Today's date _____
2. Subject initials _____
3. Code number _____
4. Date of birth _____
5. First language

   French(1) English(2) Other(3) _____
6. Gender

   Male(1) Female(2) _____
7. Marital status

   Married(1) Cohabiting(2) Single(3) Separated(4) Divorced(5) Widowed(6) _____
8. Cohabitation status

   Alone(1) Living with spouse/family(2) Living with others(3) _____
9. Highest level of education

   Less than compulsory school(1) Compulsory school(2) College(3) University(4)_____
10. Employment status

    Full time employment(1) Part time employment(2) Housewife/homemaker(3) Student(4)

    Retired(5) Unemployed(6) _____
11. Disease site

    Oral(1) Pharyngeal(2) Laryngeal(3) _____
12. T stage of tumour

    T1(1) T2(2) T3(3) T(4) _____
13. N stage of cancer

    N0(1) N1(2) N2(3) N3(4) _____
14. M stage of cancer

    M0(1) M1(2) _____
15. Overall disease stage

    I(1) II(2) III(3) IV(4) _____

16. Primary treatment

Surgery only(1) Radiotherapy only(2) Combination(3) _____

17. Date of diagnosis _____

18. Date of commencement of treatment _____

19. Date of completion of treatment _____

20. Recurrent or new UADT cancer

Yes(1) No(2) _____

21. Treatment for recurrence/new disease

Surgery only(1) Radiotherapy only(2) Chemotherapy(3) Combination(4) _____

22. Are you currently being treated for any other medical problem?

Yes(1) No(2) _____

23. Dental status at the time of diagnosis

Fully dentate(1) Partially dentate(2) Edentulous(3) _____

24. Current dental status

Fully dentate(1) Partially dentate(2) Edentulous(3) _____

25. Current prosthetic status

No prosthesis(1) Full dentures(2) Partial denture(s)(3) Combination(4) Obturator(5)_____

APPENDIX 3: EORTC QLQ-C30 and H&N37

FOR ADMINISTRATIVE USE ONLY
15941 Form 930, page 1 of 4
Step reported 67 |_|_|
Inst |_|_|_|_|
seq id |_|_|_|_|

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in your inititals: |_|_|_|_|
Your birthdate (Day, Month, Year): |_|_|_|_|_|_|
Today's date (Day, Month, Year): 66 |_|_|_|_|_|_|

		Not at All	A Little	Quite a Bit	Very Much
1.	Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2.	Do you have any trouble taking a long walk?	1	2	3	4
3.	Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4.	Do you need to stay in bed or a chair during the day?	1	2	3	4
5.	Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:		Not at All	A Little	Quite a Bit	Very Much
6.	Were you limited in doing either your work or other daily activities?	1	2	3	4
7.	Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8.	Were you short of breath?	1	2	3	4
9.	Have you had pain?	1	2	3	4
10.	Did you need to rest?	1	2	3	4
11.	Have you had trouble sleeping?	1	2	3	4
12.	Have you felt weak?	1	2	3	4
13.	Have you lacked appetite?	1	2	3	4
14.	Have you felt nauseated?	1	2	3	4
15.	Have you vomited?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
16. Have you been constipated?	1	2	3	4
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

Please go on to the next page

EORTC QLQ-H&N35

Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week	Not at all	A little	Quite a bit	Very much
31. Have you had pain in your mouth?	1	2	3	4
32. Have you had pain in your jaw?	1	2	3	4
33. Have you had soreness in your mouth?	1	2	3	4
34. Have you had a painful throat?	1	2	3	4
35. Have you had problems swallowing liquids?	1	2	3	4
36. Have you had problems swallowing pureed food?	1	2	3	4
37. Have you had problems swallowing solid food?	1	2	3	4
38. Have you choked when swallowing?	1	2	3	4
39. Have you had problems with your teeth?	1	2	3	4
40. Have you had problems opening your mouth wide?	1	2	3	4
41. Have you had a dry mouth?	1	2	3	4
42. Have you had sticky saliva?	1	2	3	4
43. Have you had problems with your sense of smell?	1	2	3	4
44. Have you had problems with your sense of taste?	1	2	3	4
45. Have you coughed?	1	2	3	4
46. Have you been hoarse?	1	2	3	4
47. Have you felt ill?	1	2	3	4
48. Has your appearance bothered you?	1	2	3	4

Please go on to the next page

During the past week:	Not at all	A little	Quite a bit	Very much
49. Have you had trouble eating ?	1	2	3	4
50. Have you had trouble eating in front of your family?	1	2	3	4
51. Have you had trouble eating in front of other people?	1	2	3	4
52. Have you had trouble enjoying your meals?	1	2	3	4
53. Have you had trouble talking to other people?	1	2	3	4
54. Have you had trouble talking on the telephone?	1	2	3	4
55. Have you had trouble having social contact with your family?	1	2	3	4
56. Have you had trouble having social contact with friends?	1	2	3	4
57. Have you had trouble going out in public?	1	2	3	4
58. Have you had trouble having physical contact with family or friends?	1	2	3	4
59. Have you felt less interest in sex?	1	2	3	4
60. Have you felt less sexual enjoyment?	1	2	3	4

During the past week:	No	Yes
61. Have you used pain-killers?	1	2
62. Have you taken any nutritional supplements (excluding vitamins)?	1	2
63. Have you used a feeding tube?	1	2
64. Have you lost weight?	1	2
65. Have you gained weight?	1	2