SYMPTOMS OF PATIENTS HOSPITALIZED BECAUSE OF MALIGNANCY: A COMPARISON OF THE PERCEPTIONS OF THE PATIENT, THE NEXT OF KIN, AND THE NURSE

BILL EATON

National Library of Canada

Acquisitions and Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada

Acquisitions et services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-42371-9

Canada

SYMPTOMS OF PATIENTS HOSPITALIZED BECAUSE OF MALIGNANCY: A COMPARISON OF THE PERCEPTIONS OF THE PATIENT, THE NEXT OF KIN, AND THE NURSE

by

BILL EATON

MD, CCFP

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree of Master of Science

Faculty of Medicine
Memorial University of Newfoundland
January 1999

St. John's Newfoundland

ABSTRACT

The aim of this descriptive study was to assess, using Cohen's Kappa statistic, the agreement between hospitalized patients with cancer and their closest family member, and between these patients and the nurse providing the bedside care, on the perceptions of cancer-related symptoms.

One hundred patients, hospitalized because of a malignancy, rated nine cancer-related symptoms using the Edmonton Symptom Assessment System, a collection of nine Visual Analogue Scales. Simultaneously, the closest family member and the nurse providing the bedside care rated the patients' symptoms using the same method. The aim of the study was to assess agreement, on these nine symptoms, between the patients and the family members, and between the patients and their nurses.

Family members demonstrated significant agreement (Cohen's Kappa > 0.35) with the patient in assessment of nausea, anxiety, drowsiness, tiredness, and appetite; but not for pain, depression, shortness of breath, or overall well-being. The nurses demonstrated significant agreement with the patient on the perception of nausea only.

Merging the nine reported similar studies on symptom agreement between patient and family with the present study there emerges no consistent evidence on agreement for pain, shortness of breath, well-being, anxiety, and appetite; while there is good evidence to suggest agreement on nausea, depression, tiredness, and drowsiness.

Similarly, reviewing the two studies involving cancer patients and nurses, there emerged no consistent evidence on the agreement for pain, nausea, tiredness, drowsiness, or appetite. While the previous studies did not assess shortness of breath, anxiety, and well-being, there was over-estimation by the nurses on the patient's ratings of depression.

CONCLUSIONS:

1) Family members agreed with the patients in their perceptions of five out of nine cancer-related symptoms: nausea, anxiety, drowsiness, tiredness, and appetite.

2) Family members did not agree with the patients on the symptoms of: pain, depression, shortness of breath, and well-being.

3) Nurses agreed with the patients in their perceptions of one out of nine (nausea) cancer related symptoms.

4) Family members were closer in their perceptions of the cancer patients' symptoms than were the nurses.

MASTER OF SCIENCE
(Clinical Epidemiology)

MEMORIAL UNIVERSITY OF NEWFOUNDLAND
St. John's, Newfoundland

TITLE: SYMPTOMS OF PATIENTS HOSPITALIZED BECAUSE OF MALIGNANCY: A COMPARISON OF THE PERCEPTIONS OF THE PATIENT, THE NEXT OF KIN, AND THE NURSE.

AUTHOR: Bill Eaton, MD, CCFP

SUPERVISOR: Dr. Graham Worrall

NUMBER OF PAGES: 112

ACKNOWLEDGEMENTS

I would like to thank all those patients and their families who participated in this study and especially those who completed the protocol. The illness of cancer can have an enormous and harsh effect on both those with the illness and those nearby, so much so, that to acquiesce to some third party inquiring about the state of their symptom perceptions must, in some, require an immense withdrawal from their bank of humanitarian largesse.

My supervisor for this thesis, Dr Graham Worrall, Associate Professor of Family Medicine at Memorial University, has been most helpful at every turn, offering up advice and wise counsel, helping focus my thoughts and energies, keeping the important issues in view, and giving freely of his time whenever requested. I offer thanks for his support and direction. Dr. Bill Bavington, Associate Professor of Community Health at Memorial University, has offered sage and timely advice on the writing-up of this thesis while Dr. Rob Foley, Assistant Professor of Medicine at Memorial University, helped with shrewd statistical suggestions.

Dr. Jennifer Hall, Assistant Professor of Family Medicine at Memorial University and Medical Director of the Palliative Care Program with the Health Care Corporation of St. John's, was instrumental in the formulation of the research issues of symptom assessment. She also assisted me by performing the second measurements of the raw data to assess the quality control of the data.

I was most fortunate for the opportunity to interact with a wonderful group of universally pleasant and personally helpful nurses. Their dedication to their patients and their comforting skill were marvellous and life-affirming to observe. The patients, doctors and the Health Care Corporation of St. John's are certainly most fortunate to have such caring and competent professionals.

I note with thanks the major assistance of the charge nurses who were able to budget time away from their regular duties to recruit subjects for this study: Shelly White of 6 E, St. Clares Hospital; and Helen Byrne, Daphne King, and Kathy-Ann Butt of 4 NA, Health Science Complex truly made this research happen. The Patient Care Managers of the three sites were, as always, facilitative and supportive. To Janet Templeton, Fatima Hewitt, and Joy Stuckless I offer grateful appreciation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

CHAPTER I

INTRODUCTION AND REVIEW OF THE LITERATURE

1.1 INTRODUCTION

Palliative, or end-of-life care; as described by Singer and MacDonald in a 1998 review of bioethics for clinicians[1]; contains 3 main elements: control of pain and other symptoms, discussions concerning the use of life-sustaining treatments, and support of those who are dying and their families. This 3-part framework was developed as a guide for physicians caring for the families and patients suffering from advanced cancer. The World Health Organization offers the following definition of palliative care: 'The active total care of patients whose disease is not responsive to curative treatment. Control of pain, of other symptoms, and of psychological, social and spiritual problems, is paramount. The goal of palliative care is achievement of the best quality of life for patients and their families. Many aspects of palliative care are also applicable earlier in the course of the illness in conjunction with anticancer treatment.'[2]

The burden of cancer-related suffering experienced by both patient and family is extensive. The incidence of cancer in Canada is rising, according to the National Cancer Institute of Canada[3], and is attributable mostly to the increasing age structure of the population. The risk of developing cancer increases with increasing age for both individuals and populations. In Canada there were approximately 129,000 new cases and 61,800 deaths from cancer in 1996. The symptoms of cancer are distressing with the most common symptoms listed in Table I. The prevalence, severity and clinical importance of these symptoms were reported in two studies done by a group of palliative care nurses and physicians working out of Cleveland, Ohio[4,5]. This group prospectively reviewed consecutive cancer patients who were referred to their service: one hundred, reported in 1991, were assessed for the presence of symptoms; and 1000, reported in 1995, were assessed for severity of symptoms. The authors discuss the paucity of research in this area.

1.2 BACKGROUND

Family members, besides experiencing their own issues during the cancer experience, suffer vicariously the distress of the cancer patient as together they live through this all too often, progressive and fatal illness. In 1994 a palliative care research group in Manitoba[6] studied 64 consecutive, consenting cancer in-patients and their families using a validated symptom distress scale. The authors found that family caregivers experience health effects, most notably deteriorations of mental health status and cognitive function, during the illness and after the death of their loved one. The authors note further that the physical burden of providing care and support to the ill patient also adds to the family stress. Using a questionnaire that was mailed out to 65 bereaved spouse care givers of recently deceased cancer patients, a nursing research group working out of California[7] in 1992, found, in the 38 respondents, that the spouses felt a lack of control over what was happening to the cancer patient and that although physical care giving was difficult at the time, on later reflection the most distress experienced by the family members was related to standing by or observing the deterioration of the cancer patient.

As the illness of cancer advances, and the number and intensity of symptoms increase; physicians may hear, either from the patient's nurse, during daily care discussions; or from the family members, during family meetings; diverse impressions about the patient's symptoms. As the patient becomes sicker and less able to communicate, family members and the nurses caring for the patient have, in the experience of the author, an increasing role in reporting on the patient's condition.

TABLE I : PREVALENCE AND SEVERITY OF SYMPTOMS IN ADVANCED CANCER: CLEVELAND, OHIO 1991 and 1995

SYMPTOM	PERCENT PREVALENCE 1991	PERCENT OF THOSE PATIENTS WHO SUFFERED THE SYMPTOM AND REPORTED IT AS SEVERE 1995
Pain	89	87
Weight Loss	58	44
Anorexia	55	67
Dyspnoea	41	46
Fatigue	40	70
Early Satiety	40	75
Constipation	40	60
Weakness	36	70
Lethargy	32	78
Nausea	32	66
Depression	31	71
Sleep Problem	28	50
Taste Change	26	77
Vomiting	25	72
Dry Mouth	n/a	27
Cough	n/a	12
Anxiety	n/a	11
Edema	n/a	11

Source: Curtis et al.[4], and Donnelly et al.[5]

1.3 RATIONALE AND OBJECTIVES OF STUDY

As a physician working in the field of Palliative Care, the author has been in situations where the patient reports one set and intensity of symptoms while the nurse reports another and the family members offer up still another.

When a distressed family member reports that the patient's symptoms have worsened the physician may either accept this information as fact, and act to assess and relieve the patient's current symptoms, or view the family observations as reflecting the family member's own distress and offer supportive statements about how as much as possible is already being done to provide comfort to the patient. The present study was undertaken because little is known about how close the ratings of the family members and hospital-based nurses are to those of the patients in regards to symptoms experienced.

1.4 OVERVIEW OF THESIS

What literature exists on the subjects of: symptom prevalence in advanced cancer, family perception of cancer symptoms, and nurse perception of these symptoms will be reviewed. What literature is available to validate the nine 100 mm visual analogue scales (VAS) of the Edmonton Symptom Assessment System (ESAS), an existing method of assessing and recording cancer symptoms, will be discussed. Finally, an observational study of 100 cancer in-patients, their family members, and the nurses providing their care, using the ESAS to rate the symptoms suffered by the cancer patients, will be reported. The primary statistical analyses of the ensuing data will be to calculate Cohen's Kappa (a measure of agreement between two observers rating the same set of symptoms) and will use the patients' ratings of their symptoms as the 'gold standard' against which the ratings of the family members and the nurses will be judged.

1.5 OBJECTIVES

The objects of this descriptive study were: (1) to identify the intensity of nine cancer-related symptoms, as rated by hospitalized cancer patients, 2) to determine the extent of agreement between the symptom ratings of these cancer patients and the ratings of their family members and those of the nurses providing the bedside care, (3) to determine over-estimation or under-estimation of these symptoms, if the family members and the nurses did not agree with the patient, and (4) to determine whether the nurses or the family members agree more closely with the patients' ratings on these nine symptoms. The nine symptoms are those rated using the Edmonton Symptom Assessment System (see section 2.6, page 18).

1.6 HYPOTHESES

Nurses either under-estimate or over-estimate somewhat their patients' symptoms, while family members over-estimate greatly the patients' symptoms. Nurses agree with the patients' assessment of the symptoms more closely than do the patients' family members.

1.7 REVIEW OF LITERATURE: OVERVIEW OF CONCORDANCE OF PERCEPTIONS OF CANCER RELATED SYMPTOMS BY PATIENT, FAMILY MEMBER, AND NURSE

Published differences (or agreement) between patient, family member, and nurse in the perceptions of the intensity and prevalence of cancer-related symptoms will be reviewed. Using WinSpirs 2.1. a Medline search was undertaken with the MESH headings of Neoplasms, Family, and Caregivers. One-hundred-thirty-seven such articles were found. Those articles that were of the anecdotal or clinical impressions variety, related to the genetic inheritance of cancer, or concerning cancer screening were not reviewed. Twelve articles in English reporting on observational surveys are reviewed. Eleven articles are listed in Table II because two of the articles concern the same cohort of subjects.

TABLE II: STUDIES ON THE RELATIONSHIP BETWEEN CANCER PATIENTS' PERCEPTIONS OF THEIR SYMPTOMS AND THE PERCEPTIONS OF THEIR FAMILY MEMBERS AND NURSES

Study	Symptom Assessed	Relation-ship to Patient	Number of Subjects	Estimation Compared With Pt.
Lobchuck et al	Multiple	Family	37	Over
Ferrell 1995	Multiple	Family	50	Over
Curtis and Fernsler	Pain and Multiple	Family	23	Under Agree
Clipp and George	Pain and Depression	Family	30 30	Agree Over
O'Brien and Francis	Pain	Family	42	Agree
Madison and Wilkie	Pain	Family	18	Over
Yeager et al	Pain	Family	86	Over
Miaskowski et al	Pain	Family	78	Over
Ferrell 1991	Pain	Family	85	Over
Holmes and Eburn	Multiple	Nurse	53	Over
Grossman et al	Pain	Nurse	103	Under

Study refers to the principal author of those studies reported in Chapter 1

Table II shows the results of this search listing the authors, number of subjects, whether nurses or family members were involved, and what type of agreement was found.

As can be seen in Table II, seven of the studies reported that family members over-estimated the patients' symptoms; while three studies found either under-estimation, or agreement, or no difference in symptom ratings. One study reported that nurses under-estimated the patients' symptoms while another described over-estimation of the patients' symptom ratings.

1.7.1 PERCEPTIONS OF PATIENT AND NURSE ABOUT CANCER RELATED PAIN

While studying consecutive in-patients with solid tumour cancers in Baltimore in 1991, Grossman et al,[8] had both patient and nurse use the 100 mm visual analogue scale (VAS) to measure the patients' mean pain experiences since admission to hospital. A total of 146 patients were admitted during the study period resulting in 103 study subjects; of the remaining: three refused, 14 failed a test of mental competence, 20 were too ill or discharged before being evaluated, while 6 had inadequate data collected. The nurses underestimated the patients' pain: patient mean score 34 mm, nurse mean score 24 mm. The Pearson correlation coefficient showed a fair relationship between patient and nurse ratings that was statistically significant ($r = 0.46$, $p < 0.0001$). For those patients who ranked their pain as low (VAS 0 - 20 mm) the percent agreement between patient and nurse was 82%, for mild pain (VAS 30 - 60 mm) the agreement was 51%, while for severe pain (VAS 70 - 100 mm) the agreement was 7%. The sub-group correlations and percent agreement were not statistically significant. Although describing little more than trends in percentage agreement (see methods section for a discussion of percent agreement) and a correlation for the whole sample but not for the sub-groups, this paper has been quoted in a reference text[9] as demonstrating a poor agreement between nurses and patients at the higher levels of pain and better agreement at lower pain intensities.

1.7.2 PERCEPTIONS OF PATIENT AND FAMILY ABOUT CANCER RELATED PAIN

O'Brien and Francis[10], working in two counties of Washington state in 1982, wrote an oft-quoted paper on family members' perceptions of the cancer patients' pain experiences. Their subjects, 42 out-patients with cancer (of either lung, pancreas, prostate or uterus) and their next

of kin, were selected from a sample of 80 cancer patients, however the sampling methods were not described. Pain was measured using a five point linear adjectival pain scale and the McGill Pain Questionnaire (a descriptive scale). Using the kappa statistic these authors concluded that the family members showed moderate agreement with the patient ($k = 0.37$ $p < 0.001$) on mean pain ratings over the past week, poor agreement with the patient on the immediate presence of pain ($k = 0.22$, $p = 0.13$), and no agreement on medications taken or on location of the pain. The family members over-estimated the patients' pain. The Kappa statistic was an appropriate indicator of agreement beyond chance, the measurement instrument has been well validated, however the poorly described sampling method weakened the study .

In 1995 Yeager and colleagues[11], working within the Oncology Nursing Research Network studied cancer out-patients and their family members in 16 sites: 15 in California and one in Montana. The subjects, selected by convenience, were 86 patients with cancer-related pain, 43% of the patients had cancer of either the breast, lung or colon while 53% had "other", not further defined, cancers. These authors used a series of 100 mm VAS to rate the patients' pain and calculated t test scores to detect any significant differences. Family members rated the pain higher (mean score 51.2 mm) than the patients (mean score 41.7 mm) did themselves ($t = 2.2$, $p = 0.03$).

As part of a larger 1996 study, which appeared to be an extension of the above study, on family coping; Maiskowski[12] and colleagues in California reported on the pain perception of 78 cancer patient/family dyads selected by convenience. The patients were recruited from 16 out-patient oncology clinics: 15 in California and one in Montana. Unfortunately neither the selection process nor the nature of the cancer was reported. The patients' pain was rated using a 100 mm VAS. The authors found that 30% of the dyads were congruent, while 18% of family members underestimated and 52% overestimated the patient's pain. The strength of the conclusions from the last two studies concerning agreement between patient and family members on pain perceptions is weakened by the non-random sample and the use of statistical tests that assess difference rather than agreement, and the use of the statistically weaker test of percent agreement.

Still out of California, Ferrell[13] et al, reported in 1991 on their study of 60 in-patients and 25 out-patients with cancer and their closest family members. Their convenience sample was chosen from patients attending either a community hospital, a national cancer centre, or a

home hospice programme. Although the cancer diagnoses were not given; the length of the illness, the analgesics used, and the nature of the family relationship were well described. Eighty-five percent of the family members lived with the patient and 63% were spouses. Pain was measured using the 100 mm VAS; the family members mean score was 69.92 (SD = 28.87) while the patients' mean score was 45.47 (SD = 28.42). Although the investigators did not apply any statistical tests to these results, the family members appeared to over-estimate the patients' pain.

As an extension of the study above, the same group, Ferrell et al[14], compared pain perceptions of the 85 cancer patients and their family members in three settings: community hospital (patient mean score 51.8, family mean score 77.8), hospice (patient 35.1, family 66.2), and a cancer centre (patient 52.8, family 67.0). The authors applied statistical tests to the difference between patients between sites but not between patient and family. Still these data appear to indicate that family members over-estimate patients' perceptions of pain. In both of these last two studies the sample was not random and the statistical tests were not applied to the difference between patient and family member; so any conclusions about agreement between patient and family must be drawn with caution.

In 1994 Madison and Wilkie[15] studied a convenience sample of 18 out-patients with lung cancer and their closest family members in Seattle, Washington. Pain was measured by the McGill-Melzack Pain Questionnaire and the 100 mm VAS. The authors used Spearman's Rank Order Correlation coefficient (rho) to compare the patients' pain as rated by the two groups. Although no correlation was found for total pain ratings (rho = 0.27, p N/S), the affective component of the pain scale showed a strong correlation (rho = 0.55 $p < 0.05$) between the patients and family members. The authors reported that 22% of families correctly estimated, 22% under-estimated, while 56% over-estimated the patients' pain ratings. The use of the rho to test for correlations between patients' and family members' pain ratings is appropriate, as used in this study, while percent agreement points only to a trend. This study, with its low number of subjects and large number of statistical tests, indicates there may be some concordance between ratings offered by patient and family.

1.17.3 SUMMARY

Of the 5 studies (six publications: the two Ferrell groups published twice on the same study) reported on above, none of the authors used a random or consecutive sampling method, 4 studies used a convenience sample while the fifth didn't report the sampling method. All authors reported on the inclusion criteria of pain and whether their subjects were in-patients or out-patients. Two studies reported on the type of malignancies involved while the remaining 3 noted only that the subjects were cancer patients. All 5 of the above studies dipped into a mixed bag of family members with 60% to 78% being spouses and the rest being offspring or parents depending on the age of the study patients.

One of the authors, appropriately, calculated Cohen's kappa while another, also appropriately, used Spearman's rho to evaluate how closely the family members agree or correlate in their perceptions of the cancer patient's suffering. The other 3 authors used the t-test to assess differences in perceptions, which seems inappropriate as a lack of difference does not equal agreement or correlation. Over-all the data reviewed indicate that family members tend to over-estimate the patient's pain.

The evidence from the one reviewed study that involved nurses indicates that they under-estimate the patients' pain.

1.17.4 PERCEPTIONS OF PATIENT AND NURSES ABOUT MULTIPLE CANCER SYMPTOMS

All of the reviewed studies reported on the symptom of pain, six studies assessed only pain, one study assessed pain and depression, leaving four studies that measured perception, by patient and nurse or by patient and family of many of the symptoms listed in Table I.

The first was conducted by Holmes and Eburn in Surrey, England and published in 1989[16]. Fifty-three cancer patients (not further described) and the nurses who were caring for them assessed the patients' symptoms using a series of 100 mm Visual Analogue Scales within a modified McCorkle & Young Symptom Distress Scale. The nurses over-estimated the patients' overall symptoms. The authors used a t-test to evaluate the statistical significance of any differences in perceptions and found that for pain, nausea, appetite, sleep, mood and

concentration there were statistically significant differences and that the nurses over-estimated these six symptoms. For the symptoms of mobility, diarrhoea, constipation, tiredness, and appearance the nurses and their patients did not demonstrate statistically significant differences in perceptions. No tests of agreement or correlation were performed. There was no description of the sampling method, or the hospital setting, nor were the raw symptom scores reported. As this study was not well described the nature of the conclusions may not be generally applicable to clinical situations.

1.17.5 PERCEPTIONS OF PATIENT AND FAMILY ABOUT MULTIPLE CANCER SYMPTOMS

The second multi-symptom study was conducted in Winnipeg, by Lobchuk et al, between August 1994 and March 1995[17]. The 37 (out of 62 eligible) subjects were lung cancer out-patients and their main family caregivers recruited consecutively from two palliative care programs and three chemotherapy out-patient programs. The authors administered the unmodified McCorkle & Young Symptom Distress Scale, which measures symptoms using a five point Likert Scale. The authors used Cohen's kappa to assess binomial agreement between the family and patient ratings, using 3 out of 5 as the cut-off point. Statistically significant agreement ($p < 0.05$) was found for fatigue ($k = 0.63$), appetite ($k = 0.60$), pain frequency ($k = 0.55$), cough ($k = 0.52$), nausea frequency ($k = 0.47$), insomnia ($k = 0.43$), breathing ($k = 0.41$), and outlook ($k = 0.33$). Symptoms on which no agreement was found (kappa values less than 0.31 and non significant p values) were: pain intensity, appearance, concentration, nausea intensity and constipation. The family caregivers over-estimated the latter five symptoms. This study appeared to be well done: the sampling method was appropriate, the subjects were well described, the raw scores were given, and the kappa statistic was used to assess agreement. This study showed that the families agreed with the lung cancer patient on 8 of the 13 symptoms assessed.

While studying the impact of a cancer pain education program on 50 cancer out-patients and their families, Ferrell et al[18], in 1995, used the 100 mm VAS to assess pain and a well validated quality of life index to assess other symptoms. The patients were all over 60 years of age and were recruited non-randomly from two cancer treating institutions in California using pharmacy and oncology clinic records that identified patients receiving analgesics. The cancer diagnoses,

analgesic use, pain sites, and spousal relationships were well described. The authors used a t test to compare mean symptom ratings and found that family members significantly ($p < 0.01$) over-estimated the following symptoms: concentration, appetite, tiredness, pain, strength, worry, and usefulness. No differences were found concerning sleep, affection given/received, disease adjustment, life enjoyment, sense of control, support, and ability to perform leisure activities. As the sample was not random and the statistical tests used evaluated differences in symptom ratings rather than agreement, the data, from this study, to indicate levels of agreement on symptom assessment, are not strong.

In 1991 Clipp and George[19] in North Carolina reported on their use of correlational computations (not further defined) and percent agreement to study the symptom perceptions of thirty cancer out-patients and their spouses. The sample was randomly selected and stratified to include equal numbers of male and female patients with either cancer of the lung or bowel. All patients were married and living with their spouse. The authors report that for pain the $r = 0.55$, with spouses under-estimating somewhat. For depression the $r = 0.53$, with the spouses over-estimating somewhat. No tests of significance were applied to these data, although this study appeared to be well done otherwise, offering a strong case for moderate concordance between patient and spouse on the cancer-related symptoms of pain and depression.

Curtis and Fernsler[20], working out of home hospice programs in Delaware in 1989, studied a convenience sample of 23 cancer patient\family member dyads. The patients were all suffering from cancer but the diagnoses were not further defined. Spouses comprised 56% of the family members. Length of time in the hospice program varied from 7 to 85 days and the ages of the patients ranged from 41 to 84 years. Multiple symptoms were assessed using a series of 100 mm VAS scales. Pain scores were different (statistical significance of $p < 0.05$) between patient (75.9) and family (63). The authors, using the t test found no statistical differences for the following symptoms: nausea, appetite, strength, work, sleep, fun, satisfaction, and quality of life. Again, lack of difference does not equal agreement, and because of the non random sampling method used the conclusions offered by these authors may not be applicable to a wider population.

1.17.6 SUMMARY

In the one study on the subject of nurses' and patients' perceptions of multiple cancer-related symptoms the nurses over-estimated: pain, nausea, appetite, sleep, mood, and concentration. No differences were found for: mobility, diarrhoea, constipation, tiredness and appearance. These findings are not more widely generizable because of the small numbers of subjects, the sampling methods, and the statistical proceedures used.

Two of the four studies involving patients and family members used the t test to test for differences between multiple symptom ratings. Within one of these studies the family members over-estimated: appetite, pain, strength, usefulness, and tiredness; while no differences were found for: sleep, affection, adjustment, control, support, and leisure activities. Within the second of these studies, the family under-estimated the patients' pain while no differences were found for nausea, appetite, strength, work, sleep, fun, satisfaction, and quality of life. A third study used correlations to show that the family members' ratings of pain and depression correlated with the patients' perceptions, while the fourth study, using Cohen's Kappa to calculate agreement, found over-estimation by the family members for: pain intensity, concentration, nausea intensity, appearance, and constipation; and agreement with the patient on fatigue, appetite, pain frequency, cough, nausea frequency, insomnia, dyspnoea. There appears to be no clear constellation of symptoms on which patient and family agree. Overall, the family members tended to over-estimate the patients' symptoms.

CHAPTER II
DESIGN, CONDUCT, AND ANALYTIC PLAN OF THE STUDY

2.1 INTRODUCTION

This descriptive study compared the perceptions of nine cancer-related symptoms, experienced by in-patients receiving cancer therapy, with the perceptions of the closest available family member and the nurse providing the bedside care. One hundred consecutive eligible patient admissions, providing the patient-family-nurse triads, were evaluated between September, 1997 and February, 1998 on three wards within the Health Care Corporation of St. John's.

2.2 ETHICAL CONSIDERATIONS

Eleven studies, reviewed in Chapter I, assessed the relationship between patient and family or between patient and nurse concerning the perceptions of cancer-related symptoms. As no clear consensus has been reached on the nature of these relationships it is ethical to study the matter further.

The author was the sole investigator and personally obtained the consent of the patient, family, and nurse subjects. All subjects were informed of the voluntary nature of their participation. Confidentiality was ensured by the investigator who stored all data in a secure place well away from the hospitals.

Recruitment methods, consent procedures, and measures to protect confidentiality will be discussed fully below.

2.2.1 HUMAN INVESTIGATION COMMITTEE

The Human Investigation Committee of Memorial University of Newfoundland, Dr. B. Younghusband, Chair, approved this study on September 11, 1997. See appendix V for a copy of the approval letter.

2.2.2 HOSPITAL APPROVAL

The Medical Advisory Committee of the Health Care Corporation of St. John's recommended and the Board of Directors of the corporation formally approved this study's commencement in October of 1997. Dr. E Parsons, Vice-President of Medical Services for the corporation was able to give conditional approval September 12, 1997 and Appendix VI shows a copy of his letter of approval.

The investigator met with the nurse managers of the Surgical, Palliative Care and Oncology programs of the Health Care Corporation of St. John's to explain the study and answer any questions, and held similar discussions with the chief of surgery at the St. Clare's site, the oncologists and haematologists at the Health Science Centre site, and the medical director of the Palliative Care Program.

None of the above managers/ directors were involved in the recruitment of subjects, collection of data, analysis of results, or subsequent write-up.

2.3 STUDY SUBJECTS

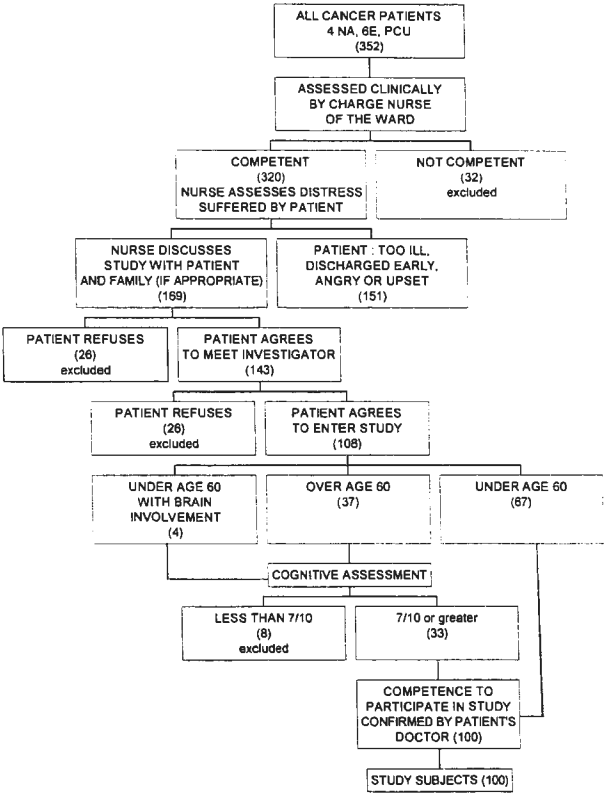
Figure I shows the flow chart summarizing the patient selection process. Please see next page for Figure I

2.3.1 INCLUSION CRITERIA

Patients were eligible to enter this study if they were: well enough to participate, suffering from a malignancy, admitted to hospital because of this malignancy, and had their closest relative available for interview. The closest relative was recruited, in descending order of priority, as per the next of kin legislation for Newfoundland: spouse, adult child, parent, sibling, grandchild, cousin, close friend. In situations of doubt the patient was asked to identify the person most appropriate.

Subjects were recruited from three hospital wards within the facilities of the Health Care Corporation of St. John's, Newfoundland: 4 North A at the Health Science Complex, the provincial referral site for oncology and haematology patients; 6 East at St. Clare's Hospital, a regional centre for cancer related surgery; and the Palliative Care Unit, a regional referral centre.

STUDY SUBJECT RECRUITMENT FLOW CHART



2.3.2 EXCLUSION CRITERIA

Patients were excluded from this study if they: refused the solicitation to participate from either the charge nurse or the investigator, were confused (as assessed by the charge nurse's clinical judgement during usual daily care or by the investigator after the administration of a short but formal test of mental status), were incapacitated due to severity of symptoms (as assessed by the charge nurse), or had no family members available or willing to participate.

2.4 RECRUITMENT OF PATIENT AND FAMILY

The procedure for recruiting subjects was as follows: the charge nurse would approach eligible patients after routine daily rounds and assess both severity of distress and mental status based on the patient's behaviour over the past 24 hours. These bedside assessments were the clinical judgements of one of the five senior, experienced nurses whose job it is to attend daily ward rounds as part of the medical team, and to supervise the nurses providing the bedside care. Patients were approached as soon as possible after admission. Those patients obviously confused, somnolent, or in great distress were not approached. If the patient appeared alert and well enough to talk the nurse would ask the following question "Would you be willing to talk to a medical researcher who is doing a study on the symptoms of people in hospital?"

If the patient agreed to such a meeting the investigator would approach the patient to explain the study procedure further and return, if the patient agreed, when the closest relative was visiting the hospital. At that time the investigator would further explain the data collection process, ask for and answer any questions, obtain verbal consent from all parties, and have both patient and family member sign consent forms.

Any patient consenting to the study who seemed cognitively intact to the charge nurse but was over 60 years of age, appeared confused to the investigator, or was suffering from brain involvement as a consequence of their disease was formally tested, by the investigator, for mental status. The test used was the Canadian Mental Status Questionnaire[21], a ten item test which was has a test-retest correlation of 0.98, a sensitivity of 0.69, and a specificity of 0.94 for moderate and severe cognitive impairment when using a cut-off score of 7 out of 10.

Verbal consent to administer this test was obtained from the patient before proceeding. Patients who scored less than 7 out of 10 were excluded from the study. No patients deemed confused by the charge nurse, and therefore excluded from the study, were formally tested for mental status.

Competence of the patient to engage in this study was confirmed by one of the patient's doctors, usually a member of the house staff. If, in the opinion of the doctor, the patient was competent to understand the voluntary nature of such research, the doctor would indicate same by signing a statement to that effect on the patient's consent form. No patients who scored 7 out of 10 or greater on formal testing were deemed incompetent by their doctors.

2.5 RECRUITMENT AND TRAINING OF NURSES

The nurses who participated as subjects in this study were actively involved in the bedside care of the patient-subjects. Sixty-four such nurses underwent a 10 to 15 minute training session on the completion of the ESAS and signed consent forms indicating their understanding of their role in the study and the voluntary nature of their participation. These nurses were told the study's aim was to compare their ratings of nine cancer-related symptoms suffered by their patient with the patient's own ratings. They were told that such assessments don't always agree but were not told whether this disagreement was in over- or under-estimation. The investigator reviewed the method of rating the patient's symptoms using the ESAS just prior to the nurse completing it.

The five charge nurses (three from 4 NA, one each from 6 E and PCU), were responsible for the recruitment of the patient-subjects, and were further instructed in ways of approaching patients as potential subjects to maximize cooperation while protecting the right of any patient to refuse participation. These nurses assessed confusion and severity of illness based on their clinical experience and expertise.

2.6 THE EDMONTON SYMPTOM ASSESSMENT SYSTEM (ESAS)

Those who work with patients in or nearing the uncomfortable phase of the cancer experience are always endeavouring to develop better

ways to assess their patient's symptoms of distress. Symptom evaluation forms, such as the ESAS, have been developed to help in the evaluation and documentation of these symptoms. Appendix I shows the actual form containing the VAS scales which are used to rate the nine symptoms assessed by the ESAS.

Of the 11 studies listed in Table II (page 6), eight used the Visual Analogue Scale (VAS) to measure symptom ratings, 2 used adjectival scales, while one used a Likert scale. The present study involves subjects similar to those in the studies listed in Table II so, if convention were to be followed, the VAS would be the apparent choice of symptom measurement instrument.

In 1990 Bruera et al[22] reported on their development and clinical usefulness of the ESAS. The authors, using t tests on the mean daily total ESAS scores of 101 consecutive admissions to their palliative care unit, reported statistically significant improvement in symptom ratings (mean distress score 410 +/-95 on day one after admission, versus 362 +/- 83 day five, $p < 0.01$) that coincided with clinically observed improvement in the patient's comfort. The form is easy to understand and quick to complete. This earlier version listed eight symptoms, more recent versions have added a scale for shortness of breath.

The ESAS is comprised of nine visual analogue scales (VAS) one for each of the common cancer-related symptoms: pain, nausea, depression, anxiety, shortness of breath, tiredness, drowsiness, appetite, and overall well being. The evidence to support the use of the VAS to measure these nine symptoms is reviewed and evaluated in Chapter III.

The specific sensitivity and specificity of the VAS could not be found in the literature; nor could the sensitivity and specificity of the ESAS, as a whole, be found. For the five symptoms of pain, nausea, shortness of breath, anxiety, and depression, the VAS appears as valid and reliable as other instruments available (see Chapter III). For the remaining four symptoms: tiredness, well being, drowsiness, and appetite; the validity and reliability of the VAS can only be inferred.

Holmes and Eburn[16], using the 11 VAS measures of the McCorkle and Young Symptom Distress Scale, calculated the reliability of the whole distress scale for 53 cancer patients and their nurses. The coefficient alpha of the nurses' data was 0.81 and 0.97 for the patients'. This scale, uses a series of 100 mm VAS as does the ESAS, making it similar enough to allow similar conclusions about the internal reliability of the ESAS.

None of the other multiple-symptom rating instruments used by the authors listed in Table II (page 6) were similar enough to the ESAS to permit comparisons of validity and reliability with these instruments.

Because of its simplicity, Canadian genesis, and its anticipated regular use with the patients referred to the palliative care service in St. John's, the ESAS was chosen as the symptom assessment tool for the present study. In the present study the subjects were asked to complete a set of nine VAS measures. Because the subjects were suffering from cancer of various types and at differing stages of their illnesses the means and standard deviations reported in the literature may be unsuitable for use in the comparative analysis of the data.

In a review of the use of the VAS (Chapter III) a difference of 10 mm or more between one rating score and another of a particular symptom is, in most of the studies using the 100 mm VAS, the standard clinically significant difference within that symptom. Similarly a difference between patient and care-giver ratings for any symptom of 10 mm or more is used as a significant difference by most of the authors who report on the use of the VAS (see Chapter III).

2.7 DATA COLLECTION

The patient and family member were asked to complete the ESAS independently from each other (usually the relative went first). The investigator was present to answer questions, ensure that the subjects' ratings were independent from each other, read the words on the form if the patient couldn't read, and physically assist those too weak to complete the form. In all cases the patient and family member completed the ESAS at the same time. Given the time issues concerning recruitment and family visiting the data collection took place during the afternoons. As such the family would have had the opportunity to sit and chat with the patient before completing the ESAS.

The investigator would then locate the patient's primary nurse (usually within five minutes) and have her complete the ESAS. All subjects in the patient/family/nurse triad completed the ESAS to rate the patient's symptoms for a similar point in time. The patient's nurse would complete the ESAS based on the regular assessments made throughout the day. As most of the data collection took place in the afternoon or early evening most of the nurses involved in this study had the opportunity to assess the patient before completing the ESAS.

Subjects were asked to complete the ESAS once only for each patient. Patients who were subjects during one admission were not included if they had a subsequent admission during the study period.

The subject's age and diagnosis were obtained verbally from the patient or the nurse while the family member's relationship and location of residence relative to the patient was obtained from either the patient or family member. The investigator did not review the patient's chart.

2.8 DATA MANAGEMENT

Each subject triad was given a study number with which the ESAS data sheets were coded, accompanied by the appropriate descriptive term (patient, family, nurse). The identity of the nurse was not recorded on the data sheets or anywhere else. The patient's age, ward location, diagnosis, mental status score (as appropriate), gender; the family member's relationship to the patient, and location of residence were recorded on the patient data collection form, a likeness of which is shown in Figure II.

A research log book was kept on each ward into which the charge nurse would record the names of patients with cancer admitted to the ward (eligible subjects). The research log book recorded those who agreed to meet the investigator, those who refused the approach of the charge nurse; those who were, in the opinion of the charge nurse, too ill or confused to take part; those who refused the approach of the investigator to take part, and those who failed the formal test of mental status. Unfortunately no record was kept to distinguish between those patients who were too ill to be approached, had no family members available, or were discharged before the charge nurse could approach them.

The investigator scored each VAS of each ESAS for each subject triad by measuring the 100 mm VAS from the left. These scores were recorded in a data book and subsequently entered into the SPSS package for personal computers. The consent forms completed by the patients, family members, and nurses; the ESAS data sheets of each member of the triad; the ward log books, data book, and personal

computer were stored securely at the home of the investigator.

FIGURE II: PATIENT DATA COLLECTION FORM

Patient Name	
Study Number	
Ward	4NA 6E PCU
Age	
Gender	Male Female
Diagnosis	
Family Member	Spouse Child Parent Sibling Other Relative Friend
Residence of Family Member in Relation to the Patient	Same House Same City Distant
Mental Status Score (as appropriate)	

2.9 QUALITY CONTROL OF THE DATA

Quality control was assessed by a second investigator (JH) measuring and scoring a random sample of 10% (270 VAS scales) of the total data set and comparing these scores with the primary investigator's (BE) measurements and scores.

Agreement between these two investigators' measurements was assessed using the Kappa statistic.

2.10 DATA ANALYSIS

The aim of the present study was to assess agreement been patient and family and between patient and nurse on the nine symptoms assessed by the ESAS. Cohen's Kappa, a statistical test of agreement between two judges measuring the same phenomena, will be applied to the symptom ratings offered by each member of the subject triad. Scatter diagrams, to show what linear relationships exist between patient and nurse and between patient and family, will be created.

The descriptive statistics of frequency, means, and standard deviations; as well as Spearman's Rank Order Correlations and t tests will be applied to the ESAS symptom data in order to compare the present study's findings with those of the authors listed in Table II. Demographic data will be presented in tables giving frequencies as appropriate.

No comparisons of agreement between patient and family member within the different sub-sets of family members (ie. spouse, adult child, sibling) or of their location of residence relative to the patient will be made because of the small numbers of subjects. For the same reason no such comparisons between the different nursing sub-groups (ward locations) will be made.

Table II (page 6) shows the results of the concordance of symptom perception between cancer patients, their family members, and their nurses as reported in Chapter I. The symptoms assessed and the statistical methods used to evaluate correlation, association, or agreement are shown for the same studies in Table XIV (page 69) and discussed in Chapter V.

Using standard texts of statistical methods[23],[24],[25] the test statistics used by the authors listed in Tables II and XIV, and the appropriate test statistics to be applied to the data of the present study will be discussed.

2.10.1 CALCULATION OF THE t SCORE

The Student's t-test was developed by the statistician, William Gossett, working in Dublin and analyzes the difference between the mean scores of two sets of observations.

As displayed in Table XIV the t-test is a commonly used statistic to assess the significance of any differences between the patient's symptoms as rated by each member of the study triad. The t score is calculated using the formula shown in Figure III.

Figure III: Calculation of the t score.

The t score

$$t=(\overline{x_1}-\overline{x_2})/S.E.$$

X_1 = mean group 1

X_2 = mean group 2

S.E. = standard error

S.E. for Unpaired Data:

$$S.E.=\sqrt{[(S_1^2+S_2^2)/n]}$$

n = sample size

S_1 = standard deviation group 1

S_2 = standard deviation group 2

S.E. for Paired Data:

$$S.E.=\sqrt{[S_d^2/(n-1)]}$$

S_d = standard deviation of difference between pairs

The computed t score is compared to a table of standard t scores and significance of the score is based on sample size. As the sample size approaches 30 the t distribution becomes normal.

For the paired t-test the t score is calculated the same way except the standard error of the differences is derived by taking the difference between the two scores for each subject, computing the mean difference for the whole sample, squaring the differences between each score and the mean difference, dividing by the degrees of freedom (sample size minus one) and taking the square root.

The difference between the t-test for unmatched subjects and the t-test for paired subjects is mathematical and has little to do with direct comparisons of two observations on the same individual. In the present study it is conceivable that the means of scores offered by the members of each triad might be close yet there be little actual agreement.

The t-test tests the null hypothesis that the difference between the means is zero. If there were perfect agreement between each member of the triad the difference between the means would be zero and the null hypothesis would not be refuted.

The paired t-test is the appropriate test statistic to test the significance of any differences between the nurses' and patients' ratings and between the family members' and patients' ratings. As some authors report on over or under-estimation of the care-giver ratings compared with the patient ratings, this test will be applied to the mean symptom scores of the 100 mm VAS scales of the raw data.

The t-test is not the appropriate test statistic to assess agreement between raters because it tests for significance of difference between groups not for agreement. The inability to disprove difference does not necessarily prove agreement.

2.10.2 CALCULATION OF THE PEARSON CORRELATION COEFFICIENT

The Pearson Product Motion Correlation Coefficient assesses the strength of the relationship between two sets of measurements at least one of which is independent. The r statistic is calculated by either of the two formulae shown in Figure IV.

Figure IV. Calculation of the Pearson Product Moment Correlation Coefficient (r)

$$r=\sum (x-\bar{x})(y-\bar{y})/\sqrt{[\sum (x-\bar{x})^2 \cdot \sum (y-\bar{y})^2]}$$

x = observations of group 1

y = observations of group 2

or

$$r=\sqrt{(\sum squares(R)/[\sum squares(R)+\sum squares(E)])}$$

R = regression

E = error

The significance of r is calculated thus: if the null hypothesis states that the probability that r = 0 the t statistic is used while if the null hypothesis states that r = some value other than zero the z distribution is used.

Pearson's r has a range of -1 to +1 the value sign indicating the direction of any relationship. Correlations from 0 to 0.25 indicate little

or no relationship, those between 0.25 and 0.50 indicate a moderate relationship, while those between 0.50 and 0.75, show a strong relationship, while those greater than 0.75 show a very strong relationship.

Although this measure appears more acceptable than the t test in that the null hypotheses (that $r = 0$, or that $r \neq 0$) can be tested, it does not take into account the pairing of observations in the present study. It is possible that disparity between ratings offered by each member of the triads might result in acceptable Pearson product moment correlation coefficients yet not reflect actual agreement.

2.10.3 CALCULATION OF SPEARMAN'S COEFFICIENT

Spearman's Rank Order Coefficient (rho) is calculated between multiple pairs (ie nurse and patient) of observations (ie patient's symptoms) by first ranking the scores for both sets of observations (usually by descending magnitude) and calculating the correlation coefficient between the rankings for each set of pairs.

The pairing and the use of ranking instead of raw data are the main differences between Spearman's rho and Pearson's r. A distillation of Spearman's formula is shown in Figure V.

FIGURE V: Calculation of Spearman's Rank Order Correlation Coefficient (rho)

$$\rho = (1-6)(\sum d^2)/n(n^2-n)$$

n = sample size

d = difference between ranks for each pair

Again the range of rho is the same as for Pearson's r and the significance is calculated using the same conversions to t or z scores

and referring to the appropriate distribution tables. Although Spearman's rho is only 91% as efficient as Pearson' r in situations where the distribution of scores is truly normal, the rho statistic does test for correlation between pairs of observations. Spearman's rho can be used to measure inter-rater concordance when ordinal data has been collected. Because the difference in paired scores is used to calculate it, rho is an appropriate test statistic for the present study.

2.10.4 CALCULATION OF COHEN'S KAPPA

Percent agreement between observers' ratings of symptoms is an often used measure reported in the studies in Table XIV and ranges from 6% to 82% depending on the symptom. There appears to be no consistency of percent agreement across these studies for the most commonly studied symptom of pain (range 10% to 82%). Percent agreement does not take into account either sample size or chance agreement. This test statistic will be used, with reservations, in this study to compare with the studies shown in Table XIV.

Cohen's Kappa was described in 1964 and measures percent or proportion agreement taking chance agreement into consideration. Cohen's k is calculated as in Figure VI using percent agreement from a cross-tabulation table of the paired observer ratings and percent chance agreement calculated by adding the margins of each cell from this cross-tabulation table.

FIGURE VI: Calculation of Cohen's Kappa (k)

$$k = \frac{\text{observed agreement - chance agreement}}{\text{1 - chance agreement}}$$

The range of k is from -1 for perfect disagreement to +1 for perfect agreement with 0 to 0.35 showing poor agreement, 0.35 to 0.5 showing moderate agreement, 0.5 to 0.75 indicating strong agreement and 0.75 and above demonstrating very strong agreement.

2.11 CHOICE OF TEST STATISTIC

Because the aim of this study is to measure agreement on perceptions of symptoms the kappa statistic will be used. Spearman's rho will also be calculated, on the whole raw data set, as a measure of correlation between pairs, and to amplify the Kappa results. Student's t-test will be calculated on the raw data to test the significance of any differences in the mean symptom scores and for comparrison with previous studies, while scatter diagrams will be drawn to search for any linear relationship between the members of the study triads.

2.12 ANALYSES OF THE DATA

In order to calculate the kappa using SPSS, for personal computer, each cell of the chi square cross-tables must be filled. Using the 100 mm scale many of the 200 cells remained unfilled. Furthermore, using the 10 cm scale, some of the sub-scales had empty cells in the cross-tables prohibiting calculation of the kappa statistic.

To facilitate the calculation of kappa the VAS raw scores will be converted from a 100 point scale of 1 mm each to a ten point scale of 1 cm each by rounding up or down from the 0.5 cm mid point between the 1 cm gradations. For example 0 to 4.9 mm = 0 cm, 5.0 to 14.9 mm = 1 cm, 95.0 to 100 mm = 10 cm. The VAS raw scores will be further converted to a three point scale; "Mild/Absent" (0 to 29 mm), "Moderate" (30 to 59 mm), and "Severe" (60 to 100 mm); as well as a binomial one; "Absent" (VAS 39 mm or less) and "Present" (VAS 40 mm or greater). Such conversions, performed in some of the studies listed in Table XIV, make clinical sense in that if the assessments of the family and nurses are greatly divergent from those of the patient, scales of a coarser nature will detect this divergence while finer, and perhaps clinically non-significant, divergences will be lost. Finally, with conversion to the coarser scales the statistical treatment will be feasible.

The means and standard deviations of the unconverted raw data will be calculated in the usual ways using SPSS for personal computer. The Student's t-test for paired data will be calculated, on the unconverted

raw data, to assess the significance of any differences in mean symptom ratings between patient and family member and between patient and nurse.

2.13 SUMMARY

The aim of this observational study is to assess the level of agreement on the perceptions of cancer related symptoms as rated by the patient and a close family member and by the patient and the nurse. The Edmonton Symptom Assessment System (ESAS) will be the instrument used to rate nine symptoms commonly encountered in the cancer experience.

All cancer patients admitted to each of three hospital wards will be eligible for recruitment into this study. Those well enough to participate will be invited by the charge nurse of the ward to meet the investigator who will subsequently invite the patient and family member to complete the ESAS. The patient's nurse will also complete the ESAS at about the same time.

Because of technical issues involving SPSS for personal computer, for ease of analysis, and following precedent displayed in the literature, the VAS scores of the ESAS will be converted from a 100 mm scale in three ways to 1) a ten point scale, 2) a three point scale, and 3) a binomial scale.

The raw data will be treated as ordinal data for the parametric tests and Spearman's rho, while the converted data will be subjected to the non parametric tests of agreement. The means and standard deviations will be calculated for the 100 mm scale data in the usual way and the t-test calculated to assess the significance of any differences in symptom assessment between the subjects. The Kappa statistic will be calculated using SPSS for the three converted scales. Although percent agreement is a simplistic measure of agreement this statistic will be calculated to allow comparison with what other studies have reported.

CHAPTER III

LITERATURE REVIEW ON THE USE OF VISUAL ANALOGUE SCALES TO MEASURE SYMPTOMS COMMONLY SUFFERED BY PATIENTS WITH CANCER AS RECORDED BY THE ESAS

3.1 INTRODUCTION

This chapter begins with a discussion of the development and research uses of the visual analogue scale prior to a more lengthy discussion of the validation of this scale as a measurement tool for the subjective phenomena of cancer-related symptoms.

Using the Silver Platter program a Medline search of the literature was performed using MESH headings of Visual Analogue Scale and each of the following nine symptoms that are commonly associated with cancer as measured by the Edmonton Symptom Assessment System (ESAS): pain, nausea, tiredness, depression, anxiety, shortness of breath, drowsiness and appetite. Dyspnoea was substituted for Shortness of Breath as the former term is a MESH heading while the latter term is part of the ESAS. The search combining drowsiness and visual analogue scales turned up no references; however the search combining insomnia and visual analogue scales did find references, but because drowsiness and insomnia differ substantially those articles were not reviewed.

Most of the articles, found during the initial search process, on symptoms other than pain, actually used the VAS to measure pain while referring to the other symptoms in passing.

The abstracts of the citations in English were scanned and only those studies that attempted to demonstrate validation of the VAS by comparing it with other measures were reviewed in depth.

Table III shows the results of this literature search.

TABLE III: RESULTS OF A LITERATURE SEARCH CONCERNING THE USE OF THE VISUAL ANALOGUE SCALES IN ASSESSING THE COMPONENTS OF THE EDMONTON SYMPTOM ASSESSMENT SYSTEM

Symptom	Total Citations	English Citations	Reviewed Articles
Pain	1,051	934	15
Shortness of Breath	27	22	8
Nausea	129	111	7
Depression	135	113	7
Anxiety	121	101	10
Tiredness	10	8	3
Drowsiness	24	20	1
Appetite	17	15	2
Well Being	0	0	0
TOTALS	1,514	1,324	53

3.2 VISUAL ANALOGUE SCALES (VAS)

The VAS was first introduced and described in the medical literature in the early 1920's by Dr. M Freyd[26]. They were re-introduced for the measurement of feelings in 1969 by Dr. R.C.B. Aitken of Edinburgh, Scotland, who commented that the VAS was superior to adjectival descriptors of feelings and stated, "Words may fail to describe the exactness of the subjective experience." Aitken[27].

The VAS has differing lengths and forms, linear or bipolar (see section on Shortness of Breath) and may be vertical or horizontal. The most common and the form used in the ESAS is horizontal and 100 mm in length. The scale is anchored at each end with the two extremes of the

symptom in question, for example when considering pain the anchors are "No Pain" and "Worst possible Pain".

To rate each symptom the subject makes a mark with a pen or pencil at a place along the line, between the two extremes, appropriate to the quantity of feeling at the time.

In 1969 Aitken reviewed some of the experiments he had conducted using the VAS. He compared aircraft pilots' descriptions of safety issues with their relative importance and had independent judges assess the same relative importance using a series of VAS ratings. He calculated significant reliability of concordance by Kendall's coefficient for each safety issue. Spearman's rank correlation coefficient was used to assess concordance between scores and areas of relative importance. Although the actual statistical results were not offered in this paper, Aitkin produced the histograms of the raw VAS scores from three of his studies to show the normal distribution, after arcsine transformation, and thereby to conclude that the scores "fulfilled the necessary requirements for analysis of variance: homogenous and independent variance, normal distribution (with arcsine transformation) and sufficient resolution in measurement to provide continuous rather than discrete scores." Aitken[27].

The debate on the analysis of VAS scores continued when in 1974 two British scientists, Bond and Lader[28], administered sixteen 100 mm VAS measures concerning feelings to 500 healthy volunteers whose ages ranged from 16 to 64 years. This heterogenous sample was made up of employees of technical colleges, universities, and hospitals. The authors presented a frequency histogram of all scores between 45 mm and 55 mm to disprove the hypothesis that subjects tended to cluster their scores about the middle. They concluded that test-retest assessments of internal validity would be inappropriate as the scales do not rate stable phenomena but measure feelings in the here and now. The distributions of the sixteen sub-scale scores were not normal in every case: those titled Clear Headed, Tranquil, Proficient, Relaxed, and Amicable were skewed to the positive; Calm, Well-coordinated, Contented, Happy and Interested were negatively skewed; while Alert, Strong, Energetic, Quick-witted, Attentive, and Proficient were normally distributed. The authors found conversion of the raw scores into natural log scores aided their factor analysis which revealed three factors (Alert, Happy, and Calm) accounted for 61% of the total variance. They could not account for the skewedness of the rating scores in their healthy population, however such distributions might occur in a study of ailing subjects because of their illness. The authors

noted the scales were quick to fill out, easy for the subjects to understand, did not require much subject motivation, and the rater is not restricted by adjectives or direct quantitative rankings.

By 1978 the VAS was becoming more common in medical research that measured subject self-rating of feelings. Maxwell[29] studied the ability of the 100 mm VAS to measure seven volumes of sound offered to 27 healthy volunteers and repeated the testing one day later. The VAS was found to be simple to explain and use and largely acceptable to the subjects. Arcsine conversion of the raw scores did not aid analysis while sensitivity was improved through conversion to proportional scores (100 x raw score/maximum score) which was done because some of the subjects rated their maximum score as less than 35. The product moment correlations of the mean scores and the volume settings were statistically significant $r = 0.96$ ($p < 0.001$) on both days. Although commenting that the sensitivity was adequate if not completely satisfactory this author found 5 of the 49 (10%) statistically significant results to be wrong in that a higher volume of sound was rated by the subjects on the second day as being lower. Maxwell concluded that concerning statistical analysis of significance, "it makes little difference most of the time whether parametric or ordinal tests are used", and that large differences (greater than 10 mm) can be tested for significance by parametric tests. Within-subject comparisons were more accurate than those between subjects.

Although the VAS has value, it is not a perfect tool, a situation alluded to above but amplified by Downie et al in 1978[10]. These authors tested the grip strength of 7 healthy volunteers and 93 out-patients suffering from arthritis in three ways: 1) clinical assessment by a physician (three finger squeeze), 2) objective assessment with a sphygmodynamometer, and 3) a 100 mm VAS rating by the subject of their subjective strength. Based on objective assessment the subjects were divided into four groups: very weak, weak, normal, and strong. Using a correlation matrix the authors found reasonable correlation between VAS and objective assessment ($r = 0.6358$, $p < 0.001$) and, offered here as contrast, between Physician and objective measure ($r = 0.9177$, $p < 0.001$). However further analysis revealed that the correlations were not as strong between the four sub groups and the VAS. For those subjects who were: Very Weak $r = 0.5651$ (p = NS), Weak $r = 0.0219$ (p = NS), Normal $r = 0.4918$ ($p < 0.001$), and Strong $r = 0.4107$ (p = NS). The authors suggest caution when using the VAS to measure subjective experience.

3.2.1 SUMMARY

Based on these studies, the VAS as a tool to measure subjective experience appears to result in scores of an ordinal nature. By selecting an arbitrary cut off (scores between 20 mm and 40 mm have been used) the scale can be employed as a binomial one with those scores above the cut off indicating the presence of the symptom; with those below, its absence. The median could also be used as the binomial cut off.

Further division of the scale into ten gradations (of 10 mm each) or into three (absent: 0 - 29 mm, mild: 30 - 59 mm, or severe: over 60 mm) will result in scores of an ordinal nature.

The VAS has been shown to measure subjective phenomena such as symptoms; is easily understood, can be taught quickly, requires little subject motivation, and results in scores of an ordinal nature whether used as a 100 mm, 10 cm, or a three point scale.

3.3 PAIN

The measurement of pain using the visual analogue scale (VAS) has become quite popular of late among pain researchers who report their findings in those medical publications listed in Medline. The VAS has been compared with other more complex and time consuming pain measures and has been found to be reliable and sensitive to change.

3.3.1 CANCER-RELATED PAIN

In 1975 Ohnhaus and Adler[31] made multiple comparisons, in six cancer patients, of the effects on pain of two different analgesics and a placebo. Pain was measured by a 100 mm VAS and a verbal rating scale. The authors reported a good correlation between scales ($r = 0.81$, $p < 0.001$); and concluded that, compared with the verbal scale, the VAS scale assessed more closely changes in the pain experience of their subjects.

Graham et al[32] studied 36 out-patients suffering cancer related pain which was assessed by the McGill Pain Questionnaire, an extensive and well validated self-rating tool to measure the cancer pain experience. Sixteen subjects also had their pain assessed using a 100 mm scale. Based on the 100 mm scale pain scores, these 16 subjects

were subsequently divided into high or low pain sub-groups. The mean pain scores of these two groups were compared using a t-test. The 100 mm scale and the Present Pain Intensity Index, part of the McGill Questionnaire, could distinguish between the two sub-groups ($p < 0.01$, 2 tailed). The authors concluded that the VAS could distinguish between patients with high or low pain intensities.

Wallenstein, et al[33] studied 34 cancer in-patients using a VAS and a set of 8 categorical pain descriptors modified from the Tursky Pain Perception Profile[34], a commonly used pain scale from the late 1970's. The measurements were done repeatedly over time (total of 1,300 measurements) in both oncology patients and those recovering from cancer related surgery. The authors found the relationship between the two scales to be a power function and that the VAS scale correlated well with the categorical scale ($r^2 = 0.65$). The VAS seemed to be more sensitive to smaller changes in pain than did the categorical scale.

Kremer et al[35] used three scales: VAS, numerical, and adjectival scales; to assess the pain of 56 patients suffering cancer related pain. Although reported but not methodologically described, the intercorrelation of the pain intensity ratings were: VAS and numerical scale, $r = 0.86$ ($p < 0.05$) and VAS and adjectival scale $r = 0.64$ ($p < 0.05$). An Anova of the three scales showed no difference in pain ratings. By the 1980's pain measurement had begun to evolve towards using simpler and shorter scales. One such scale is the Memorial Pain Assessment Card, an 8 item descriptor scale and 3 VAS measures (pain intensity, mood, and pain relief). Fishman et al[36] reported on the one time assessment of 50 cancer in-patients comparing the various components of this card with the McGill Pain Questionnaire and the Tursky assessment tool. The Spearman correlations were: VAS (pain intensity) and the Tursky, rho = 0.67 ($p < 0.001$) and VAS (pain intensity) and the McGill, rho = 0.45 ($p < 0.001$).

3.3.2 NON CANCER-RELATED PAIN

Thirty subjects suffering chronic musculoskeletal pain and 20 healthy volunteers were studied by Price, et al[37] who exposed their subjects

to six gradations of painful noxious thermal stimuli applied to the forearm. The relationship between skin temperature and the sensation intensity as measured by the VAS was curvilinear while the relationship between log skin temperature and log VAS intensity was linear. The authors concluded the VAS to be a valid instrument to measure, within subjects, changes in pain intensity.

Littman et al[38] reviewed 23 studies that compared the VAS and a verbal pain intensity scale in clinical trials of analgesic use with 14 protocols and 1,497 subjects suffering a variety of pain syndromes. Pearson product moment correlation coefficients were determined for the sum of VAS rated differences of pain (pre and post analgesic) and the equivalent sum rated by the verbal pain intensity scale. The correlation for 1330 subjects in whom all data could be found was so high ($r = 0.933$) that the authors conclude there was no difference in sensitivity between the VAS and the verbal descriptor scale.

Revill, et al[39] in 1976 studied the pain recall of twenty women who had recently experienced childbirth. The investigators first studied the visual motor skills needed to complete the VAS finding no difference in ability to complete the VAS whether or not the subjects had been given the sedating analgesic, pethidine. These authors reported stable scores of memory recall after five minutes and 24 hours ($r = 0.967$, p 'very highly significant'). They also studied four lengths of scale: 5, 10, 15, and 20 cm; finding significantly greater variability in the 5 cm scale. They concluded that scales of 10, 15, and 20 cm rating a constant pain stimulus are reproducible, and for different pain stimuli, measure real change in their subjects' pain perception.

Downie et al.[40] reported on 100 rheumatic patients who rated their pain with three commonly used scales: (1) the VAS, (2) an eleven block vertical column rating pain from 0 to 10, and (3) a simple descriptive scale (nil, mild, moderate, severe, and very severe). The subjects were tested once with each scale after a 'washout' period. Correlation coefficients were: VAS and numerical scale, $r = 0.64$; VAS and descriptive scale, $r = 0.705$. Although no measures of significance were supplied the correlation coefficients were acceptable indicating that the VAS measures pain almost as well as the other two methods.

While studying episiotomy pain in 26 women, Reading[41] compared the VAS with other commonly used pain scales: a verbal scale (mild, discomforting, distressing, horrible, and excruciating), a 10 point numerical scale, and a variant of the verbal descriptive component of the McGill Pain Questionnaire. Using Pearson Product Moment

correlations the authors found the verbal scale and the numerical scale did not correlate with each other ($r = 0.26$), neither did the verbal scale correlate with the VAS ($r = 0.29$). The numerical scale and the VAS correlated somewhat ($r = 0.46$, $p < 0.01$).

Linton and Gotestam[42] studied 15 patients with chronic pain comparing, over repeated measurements, the pain intensity rated by the VAS and a verbal descriptive scale. Five patients rated their pain exactly the same each day and were thus excluded from the Pearson product moment correlation analysis which showed a correlation for the remaining ten subjects of 0.68 ($p < 0.05$). It would appear that five subjects either didn't have day to day pain fluctuations (as did the other ten) or were unable to rate pain using either scale.

Jensen et al[43] studied the pain reported and rated by 75 patients with chronic pain using six methods: (1) the VAS, (2) a 101 point numerical scale, (3) an eleven item box scale, (4) a 6 point categorical descriptive behaviour scale, (5) a five item verbal categorical scale, and (6) a four item verbal categorical scale. The latter five scales measuring current pain correlated well with the VAS: $r = 0.71$ ($p < 0.001$, 2 tailed).

Hurst et al[44] compared in 233 rheumatoid arthritis patients the use of the VAS with the Stanford Health Assessment Questionnaire and a newly developed scale. For moderate or severe pain as measured by the Stanford scale the mean VAS scores were associated ($p < 0.001$) using a t test. Such an association was not noted at the mild discomfort level. These authors, among others, used the VAS for pain as a part of the gold standard to validate their new scale.

By the 1990's the assessment of pain in research protocols had begun to include a VAS scale even when other pain assessment measures were used. Beattie et al[45] employed a double blind randomized controlled trial design to study the effects of ketorolac use in 69 (vs 61 placebo controls) post operative subjects measuring the pain with a VAS as well as recording pulse and BP. All three measures were higher ($p < 0.001$ to 0.05) in the control subjects compared with those receiving the drug. This is indirect evidence of the relationship of VAS scores to the pain experience.

3.3.3 SUMMARY

The VAS used either once or repeatedly to measure the pain experience appears to be a tool that is valid, reproducible, and sensitive to change. The scale is easy to use, quick and clinically appropriate as well as being acceptable for use in the research of pain.

3.4 SHORTNESS OF BREATH

The use of the VAS to assess Shortness Of Breath (SOB) has a much shorter history than does that for pain assessment. Two versions of the VAS are described in the literature; a 100 mm vertical or horizontal line anchored at the ends by the phrases "no shortness of breath" and "shortness of breath as bad as it can be", and a bipolar type comprised of a 200 to 400 mm line anchored at the ends by the phrases "very much worse" and "very much better" with "no change" in the middle of the line. I could find no studies comparing the two types of scales.

3.4.1 CANCER-RELATED SHORTNESS OF BREATH

Booth and colleagues[46] studied the use of oxygen or compressed air delivered in a randomized, blinded fashion via nasal prongs at 4 l/min in 38 subjects with advanced cancer in the lungs. Shortness of breath was measured by the Borg scale (a 10 category scale to measure dyspnoea), a 100 mm horizontal VAS, and a qualitative assessment by the subject. A pulse oximeter was used to assess arterial oxygenation. The authors reported the pattern of responses to be the same for the Borg scale and the VAS but did not supply statistical testing results to support this conclusion. Dyspnoea, as measured by the VAS, was significantly reduced ($p < 0.001$) by breathing compressed air or oxygen when compared with breathing room air. These authors appear to have used the VAS as a measurement standard.

3.4.2 NON CANCER-RELATED SHORTNESS OF BREATH

Gift[47] studied 16 asthmatic and 30 COPD recently admitted in-patients using both the vertical and horizontal VAS and a test of peak expiratory flow rate to compare the dyspnea perceived by her subjects with an objective measure of dyspnea. The ratings were

done on admission and those subjects who demonstrated less than 150 lpm of peak flow were assessed again prior to discharge. The construct validity (does the VAS measure shortness of breath?) was assessed by using t tests between the subjects' mean VAS scores during the times when peak flow was higher (not dyspnoeic) or lower (dyspnoeic). Concurrent validity (do the tests of dyspnoea concur?) was tested between the two VAS ratings and the peak flow measurements using Pearson product moment correlation's. The correlation between vertical and horizontal scales was - 0.95, between vertical scale and peak flow - 0.85 and between horizontal scale and peak flow - 0.71. No tests of significance were offered. Statistically significant differences were recorded in the VAS rated shortness of breath between times of lower and higher peak flow: t = 12.35 for asthmatics ($p < 0.01$) and t = 9.73 for COPD subjects ($p < 0.01$). The vertical VAS was simple to use and valid for measuring and monitoring patients' perceptions of dyspnoea. The horizontal scale was almost as good a measure as the vertical scale.

Noseda et al[48] studied perceptions of SOB in 12 asthmatic and 12 COPD subjects using the VAS after inhalation of the bronchoconstricting agent histamine. The VAS was of the bipolar variety. The subjects' pulmonary functions were also assessed using: forced expiratory volume in one second, specific inspiratory airway resistance, functional residual capacity, and inspired vital capacity. Five of the 12 COPD subjects demonstrated little or no induced change in respiratory function and were classified as "low perceivers". These subjects were found to have poor correlations between VAS ratings and objective measures of lung function. The authors used the Pearson correlation coefficient and a stepwise multiple regression analysis to test relationships between the VAS and the five physical measures. They found that the VAS demonstrated a statistically significant relationship with specific airway resistance ($r = .953$, $p = .022$) and FEV1 ($r = .934$, $p = .002$) in asthmatics and with inspiratory vital capacity ($r = .81$, $p = .01$) and maximal inspiratory flow at 50% forced vital capacity ($r = .805$, $p = .003$) in COPD subjects. Although histamine induces bronchoconstriction it does so inconsistently and these authors could not prove a dose response association. However the association between the VAS-assessed SOB and objective measures of lung function demonstrates the usefulness of the VAS type of measure: in asthmatics the measure is more consistant than those with COPD as the former subjects appeared more homogeneous in their response to shortness of breath.

The same group of investigators[49] studied perceived VAS-measured SOB in 16 asthmatic and 16 COPD out-patient clinic attenders who received inhalations of saline followed by terbutaline. The VAS used was again the bipolar 400 mm variety described above. The investigators measured lung functioning six times, at baseline and once after each of three saline and two terbutaline inhalations. The tests were: specific inspiratory resistance, specific expiratory resistance, functional residual capacity, FEV1, inspired vital capacity, total lung capacity, and the VAS. Eight of the sixteen COPD subjects were classified as "low perceivers" and compared to baseline assessments showed little variation of either lung function or VAS ratings following the inhalations. The investigators used the Spearman correlation coefficient and the Wilcoxon signed rank test to evaluate the significance of concordance between the post-terbutaline and the post-saline inhalations. All rankings listed were statistically significant ($p < 0.05$ with most $p < 0.01$). In the asthmatics and COPD "high perceivers" the specific inspired resistance, specific expiratory resistance, and inspired vital capacity all correlated well with the VAS. Only the asthmatics showed a statistically significant correlation between VAS ratings and functional residual capacity, FEV1 demonstrated a weak, non significant relationship with VAS ratings.

The same group[50] compared terbutaline-induced bronchodilation as measured by the lung function studies described above with decreasing shortness of breath as rated by the bipolar VAS. Their subjects were 36 out-patients, 16 with asthma, 20 with COPD; 2 asthmatic and 7 COPD subjects were "low perceivers".The investigators tested the relationships using the squared correlation coefficient within a linear regression analysis. They found statistically significant relationships between VAS ratings and specific inspiratory resistance in the asthmatic ($r^2 = .831$, $p < 0.01$) and "high perceiver" COPD subjects ($r^2 = .760$, $p < 0.05$). Given the large numbers of lung function tests done and the only slightly smaller number of negative results, one might expect that some positive results might occur by chance alone.

Noseda and colleagues[51] used the bipolar VAS to measure dyspnoea in 19 COPD patients as they exercised in a lab against both progressive and high intensity resistance. Ventilation, as measured by VO2Max, VEMax, and heart rate, correlated well with dyspnoea ($r > 0.90$) in 15 subjects and well enough in the remaining 4 subjects ($r = 0.747$, 0.840, 0.865, and 0.888).

Wilson and Jones[52] also compared the Borg Scale and the VAS to measure dyspnoea while testing the minute ventilation in 10 healthy subjects who were exercised in a lab. The authors found statistically significant relationships ($p < 0.01$) between the minute ventilation and the VAS score ($r^2 = 0.68$) and the Borg score ($r^2 = 0.75$) using analyses of variance and covariance. The VAS scores displayed a wider range than the Borg and the coefficient of determination for the two scales was 0.71 (no tests of significance reported). These authors concluded that both scales measure dyspnoea but suggest the Borg scale seems somewhat better as a test of dyspnoea in healthy subjects.

Gift et al[53] made multiple measurements of VAS-rated SOB, serum cortisol, pO2, pCo2, anxiety, use of accessory muscles, depression, and somatization in 6 male COPD subjects of mean age 64 years. The investigators divided their subjects into high or low dyspnoea perceivers based on their VAS ratings and compared the remaining measurements between these two groups. The t test was used to compare the means of each measurement between low and high dyspnea levels. The differences between the two sub-groups of subjects were statistically significant for measurements of anxiety ($t = 2.8$, $p = .01$), cortisol concentration ($t = 2.6$, $p = 0.02$), depression ($t = 2.1$, $p = .04$), PC02 ($t = 2.1$, $p = .04$), and use of accessory muscles ($t = 5$, $p = .002$); while somatization and p02 were not different. These results indicate a physiologic difference within subjects when short of breath compared to when they were not, and that the VAS can be used to measure these differences.

3.4.3 SUMMARY

Although not as fully studied nor as well validated as the VAS for pain, the VAS for shortness of breath appears to be valid and sensitive to change in healthy, asthmatic, and high perceiver COPD subjects. The one study comparing the VAS to other objective measures of dyspnoea in cancer subjects indicates this measure is useful in this group.

3.5 NAUSEA

Nausea is a subjective experience as well as an observable physiologic phenomenon, with sweating, pallor, and sighing; and

has been difficult to reliably quantify. Morrow[54] reviewed 144 articles on the assessment of nausea noting a vast variety of self and observer rating scales including: binomial ("present / absent" or "yes / no" assessments), 3 -point, 4 - point, 5 - point, 6 - point, 10 - point scales and the 100 mm VAS. Because he felt none of these scales to be sufficient he developed his own which he used to test 20 chemotherapy subjects before and after each of four chemotherapy sessions. He offers acceptable mean test-retest reliability coefficients but did not compare his scale with the 100 mm VAS. He concluded, although he did not offer much statistical testing to strengthen his argument, that self-rating of nausea was quite appropriate.

3.5.1 CANCER-RELATED NAUSEA

By 1990 the competition between the commonly used nausea rating scales had thinned to the VAS, a VAS variant, and discrete descriptor scales. Del Favero et al[55] were involved in four double-blind randomized clinical trials and two observational trials of chemotherapeutic agents and nausea. They compared the self-ratings of nausea offered by 849 subjects using three scales: a 100 mm VAS anchored by "no nausea" and "worst nausea I've ever felt", a 100 mm analogue chromatic continuous scale (descriptively anchored the same as the VAS but graded in colour from pale pink to dark red) and a discrete descriptive scale (none, mild, moderate, severe). The authors reported that 95% of subjects were able to complete the rating scales. Spearman's correlation coefficients were computed for maximum intensity, presence, and quantity of nausea. Correlation of the discrete scale vs VAS revealed the following: for intensity $r = 0.68$, for presence $r = 0.80$, and for quantity $r = 0.95$ (all $p < 0.05$); while correlations between the VAS and the chromatic scale were: for intensity $r = 0.74$, for presence $r = 0.87$ (both $p < .001$), while for quantity the correlation was poor. There appears to be no advantage of one of these three scales over the others.

Bruera et al[56] studied the use of controlled and immediate release metoclopramide in twenty-nine patients with advanced cancer and severe nausea. They used the 100 mm VAS anchored as above and a three point descriptive scale (mild, moderate, and severe) with appropriate amplifiers further describing the categories. They also recorded episodes of vomiting and patient preference between the short or long acting anti-vomiting agent. The patients who had

higher vomiting scores also had higher VAS scores and higher categorical scores but no direct comparisons between the latter two measures were reported.

In Sweden, Borjeson and partners[57] studied the nausea suffered by 104 women with cancer over four chemotherapy sessions. Both the 100 mm VAS and a four category scale (none, mild, moderate, and severe) were used to assess nausea resulting in 348 simultaneous ratings. The authors reported the mean VAS scores for those subjects who rated their nausea on the categorical scale as follows none = 0.7 mm, mild = 24.8, moderate = 48.3, and severe = 75.1. Neither tests of correlation nor association were reported.

3.5.2 NON CANCER-RELATED NAUSEA

In an experimental design, Muth et al[58] invited 50 healthy subjects to sit in an optokinetic rotating drum for sixteen minutes of nausea-inducing gyrations. Using a newly designed descriptor scale (modified from Morrow as described above) and a 300 mm VAS anchored similarly as above, they recorded the self rated nausea of their subjects. Their scale correlated well with the VAS ($r = .71$, $p < 0.01$) while both scales could distinguish those susceptible and those not susceptible to motion sickness ($t = -6.77$, $p < .001$). They concluded their scale had an acceptable construct validity because of its high correlation with their VAS. They did not test their subjects with the 100 mm VAS.

The following two studies are reported here as examples of how the VAS is being currently used to measure nausea in clinical trials. Miguel et al[59] studied the effects of sufentanil given either intravenously or via the epidural route on narcotic requirements of 50 subjects in the post operative phase of abdominal surgery. The investigators used the 100 mm VAS for nausea finding no differences between the two groups. The average nausea score for the epidural group was 10 mm: for the intravenous group, 3 mm. Sohi and associates[60] used the 100 mm VAS to assess post operative nausea of 125 patients undergoing cholecystectomy in a double blind controlled trial on the use of preoperative scopolamine. They found no difference between the drug and control group. The highest mean level of nausea was 25 mm.

3.5.3 SUMMARY

By the mid 1990's the 100 mm VAS was being used as the most common measurement of nausea in studies of anti-emetic agents and of those causing nausea such as cancer chemotherapeutic agents. Despite the lack of correlation and agreement analysis I conclude the VAS appears to be as valid a tool to measure patient experienced nausea as any other that exists.

3.6 DEPRESSION

3.6.1 CANCER-RELATED DEPRESSION

Ahles et al[61] studied 29 out-patients and 11 in-patients who were suffering from cancer related pain. They administered the VAS daily and once for each of the Beck Depression Inventory (a 21 item self report questionnaire) and the Symptom Checklist -90, two commonly used screening tools for depression of mood. The correlation coefficients (not further defined) for VAS vs the Beck Depression Inventory was 0.51 ($p < .05$) and for VAS vs the Symptom Checklist $r = 0.41$ ($p < 0.05$). The authors suggest their data support the validity of the VAS as a practical instrument for the measurement of mood in patients suffering cancer related pain, however the correlation coefficients were not really high enough to support this conclusion with any degree of certainty.

Chochinov and cohorts[62] interviewed 197 in-patients suffering from advanced cancer. They conducted a structured clinical interview to determine the presence or absence of depression and then tested each subject for depression using three methods: the 100 mm VAS, the Beck Depression Inventory-short form, a one item interview (Are you depressed?), and a two item interview (one question each about depressed mood and loss of interest). A total of 24 subjects met the interview criterion for depression. The investigators compared the structured interview with the other four measures. The sensitivity and specificity were found to be: single item interview 1.00 and 1.00, two item interview 1.00 and 0.98, Beck 0.79 and 0.71, and VAS (< 55 mm) 0.72 and 0.50. These results demonstrate that either: the VAS is inferior to the other tools as a measure of mood in patients with advanced cancer, or that depression is not well defined by the VAS score cutoff of 55 mm. Although the authors did not discuss the range of the VAS mood

scores, a cut-off of 55 mm seems quite high. People with advanced disease may rate their mood as low because of the sheer immensity of their illness but not consider themselves clinically depressed.

3.6.2 NON CANCER-RELATED DEPRESSION

Luria[61] studied 62 recent in-patients with functional psychiatric disorders. The subjects were grouped by diagnostic categories and mood was tested with repeated 100 mm VAS measurements, a one time assessment using the Self-rating Depression Scale (multiple choice descriptive scale) and the Clyde Mood Scale (multiple choice scale), a structured clinical interview, and clinical impression of mood while on the ward. Correlation was tested using the Pearson correlation coefficient between the VAS and the Self-rating Depression Scale: $r = -0.56$ ($p < 0.001$) and between the VAS and the "Unhappy" sub scale of the Clyde Mood Scale: $r = -.65$ ($p < 0$ .001). Clinical impression of mood correlated well with VAS scores in non schizophrenic subjects (mean r ranged from 0.61 to 0.56 , $p < .001$). Reliability using a within-patient test-retest methodology showed a mean r of 0.59 for those subjects with affective psychoses while for schizophrenic subjects mean $r = 0.56$ and for personality disordered subjects mean $r = 0.74$ (all $p < .001$). Across-patient test-retest reliability coefficients at 2 hours ranged from 0.73 to 0.91 ($p < 01$). The author concluded the VAS was valid and correlated well with the other two often used scales of mood assessment.

Folstein and Luria[64] studied 133 military personnel recently admitted because of psychiatric or orthopaedic illness and 33 patients hospitalized because of psychiatric illness. Their subjects were tested for mood by a) a 100 mm VAS titled "How is your mood right now?" anchored by "Your worst mood" and "Your best mood", b) the Self Rating Depression Scale, and c) the Clyde Mood Scale. The VAS was completed daily while the latter two scales were administered once each. The VAS scores were compared with each subject's scores on the other two scales. The correlation coefficients for the military personnel group were VAS vs Self Rating Depression Scale $r = -0.64$ ($p < 0.001$) and VAS vs Clyde Unhappy Scale $r = -0.38$ ($p < .01$) and for the hospitalized psychiatric subjects: VAS vs Self Rating Depression Scale $r = -.67$ ($p < .001$) and VAS vs Clyde Unhappy Scale $r = -0.51$ ($p < 0.001$). The VAS appeared to measure true differences in mood between subjects whether they were admitted with psychiatric or orthopaedic diagnoses. The VAS correlated well enough with the other two

scales to indicate its validity as a true reflection of how a subject perceives mood. The subjects in this study were, overall, quite young: mean age of the military group was 22 years, and no gender controls were available.

Rosenberg et al,[65] recruited 472 out-patients who were depressed enough as to require antidepressants. The investigators assessed the mood of their subjects using the 100 mm VAS, the Hamilton Depression Scale, a commonly used test of mood that is administered by a trained operator, and the Clinical Global Impression Scale. Although no formal correlation testing was performed the authors noted their subjects' VAS ratings showed similar improvements to those of the other two scales. This gives only indirect evidence that the VAS measures mood as well as the other two scales.

While studying the effects of danazol on the symptoms of premenstrual tension in 28 women Hahn[66] and colleagues used the VAS and Beck Depression Inventory to assess mood. The authors found statistically nonsignificant differences in mood between placebo and drug group using these scales. No comparisons were made concerning the Beck scale and the VAS. The author's use of the VAS as a complement to the Beck scale indicates their confidence of the VAS but little else. Bloch et al[67] used the VAS, the Clinical Global Impression Scale, and the Hamilton Rating Scale for Depression to assess the mood of 33 depressed out-patients undergoing a trial of antidepressant plus either placebo or lithium. All three scales measured statistically significant ($p < .001$) improvement in mood as the treatment progressed. No formal testing of correlations between tests was made. These authors' findings give indirect evidence that the VAS is as good as the other two scales and indicate the VAS does measure change in mood as depression lifts.

To study the effects of antidepressant medication combined with sleep deprivation in 51 depressed patients, Kuhs et al[68] measured each subject's mood using the VAS and the Hamilton Rating Scale for Depression. The mood scores improved after treatment as measured by the Hamilton scale but not by the VAS. Although no formal correlations were made between the two scales this study casts doubt on the suitability of the VAS to measure changes in mood.

3.6.3 SUMMARY

Overall the body of evidence supports the view that the VAS for depression is able to distinguish between subjects with low or high mood and to measure improvement in the mood of depressed individuals as they respond to anti-depressant therapy.

3.7 ANXIETY

3.7.1 CANCER-RELATED ANXIETY

Miller et al[69] studied the pre-operative anxiety of 44 women about to undergo surgery for breast cancer using the Hospital Anxiety and Depression Scale, the Spielberger State Trait Anxiety Inventory and the 100 mm VAS anchored by "Not anxious at all" and "Most anxious I can imagine." The Pearson product moment correlations for the anxiety component of the scales were: VAS vs Hospital Scale 0.74, VAS vs Spielberger Inventory 0.62 and between the Hospital Scale and Spielberger Inventory 0.81 (all $p < .001$). The mean VAS score for anxiety was 50.3 with a standard deviation of 28.35. This study shows the VAS is acceptable and nearly as good at measuring anxiety as the other two commonly used scales.

3.7.2 NON CANCER-RELATED ANXIETY

In 1975 Bond and Lader[70] used a series of 16 (100 mm) VAS ratings to measure the calming effects of either placebo, butobarbitone, flurazepam 15 mg, or flurazepam 30 mg on the subjective feelings of 500 healthy volunteers. The authors performed a factor analysis of their subjects' 16 VAS ratings at 12, 15 and 18 hours after administration of the medication or placebo. The first factor was not significant but showed a trend while the second and third factors showed statistically significant ($p < .05$) and clear drug effects, indicating that the VAS sub-scales for "Calm/Excited" and "Contented/ Discontented" could measure drug effect. Although these authors did not use the more modern form of VAS their work implies an ability of the VAS to measure anxiety.

Brown[71] studied 66 patients before and after either open surgery, percutaneous nephrolithotomy, or extracorporeal shockwave lithotripsy for renal calculi. She measured anxiety by palmar sweat

response, Spielberger State - Trait Anxiety Inventory (20 statements concerning present anxiety feelings ranked from 1 'not at all' to 4 'very much so'), and the 100 mm VAS anchored at one end by a positive statement and at the other by a negative one. Using a t test a statistically significant decrease in palmar sweat response was noted between the open surgery group pre- and post-op ($p < .001$) coinciding with a similar statistically significant decrease in VAS ratings ($p < 0.05$). There was no such change in the Spielberger scores. There were non-significant reductions in scores for the three rating measures before and after the remaining two surgical treatments. The correlations between sweat responses and the other two measures of anxiety were poor. One conclusion was that palmar sweat responses measured state anxiety while the other two scales measured trait anxiety. This study hints at the use of the VAS to measure anxiety.

Egan et al[72] in 1992 reported on the effects of midazolam vs placebo on pre- and post-operative anxiety of forty women undergoing hysterectomy. The subjects were assessed for anxiety using the Spielberger Inventory and the 100 mm VAS for anxiety anchored by "No worry" and "Worst possible worry". They found lower scores in the midazolam group vs the placebo group as assessed by both scales (ANOVA $p < .01$). They did not report any direct correlations between the two scales. This indicates the VAS can measure drug-induced decrease in anxiety but offers indirect evidence that the VAS correlates with the Spielberger Index.

Frattola et al,[73] studied the effects of an anxiolytic agent vs placebo in a double-blind randomized controlled trial with 40 anxious elderly out-patients. During the course of the study the subjects were repeatedly assessed for anxiety using the Hamilton Rating Scale for Anxiety (an interview style measure), the Spielberger State Trait Anxiety Inventory, and a 100 mm VAS anchored by "I feel alright and perfectly relaxed" and "I feel anxious and awfully tense". All three scales registered a decrease in anxiety scores for the drug vs the placebo groups ($p < 0.0023$ to < 0.016). The mean VAS score differences between the two groups were from 68.2 for placebo to 51.3 for the drug. The authors did not perform direct correlations between the VAS changes and those for the other measures. This study indicates the VAS can measure change in anxiety within subjects.

Scott and Cadden[74] measured anxiety using the 100 mm VAS anchored by "No anxiety" and "As anxious as I can imagine" and

heart rate in 15 healthy volunteers who were part of a study on experimental pain induced by an electrical stimulus to the upper lip. The subjects recorded their anxiety after a block of ten such stimuli. The authors also recorded the reduction of masseter muscle inhibition induced by the electrical stimuli. Using a t test the authors found statistically significant increases in both heart rate and VAS scores after the electrical stimuli as compared with the control period ($p < .00005$). They found a correlation between the reduction in masseter muscle inhibition and heart rate ($r = -0.498$; $p < 0.01$) but not between muscle response and the VAS ($r = -0.196$; $p = 0.3$). The mean VAS scores increased from 16.4 mm to 56.7 mm while the heart rate increased from 70 to 80.7 beats per minute. The authors did not report any correlations between the heart rate and the VAS. This paper indicates that in experimental pain situations the VAS can measure change in anxiety, and correlates acceptably with some of the physical concommitants of anxiety.

Using the Hamilton Rating Scale for Anxiety, the Self-rating Depression Scale and the 100 mm VAS anchored by "Well being" and "Uneasiness" Laakman et al[75] studied the response of amitriptylline, alprazolam, lorazepam, or placebo on the anxiety of 197 depressed out-patients. All three scales measured a decrease in anxiety ($p < 0.05$) in the drug groups vs the placebo group. No correlations between the scales and the VAS were reported. The mean VAS scores decreased between 30 (SD 16) and 34 (SD 27) mm for those subjects treated with active medications. This study offers evidence that the VAS measures change in anxiety within subjects who are depressed but does not indicate how the VAS correlates with other well used anxiety measures.

Penttila et al[76] compared the reduction in pre-operative anxiety induced by either diazepam or triazolam in 81 subjects awaiting elective surgery. The authors measured this reduction using a 50 mm VAS card and the anesthesiologists' evaluation on a 4 point scale (relaxed, minor, moderate, or severe anxiety). The authors used t tests to establish significant reductions in anxiety for the triazolam group from baseline as recorded by the anesthesiologists ($p = .004$) and the VAS ($p = .002$). No correlations of scores between the two scales were recorded. This study indicates the ability of the VAS to measure change in anxiety within subjects.

As part of a larger study on medications to reduce experimental anxiety Hetem et al[77] induced anxiety in 43 healthy subjects by two methods; public speaking (a well known elevator of anxiety)

simulated by talking to a video camera, and by an aversive white noise. The investigators measured heart rate, blood pressure and anxiety using the Spielberger State-Trait Anxiety Inventory and a collection of 16 100 mm VAS measures. Using a MANOVA procedure the authors found both tests measured significant increases in induced anxiety, VAS of anxiety $F = 17.46$, $p < .001$), Spielberg ($F = 37.26$, $p < .001$). They also found increases in systolic ($F = 11.59$, $p < .001$) but not diastolic blood pressure and no significant changes in pulse rate in response to the induced anxiety. This study offers evidence that the VAS measures change in induced anxiety within subjects.

While trying to evaluate the predictors of cooperation during gastrointestinal endoscopy in 251 outpatients, Mahajan and company[47] measured anxiety using five procedures: 1) a single question to the subject about immediate anxiety scored as: not at all, slightly, moderately, or excessively, 2) a modified Hospital Anxiety and Depression Scale, 3) a 100 mm VAS anchored by "No anxiety" and "Maximal anxiety", 4) an endoscopist's rating of cooperation (excellent, good, fair, or poor) and 5) two Patient Comfort Scales (patient satisfaction questionnaires completed 24 hours after the procedure). The Spearman correlation coefficients of the endoscopist's ratings and the various anxiety measures were as follows: single question, rho = 0.139 ($p < 0.03$); VAS rho = 0.147, $p < 0.02$), the modified Hospital Scale, rho = 0.117 (p NS); and Patient Comfort Scales were not statistically significant. The authors found statistically significant, but weak, correlations between the patient's willingness to undergo a repeated endoscopy (part of the Patient Comfort Scale) and the single question ($r = 0.148$, $p < .02$), the VAS ($r = 0.195$, $p < .002$), and the modified Hospital Scale ($r = 0.170$, $p < .007$). These authors' statistical computations show very weak although statistically significant correlations between trained observer ratings and VAS measured anxiety, and between the Patient Comfort Scale and the VAS and may indicate the VAS as an appropriate scale for measuring anxiety between subjects.

3.7.3 SUMMARY

Given that some of the above authors offer acceptable correlations between this scale and other measures, the VAS appears acceptable as a measure of anxiety.

3.8 TIREDNESS

3.8.1 CANCER-RELATED TIREDNESS

Smets et al[78] developed the Multidimensional Fatigue Inventory, a twenty item self-report questionnaire consisting of the following sub-scales: general fatigue, physical fatigue, reduced anxiety, reduced motivation, and mental fatigue, which they compared with a 100 mm VAS anchored by "Not at all tired" and "Extremely tired." They studied 111 cancer out-patients receiving radiotherapy, 357 out-patients suffering from a chronic fatigue state, 481 healthy university students, 158 healthy medical students, 46 healthy junior physicians, and 160 healthy army recruits. They calculated a significant internal consistency of their inventory (Cronbach's alpha < 0.80 in most sub-scales), and significant construct validity (univariate analyses of variance between sub groups of subjects as $p < .001$). They concluded their inventory measures fatigue and can distinguish between subjects based on levels of fatigue. Significant correlation coefficients between all the sub-scales and the VAS were reported including general fatigue ($r = 0.77$, $p < .001$).

Using their inventory, as reported above, and a 100 mm VAS, the same group[79] studied 134 cancer out-patients who were receiving radiotherapy. They calculated significant correlation coefficients between the VAS scores and all the separate sub-scales including general fatigue ($r = 0.83$, $p < 0.001$). Both these works offer strong support for the use of the VAS as a valid screening tool to measure fatigue and to distinguish between subjects based on their level of fatigue.

3.8.2 NON CANCER-RELATED TIREDNESS

Wigers[80] reported on a study of 44 fibromyalgia out-patients who were followed for four and one half years after the invitation to

participate in an exercise program. The author assessed pain, disturbed sleep, depression, and lack of energy using separate 100 mm VAS ratings anchored by "Nothing" and "The worst you have ever experienced," as well as physical activity level (adequate/other), disability pension (full/other), symptom duration, and negative life events (> 2/others). By multiple regression analysis the VAS for lack of energy was compared with activity level (standardized correlation coefficient = 0.6, $p < .0001$), with disability pension (standardized correlation coefficient = 0.3, $p < .01$), and with negative life events (standardized correlation coefficient = 0.3, $p < 0.05$), the regression $r = .58$ while $F = 8.6$. Symptom duration demonstrated poor correlations. The correlation between the VAS and physical activity is moderate indicating the usefulness of the VAS to measure tiredness.

3.8.3 SUMMARY

Although not as well studied as the VAS for pain, nausea, depression, and dyspnoea the VAS for tiredness appears to be an acceptable and valid means of measuring tiredness.

3.9 DROWSINESS

There appear to be numerous articles available wherein the VAS is used to measure sleep either in quantity or duration but these do not truly reveal much about drowsiness. The term drowsiness itself is somewhat confusing when juxtaposed with tiredness: the former might suggest sleepiness or somnolence, while the latter could imply apathy or listlessness. Still, a review of the available literature revealed the vast majority of the research articles that discuss the use of the VAS for drowsiness do so in its context as a measurement tool with very little in the way of justification through referenced validation studies.

3.9.1 CANCER-RELATED DROWSINESS

Faithfull[81] described, in 7 cancer patients who had undergone cranial radiotherapy, the somnolence syndrome consisting of excessive sleep, lethargy and anorexia. The drowsiness was rated using a series of 100

mm VAS scales anchored by the following seven descriptor pairs: "Alert/ Drowsy", "Muzzy/ Clear-headed", "Lethargic/ Energetic", "Attentive/ Dreamy", "Mentally Slow/ Quick-witted", "Strong/ Feeble", and "Interested/ Bored". The subjects also kept daily diaries and gave open ended interviews of their feelings. Because of the small numbers and qualitative nature of the diaries and interviews, quantitative analyses were not possible. She did offer a review of the use of the VAS suggesting that it has been well validated.

3.9.2 SUMMARY

The use of the VAS to quantify drowsiness has not been well studied.

3.10 APPETITE

Rabin and colleagues[82] measured hunger and satiety with a 10 cm VAS, food consumption, and serum glucose and lactate levels in 9 healthy subjects before and after two test meals. Using coefficients of repeatability they found large variations between the mean scores for all their measures before and after the test meals. They concluded that the wide variations of appetite ratings by the VAS should not be interpreted as a failure of the VAS but rather that hunger and satiety were not solely dependant on serum glucose levels.

Indirect evidence of the ability of the VAS to measure change within subjects comes from Beal et al[83] who offered 139 AIDS patients with weight loss either Dronabinol or placebo. Those taking the active agent gained weight, felt better and demonstrated, with the use of t tests, statistically significantly higher scores ($p < 0.015$) for VAS measured appetite.

One must proceed here with caution; the VAS was used to measure the difference of any effect exerted by the active agent over placebo so conclusions about the validity of the VAS being able to measure change in appetite based on the lack of response to placebo could be problematic.

3.10.1 SUMMARY

The use of the VAS to measure appetite has not been well studied.

3.11 WELL BEING

There appears to be little or no research into the use of a single VAS to rate overall well being.

3.12 OVERALL SUMMARY

The VAS is well validated as a tool to measure subjective feelings. Six symptoms: pain, mood, dyspnoea, anxiety, nausea, and tiredness have been well studied and the VAS has been validated as an appropriate measurement tool for their quantification; while the three symptoms of appetite, overall well being and drowsiness are not well enough studied, according to the literature search reported here, to determine if they are accurately rated by the VAS.

3.13 CHOICE OF MEASUREMENT INSTRUMENT

In any study involving sick people a measurement tool must have three qualities: be easy to understand and use, be able to measure what it reports to measure, and be quickly done. As has been demonstrated above the VAS meets these three criteria.

Of the symptoms listed in Table I (page 3) the following nine will be studied as they appear, to the author, to be the most clinically important and are those measured by the Edmonton Symptom Assessment System (ESAS): pain, nausea, tiredness, drowsiness, dyspnoea, appetite, depression, anxiety, and well being.

The instrument used was the Edmonton Symptom Assessment System or ESAS (Section 2.6, page 18) which is comprised of nine 100 mm VAS defining the nine symptoms listed above. Because of its ease of comprehension, the short time taken to complete, the validity of most of the sub-scales, and its clinical application this instrument appeared the most appropriate for this study.

CHAPTER IV

RESULTS OF THE STUDY

4.1 QUALITY CONTROL OF DATA

Quality control was assessed by comparing the measurements of the raw ESAS scores, made by the principal investigator (BE), with those of a second investigator (JH). A random sample was selected of 10 patient\family member\nurse triads, consisting of 9 VAS measures per subject and three subjects per triad for a total of 270 measurements.

Kappa could not be computed using the 100 mm scale because at least one of the cells of the crosstable was empty and SPSS for personal computer could not therefore determine this statistic.

When the raw scores were converted to a ten point scale the Kappa between the two investigators' measurements of the VAS data was high: $k = 0.96$ ($p < 0.001$), indicating that the primary investigator's measurements of the raw VAS data were accurate.

4.2 DEMOGRAPHIC RESULTS

352 patients were recorded in the ward log book, between September, 1997 and February, 1998 as eligible to enter this study, with 100 patient/family/nurse triads completing the protocol. As can be seen in Table IV only 28.4% of apparently eligible patients were enrolled while 17.3% refused and 11.3% were deemed too confused to participate. Of those who did not participate 43% were in the 'other' category meaning they were either too ill to be approached by the charge nurse, were too busy to be interviewed by the investigator, had no family members available, or were discharged before they could be interviewed.

Sixty-four nurses participated as subjects in this study. All except one were female. No other information about these nurses, such as years since graduation, status concerning certification from the Canadian Association of Nursing Oncology, age, or years in present job was recorded. The five charge nurses have all been employed as nurses for at least 15 years each and are considered senior and experienced.

TABLE IV: DISTRIBUTION OF ELIGIBLE AND ENROLLED PATIENTS BY HOSPITAL LOCATION

Ward	Total Eligib.	Con-fusion RN Assess	Refuse To RN	Con-fusion Invest Assess	Refuse To Invest-igator	Other	Total En-rolled
4NA	227	12	10	4	23	116	62
6 East	89	10	11	2	9	26	31
PCU	36	10	5	2	3	9	7
Total	352	32	26	8	35	151	100
%Total	100	9	7.3	2.3	10	43	28.4

The breakdown of this 'other' category by percent of total patients potentially eligible and by hospital ward reveals the following: PCU 25%, 6 East 29%, and 4NA 51%. Many of the patients on 4NA were admitted for the administration of chemotherapy and were consequently often too ill to participate. Once the therapy had finished the patients were quickly discharged and therefore unavailable for interview. Also 4NA is the major provincial referral site for haematological malignancies and many of the patients came from far away without any family members. The breakdown into these three categories: too sick, early discharge, no family present, was not recorded during the data collection.

TABLE V: NATURE OF THE RELATIONSHIP BETWEEN THE FAMILY MEMBER AND THE PATIENT

Family Member	Number of Subjects (n = 100)
Spouse	57
Adult Child	16
Sibling	14
Parent	7
Other Relative	4
Friend	2

TABLE VI: LOCATION OF RESIDENCE OF FAMILY MEMBERS IN RELATION TO THE PATIENT

Location of Residence of Family Member	Number of Subjects (n = 100)
Same House	69
Same City	16
Distant Community	15

The mean age of the patient subjects was 54.9 years with a range of 22 to 88 years, median age was 57 years. Sixty-one percent of the patient subjects were female: 39% male. The ward locations of the patients are shown in Table IV. Table V displays the relationship between the patients and their next of kin, while Table VI indicates where the next of kin lived in relation to the patient. Table VII lists the frequency of diagnoses of the patient subjects.

TABLE VII: TYPE AND FREQUENCY OF MALIGNANCY IN THE STUDY SUBJECTS

TYPE OF MALIGNANCY	FREQUENCY (%)
Leukaemia	24
Colon	17
Lymphoma	14
Breast	11
Lung	8
Stomach	6
Gynaecologic	7
Other G.I.	4
Other	6
Missing	3

4.3 STATISTICAL RESULTS

Table VIII shows the means, standard deviations, and t scores with significance testing, of any differences of the ratings, for each of the ESAS sub-scales done by the patient and family member. At a confidence level of 95% the family members rated the patients' sensation of well-being as worse than did the patients, while no differences were detected for ratings of the remaining 8 symptoms.

TABLE VIII: MEAN E.S.A.S SCORES (in mm) AS RATED BY THE PATIENT AND FAMILY,STANDARD DEVIATIONS, t SCORES, AND STATISTICAL SIGNIFICANCE LEVELS

Symptom	Patient (SD)	Family (SD)	t score	Statistical Significance
Appetite	42.18 (34.82)	45.49 (34.86)	0.99	0.32
Well Being	37.36 (28.17)	44.95 (34.86)	2.32	0.02 *
Tiredness	34.01 (27.49)	37.73 (27.44)	1.53	0.13
Drowsiness	26.98 (26.76)	24.94 (24.91)	-0.92	0.36
Anxiety	26.32 (26.76)	30.77 (27.12)	1.67	0.1
Dyspnoea	21.28 (24.65)	22.96 (26.94)	0.59	0.6
Depression	18.17 (22.79)	20.13 (22.8)	0.76	0.45
Nausea	17.66 (25.27)	14.65 (19.88)	-1.73	0.09
Pain	16.32 (15.22)	19.88 (20.04)	1.78	0.08

* significant difference ($p < 0.05$)

Appendix VIII contains the scatter diagrams, with line of best fit, that show any linear relationship that might exist between the perceptions of the patient and family member and between patient and nurse for each of the nine symptoms assessed. These scatter diagrams illustrate that the statistically significant or non-significant differences, as demonstrated by the t - test, are true differences; and that any agreements, as demonstrated by Cohen's Kappa, are true agreements; and that no other relationship exists between the symptom perceptions of the patient\ family\ nurse triads.

TABLE IX: MEAN E.S.A.S. SYMPTOM SCORES (in mm) AS RATED BY PATIENT AND NURSE, STANDARD DEVIATIONS, t SCORES, AND STATISTICAL SIGNIFICANCE LEVELS

Symptom	Patient (SD)	Nurse (SD)	t score	Statistical Significance
Appetite	42.18 (34.82)	47.21 (24.31)	1.36	0.18
Well Being	37.36 (28.17)	47.45 (21.14)	2.71	0.01 *
Tiredness	34.01 (27.49)	40.12 (23.00)	2.02	0.05*
Drowsiness	26.98 (26.76)	24.67 (24.39)	-0.74	0.46
Anxiety	26.32 (26.76)	32.28 (23.06)	1.79	0.08
Dyspnoea	21.28 (24.65)	12.56 (18.75)	-3.18	0.002*
Depression	18.17 (22.79)	28.44 (22.98)	3.45	0.001*
Nausea	17.66 (25.27)	21.74 (25.94)	1.68	0.1
Pain	16.32 (15.22)	26.92 (23.42)	4.11	0.01*

* significant difference ($p < 0.05$)

Table IX shows the means, standard deviations, and significance of any differences of the ratings for each of the ESAS sub-scales done by the patient and nurse. At a confidence level of 95% the nurses' ratings were different from those of the patient for five of the nine symptoms: over-estimations for the lack of well being, tiredness, pain, and depression; and under-estimation of the patients' perceptions of dyspnoea.

Using the ten point (10 cm) scale, 7 out of the 18 kappa calculations could not be computed using SPSS for personal computer.

TABLE X: AGREEMENT (KAPPA) BETWEEN PATIENT AND FAMILY MEMBER AND BETWEEN PATIENT AND NURSE ON ESAS SCORES FOR A THREE POINT SCALE

Symptom	Kappa (significance) Patient and Family	Kappa (significance) Patient and Nurse
Appetite	0.43 (0.001)	0.05 (0.4)
Well Being	0.19 (0.007)	0.06 (0.38)
Tiredness	0.35 (0.001)	0.10 (0.17)
Drowsiness	0.32 (0.001)	0.21 (0.006)
Anxiety	0.25 (0.005)	0.03 (0.72)
Dyspnoea	0.16 (0.03)	0.19 (0.01)
Depression	0.45 (0.001)	0.08 (0.28)
Nausea	0.43 (0.001)	0.28 (0.001)
Pain	N/C	N/C

See text for definition of the three point scale

Kappa: 0 to 0.35 = poor agreement, 0.35 to 0.5 = moderate agreement, 0.5 to 0.75 = strong agreement

Table X shows the Kappa as calculated using the three point scale and demonstrates that family members agreed with the patient on four of the eight symptoms (appetite, tiredness, depression and nausea) for which this statistic could be calculated, while the nurses showed poor agreement for all of these eight symptoms. The family members did not agree with the patient on the ratings of well being, drowsiness, anxiety, and dyspnoea. Because no patients rated their pain as 'severe' (over 60 mm on the Pain VAS), those cells of the cross-table were left empty which precluded the calculation of the Kappa statistic by SPSS for personal computer.

TABLE XI: AGREEMENT (KAPPA), AND PERCENT AGREEMENT BETWEEN PATIENT AND FAMILY MEMBER AND BETWEEN PATIENT AND NURSE ON ESAS SCORES FOR A BINOMIAL SCALE

Symptom	Kappa (significance) and Percent Agreement Patient and Family		Kappa (significance) and Percent Agreement Patient and Nurse	
Appetite	0.44 (0.001)	71%	0.11 (0.26)	55%
Wellness	0.29 (0.003)	54%	-0.07 (0.44)	45%
Tiredness	0.47 (0.001)	75%	0.11 (0.24)	55%
Drowsiness	0.44 (0.001)	77%	0.17 (0.08)	66%
Anxiety	0.44 (0.001)	75%	0.13 (0.16)	62%
Dyspnoea	0.28 (0.004)	75%	0.11 (0.26)	75%
Depression	0.34 (0.001)	78%	0.05 (0.59)	63%
Nausea	0.56 (0.001)	90%	0.42 (0.001)	82%
Pain	0.3 (0.001)	84%	0.04 (0.6)	71%

See text for definition of the binomial scale

Kappa: 0 to 0.35 = poor agreement, 0.35 to 0.5 = moderate agreement, 0.5 to 0.75 = strong agreement

Table XI displays the results using the binomial scale, the family members' ratings demonstrated moderate or strong agreement with those of the patients on five of the nine symptoms: appetite, tiredness, drowsiness, anxiety, and nausea. The nurses agreed with the patient on the perception of only one of the nine symptoms: nausea.

The family members did not agree with the patient on the ratings of well being, dyspnoea, depression, and pain. Table XI also shows the percent agreement between patient and family member and between patient and nurse. This indicator ranged from 45% to 90%.

TABLE XII: CORRELATION (SPEARMAN'S rho) BETWEEN SYMPTOM RATINGS OF PATIENT AND FAMILY AND BETWEEN PATIENT AND NURSE ON ESAS SCORES FOR A 100 POINT SCALE

Symptom	rho (significance) Patient and Family	rho (significance) Patient and Nurse
Appetite	0.54 (0.001)	0.23 (0.2)
Wellbeing	0.37 (0.001)	-0.2 (0.86)
Tiredness	0.61 (0.001)	0.29 (0.003)
Drowsiness	0.55 (0.001)	0.23 (0.02)
Anxiety	0.49 (0.001)	0.19 (0.06)
Dyspnoea	0.42 (0.001)	0.31 (0.002)
Depression	0.47 (0.001)	0.23 (0.02)
Nausea	0.56 (0.001)	0.37 (0.001)
Pain	0.46 (0.001)	0.28 (0.005)

Rho: 0 to 0.25 = no correlation, 0.25 to 0.50 = moderate correaltion
0.50 to 0.75 = strong correaltion, > 0.75 = very strong correlation

Table XII displays the results of the Spearman's Rank Order Correlation (rho) calculations for the full data set using the 100 mm scale. SPSS for personal computer can calculate rho without the technical problems found with the calculation of Cohen's Kappa; so the whole, unmodified, data can be used for the determination of this statistic. There was strong correlation (rho > 0.50) between patient and family members on the ESAS scores for four of the nine symptoms: appetite, tiredness, drowsiness and nausea, and moderate correlation (rho > 0.25) for the remaining five symptoms: well-being, anxiety, dyspnoea, depression, and pain.

Also Table XII shows that, for the full data set, the nurses' ESAS scores showed four of the nine symptoms: nausea, pain, tiredness, and dyspnoea to have moderate correlations (rho > 0.25) with those of the patient, and none to have strong correlations (rho > 0.50).

TABLE XIII: CORRELATION (SPEARMAN'S rho) BETWEEN SYMPTOM RATINGS OF PATIENT AND FAMILY AND BETWEEN PATIENT AND NURSE ON ESAS SCORES FOR A BINOMIAL SCALE

Symptom	rho (significance) Patient and Family	rho (significance) Patient and Nurse
Appetite	0.51 (0.001)	0.11 (0.3)
Wellbeing	0.30 (0.003)	-0.08 (0.4)
Tiredness	0.48 (0.001)	0.12 (0.2)
Drowsiness	0.44 (0.001)	0.17 (0.08)
Anxiety	0.46 (0.001)	0.14 (0.2)
Dyspnoea	0.29 (0.004)	0.11 (0.3)
Depression	0.35 (0.001)	0.05 (0.6)
Nausea	0.58 (0.001)	0.43 (0.001)
Pain	0.32 (0.001)	0.05 (0.6)

Rho: 0 to 0.25 = no correlation, 0.25 to 0.50 = moderate correaltion
0.50 to 0.75 = strong correaltion, > 0.75 = very strong correlation

Table XIII displays the Spearman's rho correlations between patient and family and between patient and nurse using the ESAS data manipulated to produce scores of a binomial nature: using, as in Table XIII, the VAS cut-off of 40 mm within the 100 mm scale.

The family members' ESAS scores showed strong (rho > 0.50) correlations with those of the patient for the two of the nine symptoms: appetite and nausea, and moderate correlations (rho > 0.25) the remaining seven symptoms. The nurses' ESAS scores showed a moderate correlation (rho > 0.25) with those of the patient for only one of the nine symptoms: nausea, and no strong correaltions.

CHAPTER V: DISCUSSION OF RESULTS

5.1 INTRODUCTION

This chapter will begin with a discussion of recruitment concerns, followed by discussions of demographic matters and statistical considerations. The results of the present study will be compared with the findings of previous studies, those discussed in Chapter I and summarized in Table II (page 6). Table XIV (page 69) shows those same studies referred to in Table II with the addition of statistical methods used.

5.2 RECRUITMENT ISSUES

The charge nurses for each ward kept a study log book into which the names of the patients currently under their care were recorded. Although the investigator presented himself to the ward or spoke to the charge nurse every week-day of the study period not all eligible patients were recorded either because they were admitted and discharged before the charge nurse met them (ie. weekend or short stay admissions, or transfers off the ward soon after admission), or the charge nurse was too busy to approach the patient concerning participation.

Although the numbers of those patients who were missed are not known, 4 North A and the PCU had discharge books, completed after the patient was sent home, which the investigator reviewed in order to ensure all admitted, eligible patients were recorded in the research log books. On 6 E the charge nurse recorded only those patients who were assessed at the bedside; which included those too ill to participate, those who refused, and those who agreed to meet the investigator. The investigator reviewed the patient list, daily, to capture the names of those eligible patients who were over-looked by the charge nurse. As the medical records departments list discharges from the hospital but not transfers to other wards, these departments could not further define the numbers of eligible patients any better than the investigator's methods of recording all eligible subjects.

The recruitment process was open to selection bias in that the charge nurse's primary duty was to protect the privacy and welfare of her often quite ill patients rather than collect subjects for research

purposes. If a patient or family member was upset or angry the nurse would exclude the patient from the study by not asking the patient and family to consider participation. The threshold in clinical nurses for making these judgements will vary among and between nurses, and was not assessed formally during this research. If such patients were angry or upset because of family issues the overall study agreement between patient and family on symptom assessment might be less because, or might be an indication of this conflict.

On some wards on some days all of the patients except those who were markedly confused or in severe distress were invited to participate while on other days the situation was less advantageous to recruitment. The work load of the charge nurses seemed to be the predominant factor here, although this was not formally assessed.

In preparation to begin this study, the investigator calculated, from data supplied by the medical record departments of the two hospital sites, the average weekly discharge rate for eligible patients. That rate was twenty-two patients per week. Over the study period of sixteen weeks, between September, 1997 and February, 1998; the expected number of discharges was 352, the same number as those recorded as eligible. As the expected number of patients was so close to the recorded eligible patients there appeared to be few eligible patients who missed the opportunity to consider participation in this study.

Table IV (page 57) shows that of the 352 patients available for recruitment 151 were recorded as "other" indicating they were, in the opinion of the charge nurses, either; too ill, too distressed, or had families that were too distressed; or just over-looked. The percent breakdown of those eligible, by hospital ward location, of this "other" category reveals the following: PCU 25%, 6E 29%, and 4NA 51%. The intense work load of the charge nurses mitigated against a further breakdown of this "other" category, into the sub-groups of too distressed, no relatives, over-looked, or discharged too soon to be approached.

The high frequency of patients, on 4 NA, in this "other" category deserves mention: Table VIII (page 59) shows that 38% of the study subjects (73% of those subjects from 4NA) were suffering from leukaemia or lymphoma. As 4NA is the provincial referral centre for the treatment of haematological malignancies many of the patients there are repeatedly admitted for the administration of chemotherapeutic agents which are likely to have side effects unpleasant enough to inhibit patients from volunteering to enter a

study such as the present one. By the time such patients begin to recover they are often quickly discharged to make room for other patients waiting for similar treatment. Although many patients were recorded as eligible to participate they were either too ill or were discharged before they could be approached to participate. Also once the first couple of treatments of chemotherapy have been completed the out-of-town patient may feel comfortable enough with the situation to arrive alone for subsequent treatments. In other words, because the family members were not available for interview the patient was excluded.

Some of the family members of eligible patients appeared, in the opinion of the charge nurse, unduly distressed by anger, frustration, or sadness. In such situations the patient was not approached in an attempt to decrease, or at least not inflame, the family distress. The numbers of such situations were not recorded.

Table IV (page 57) demonstrates that as many patients as were studied (100) were either confused (40) or refused (61). This is not surprising as many of the eligible patients were in physical distress, emotionally frail, or quite incapacitated due to advanced disease or as a result of treatment; none of which is conducive to voluntarism in even the shortest of research projects. Of those confused, 32 were so deemed by the charge nurse while 8 more were rejected because of low scores on formal mental status testing.

It is possible that the sicker more distressed patients were not invited to participate based on the feelings of the charge nurses; certainly for pain, nausea, shortness of breath, tiredness, and drowsiness the subjects in the studies listed in Table XIV (page 69) rated their symptoms as more severe than did the subjects in the present study.

The conclusion: that the family members agree with the patient on 6 of the 9 symptoms of the ESAS, may be limited to those patients with low symptom severity ratings. However there was strong agreement between patient and family on the subject of poor appetite where the mean symptom severity rating was relatively high at 42.18 mm.

5.3 DEMOGRAPHIC ISSUES

As shown in Table VI (page 58), 57% of the family members were spouses, which is similar to the 60% to 78% noted in the studies listed in Table II, while the proportion who lived with the patient (Table VII,

page 58) was similar as well.

As shown in Table VIII (page 59) 38% of the subjects were afflicted by haematological malignancies which is much higher than those studies listed in Table II that report on the malignancies suffered by their subjects. This is not a major concern as the object of this study was to compare agreement of symptom perceptions and haematological malignancies create a constellation of symptoms as unpleasant as other malignancies.

Apart from the high proportion of patients with haematological malignancies the subjects in the present study appear comparable with those subjects in previous studies, indicating that the results and conclusions from the present study are applicable to hospitalised patients suffering from cancer.

5.4 STATISTICAL ISSUES

The means and standard deviations of the ESAS data, shown in Tables VIII and IX (pages 59, 60), are listed in descending order. The means are unique to this group of subjects and can vary from those means of other subject cohorts depending on the severity, duration, and stage of the cancer experience. Table XIV (page 69) shows the mean symptom scores, where the VAS was used and where the scores were provided, for those studies listed in Table II. The wide standard deviations of the VAS scores of the present study indicate that the raw scores are not normally distributed, and that the subjects offered a wide range of responses. The standard deviations form the present data are not unlike those found in the studies listed in Tables II and XIV.

Ideally the kappa statistic should have been calculated using the 100 mm scale but, as mentioned above, SPSS for personal computer will not compute this for the present data set. There were no patient ratings of pain above 60 mm so even using the three point scale the kappa statistic could not be calculated for this symptom.

The kappa and rho for the binomial ESAS data were calculated to be almost the same; indicating, perhaps, that both are measuring the same thing. Both tests confirmed the same symptoms on which the family (5 of 9 for kappa and 6 of 9 for rho) and nurses (1 of 9) concorded with the patient. Spearman's rho for the 100 mm raw data is, perhaps, a closer reflection of the true state of things than the rho calculated using the binomial data because of the greater sensitivity of the larger scale

to measure correlation in symptom ratings.

TABLE XIV: STUDIES ON THE RELATIONSHIP BETWEEN CANCER PATIENTS' PERCEPTIONS OF THEIR SYMPTOMS AND THE PERCEPTIONS OF THEIR FAMILY MEMBERS AND NURSES: STATISTICS USED, AND MEAN PATIENT SYMPTOM SCORES

Study	Relation to Patient	Symptom Assessed	Statistical Tests	Pt Mean Symptom Score (SD)	% Agree
Lobchuck	Family	Multiple	Kappa	59 (24) to 34 (17)	44% to 65%
Ferrell	Family	Pain	t-test	53 (29) to 35 (26)	N/A
Clipp	Family	Pain Depress	Correl-ations	N/A	48% 43%
O'Brien	Family	Pain	Kappa	N/A	54%
Madison	Family	Pain	Rho	32.1 (30.8)	22%
Yeager	Family	Pain	t-test	41.7 (40)	N/A
Miaskowski	Family	Pain	t-test	N/A	29.5 %
Ferrell	Family	Pain	t-test	52	N/A
Holmes	Nurse	Multiple	t-test Correl-ations	N/A	6% to 20%
Grossman	Nurse	Pain	Pearson	34	7% to 82%
Eaton	Family Nurse	Multiple	Kappa rho, t-test	16 (15) to 42 (35)	45% to 90%

Although the values for rho and k for the binomial data are virtually the same, correlation does not always mean agreement and, because the purpose of the present study was to measure agreement, the kappa from the binomial data was used to draw the final conclusions about the way cancer patients, their families, and their nurses perceive symptoms. The calculated rho's were used to confirm or buttress the kappa statistical results.

The scatter diagrams found in Appendix VIII would show a slope of 45° and a y intercept of 0 if perfect one-to-one agreement existed. The further away from these ideals the less agreement. The scatter diagrams also show that no other linear relationships exist within those symptoms where the t-test shows statistically significant difference between raters.

5.5 COMPARISONS WITH PREVIOUS STUDIES

As shown in Table XIV the percent agreement between nurse and patient, in previous studies, on symptom rating was between 6% and 82% while Table XI (page 62) shows that for the present study the agreement was between 45% and 82%. For family members the equivalent numbers are between 22% and 65%, for previous studies, and between 54% and 90% for this study. Although the family members in the present study are closer in percent agreement with the patient than has been reported by other authors, percent agreement has little meaning really because the range of percent agreement for the same Kappa can vary widely, as can the range of Kappa for any given percent agreement.

5.6 PAIN

5.6.1 PATIENT AND FAMILY

All of the nine studies listed in Table XIV involving cancer patients and their family members compared the subjects' perceptions of pain. Two researchers (Lobchuck, O'Brien) used the Kappa statistic to assess agreement between family members' and patients' perceptions of pain. The Kappas of 0.31 and 0.24, indicating low levels of agreement, closely correspond to the Kappa of 0.30, using the binomial data set, found in the present study.

Two groups (Madison, Clipp) used correlation statistics to assess concordance of pain perceptions. Rho in the first study was 0.27 for total pain scores; while, in the second, r was 0.55 with 48% agreement, indicating low to moderate agreement which mirrors the findings of the present study where rho was 0.46 for the full data set and 0.32 for the binomial data with 84% agreement.

Five authors (Ferrell 1995, Curtis, Yeager, Miaskowski, Ferrell 1991) reporting on the difference between patient and family perceptions, used the t-test to provide statistical demonstration of this difference. The mean pain VAS scores, scored out of 100 mm, of the patients ranged from 24.1 mm to 59 mm. Three of the studies (221 subject dyads) found statistically significant over-estimation of the patients' pain by the family members; while two studies (101 subject dyads) found statistically significant under-estimation by the family. In the present study the mean patient pain score was 16.32 while the mean family pain score was 19.88, however the t-test demonstrated no statistically significant difference.

The subjects in the present study reported much less pain than did those from other studies indicating, perhaps, that the patients in the present study were not as ill. With poor agreement, conflicting correlations and mixed differences in perceptions reported in previous studies, the low kappa and non significant t-test found in the present study data are not surprising. Taking all the results into consideration, there appears to be no consistent evidence that family members perceive pain the same as the patients.

5.6.2 PATIENT AND NURSE

Two studies compared the ratings of cancer-related pain by the patient and the nurse. Grossman (103 subjects) reported a correlation of 0.46 ($p < 0.0001$) between both groups, the mean pain scores (out of 100 mm) were 34 mm for the patient and 24 mm for the nurses but no tests of significance for this difference were carried out. Holmes (53 subjects) reported that 14 nurses under-estimated, 5 agreed with, and 34 over-estimated the patients' ratings: the t-test was significant ($p < 0.05$) for over-estimation by the nurses. In the present study rho for the full data set was 0.29 ($p < 0.005$) and 0.05 ($p < 0.6$) for the binomial data set. While rating symptoms on a 100 mm scale, the nurses (26.92 mm) over-estimated the patients' (16.32 mm) pain, ($p < 0.01$). Again, as with the family members, there appears to be no consistent evidence that nurses perceive pain the same as the patient.

5.7 SHORTNESS OF BREATH

5.7.1 PATIENT AND FAMILY

Lobchuk (37 subjects) calculated the agreement, using the Kappa statistic, between patient and family for ratings of dyspnoea as 0.41 ($p = 0.01$) indicating moderate agreement; while the mean pain ratings were not significantly different, patient 2.22 and family 2.47 (scored out of 5 using a Likert scale). The equivalent Kappa found in the present study was 0.28 ($p = 0.004$), indicating weak agreement; while the difference in mean symptom ratings, between the patient, 21.28 mm and the family, 22.96 mm (scored out of 100 mm), was not statistically significant. When converted to the same scale the Lobchuk subjects rated their dyspnoea as being twice as severe as the subjects in the present study. Lobchuck calculated the Kappa statistic from a binomial scale between 'low' (1/5 or 2/5) and 'high' (3/5, 4/5, or 5/5) symptom ratings, which was equivalent to the 40 out of 100 cut point used in the present study. The subjects in the Lobchuk study were out-patients with lung cancer who appeared to be sicker than the in-patient subjects from the present study and, unlike the present study subjects, they had family members who agreed with their perceptions of shortness of breath. The findings of the present study are at variance with the Lobchuk study meaning, perhaps, that at lower levels of dyspnoea family members agree less often with the patient than they do when this symptom is more distressing. There is mixed evidence that patients and family members agree on the ratings of this symptom.

5.7.2 PATIENT AND NURSE

No previous studies compared the perceptions of dyspnoea between patient and nurse. No agreement was found between nurses and patients in the present study ($k = 0.11$, $p = 0.26$); while the difference between the mean symptom scores of the patient (21.28 mm) and the nurse (12.56 mm) was found to be statistically significant ($p < 0.002$). The present findings: that the nurses did not agree with, and under-estimated, their patients' shortness of breath may not be applicable to all clinical situations because of diverse nature of the sample and the lack of studies in the literature to confirm these findings.

5.8 NAUSEA

5.8.1 PATIENT AND FAMILY

Lobchuk measured two aspects of nausea and calculated the Kappas between patient and family member as 0.47 ($p < 0.05$) for nausea frequency and 0.26 ($p = 0.08$) for nausea intensity. The raw symptom scores, using a 1 to 5 Likert scale, were 1.78 for the patients on both aspects, and 1.81 (frequency) and 2.06 (intensity) for the family members. Curtis (23 subjects) found a difference between the mean nausea score (out of 100 mm) of the patient (15.3 mm) and of the family members (24.4 mm) that was not statistically significant. In the present study there was moderate agreement ($k = 0.56$, $p < 0.00$), while the mean symptom scores showed a difference between patient (17.66 mm) and family (19.88 mm) that was not statistically significant. Again, when converted to the same scale, the patient subjects in the Lobchuk study rated their nausea as being almost twice as severe as the present study subjects, while the Curtis subjects' ratings were about the same as those of the present study's subjects. On the available evidence, a conclusion that family members and patients agree on symptom ratings of nausea appears justified.

5.8.2 PATIENT AND NURSE

Holmes (53 subjects) found a significant difference between the perceptions of nausea by patients and nurses reporting that 9 nurses under-estimated, 9 agreed with, and 35 over-estimated the patients' nausea ratings. Unfortunately the raw scores were not reported. The kappa found in the present study was 0.42 ($p < 0.00$) while the difference between the mean nausea scores of the patient (17.66 mm) and the nurse (21.74 mm) were not significantly significant. The sample in the Curtis study was poorly described as "53 cancer patients" so the results there might not be as generally applicable as those from the present work. There appears to be mixed evidence that cancer patients and nurses agree on the ratings of nausea.

5.9 DEPRESSION

5.9.1 PATIENT AND FAMILY

Clipp (30 subjects) reported the inter-spouse correlation for depression as 0.53 with; 43% agreement, 46% over-estimation, and 11% under-estimation by the spouses. Curtis (23 subjects) found no differences on

the ratings of 'fun', 'life satisfaction', and 'quality of life' between the patients and family, however this study did not specifically assess depression or changes in mood. Unfortunately, neither study reported the raw scores. In the present study there was poor agreement between patient and family ($k = 0.34$, $p < 0.01$), while the inter-spouse correlation was strong (rho = 0.47 $p < 0.01$) for the complete data set, and fair (rho = 0.35, $p < 0.01$) for the binomial data Percent agreement was 78%. As a trend, there appears to be strong concordance, but poor agreement, between patient and family on the presence of cancer-related depression.

5.9.2 PATIENT AND NURSE

Holmes found a statistically significant difference in mood ratings between patients and their nurses ($p > 0.001$), with 7 nurses under-estimating, 3 agreeing and 43 over-estimating. The raw scores were not reported. The present study's statistically significantly different ($p = 0.001$) depression scores of the patient (18.17 mm) and the nurse (28.44 mm) and Kappa of 0.05 ($p = 0.59$) indicates low levels of agreement between the patients and their nurses on the presence of depression, with strong evidence that the nurses over-estimated this symptom.

5.10 ANXIETY

5.10.1 PATIENT AND FAMILY

No previous studies evaluated the level of agreement or difference in perceptions of anxiety between patient and family members. There was, in the present study, moderate agreement ($k = 0.44$, $p < 0.001$), fair correlation (rho = 0.49, $p < 0.001$) on the whole data set, and fair correlation (rho = 0.46, $p < 0.001$) on the binomial data set. Despite these findings the conclusion that patients and families agree on the presence of anxiety must be made with caution because of the absence of previous evidence.

5.10.2 PATIENT AND NURSE

No previous studies evaluated the level of agreement or difference in perceptions of anxiety between patient and nurse. In the present study, there was poor agreement ($k = 0.13$, $p < 0.16$), and no correlation

(rho = 0.19, $p < 0.06$) for the full data set or (rho = 0.14, $p < 0.20$) for the binomial data set. As well the difference between patient (26.32 mm) and nurse (32.28) anxiety scores were not statistically significant. Again, the over-all conclusion that nurses and patients do not agree on the ratings for anxiety, must be made with caution because of the lack of previous data.

5.11 TIREDNESS

5.11.1 PATIENT AND FAMILY

Lobchuk reported a Kappa of 0.63 ($p < 0.05$) indicating a strong agreement between patients and family members on the ratings of tiredness. Scored out of 5 on a Likert scale, the difference in the mean fatigue scores of 2.95 (patient) and 3.19 (family) were not statistically significant. The equivalent findings from the present study were a moderate agreement between patient and family ($k = 0.47$, $p < 0.01$) and a difference between the mean symptom scores of the patient (34.01 mm) and the family members (37.73 mm) which was not statistically significant. When converted to an equivalent scale the Lobchuk subjects rated tiredness as being twice as severe as did the subjects in the present study. The high Kappas from both studies offer strong evidence that family members perceive tiredness the same as the patients.

5.11.2 PATIENT AND NURSE

Holmes (53 subjects) found difference (p non significant) between nurses' and patients' perception of tiredness that was not statistically significant. Unfortunately the raw scores were not given nor was the breakdown of the nurses' over- or under-estimations. The equivalent findings from the present study included a Kappa of 0.11 ($p = 0.24$) and a significant difference ($p < 0.05$) between the patients (34.01 mm) and the nurses (40.12 mm) indicating the opposite of the Holmes study. There is no consistent evidence to indicate that patients and their nurses agree on the ratings of tiredness.

5.12 DROWSINESS

5.12.1 PATIENT AND FAMILY

Lobchuk evaluated perceptions of insomnia, which may parallel the drowsiness assessed in the present study, and found moderate agreement ($k = 0.43$, $p > 0.05$) with raw scores (out of 5) of 2.22 (patient) and 2.65 (family) which was a statistically significant ($p > 0.05$) difference. Curtis looked at sleep, which may parallel the drowsiness assessed in the present study, and reported a difference between the ratings (scored out of 100) of the patients (33.1 mm) and family members (36.3 mm) which was not statistically significant. The present study found a moderate agreement ($k = 0.44$, $p < 0.01$) and a difference in symptom scores between patient (26.89 mm) and family (24.67 mm) that was not statistically significant. Although the Lobchuk subjects rated insomnia higher than the present subjects rated drowsiness, there was equivalent agreement on symptom ratings between the two studies. If insomnia and drowsiness are equivalent symptoms, the conclusion that family members and patients agree on their perceptions of drowsiness may be justified.

5.12.2 PATIENT AND NURSE

Holmes (53 subjects) assessed the perceptions of sleep, which may parallel the drowsiness assessed in the present study, reporting a statistically significant ($p > 0.05$) difference between nurses and patient on this symptom rating. Twelve nurses under-estimated, 7 agreed with, and 34 over-estimated the patients' perceptions of their sleeping. The present study findings of a low Kappa; 0.17 ($p = 0.08$) and a non significant difference ($p = 0.46$) in symptom ratings between patient (26.98 mm) and nurse (24.67 mm), when compared with the Holmes study, suggests that no conclusions may be drawn on the subject of agreement between nurses and their patients on the perception of drowsiness.

5.13 APPETITE

5.13.1 PATIENT AND FAMILY

Lobchuk found a Kappa of 0.60 ($p < 0.001$) with a non significant difference in appetite ratings (scored out of five) between patients (2.14) and family (2.41), while Curtis reported on a non significant difference between patient (62.0 mm) and family (59.4 mm) on this symptom. The present study found a Kappa of 0.44 ($p < 0.01$) and a

non significant difference ($p = 0.32$) on ratings by patients (42.18 mm) and family (45.49 mm). When converted to equivalent scales the symptom ratings for appetite are about the same between the three studies. The lack of significant difference and the moderate to high Kappas promote the conclusion that family members agree with the patient on the rating of appetite.

5.13.2 PATIENT AND NURSE

Holmes found a significant ($p > 0.001$) difference between patient and nurse on the ratings of appetite with 9 nurses under-estimating, 10 agreeing with, and 34 over-estimating the patients' ratings. Unfortunately the raw scores were not provided. The present study's Kappa of 0.11 ($p = 0.26$) and the non significant ($p = 0.18$) difference between patient (42.18 mm) and nurse (47.21 mm) indicates that no conclusions should be drawn about agreement on appetite ratings between patients and their nurses.

5.14 WELL BEING

5.14.1 PATIENT AND FAMILY

Curtis found a difference that was not statistically significant ($p < 0.71$) between patient (59.1 mm) and family (56.1 mm) on the ratings, scored out of 100 mm, of 'quality of life' which may parallel the symptom of 'well being' of the present study. The low Kappa of 0.29 and the statistically significant ($p = 0.02$) difference between patient (37.36 mm) and family (44.95 mm) in the present study is at variance with the Curtis study and indicates that no conclusions should be drawn on the abilities of families to rate the patients' sense of well being.

5.14.2 PATIENT AND NURSE

There were no studies that assessed the nurses' ratings of their patients' sense of well being. In the present study there was poor agreement ($k = -0.07$, $p = 0.26$) between patient and nurse on the rating of well-being as well as poor correlation ($rho = -0.20$, $p = 0.86$) using the full data set and ($rho = -0.08$, $p = 0.40$) when using the binomial data set. The difference between the mean symptom ratings of the patient (37.36) and the nurse (47.45) was statistically significant ($p < 0.01$), however the conclusion that nurses consistently over-estimate well-

being must be made with caution because of the lack of previous data.

5.15 SUMMARY

5.15.1 PATIENT AND FAMILY

There was agreement, in the present study, between patient and family on the ratings of nausea (k = 0.56), tiredness (k = 0.47), drowsiness (k = 0.44), appetite (k = 0.44), and anxiety (k = 0.44). Except for anxiety, for which there were no previous studies with which to compare, the levels of agreement agreement found in the present study confirm those findings from previous studies.

There was poor agreement, in the present study, between patient and family on the rating of shortness of breath (k = 0.28) which was is at variance with the one previous study that assessed this symptom.

There was not agreement, in the present study, between patient and family on the ratings of pain (k = 0.30), depression (k = 0.34), and well being (k = 0,29). There was one study comparing quality of life that was at variance with the present findings related to well being, while studies comparing ratings of pain offered mixed results.

5.15.2 PATIENT AND NURSE

There was agreement, in the present study, between patient and nurse on the rating of one symptom: nausea (k = 0.42). This finding is at variance with the one previous study that assessed this symptom.

There was over-estimation, in the present study, by the nurses for the symptoms of depression, pain, tiredness, and well being; which confirms the findings of over-estimation found in previous studies for depression. The present findings for pain confirm the mixed levels of agreement, correlation, and lack of difference reported from previous studies. There were no previous studies that assessed ratings of tiredness or well being.

There was under-estimation, in the present study, by the nurses for shortness of breath. There were no previous studies that rated this symptom.

There was neither agreement nor statistically significant difference, in the present study, between patient and nurse on the ratings of anxiety ($k = 0.13$), drowsiness ($k = 0.17$), and appetite ($k = 0.11$). A previous study found over-estimation by the nurses on ratings of drowsiness and appetite, while there were no previous studies which assessed the ratings of anxiety.

5.16 CLINICAL ISSUES

It appears from the analysis of the study results that the family members recognize the patient's distress more closely than do the nurses. Although it is more likely that the nurses are more knowlegable of the meaning and clinical importance of the patient's symptoms than are the family, the family members have a much closer relationship with the patient and therefore are closer in their ratings of the patient's symptoms.

As most of the study data was collected in the afternoon the nurses who completed the ESAS would have had the chance to assess the patient in their usual way. However the family members would have had the chance to sit with the patient and chat for up to many hours before completing the ESAS. Time spent with the patient, immediately prior to completing the ESAS, might be another factor in the difference between family and nurse perceptions of the patient's symptoms.

Doctors who may rarely meet the family members, and who rely on the nursing staff for observations of their hospitalized patients' distress might reflect on the nature of this information. Certainly taking the time to sit and listen to the patient directly would be the best way of all. Nonetheless, doctors working in hospitals have more contact with, and may listen more often to their nurse colleagues than to the patients' family members.

5.17 LIMITATIONS

5.17.1 SAMPLE

The sample, although drawn from consecutive patients admitted to three hospital wards where cancer is treated, may have been biased against those patients and families who displayed high levels of distress. One possible reason for this distress might be pre-existing family pathology exacerbated by the cancer illness; exclusion of such

family-subjects might produce a study sample with more contented patients surrounded by more closely knit families who might therefore be closer in symptom ratings. This sample may only be representative of those hospitalized patients with low levels of symptoms making the results not applicable to those hospitalized cancer patients who are quite ill.

5.17.2 TYPE OF MALIGNANCY

The sample, with its high proportion of patients suffering haematological malignancies, may not be representative of all hospitalized cancer patients. However, as the aim of the study was to assess agreement between the study subjects the nature of the malignancies suffered by the patient-subjects may be less important than the nature of the relationship between patient and family member or between patient and nurse.

5.17.3 LOCATION OF STUDY

The study subjects were hospitalized so the results may not be generally applicable to the out-patient population with cancer.

5.17.4 STATISTICS

Unfortunately the SPSS could not calculate the Kappa on the 100 mm scale so the data catagories were collapsed to a binomial scale (using the 40 mm cut-off, as have other authors) which lowers the strength of the statistical calculation. The calculation of a weighted Kappa was beyond the expertise of the author.

5.18 INSTRUMENT

In reviewing the raw data the investigator noted that one patient, one family member and one nurse (not of the same study triad) rated all symptoms as exactly the same, indicating that either all nine symptoms were perceived as being of equal intensity, or the subjects didn't understand the process, or were too busy or distracted to be bothered with careful completion of the ESAS.

While instructing and observing the study subjects in the completion

of the ESAS the investigator noted all subjects completed the forms within two minutes. Subjects understood the meaning and method of completing the ESAS within five minutes. Those who could not read were helped by investigator who read out the descriptive anchors; no subjects objected to this and were able to complete the ESAS within two minutes. Those subjects who were too weak to write were invited to make an imaginary mark along the line with their finger where the investigator then made the mark with a pencil; no subjects objected to this.

The nurses were instructed, in groups of three or four; for ten to fifteen minutes, in the use of the ESAS before the study began and again briefly before the actual data collection. To preserve nurse confidentiality no records were kept as to which nurse completed which ESAS form; however the investigator noted that rarely did the same nurse become involved as a subject more than twice and most nurses completed only one ESAS form. Practice bias, therefore, did not appear to exist in this study.

CHAPTER VI:

CONCLUSIONS, RECOMMENDATIONS, AND AREAS FOR FURTHER STUDY

6.1 CONCLUSIONS

1) Family members agreed with the patients in their ratings of 5 out of 9 cancer-related symptoms.

2) Nurses agreed with the patient on the rating of 1 out of 9 cancer-related symptoms.

3) Family members over-estimated the patients' ratings of well being.

4) Nurses over-estimated the patients' ratings of well being, tiredness, depression, and pain.

5) Nurses under-estimated the patients' ratings of shortness of breath.

6) Family members were closer, than were the nurses, in their perceptions of the patients' cancer-related symptoms.

6.2 RECOMMENDATIONS

1) Nurse who treat hospitalized patients suffering from a malignancy should obtain information regarding cancer-related symptoms by listening first to the patient. If the patient is not able to communicate with the nurse, the family should be the next source of this information.

2) Because of ambiguity in the meaning, low levels of agreement between the dyads, and lack of evidence from

previous studies, the VAS sub-scale for well being should be deleted from the ESAS.

3) Because of ambiguity in meaning, confusion with the symptom of tiredness, and lack of evaluation of this symptom in previous studies the symptom of drowsiness should be replaced, in the ESAS, by a VAS for sleep. The anchoring statements might read: "Worst possible sleep last night" and "Best possible sleep last night".

6.3 AREAS FOR FURTHER STUDY

This study evaluated a one time use of the ESAS in patients who were not so ill as to refuse the study because of physical or emotional distress. Would once or twice daily assessments show the same levels and distribution of agreement? As the patients became sicker with progression of disease would agreement be the same when the symptom ratings were higher?

Patient outcome and the question of whether using the ESAS to assess patient symptoms is of value in increasing patient comfort? was not addressed in this study. One area of further study could be to regularly record the patients' symptom perceptions using the ESAS and record what comfort measures were instituted based on these scores.

The main findings of the present study: the family members are closer in their perceptions of the patients' distress than are the nurses could be further assessed by a longitudinal study using, perhaps daily, repeated measures by family members patients and nurses.

REFERENCES:

1.Singer PA, MacDonald N. Bioethics for Clinicians: 15. Quality end-of-life Care. Canadian Medical Association Journal 1998; July 28, 159 - 162

2.Cancer Pain Relief and Palliative Care. Technical Report Series 804. Geneva: World Health Organization, 1990

3.3.3.Canadian Cancer Statistics 1996, National Cancer Institute of Canada, Toronto Canada, 1996

4.Curtis EB, Krech R, and Walsh TD: Common Symptoms in Patients With Advanced Cancer, Journal of Palliative Care 1991: 7 (2) 25 - 29

5.Donnelly S, Walsh TD, and Rybicki L The Symptoms of Advanced Cancer: Identification of Clinical and Research Priorities by Assessment of Prevalence and Severity, Journal of Palliative Care 1995: 11 (1) 27 - 32

6.Kristjanson LJ, Sloan JA, Dudgeon D, and Adaskin E, Family Members' Perceptions of Palliative Cancer Care: Predictors of Family Functioning and Family Members' Health. Journal of Palliative Care 1996, 12 (4) 10 - 20

7.Stetz KM, and Hanson WK Alterations in Perceptions of Caregiving Demands in Advanced Cancer During and After the Experience. The Hospice Journal 1992: 8 (3) 21 - 34

8.Grossman SA, Sheilder VR, Sweden K, Mucenski J, and Piantadosi S. Correlation of Patient and Caregiver Ratings of Pain. Journal of Pain and Symptom Management 1991; 6 (2) 53 - 57

9.Foley KM. Pain Assessment and Cancer Pain Syndromes. Pages 148 - 166 In The Oxford Textbook of Palliative Medicine: Doyle D, Hanks GWC, and McDonald N, editors: Published by Oxford University Press, 1993; Oxford, England

10.O'Brien J, and Francis A. The Use of Next of Kin to Estimate Pain in Cancer Patients. Pain 1988 (35) 171 - 178

11.Yeager KA, Miaskowsli C, Dibble SL, and Walhagen M. Differences in Pain Knowledge and Perception of the Pain Experience Between Outpatients With Cancer and Their Family Caregivers. Oncology Nursing Forum 1995; 22 (8) 1235 - 1241

12.Maiskowski C, Zimmer EF, Barrett KM, Dibble SL, and Walhagen M. Differences in Patients' and Family Caregivers' Perceptions of the Pain Experience Influence Patient and Caregiver Outcomes. Pain 1997; 72, 217 - 226

13.Ferrell B, Rhiner M, Cohen MZ, and Grant M. Pain as a Metaphor for Illness Part I: Impact of Cancer Pain on Family Caregivers. Oncology Nursing Forum. 1991; 18 (8) 1303 - 1308

14.Ferrell BR, Ferrell BA, Rhiner M, and Grant M. Family Factors Influencing Cancer Pain Management. Postgraduate Medical Journal 1991 67 (suppl 2) S64 - S69

15.Madison JL, and Wilkie DJ. Family Members' Perceptions of Cancer Pain: Comparisons with Patient Sensory Report and by Patient Psychological Status. Nursing Clinics of North America. 1995; 30 (4) 625 - 645

16.Holmes S, and Eburn E. Patients' and Nurses' Perceptions of Symptom Distress in Cancer. Journal of Advanced Nursing, 1989; 14, 840 - 846

17.Lobchuck MM, Kristjanson L, Degner L, Blood P, and Sloan J. Perceptions of Symptom Distress in Lung Cancer Patients: I. Congruence Between Patients and Primary Family Caregivers. Journal of Pain and Symptom Management, 1997, 14 (3) 136 - 146

18.Ferrell BR, Grant M, Chan J, Ahn C, and Farrell B. The Impact of Cancer Pain Education on Family Caregivers of Elderly Patients. 1995 Oncology Nursing Forum. 1995; 22 (8) 1211 - 1218

19.Clipp EC, and George LK. Patients With Cancer and Their Spouse Caregivers: Perceptions of the Illness Experience. Cancer; 1992 69 (4) 1074 - 1079

20. Curtis AE, Fernsler JI. Quality of Life of Oncology Hospice Patients: A Comparison of Patient and Primary Caregiver Reports. Oncology Nursing Forum 1989, Vol 16, No 1, 49 - 53

21. Robertson D, Rockwood K, Stolee P. A Short Mental Statue Questionaire. Canadian Journal on Aging; 1982, 1, 16 - 20

22.Bruera E, Kuehn N, Miller MJ, Selmser P, and MacMillan. The Edmonton Symptom Assessment System (ESAS): A Simple Method for the Assessment of Palliative Care Patients. Journal of Palliative Care 1991; 7 (2) 6 - 9

23.Coulton T, Statistics in Medicine Little, Brown and Company, 1974

24.Norman GR and Streiner DL. PDQ Statistics. D.C. Becker Inc. 1986

25.Daniel WW. Biostatistics, A Foundation for Analysis in the Health Sciences. Sixth Edition. John Wiley and Sons.

26.Freyd M. Journal of Educational Psychology 1923; 14, 83 -

27.Aitken RCB. Measurement of Feelings Using Visual Analogue Scales. Proceedings of the Royal Society of Medicine. 1969; 62, 989 - 993

28.Bond A, and Lader M. The Use of Analogue Scales in Rating Subjective Feelings. British Journal of Medical Psychology. 1974; 47, 211 - 218

29.Maxwell C. Sensitivity and Accuracy of the Visual Analogue Scale: A Psychophysical Classroom Experiment. British Journal of Clinical Pharmacology. 1978; 6, 15 - 24

30.Downie WW, Leatham PA, Rhind VM, Pivkup ME, and Wright V. The Visual Analogue Scale in the Assessment of Grip Strength. Annals of Rheumatic Diseases 1978; 37, 382 - 384

31.Ohnhaus EE, and Adler R. Methodological Problems in the Measurement of Pain: A Comparison Between the Verbal Rating Score and the Visual Analogue Scale. Pain 1975; 1, 379 - 384

32.Graham C, Bond SS, Gerkovich MM, and Cook MR. Use of Pain Questionnaire in the Assessment of Cancer Pain: Replicability and Consistency. Pain 1980; 8, 377 - 387

33.Wallenstein SL, Heidrich G, Kaiko R, and Houde RW Clinical Evaluation of Mild Analgesics: The Measurement of Clinical Pain British Journal of Pharmacology 1980; 10, 319S - 327S

34.Tursky B. The Development of a Pain Perception Profile: A Psychological Approach. In Pain: New Perspectives in Therapy and Research. Editors Weinsenberg M, and Tursky B. 1973; 171 - 194 New York, Plenium Press

35.Kremer E, Atkinson JH, and Ignelzi RJ Measurement of Pain: Patient Preference Does Not Confound Pain Measurement Pain 1981; 10, 241 - 248

36.Fishman B, Pasternack S, Wallenstein SL, Houde RW, Holland JC, and Foley KM. The Memorial Pain Assessment Card, A Valid Instrument for the Evaluation of Cancer Pain Cancer 1987; 60, 1151 - 1158

37.Price D, McGrath PA, Rafii A, and Buckingham B. The Validation of Visual Analogue Scales as Ratio Scale Measures for Chronic and Experimental Pain. Pain 1983; 17, 57 - 65

38.Littman GS, Walker BR, and Schneider BE. Reassessment of Verbal and Visual Analog Ratings in Analgesic Studies. Clinical Pharmacology and Therapeutics. 1985; 38 (1) 16 - 23

39.Revill SI, Robinson M, and Hogg MIJ. The Reliability of a Linear Analogue for Evaluating Pain. Anesthesia 1976; 31, 1191 - 1198

40.Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, and Anderson JA. Studies with Pain Rating Scales. Annals of the Rheumatic Diseases. 1978; 37, 378 - 381

41.Reading AE. A Comparison of Pain Rating Scales. The Journal of Psychosomatic Research 1980; 24, 119 - 124

42.Linton SJ, and Gotestam A Clinical Comparison of Two Pain Scales: Correlation, Remembering Chronic Pain, and a Measure of Compliance Pain 1983; 17, 57 - 65

43.Jensen M, Karoly P, and Bracer S. The Measurement of Pain Intensity: a Comparison of Six Methods Pain 1986; 27, 117 - 126

44.Hurst NP, Kind P, Ruta D, Hunter M, and Stubbings A. Measuring Health-Related Quality of Life in Rheumatoid Arthritis: Validity, Responsiveness and Reliability of EUROQOL (EQ-50) British Journal of Rheumatology 1997; 36, 551 - 559

45.Beattie W, Warriner CB, Etches R, Badner NH, Parsons D, Buckley N, Chan V and Girard The Addition of Continuous Intravenous Infusion of Ketorolac to a Patient-Controlled Analgesic Morphine Regime Reduced Postoperative Myocardial Ischemia in Patients Undergoing Elective Total Hip or Knee Arthroplasty Anesthesia and Analgesia 1997; 84, 715 - 722

46.Booth S, Kelly MJ, Cox NP, Adams L, and Guz A. Does Oxygen Help Dyspnea in Patients With Cancer? American Journal of Respiratory and Critical Care Medicine. 1996; 153: 1515 - 1518

47.Gift AG. Validation of a Vertical Visual Analogue Scale as a Measure of Clinical Dyspnoea. Rehabilitation Nursing 1989; 15 (6) 323 - 325

48.Noseda A, Schmerber J, Prigogine V, deMaertelaer V, and Yernault JC. Perception of Dyspnoea During Acute Changes in Lung Function in Patients with Either Asthma or COPD. Respiratory Medicine 1995; 89: 477 - 485

49.Noseda A, Schmerber J, Prigogine T, and Yernault JC. Perceived Effect on Shortness of Breath of an Acute Inhalation os Saline or Terbutaline: Variability and Sensitivity of a Visual Analogue Scale in Patients with Asthma or COPD. European Respiratory Journal 1992; 5: 1043 - 1053

50.Noseda A, Schmerber J, Prigogine T, and Yernault JC. How Do Patients With Either Asthma or COPD Perceive Acute Bronchodilation. European Respiratory Journal 1993; 6: 636 - 644

51.Noseda A, Carpiaux J-P, Schmerber J, and Yernault JC. Dyspnoea Assessed by Visual Analogue Scale in Patients with Chronic Obstructive Lung Disease During Progressive and High Intensity Exercise . Thorak 1992; 47: 363 - 368

52.Wilson RC, and Jones PW. A Comparison of the Visual Analogue Scale and Modified Borg Scale for the Measurement of Dyspnoea During Exercise. Clinical Science 1989; 76: 277 - 282

53.Gift AG, and Cahill CA. Psychophysiologic Aspects of Dyspnea in Chronic Obstructive Pulmonary Disease: A Pilot Study. Heart and Lung. 1990; 19 (3): 252 - 257

54.Morrow GR. The Assessment of Nausea and Vomiting. Past Problems, Current Issues, and Suggestions For Further Research. Cancer 1984 May 15 Supplement: 2267 - 2278

55.Del Favero A, Roila F, Basurto C, Minotti V, Ballatori E, Patoia L, Tonato M, and TognoniG. Assessment of Nausea. European Journal of Clinical Pharmacology 1990; 38: 115 - 120

56.Bruera ED, MacEachern TJ, Spachynski KA, LeGatt DF, MacDonald RN, Babul N, Harsanyi Z, and Drake AC. Comparison of the Efficacy, Safety, and Pharmacokinetics of Controlled Release and Immediate Release Metoclopramide for the Management of Chronic Nausea in Patients with Advanced Cancer. Cancer 1994; 74 (12): 3204 - 3211

57.Borjeson S, Hursti TJ, Peterson C, Fredikson M, Furst AJ, Avall-Lindvist and Steineck G. Similarities and Differences in Assessing Nausea on a Verbal Category Scale and a Visual Abalogue Scale. Cancer Nursing 1997; 20 (4): 260 - 266

58.Muth ER, Stern RM, Thayer JF, and Koch KL. Assessment of the Multiple Dimensions of Nausea: The Nausea Profile. Journal of Psychosomatic Research. 1996; 40 (5): 511 - 520

59.Miguel R, Barlow I, Morrell M, Scharf J, Sanusi D, and Fu E. A Prospective Randomized Double-Blind Comparison of Epidural and Intravenous Sufentantil Infusions. Anesthesiology 1994; 81: 346 - 352

60.Sohi HS, Heipel J, Inman KJ, Chinnick B, Cunningham DG, Holliday RL, and Girotti MJ. Preoperative Transdermal Scopolamine Does Not Reduce the Level of Nausea and Frequency of Vomiting after Laparoscopic Cholecystectomy. Canadian Journal of Surgery 1994; 37 (4): 307 - 312

61.Ahles TA, Ruckeschel JC, and Blanchard EB. Cancer Related Pain II. Assessment With Visual Analogue Scales. Journal of Psychosomatic Research 1984; 28 (2): 121 - 124

62.Chochinov HM, Wilson KG, Enns M, and Lander S. "Are You Depressed?" Screening for Depression in the Terminally Ill.

63.Lauria RE. The Validity and Reliability of the Visual Analogue Mood Scale. Journal of Psychiatric Research. 1975; 12: 52 - 57

64.Folstein MF, and Lauria RE. Reliability, Validity, and Clinical Application of the Visual Analogue Mood Scale. Psychological Medicine 1973; 3: 479 - 486

65.Rosenberg C, Damsbo N, Fuglum E, Jacobson LV, and Horsgard S. Citalopram and Imipramine in the Treatment of Depressive Patients in General Practice. A Nordic Multicentre Clinical Study. International Clinical Pharmacology 1994; 9 Supplement (1): 41 - 48

66.Hahn PM, Van Vugt DA, and Reid RL. A Randomized Placebo Controlled Crossover Trial of Danazol for the Treatment of Premenstrual Syndrome. Psychoneuroendocrinology 1995; 20 (2): 193 - 209

67.Bloch M, Schwartzman Y, Bonne O, and Lerer B. Concurrent Treatment of Nonresistant Major Depression With Desipramine and Lithium: A Double Blind Placebo Controlled Study. Journal of Clinical Psychopharmacology 1997; 17 (1): 44 - 48

68.Kuhs H, Farber D, Borgstadt S, Moresk S, and Tolle R. Amitriptyline in Combination With Repeated Late Sleep Deprivation Versus amitriptyline Alone in Major Depression. A Randomized Study.

69.Miller K, Jelicic M, Bonke B, and Asbury AJ. Assessment of Preoperative Anxiety: Comparison of Measures in Patients Awaiting Surgery for Breast Cancer. British Journal of Anesthesia 1995; 74: 180 - 183

70.Bond A, and Lader M. The Use of Analogue Scales in Rating Subjective Feelings. British Journal of Medical Psychology 1974; 47: 211 - 218

71.Brown SM. Quantitative Measurement of Anxiety in Patients Undergoing Surgery for Renal Calculus Disease. Journal of Advanced Nursing 1990; 15: 962 - 970

72.Egan KJ, Ready LB, Nessly M, and Greer BE. Self Administration of Midazolam for Postoperative Anxiety: A Double Blind Study. Pain 1992; 49: 3 - 8

73.Frattola L, Poilti R, Bassi S, Albizzati MG, Cesana BM, Bottani MS, Priore P, Borghi C, and Morselli PL. Effects of Alpidem in Anxious Elderly Outpatients: A Double Blind Placebo controlled Trial. Clinical Neuropharmacology ;15 (6): 477 - 487

74.Scott AJ, and Cadden SW. Suppression of an Inhibitory Jaw Reflex by the Anticipation of Pain in Man. Pain 1995; 125 - 131

75.Laakman G, Faltermaier-Temizel M, Bossert-Zaudig S, Baghai T, and Lorkowski. Treatment of Depressive Outpatients with Lorazepam, Alprazolam, Amitriptyline, and Placebo. Psychopharmacology 1995; 120: 109 - 115

76.Pentilla HJ, Nuutinen LS, Kiviluoma KT, and Partanen AM. Sublingual Triazolam Versus Peroral Diazepam as a Premedication for General Anesthesia. Canadian Journal of Anesthesia 1995; 42 (10): 863 - 868

77.Hetem LAB, deSousa CJ, Guimaraes FS, Zuardi AW, and Graeff FG. Effects of d-Fenfluramine on Human Experimental Anxiety. Psychopharmacology 1996; 127: 276 - 282

78.Smets EMA, Garssen B, Bonke B, and de Haes JCJM. The Multidimensional Fatigue Inventory (MFI): Psychometric Qualities of an Instrument to Assess Fatigue. Journal of Psychosomatic Research 1995; 39 (5): 313 - 325

79.Smets EMA, Garssen B, Cull A, and de Haes JCJM. Application of the Multidimensional Fatigue Inventory (MFI-20) in Cancer Patients Receiving Radiotherapy. British Journal of Cancer 1996; 71: 241 - 245

80.Wigers SH. Fibromyalgia Outcome: The Predictive Values of Symptom Duration, Physical Activity, Disability, Pension, and Critical Life Events - A 4.5 Year Prospective Study

81. Faithfull S, Patients' Experiences Following Cranial Radiotherapy: a Study of the Somnolence Syndrome. Journal of Advanced Nursing. 1991; Aug; 19 (8): 939 - 946

82.Raben A, Tagliabue A, and Astrup A. The Reproducibility of Subjective Appetite Scores. British Journal of Nutrition. 1995; 73 (4): 517 - 530

83.Beal JE, Olson r, Laubenstein L, Morales JO, Bellman P, Yangco B, Lefkowitz L, Plasse TF, Shepard KV. Dronabinol as a Treatment for Anorexia Associated with Weight Loss in Patients with AIDS. Journal of Pain and Symptom Management 1995; 10 (2) 89 - 97

APPENDIX I

THE EDMONTON SYMPTOM ASSESSMENT SYSTEM

(ESAS)

Name: _____

Room #: _____ Time: _____

Please cross the line at the point that best describes: *(For coding)*

No pain	_____	Worst possibl pain
Not tired	_____	Worst possibl tiredness
Not nauseated	_____	Worst possibl nausea
Not depressed	_____	Worst possibl depression
Not anxious	_____	Worst possib anxiety
Not drowsy	_____	Worst possib drowsiness
Best possible appetite	_____	No appetite
Best possible sensation of wellbeing	_____	Worst possi sensation o wellbeing
No shortness of breath	_____	Worst poss shortness c

_____

Assessed by: _____

APPENDIX II

PATIENT CONSENT FORM

FACULTY OF MEDICINE
MEMORIAL UNIVERSITY OF NEWFOUNDLAND
ST. JOHN'S, NEWFOUNDLAND A1B 3V6

CONSENT TO PARTICIPATE IN BIO-MEDICAL RESEARCH

TITLE: Symptom assessment of hospital patients: a comparison of rating scores supplied by the patient, the family, and the nurse.

INVESTIGATOR: Dr. Bill Eaton
Department of Family Practice
Health Science Centre
737 6744

You, your nurse and one of your family, have been asked to participate in a research study. Your participation in this study is entirely voluntary. You may decide not to participate or you may withdraw from the study at any time without affecting your usual treatment.

Confidentiality of information concerning you will be maintained by Dr. Eaton.

Dr. Eaton will be available during the study should you have any problems or questions about the study.

This study will compare your evaluation of your symptoms (how you are feeling) with the evaluation done by one of your family and by your nurse. Other researchers have found that nurses, patients, and family tend to evaluate the patient's symptoms differently. By finding the special ways that patient, family, and nurses evaluate the patient's symptoms the investigator hopes to develop strategies to assist patient and family cope with serious illness.

1. **Purpose of study**: To compare your estimation of your symptoms with the estimations of your symptoms done by your nurse and by a close family member.

2. **Description of procedures and tests:** You, your next of kin or close family member, and your nurse will be asked to complete a simple form that rates nine symptoms commonly suffered by people with serious illness such as cancer.

3. **Duration of subject's participation:** Less than one half hour to discuss the study, sign a consent form and complete the symptom assessment form.

4. **Forseeable risks, discomforts, or inconveniences:** Completion of the symptom rating form take between 20 and 90 seconds.

5. **Benefits the subject may receive:** There are no guaranteed benefits from participation.

6. **Alternative procedures or treatment for those not entering the study:** Patient care will not be affected whether or not you choose to enter the study.

7. **Liability statement.** **"Your signature on this form indicates that you have understood to your satisfaction the information regarding your participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the investigator, sponsors, or involved institutions from their legal and professional responsibilities."**

I,_____, the undersigned, agree to my participation in the study of symptom assessment and to the participation of my next of kin or close family member _____(name of family member), and to the participation of my nurse.

Any questions have been answered and I understand what is involved in the study. I realize that participation is voluntary and that there is no guarentee that I will benefit from my involvement. I acknowlege that a copy of this form has been given to me.

_____ _____
(Signature of Participant) (Date)

_____ _____
(Witness Signature) (Date)

To be signed by investigator

To the best of my ability I have fully explained to the subject the nature of this research study. I have invited questions and provided answers. I believe that the subject fully understands the implications and voluntary nature of the study.

_____ _____
(Signature of Investigator) (Date)
Telephone number: 737 6477

To Be Signed By A Physician
(not a co-investigator or otherwise involved in this trial)

"I believe that this subject fully understands the implications and voluntary nature of the study and is competent to enroll in this research study."

_____ _____
(signature of second physician) **(Date)**

APPENDIX III

FAMILY MEMBER CONSENT FORM

FACULTY OF MEDICINE
MEMORIAL UNIVERSITY OF NEWFOUNDLAND
ST. JOHN'S, NEWFOUNDLAND A1B 3V6

CONSENT TO PARTICIPATE IN BIO-MEDICAL RESEARCH

TITLE: Symptom assessment in hospital patients: a comparison of rating scores supplied by the patient, the family, and the nurse.

INVESTIGATOR: Dr. Bill Eaton
Department of Family Practice
Health Science Centre
737 6744

You and your family member, presently a patient in this hospital, have been asked to participate in a research study. Your participation in this study is entirely voluntary. You may decide not to participate or you may withdraw from the study at any time without affecting the usual treatment.

Confidentiality of information concerning you will be maintained by Dr. Eaton.

Dr. Eaton will be available during the study should you have any problems or questions about the study.

If you agree you will be asked to evaluate the symptoms suffered by your family member by filling out a short form. The nurse caring for your family member will also complete a similar form. The investigator will compare your evaluation with those done by your family member and by the nurse. Other researchers have found that nurses, patients, and family tend to evaluate the patient's symptoms differently. By finding the special ways that patient, family, and nurses evaluate the patient's symptoms the investigator hopes to develop strategies to assist patient and family cope with serious illness.

1. **Purpose of study**: To compare your estimation of the patient's symptoms with the estimations of these symptoms done by the patient and by the nurse..

2. **Description of procedures and tests:** You, your family member presently a patient in this hospital, and the nurse will be asked to complete a simple form that rates nine symptoms commonly suffered by people with serious illness such as cancer.

3. **Duration of subject's participation:** Less than one half hour to discuss the study, sign a consent form and complete the symptom rating form.

4. **Forseeable risks, discomforts, or inconveniences:** Completion of the symptom rating form take between 20 and 90 seconds.

5. **Benefits the subject may receive:** There are no guaranteed benefits from participation.

6. **Alternative procedures or treatment for those not entering the study:** Patient care will not be affected whether or not you choose to enter the study.

7. **Liability statement.** **"Your signature on this form indicates that you have understood to your satisfaction the information regarding your participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the investigator, sponsors, or involved institutions from their legal and professional responsibilities."**

I,_____, the undersigned, agree to my participation in the study of symptom assessment as described.

Any questions have been answered and I understand what is involved in the study. I realize that participation is voluntary and that there is no guarentee that I will benefit from my involvement. I acknowlege that a copy of this form has been given to me.

_____ _____

(Signature of Participant) (Date)

_____ _____

(Witness Signature) (Date)

To be signed by investigator

To the best of my ability I have fully explained to the subject the nature of this research study. I have invited questions and provided answers. I believe that the subject fully understands the implications and voluntary nature of the study.

_____ _____

(Signature of Investigator) (Date)

Telephone number: 737 6477

APPENDIX IV

NURSE CONSENT FORM

FACULTY OF MEDICINE
MEMORIAL UNIVERSITY OF NEWFOUNDLAND
ST. JOHN'S, NEWFOUNDLAND A1B 3V6

CONSENT TO PARTICIPATE IN BIO-MEDICAL RESEARCH

TITLE: Symptom assessment in cancer patients: a comparison of rating scores supplied by the patient, the family, and the nurse.

INVESTIGATOR: Dr. Bill Eaton
Department of Family Practice
Health Science Centre
737 6744

You, a nurse in this hospital, have been asked to participate in a research study. Your participation in this study is entirely voluntary. You may decide not to participate or you may withdraw from the study at any time without affecting the usual treatment.

Confidentiality of information concerning you will be maintained by Dr. Eaton.

Dr. Eaton will be available during the study should you have any problems or questions about the study.

If you agree you will be asked to evaluate, by filling out a short form, the symptoms suffered by some of your cancer patients if they agree to take part in this study. The investigator will compare your evaluation with those done by the patient and by a family member. Other researchers have found that nurses, patients, and family tend to evaluate the patient's symptoms differently. By finding the special ways that patient, family, and nurses evaluate the patient's symptoms the investigator hopes to develop strategies to assist patient and family cope with serious illness.

1. **Purpose of study**: To compare your estimation of the patient's symptoms with the estimations of these symptoms done by the patient and by a family member.

2. **Description of procedures and tests:** You, your patient and a close family member will be asked to complete a simple form that rates nine symptoms commonly suffered by people with serious illness such as cancer.

3. **Duration of subject's participation:** Less than one half hour to discuss the study, sign a consent form and complete the symptom rating form.

4. **Forseeable risks, discomforts, or inconveniences:** Completion of the symptom rating form take between 20 and 90 seconds.

5. **Benefits the subject may receive:** There are no guaranteed benefits from participation.

6. **Alternative procedures or treatment for those not entering the study:** Patient care will not be affected whether or not you choose to enter the study.

7. **Liability statement.** **"Your signature on this form indicates that you have understood to your satisfaction the information regarding your participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the investigator, sponsors, or involved institutions from their legal and professional responsibilities."**

I,_____, the undersigned, agree to my participation in the study of symptom assessment as described.

Any questions have been answered and I understand what is involved in the study. I realize that participation is voluntary and that there is no guarentee that I will benefit from my involvement. I acknowlege that a copy of this form has been given to me.

_____ _____
(Signature of Participant) (Date)

_____ _____
(Witness Signature) (Date)

To be signed by investigator

To the best of my ability I have fully explained to the subject the nature of this research study. I have invited questions and provided answers. I believe that the subject fully understands the implications and voluntary nature of the study.

_____ _____
(Signature of Investigator) (Date)
Telephone number: 737 6477

nurse 3

APPENDIX V

HUMAN INVESTIGATION COMMITTEE APPROVAL

APPENDIX VI

HEALTH CARE CORPORATION OF ST. JOHN'S APPROVAL

APPENDIX VII

CANADIAN MENTAL STATUS EXAMINATION FORM

THE CANADIAN MENTAL STATUS QUESTIONAIRE

(MSQ)

This test was developed by Robertson and Rockwood in Saskatchewan in the early 1980's and has been validated to assess cognitive impairment. Scores below 2/10 indicate sever cognitive impairment while scores between 2/10 and 7/10 indicate moderate impairment. Those who score above 7/10 are either not impaired or mildly impaired. This test is used because it is very quick and non-intrusive. The test is given orally.

MSQ

1) Full name

2) Address

3) Age

4) Present day

5) Present month

6) Present year

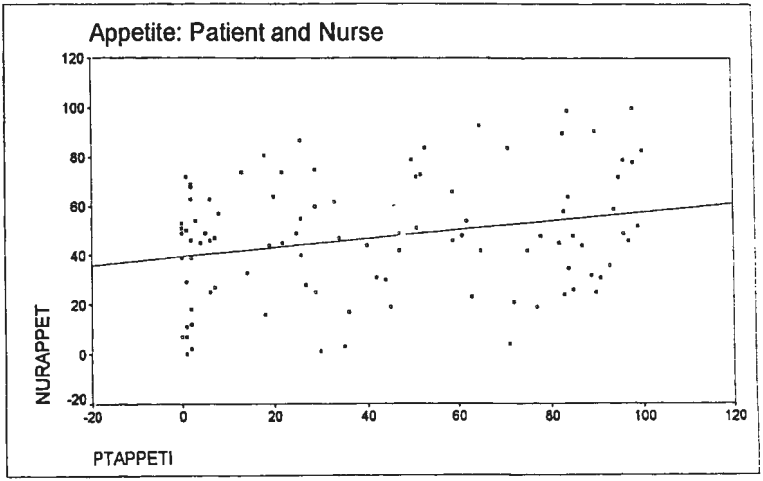
7) Year WW I began

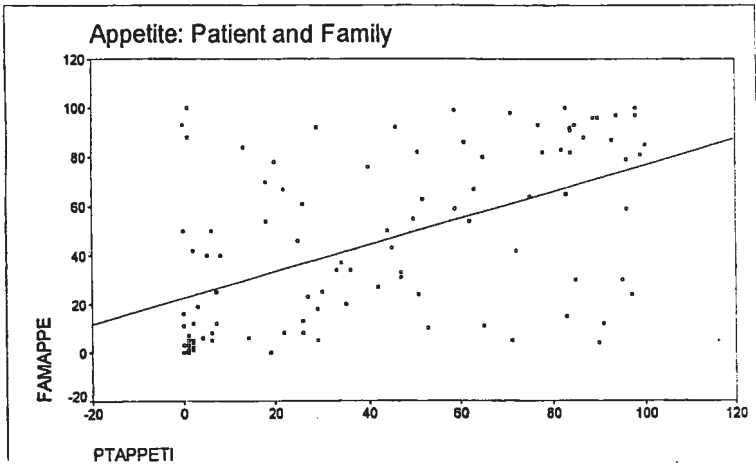
8) Name of Canada's Prime Minister

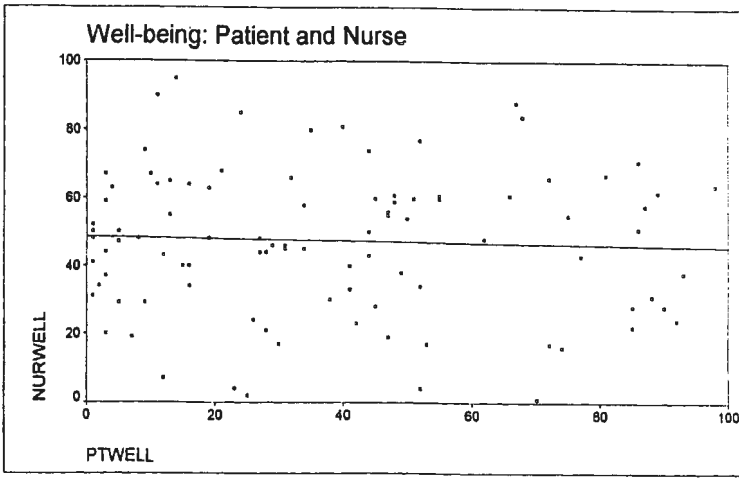
Ask subject to listen to and repeat and recall three items:
Bed, Chair, Window

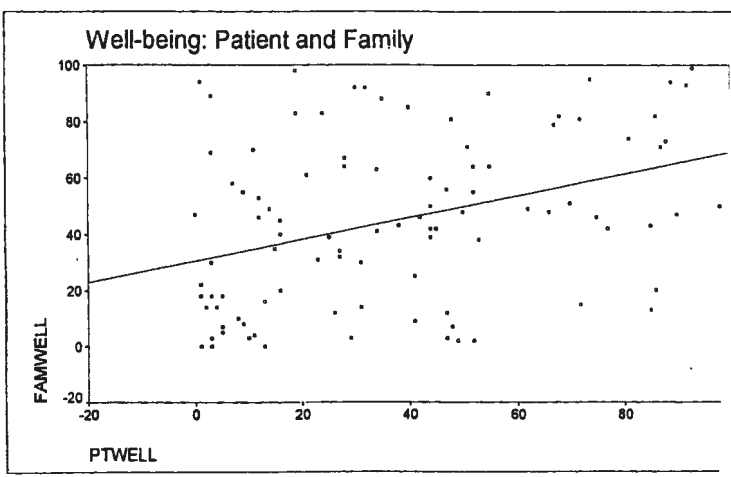
9) Count backwards from twenty to zero

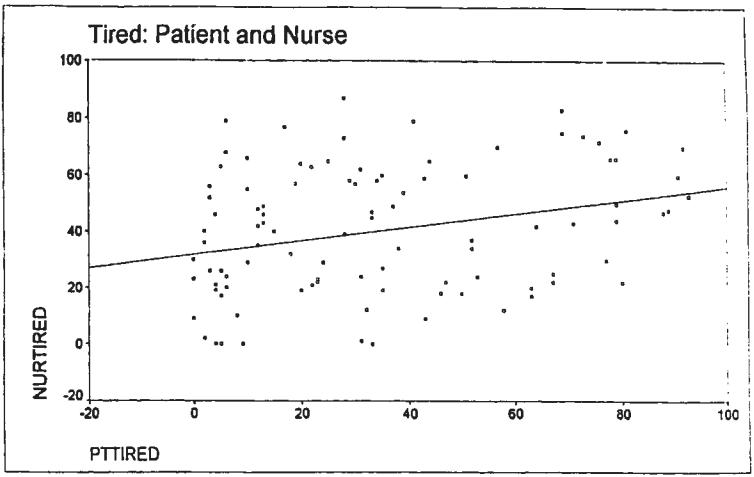
10) Recall the three items

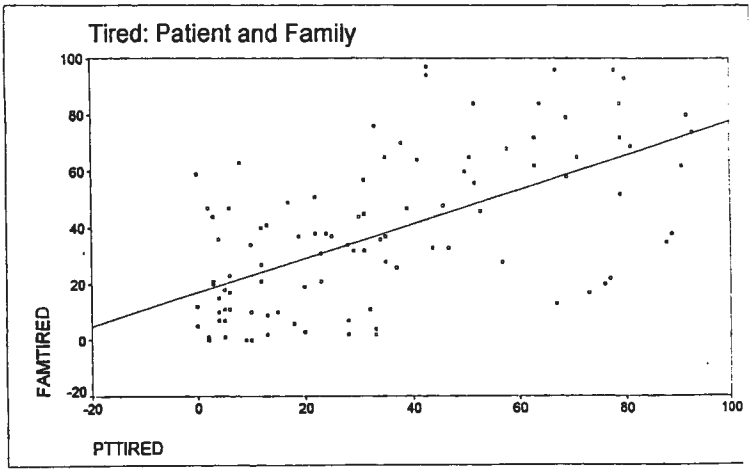

Appetite: Patient and Nurse
NURAPPET
PTAPPETI
120
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
120

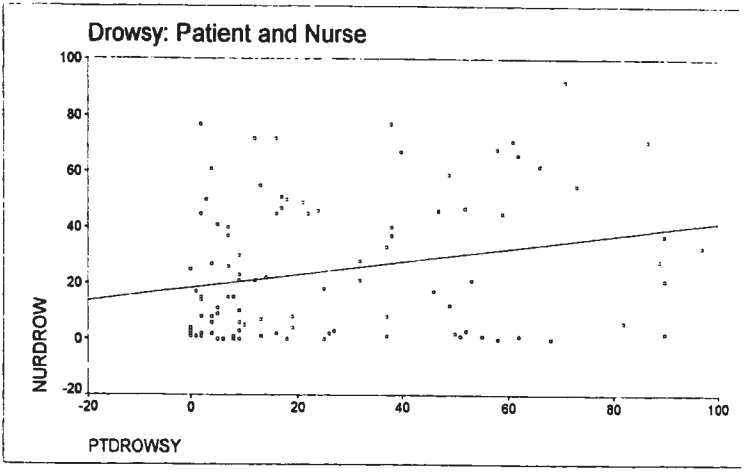

Appetite: Patient and Family
FAMAPPE
PTAPPETI
120
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
120

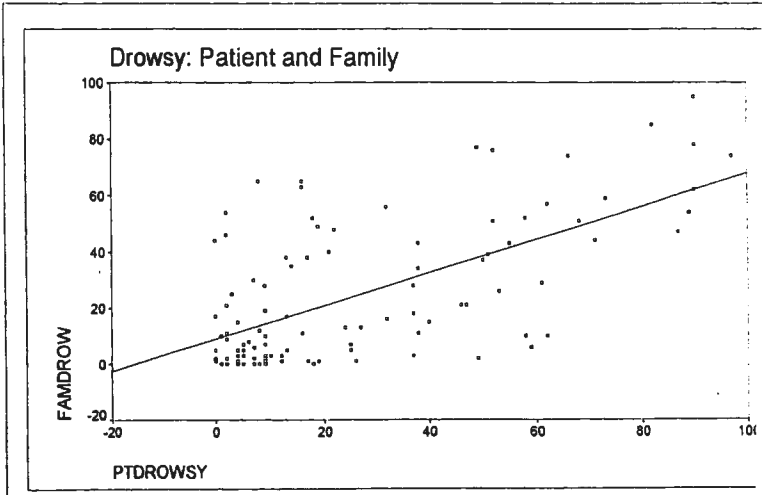

Well-being: Patient and Nurse
NURWELL
PTWELL
100
80
60
40
20
0
0
20
40
60
80
100

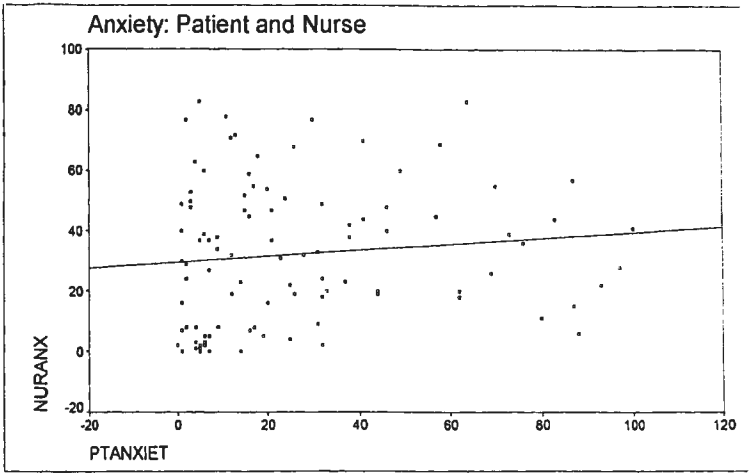

Well-being: Patient and Family
FAMWELL
PTWELL
100
80
60
40
20
0
-20
-20
0
20
40
60
80

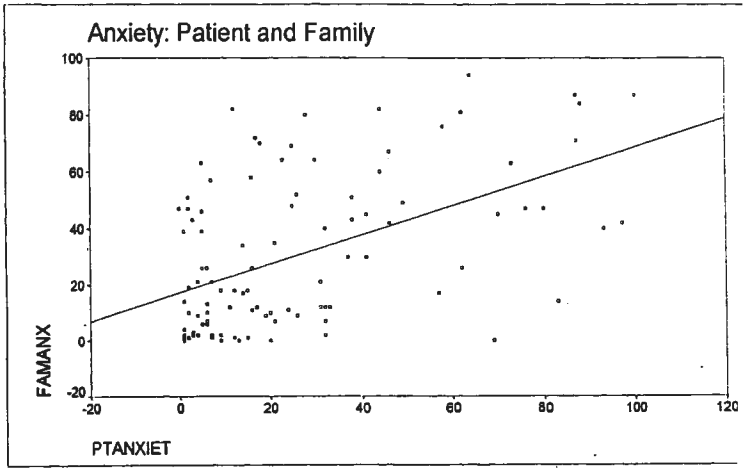

Tired: Patient and Nurse
NURTIRED
PTTIRED
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100

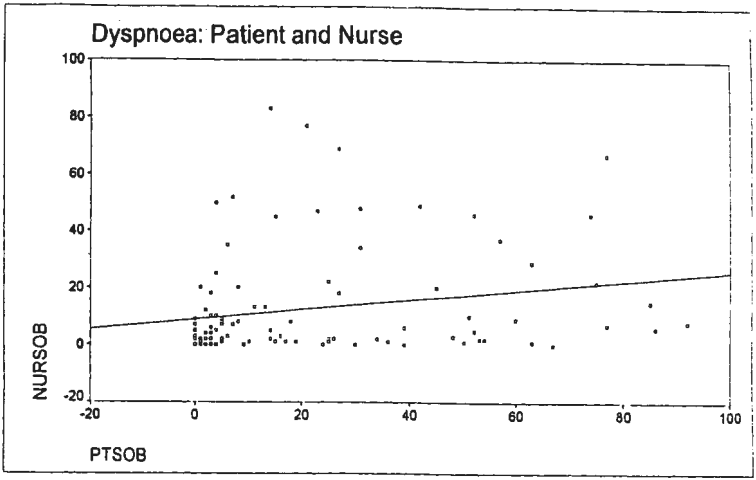

Tired: Patient and Family
FAMTIRED
PTTIRED
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100

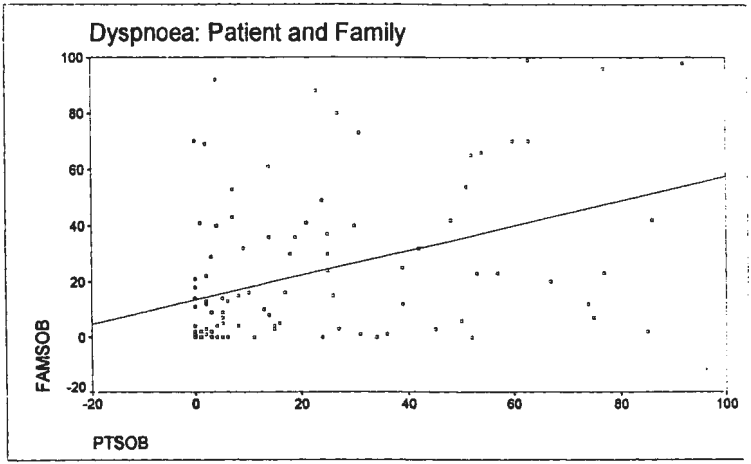

Drowsy: Patient and Nurse
NURDROW
PTDROWSY
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100

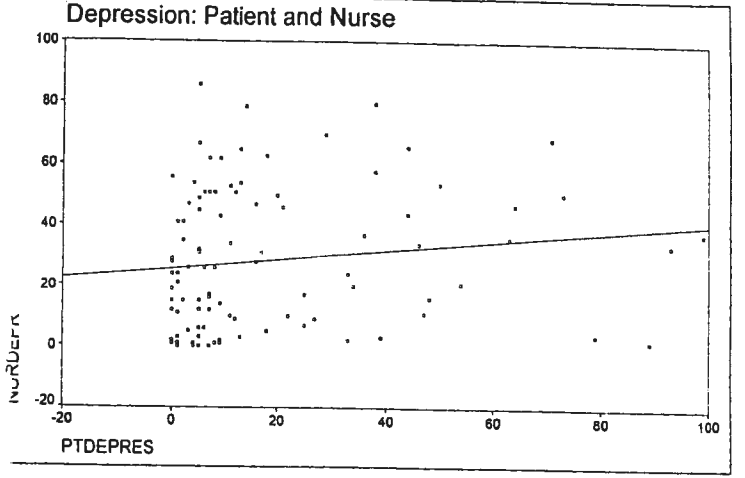

Drowsy: Patient and Family
FAMDROW
PTDROWSY
100
80
60
40
20
0
-20
-20
0
20
40
60
80

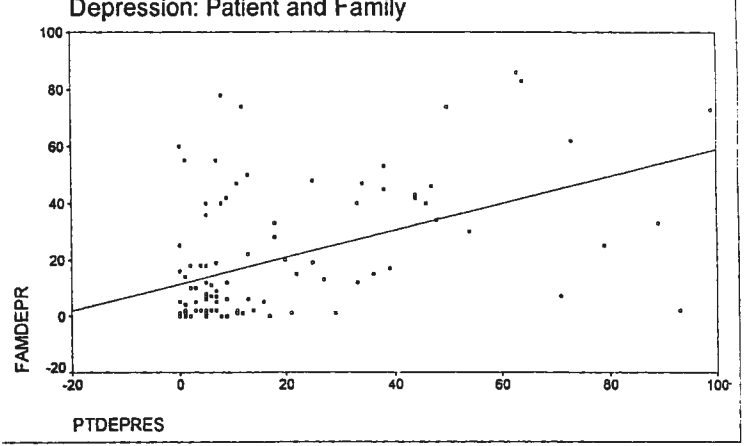

Anxiety: Patient and Nurse
NURANX
PTANXIET
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
120

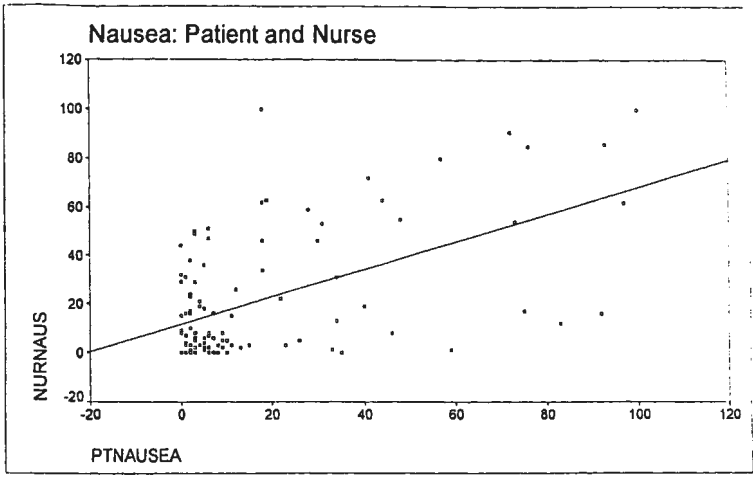

Anxiety: Patient and Family
FAMANX
PTANXIET
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
120

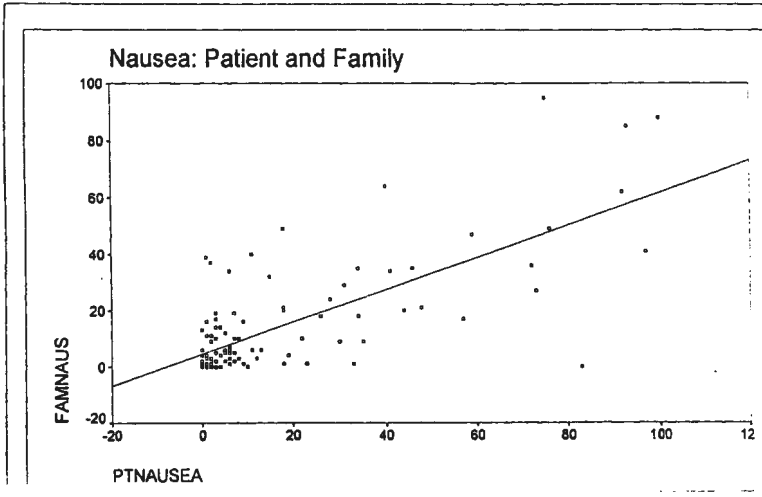

Dyspnoea: Patient and Nurse
NURSOB
PTSOB
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100

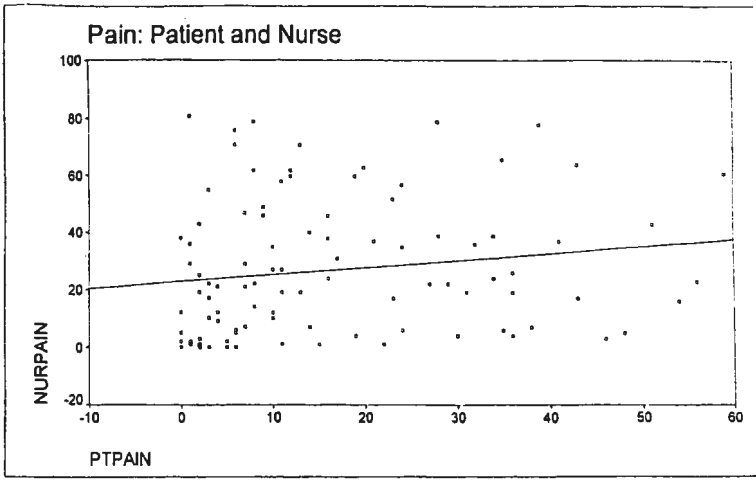

Dyspnoea: Patient and Family
FAMSOB
PTSOB
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100

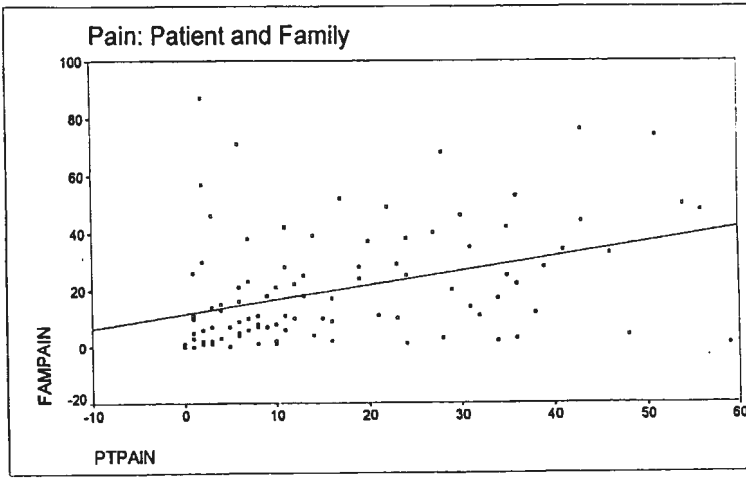

Depression: Patient and Nurse
NURDEPR
PTDEPRES
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100


Depression: Patient and Family
FAMDEPR
PTDEPRES
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100


Nausea: Patient and Nurse
NURNAUS
PTNAUSEA
120
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
120


Nausea: Patient and Family
FAMNAUS
PTNAUSEA
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100


Pain: Patient and Nurse
NURPAIN
PTPAIN
100
80
60
40
20
0
-20
-10
0
10
20
30
40
50
60


Pain: Patient and Family
FAMPAIN
PTPAIN
100
80
60
40
20
0
-20
-10
0
10
20
30
40
50
60