# An Evaluation of the Psychometric Properties of the Edmonton Symptom Assessment System in a New Zealand Community Hospice Setting

Emma Sprague

MHSc

2017

# An Evaluation of the Psychometric Properties of the Edmonton Symptom Assessment System in a New Zealand Community Hospice Setting

Emma Sprague

A dissertation submitted to
Auckland University of Technology
In partial fulfilment of the requirements for the Masters of
Health Science (MHSc)

2017

School of Nursing

Auckland University of Technology

# Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Full Name: Emma Alice Sprague

# Acknowledgements

I would like to give my heartfelt thanks to my supervisors for their patience, advice and support in completing this dissertation. Firstly, Dr Margaret Roberts who inspired this research through her passion and experience in psychometrics. Margaret also kindly lent her impeccable proofreading skills, which has helped me grow considerably as an academic writer. Secondly, Professor Richard Siegert whose experience in data analysis (specifically SPSS and Rasch), was invaluable. I would also like to thank Dr Oleg Medvedev for supporting me through the Rasch analysis process. Without their combined expertise and contagious enthusiasm, this project would not have been possible.

I would like to give my thanks to my family and friends for their encouragement and positive thoughts. They helped to keep me on track and persevere when completing this dissertation felt almost impossible.

I thank my colleagues and managers for their compassion, understanding, and unwavering belief in me, supporting me to complete this study while working full-time.

I also wish to acknowledge Hospice West Auckland and the many patients who contributed data to the study.

A final thanks to Geraldine Peters for her prompt professional proofreading service.

# Abstract

Palliative patients experience substantial and distressing symptoms that impact on their quality of life. To date, a 'gold standard' measure is it yet to be identified for symptom assessment in palliative care. The Edmonton Symptom Assessment System (ESAS) is the most commonly used measure to assess symptoms within palliative care throughout the world. Various studies have investigated the reliability and validity of the ESAS. However, these studies have revealed inconsistent validity and have predominately focused the measures use in populations with cancer only diagnoses.

The cases of all patients admitted to a community hospice setting in New Zealand over the six month period December 2015-May 2016, were reviewed for this study (*n*=229). As part of the usual admitting process, the registered nurse conducts several health related assessment measures; including the ESAS and a performance status measure, the ECOG. The scores of the ESAS and the ECOG were anonymised and collated for data analysis.

The psychometric properties of the 12-item, numeric rating scale version of the ESAS were evaluated using descriptive statistics, exploratory factor analysis, and Rasch analysis. Exploratory factor analysis found three factors. One: drowsiness, tiredness, appetite, wellbeing, and complexity, two: anxiety, depression, and shortness of breath, three: constipation, nausea, insomnia, and pain. Rasch analysis confirmed that the ESAS is a unidimensional scale. Uniform rescoring of the ESAS indicated that the validity of item scoring could be improved by collapsing scoring options from 11 options to four options.

The ESAS is shown to have some evidence of validity and reliability for assessing symptoms within the present research setting of New Zealand, community palliative care. However, the present study identified inconsistent factor structures. Therefore, several areas of key focus for future research has been identified to further validate the use of the ESAS within the New Zealand community palliative care setting.

# Contents

# List of Figures

# List of Tables

# Chapter 1 - Introduction

## 1.1 Background

Palliative care aspires to be a holistic approach to healthcare that focuses on improving quality of life for patients with a terminal illness, and their families (MacLeod, Vella-Brincat, & MacLeod, 2016). MacLeod et al. (2016) states holistic care is achieved through undertaking accurate and careful assessment of the physical, psychosocial, emotional, and spiritual domains. The individual's needs related to each domain are then addressed where able through providing care that decreases symptom burden, and maintains the patients ability to engage in their usual actives of daily living (performance status) (MacLeod et al., 2016).

Gapstur (2007) defines symptom burden as the subjective, quantifiable prevalence, of frequency and severity of symptoms, which produce multiple, negative physiologic and psychosocial drains on the patient. Consequences of increased symptom burden can include decreased length of survival, decreased performance status and lowered quality of life (Cleeland et al., 2000; Ferreira et al., 2008; Gapstur, 2007; Paice, 2004; Weisbord et al., 2003). As quality of life is the main goal of palliative care it is essential that symptoms be recorded accurately, and that frequent reassessment is undertaken to ensure timely and appropriate management (Kaasa & Loge, 2003; MacLeod et al., 2016; Paice, 2004).

There are many different ways a holistic assessment of a patient can be undertaken. It is common practice to employ a 'head to toe' body systems approach combined with the use of appropriate structured assessment measures. A structured assessment measure is any measure that is administered and scored in a standard, predetermined manner (Porter, Larsson, & Lee, 2016). Homsi et al. (2006) found that structured approaches to assessment yields ten times the amount of symptom reporting than that obtained by volunteered reporting alone. Patient information collected using structured assessment measures has several benefits in clinical practice including: the detection of physical and psychological problems that may otherwise be overlooked; monitoring of disease progression; evaluating the efficacy of prescribed treatments or interventions; and facilitating both patient-clinician and clinician-clinician communication (Porter et al., 2016; Valderas et al., 2008). Furthermore, structured

measurement assists clinicians to make decisions on the basis of scientific data rather than by clinical judgement alone, ensuring professional accountability and quality assurance (Haynes, Smith, & Hunsley, 2011).

The Edmonton Symptom Assessment System (ESAS) (Bruera, Kuehn, Miller, Selmser, & Macmillan, 1991) is the most common multi-symptom assessment structured assessment measure used internationally throughout the cancer and palliative care settings (Richardson & Jones, 2009). The ESAS (Bruera et al., 1991) was initially designed to measure eight common symptoms in palliative cancer patients, with the option for an additional rater-selected symptom. However, as the initial study reported 'shortness of breath' being commonly added as the additional symptom, they proposed this symptom was included as a permanent item for future use of this scale, and consequently adapted the ESAS (Bruera et al., 1991) to a nine symptom version: pain, activity, nausea, depression, anxiety, drowsiness, appetite, sensation of well-being and shortness of breath (Bruera et al., 1991). It is common practice for individual settings to modify the ESAS (Bruera et al., 1991) to their area of clinical practice. These modifications include removing the response format; replacing the visual analogue scale with a numeric rating scale; duration of symptoms e.g. assessing the symptom over the past 24hrs rather than at the exact moment of assessment; and number of symptoms: with versions having between eight and 12 items. The impact of these modifications on the scales reliability and validity has not yet been established, and may therefore be problematic. Furthermore, use of the ESAS (Bruera et al., 1991) within a New Zealand context is yet to be researched.

# Chapter 2 - Literature Review

## 2.1 Psychometric Properties of Patient-Reported Outcome Measures

Psychometrics concerns the validation of scale or instruments use, by assessing if these measures are reliable and valid forms of measurement (Ginty, 2013). Clinicians use a variety of assessment measures daily in their clinical practice. However many do not have the background in psychometrics to be able to determine which is the most reliable and valid measure for use within specialised practice areas. Consequently, clinicians rely on measures that are in common use, many of which do not have established psychometric properties. This has implications for clinical utility of the measure itself, and for patient care. Whether based on valid evidence or not, the clinician must ultimately make decisions about the condition of a patient and the services they will receive (Haynes et al., 2011). For decisions to be valid and beneficial to the patient, judgements should be based on sound clinical assessment evidence guided by scientific principles (Haynes et al., 2011). Given this, investigation of the appropriateness, meaningfulness and usefulness (validity) of scales being utilized in current clinical practice is essential to ensure the accuracy of the measure and to minimise negative clinical consequences if the construct (symptom distress) is underrepresented (Haynes et al., 2011). Potential and serious consequences include the under or over treatment of a patient.

Aktas, Walsh and Kirkova (2015) found 57 different symptom assessment measures in use within the cancer setting alone. These measures varied considerably regarding the extent of the psychometric validation undertaken; furthermore, the research completed was methodologically limited. Additionally, Aktas et al. reported a number of instrument drawbacks, including the use of limited numbers of symptoms with meaningful differences, and a varied degree of reliability. Therefore, further research on multi-symptom assessment measures is required to establish the sufficient reliability and validity required for clinical decision-making. Furthermore, thorough psychometric testing may be able to lead to the identification of a ‘gold standard’ measure for symptom assessment, which is currently lacking.

### 2.1.1 Reliability

Reliability explores the idea of when clinicians measure the same person using the same test multiple times, the outcome is consistent i.e. that the clinician is receiving an accurate reflection of their true score (Kuder & Richardson, 1937). Spearman (1904) states that the reliability of a test is defined by correlation of the observed score. The observed score equals the true score, less measurement error (Spearman, 1904). Clinicians who are utilising specific measures, need to have confidence the scores will be consistent when used in their patient population and setting. It is essential that all health care measures be assessed for reliability and measurement error within the population it is being utilised within, as clinical assessment measures are of little value if the data obtained from the patient is not reliable (Haynes et al., 2011; Polit & Yang, 2016).

There are several indicators of reliability: internal consistency, test-retest reliability, and inter-rater reliability (Haynes et al., 2011). *Internal consistency* evaluates the degree of uniformity of the scores from all items on a measure (Haynes et al., 2011). Where item scores are all positively inter-related, i.e. uniform clinicians can have confidence that all items on a measure are in fact measuring the same attribute (Haynes et al., 2011; Polit & Yang, 2016).

*Test-retest reliability* evaluates the stability of the measure, e.g. when a patient is given the measure to complete, then asked to repeat the test again, a similar score would be expected (de Vet, Terwee, Mokkink & Knol, 2011). The challenge in measuring and interpreting test-retest reliability is determining the appropriate time frame between retest (Haynes et al., 2011). Kline (2000) states that changes in the patient and time frames of retesting can negatively impact on test-retest. Furthermore, external factors such as poor test instruction, the use of a proxy reporter or changes in the environment can lead to measurement error (Kline, 2000).

*Inter-rater reliability* is the degree of agreement or consistency between two or more raters completing the measure; i.e. evaluating the scores of a patient when the measure is completed by themselves, a clinician, or their carer (Cohen & Swerdlik, 2005). This is an important consideration within the health setting, where carers and clinicians often assist with, or completes measures on behalf of patients. A reliable measure should display high levels of inter-rater reliability. Thus, allowing the

clinicians using the measured data to make clinical decisions and to be confident that scores are derived in a systematic, consistent way by the various raters using the measure (Cohen & Swerdlik, 2005).

Although the three types of reliability are assessed separately, they are closely related and all should be taken into account when evaluating the reliability of a measure (Kline, 2000). Reliability is a critical aspect of any measure and is essential for the clinician to make informed choices when using assessment measures (Haynes et al., 2011). Reliability must be established before evaluating the validity of a scale as it is not possible for a measure to have validity without reliability (Kline, 2000).

### 2.1.2 Validity

To ensure ethical measurement of health information clinicians need to be confident the practices they are undertaking and the measures they are utilising are valid. Messick (1989) stated that "validity is an overall evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment" p. 5. This means that thorough investigation into validities several components is necessary (Haynes et al., 2011). There are ethical and clinical consequences for using measures in palliative care without having established validity, such as impairing the clinician's ability to track patient functioning, and assessments for change (Haynes et al., 2011).

*Face validity* relates to the perception of validity of the measure according to the user (Cohen & Swerdlik, 2005). While face validity does not directly impact on statistical validity it has important implications in terms of the utility of the measure (Kline, 2000). Meaning that if the measure appears valid to the user, raters may be motivated to complete it.

*Criterion validity* judges how accurately the scores on an assessment measure can be used to infer a likely outcome on another measure of interest (Cohen & Swerdlik, 2005). There are two aspects of criterion validity to consider, concurrent validity and predictive validity. Concurrent validity observes for high levels of correlation with another test that measures the same variable, e.g. pain, shortness of breath, etc. (Kline, 2000). Scores from the measure being validated should be correlated with an external

"gold standard" measure taken at the same time (Polit & Yang, 2016). Strong concurrent validity means that the user can be confident the scale is measuring what it is designed to measure; i.e. symptom distress. Predictive validity assesses the measures ability to predict a certain criterion (Kline, 2000). A clinical example would be the ability of the measure to predict with certainty the level of performance status on the basis of the symptom burden assessment scores of the ESAS (Bruera et al., 1991).

*Content validity* assessed the degree to which a multi-symptom assessment measure has an appropriate set of relevant items, reflecting the full degree of the construct measured (Haynes, Richard, & Kubany, 1995). In the context of a health measure clinicians observe and provide feedback on the critical content required in order to reflect the full construct e.g. symptom distress.

Furthermore, it is equally important that the clinician using the measure understands the response processes of the rater using the scale (American Educational Research Association, 1999). This means that the rater's interpretation of the item is equally as important as the clinicians perspective. Content validity is a critical concern in practice as clinical judgements can be adversely affected by variance in irrelevant behaviour, and insufficiently informed by variance in relevant behaviour (Haynes et al., 2011).

*Construct validity* refers to the degree to which individual differences in a measure reflect the individual differences of the underlying construct (Haynes et al., 2011). This suggests that the test indeed measures what it claims it measures. Cronbach and Meehl (1955) identified that it was necessary to understand the existence of unobservable constructs to develop and validate theories of patient functioning. Furthermore, to validate measures of unobservable constructs, such as depression and anxiety it is necessary to show that a measures construct relates to measures of other related constructs in a predictable way (Cronbach, 1951). As health professionals use information gathered from these scales to make clinical decisions evaluating whether or not the measurement is valid for its intended use is an important to consider when choosing to implement a measure in clinical practice.

*Convergent validity* is a type of construct validity concerning the degree to which scores on a measure are correlated with scores of another measure in which there is a hypothesized correlation (Polit & Yang, 2016), i.e. patients with high symptom burden should also experience poor levels of performance status.

## 2.2 Importance of Psychometric Evaluation of Measurement Tools in Palliative Care

Accurately assessing for changes in symptom burden is a vital component to providing effective palliative care that manages symptoms and maintains quality of life. A major implication of using a scale in palliative care for which reliability and validity have not been established is that the responses may not accurately reflect the true level of symptom burden. This means that clinicians may make decisions regarding medication and interventions on an inaccurate representation of symptoms. There is the potential for patients to be left undertreated and suffer, or over treated and experiencing high levels of sedation.

Although the ESAS (Bruera et al., 1991) was designed to be adapted by clinicians, the practice of adding and removing items based on their perceived clinical need may negatively affect the measure's reliability and validity. These alterations create an unknown variety of formats, making it challenging to track all versions currently in clinical use. Furthermore, the clinicians frequent adaption of the ESAS (Bruera et al., 1991) based on perceived need, rather than researched findings, creates a substantial lack in regularisation and it becomes challenging for clinicians to confirm the measure's psychometric properties. The investigation into these modified versions of (ESAS) (Bruera et al., 1991) has reported some evidence of its psychometric properties when used in a cancer setting, see Table 1 and 2 below. However, many of these studies have flawed methodological approaches (Aktas et al., 2015). Furthermore, to date, no research has been undertaken on the 12-item version that is being used in hospices throughout New Zealand and Australia. Additionally, very few studies have investigated the measures use outside of the advanced cancer settings. The many adaptations of the ESAS (Bruera et al., 1991) suggest that further studies are required to establish the ESAS's psychometric properties (Hearn & Higginson, 1997; Nekolaichuk, Watanabe, & Beaumont, 2008).

The ESAS (Bruera et al., 1991) was not designed as a clinician-rated scale. However, in practice, clinicians may commonly be *proxy reporting* (Polit & Yang, 2016); this is, completing the measure without any direct client input. Meaning they are estimating symptoms scores based on their perception of the patients' experiences, rather than directly asking patients to indicate a score. Proxy reporting has implications regarding the validity of the scale due to the fact the proxy reporter is only able to estimate the severity of the symptoms, providing merely a 'best guess' score rather than a true representation of the severity of symptoms. For this reason, evidence suggests that proxy assessments are not substitutes for the 'gold standard' of patients self-reporting symptoms (Richardson & Jones, 2009).

A further area of consideration is the variance in rates of symptom assessment and re-assessment within a palliative care setting. Presently, the frequency of measurement using the ESAS (Bruera et al., 1991) to observe change is not fully established and varies between settings. Daily assessment is recommended in most inpatient palliative care settings (Richardson & Jones, 2009). However, patients within a community setting are visited less often; therefore, the frequency of assessments would be reduced. It is important for clinicians to know the optimal frequency of reassessment to therefore establish optimal results and standardise measurement frequency across clinicians within the same settings.

### 2.2.1 Reliability of the ESAS

Previous studies have evaluated the reliability of the ESAS (Bruera et al., 1991), see Table 1 below.

Table 1
*Internal Consistency and Test-Retest Reliability of ESAS*

| Study | Version of ESAS | Sample Characteristics | Cronbach's α | Test-retest r | Test-retest freq. |
|---|---|---|---|---|---|
| Chang, Hwang, and Feuerman (2000) | 9-item VAS | American Oncology services; inpatient and outpatient (*n*=242, aged 59-73, *M*=69) | .79 | .86 | 24hrs |
| Davison, Jhangri, and Johnson (2006) | Modified[1], 10-item NRS | Canadian dialysis patients (*n*=560, age range not specified, M=63.5) | - | .70 | 1 week |
| Moro et al. (2006) | 9-item NRS | Italian Palliative care services; inpatient and home palliative care, cancer diagnosis only (*n*=241, age range not specified, M=66.5) | - | .61 | 24hrs |
| Carvajal, Centeno, Watson, and Bruera (2011) | Modified[2], 10-item NRS, translated into Spanish | Spanish Oncology services; inpatient and outpatient settings (*n* =171, aged 23-87, *M*=58) | .75 | .65 | 4-6hrs |
| Hannon et al. (2015) | Modified[3], 11-item NRS | Canadian Oncology outpatient clinic (*n*=202, aged 25-83, *M*=58) | .83 | .62 | 24hrs |
| Yokomichi et al. (2015) | 9-item NRS, translated into Japanese | Japanese Oncology services; inpatient and outpatient settings (*n*=292, age range not specified, *M*=65) | .87 | .90 | 30mins |

Note: NRS = Numeric rating scale, VAS = Visual analogue scale

ESAS Modifications: [1]Modified anchor (worst possible to severe), added item – 'pruritus'
[2]Replaced item 'activity' with 'fatigue', added item – 'difficulty in sleeping'
[3]Replaced the item 'activity' with 'tiredness'

*Internal consistency* of the ESAS (Bruera et al., 1991) has been reported as ranging between .75-.87 (see Table 1). A Cronbach's alpha of greater than .70 is considered good (Pallant, 2016). The variance seen in the reported Cronbach's alphas signifies measurement error. Tavakol and Dennick (2011) state Cronbach's alpha is affected by the length and dimensionality of a measure. Previous studies have used a variety of versions of the ESAS (Bruera et al., 1991), which have reported numerous factor structures. This is likely responsible for the varied internal consistency of this measure.

*Test-retest reliability* of the ESAS (Bruera et al., 1991) was investigated using Spearman's rho – this is a correlation statistic that assesses change in item scoring between an initial and repeated test. A Spearman's rho of greater than .50 is considered acceptable (Pallant, 2016). Test-retest reliability was reported as ranging from .62-.90 (see Table 1). Several factors that may be responsible for the large range of reported Spearman's rho (Kline, 2000). Firstly, changes in the patient; for example does the mere fact of asking a palliative patient who is short of breath and fatigued to complete a measure impact on their reported levels of symptom distress? Secondly, external factors contributing to measurement error such as, poor test instruction, the use of a proxy reporter or a change of environment (e.g. home to clinic) between assessments. Thirdly, the time frame of retesting; for example, if there is a long period between test-retest within a palliative care setting it would be logical for scores to change as a result of disease progression and worsening performance status, or equally from an improvement in symptoms following intervention. Finally, the population for test-retest should be a large and representative sample of the population for whom the test is intended (Kline, 2000).

Although the ESAS (Bruera et al., 1991) is intended as a self-report measure, it is common practice for clinicians and carers to complete the measure on their behalf. Therefore, *inter-rater reliability* of the ESAS (Bruera et al., 1991) has been investigated by several researchers. Nekolaichuk, Bruera, et al. (1999), Nekolaichuk, Maguire, Suarez-Almazor, Rogers and Bruera (1999) and Pautex, Berger, Chatelain, Herrmann and Zulian (2003) report that both physicians and nurses underestimated physical symptoms such as pain, drowsiness and shortness of breath, and overestimated psychological symptoms such as depression and anxiety. Inter-rater

reliability did not improve over time when retested, highlighting the importance of patient self-report (Nekolaichuk, Bruera, et al., 1999; Nekolaichuk, Maguire, et al., 1999; Pautex et al., 2003). This is an important issue, particularly when assessments are relied upon to review outcomes of treatment or interventions. For example, a reduction in scores because of a change in rater rather than the implementation of a new medication could have serious clinical implications.

### 2.2.2 Validity of the ESAS

Several aspects of the validity of the ESAS (Bruera et al., 1991) has been explored, using a variety of versions of this scale, predominantly within a cancer setting.

*Face validity* of both the traditional and the translated versions of the ESAS (Bruera et al., 1991) have been investigated (Watanabe, Nekolaichuk, Beaumont, & Mawani, 2009). Patients reported that the ESAS (Bruera et al., 1991) was easy to complete and covered symptoms to meet their requirements (Chinda, Jaturapatporn, Kirshen, & Udomsubpayakul, 2011; Claessens, Menten, Schotsmans, & Broeckaert, 2011). From these authors, it appears there is some face validity of the ESAS (Bruera et al., 1991). Furthermore, uptake is extensive across the palliative care setting in New Zealand and Australia, indicating clinicians' perception of face validity.

*Concurrent validity* of the 9-item version of the ESAS (Bruera et al., 1991) have been investigated using other established cancer symptom measures: the MD Anderson Symptom Inventory (MDASI) (Cleeland et al., 2000), the Memorial Symptom Assessment Scale (MSAS) (Portenoy et al., 1994), the Rotterdam Symptom Checklist (RSCL) (de Haes, 2012), and the Symptom Distress Scale (SDS) (McCorkle & Young, 1978). A 'gold standard' measurement tool for assessing symptoms has yet to be identified.

Overall the ESAS (Bruera et al., 1991) showed moderate to strong correlation with these alternate symptom assessment measures (see Table 2). Correlations above .75 are regarded as strong support for the concurrent validity (Kline, 2000). Determining concurrent validity in a setting where there is no identified 'gold standard' to measure against is problematic; as we are then required to correlate against measures of dubious validity (Kline, 2000). Kline (2000) states in such circumstances moderate correlations of .5 are acceptable.

Table 2
*Concurrent Validity of the individual ESAS Items and Corresponding Items on Alternative Symptom Assessment Scales*

| ESAS Items | MDASI[1] | MSAS[2] | RSCL[3] | SDS[4] |
|---|---|---|---|---|
| Pain | .73 | .83 | - | - |
| Tiredness | .76 | - | - | - |
| Drowsiness | .74 | - | - | - |
| Nausea | .61 | .62 | .46 | .88 |
| Lack of appetite | .80 | .75 | .46 | .74 |
| Shortness of breath | .56 | .83 | .58 | .84 |
| Depression | .68 | .44 | .45 | .64 |
| Anxiety | .61 | .45 | .48 | - |
| Wellbeing | - | .75 | - | - |
| Total symptom distress | .79 | .72 | - | .77 |

Note: NRS = Numeric rating scale, VAS = Visual analogue scale
Version of ESAS used and any modifications:
[1]9-item, NRS, translated into Japanese (Yokomichi et al., 2015)
[2]Original 9-item, VAS (Chang et al., 2000)
[3]Original 9-item, VAS (Philip, Smith, Craft, & Lickiss, 1998)
[4]9-item, NRS, translated into Italian (Moro et al., 2006)

*Predictive validity* was investigated by examining whether scores from the ESAS (Bruera et al., 1991) were indicative of scoring on performance status measures; i.e., do high levels of symptom distress predict someone will therefore also have a measure of poor performance. Carvajal et al. (2011) used the Karnofsky Performance Score (KPS) (Karnofsky & Burchenal, 1949) and Yokomichi et al. (2015) used the ECOG (Oken et al., 1982). Both researchers reported that when performance status was rated poorly, ESAS (Bruera et al., 1991) scores were higher, i.e. those with a poorer performance status were more likely to be experiencing symptoms and symptom burden.

*Content validity.* Observing for content validity assesses the degree to which a multi-symptom assessment measure, such as the ESAS (Bruera et al., 1991) has an appropriate set of relevant items, reflecting the full degree of the construct being measured (Polit & Yang, 2016). Specifically, do the individual items on the ESAS (Bruera et al., 1991) reflect the full content of "symptom distress" for patients within a palliative care setting. Strömgren, Groenvold, Pedersen, Olsen and Sjogren (2002)

conducted an extensive literature search on the prevalence of symptoms experienced by palliative care patients. The researchers reported 63 different symptom problems; the top 12 common symptoms were: pain, fatigue, loss of appetite, nausea, depression, poor physical function, concentration, constipation, breathlessness, vomiting, drowsiness and anxiety (Strömgren et al., 2002). The 12-item version of ESAS (Bruera et al., 1991) covers nine of these symptoms; not including poor physical function, vomiting, and concentration. However physical function is accounted for by co-administration of the ECOG (Oken et al., 1982) within the clinical setting of the present research. While this version of the ESAS (Bruera et al., 1991) does not cover all of the top rating symptoms, the clinician can be confident it has strong content validity as it covers most of the common symptoms exhibited within a palliative care setting.

*Construct validity.* Reviewing construct validity assists in establishing evidence that the test indeed measures what it claims it does; i.e. that the ESAS (Bruera et al., 1991) measures the construct of symptom distress and not a different construct. Kline (2000) defines two key aspects to consider the construct validity of a measure; scores on the ESAS (Bruera et al., 1991) should correlate with scores on other symptom distress measures (Kline, 2000), and scores should not correlate with tests that are not supposed to be symptom distress measures (Kline, 2000). As outlined in Table 2 above, the ESAS (Bruera et al., 1991) has shown moderate to strong correlation with other symptom distress scales.

### 2.2.3 Responsiveness of the ESAS

Health care professionals are particularly concerned with the measurement of change in patient status. This includes evaluating treatment efficacy or identifying a deterioration in the patients wellbeing, i.e. the responsiveness of a scale (Polit & Yang, 2016). The area under the receiver operating characteristic (AUC ROC) curve is a statistical measure that has relevance in assessing responsiveness of a scale (Polit & Yang, 2016). Polit and Yang (2016) state values for the AUC ROC range from .50 (no better than random classification) to 1.0 (perfect classification), the aim is for scores of .70 and above.

Two recent studies have investigated the responsiveness of a numeric response scale version of the traditional 9-item ESAS (Bruera et al., 1991); in an outpatient

haematology setting (Lopez, Liles & Knupp, 2014) and an outpatient oncology setting (Hui et al., 2016). Lopez et al. (2014) reported pain ($P = 0.03$) as the only symptom score that changed significantly between the initial and follow-up visits in patients with a self-reported pain crisis, pain ($P < 0.00$), fatigue ($P = 0.00$), depression ($P = 0.05$), nausea (P = 0.04) (Lopez et al., 2014). Furthermore, Hui et al. (2016) reported the ESAS (Bruera et al., 1991) scores had moderate responsiveness, with the AUC ROC between 0.69 and 0.76.

### 2.2.4 Dimensionality of the ESAS

The characteristic a measure is designed to assess (construct) can be either unidimensional, i.e. measuring one overall construct, e.g. symptom distress, or multidimensional, measuring several constructs (de Vet et al., 2011; Messick, 1995). The ESAS (Bruera et al., 1991) is intended to be unidimensional as it is a scale where scores can be summed to provide a total score of symptom distress; ESAS-SDS (Bruera et al., 1991). Therefore, investigation into the characteristics or dimensions of the ESAS (Bruera et al., 1991) measure is needed to confirm the scale is indeed unidimensional, as it was designed. However, this is yet to be confirmed by applying statistical procedures such as factor analysis and Rasch analysis.

*Factor analysis* allows researchers to identify the strength and direction of the relationship between items of a measure (Ferrero, Petri, Carbone, & Catelani, 2015). This allows the researcher to identify if the scale is measuring the intended construct or if there a is measurement error affecting the observed score.

Exploratory factor analysis (EFA) explores the dimensionality of a measure in the absence of any hypotheses of the strength and direction of relationships between items (Polit & Yang, 2016). Kline (2000) states EFA is a preliminary step in establishing important variables of the measure. Furthermore, EFA is particularly useful when measures have many items and may require condensing (Kline, 2000). In the case of the present research, EFA will identify if using additional symptoms in this 12-item version of the ESAS (Bruera et al., 1991) affects the measurement of overall symptom distress and potentially identify items that should be removed. In contrast, confirmatory factor analysis (CFA) tests explicit hypotheses of relationships between items that have been generated by previous research, for example through EFA and theory testing model (Polit & Yang, 2016).

Several studies have investigated the factor structure of the ESAS using EFA (Bruera et al., 1991) (see Table 3 below). Chow, Fan, Hadi and Filipczak (2007) concluded that the ESAS-SDS (total symptom distress score) is consistent with a latent construct of well-being. Chow et al. reported three main factors (see Table 3), within a radiotherapy clinic setting for advanced cancer patients. However, during follow-up post-radiotherapy treatment, the factor structure changed, except 'fatigue and drowsiness' and 'depression and anxiety' (Chow et al., 2007). This change in factor structure over the 12-week period between testing suggests measurement error. Potential sources may be changes in the patients' status, e.g. disease progression; changes in the environment; or changes in the rater when scores were collected face to face before treatment, then later assessments were made by phone interview.

Carvajal et al. (2011) also found three main factors of the ESAS (Bruera et al., 1991) within an advanced cancer setting (see Table 3). They named their factors: "soft physical symptoms", "emotional components", and "hard physical symptoms" (Carvajal et al., 2011).

Barata et al. (2016) conducted EFA of the ESAS (Bruera et al., 1991) in a palliative care setting that has patients with a cancer diagnosis, using a co-administered a 10-item Likert scale of additional symptoms. These additional symptoms were: vomiting, constipation, weight loss, dysphagia, dry mouth, sweating, hiccups, insomnia, sleep disturbance and lack of memory (Barata et al., 2016). They also reported three main factors (see Table 3), naming them; "neuro-psycho metabolic", "gastro-intestinal" and "sleep impairment" (Barata et al., 2016).

Most recently, Ganesh et al. (2017) undertook EFA of the ESAS (Bruera et al., 1991), within a radiotherapy clinic setting for advanced cancer patients. Their research also reported three factor (see Table 3).

Only two items, depression and anxiety, occurred consistently within the same factor structure across the various studies. Gorsuch (2014) identified several reasons for a change in factor structure. Firstly, the procedure used to conduct factor analysis may vary between studies, the population using the measure is different, and the domain or construct being measured is different. This suggests that an array of items that may provide good construct validity regarding measuring symptom distress in cancer

patients, may not provide a good representation of symptom distress in non-cancer patients, thus altering the factor structure between groups. More research is required to determine whether a clear and stable factor structure exists and if factorial invariance is found across clinical settings (Richardson & Jones, 2009).

Table 3

*Factor Analysis of the ESAS*

| Study | ESAS Version | Factor One | Factor Two | Factor Three |
|---|---|---|---|---|
| Chow et al. (2007) | Modified[1] 9-item, NRS | | | |
| - Admission | | | | |
| | | Fatigue | Anxiety | Shortness of breath |
| | | Pain | Depression | Nausea |
| | | Drowsiness | | Appetite |
| | | Well-being | | |
| - 12 weeks post treatment | | Pain | Anxiety | Nausea |
| | | Fatigue | Depression | Shortness of breath |
| | | Appetite | Well-being | Drowsiness |
| Barata et al. (2016) | Modified [2] 9-item, NRS | Tiredness | Nausea | Insomnia |
| | - Alongside a 10-item 'Likert scale' | Appetite | Vomiting | Sleep disturbance |
| | | Dyspnoea | Constipation | |
| | | Depression | Dry mouth | |
| | | Anxiety | Hiccups | |
| | | Well-being | Nausea | |
| | | | Vomiting | |
| Carvajal et al. (2011) | Modified[3] 10-item, NRS | | | |
| | | Nausea | Anxiety | Pain |
| | | Appetite | Depression | Shortness of breath |
| | | Drowsiness | Well-being | |
| | | | Sleep | |
| Ganesh et al. (2017) | Modified[4] 9-item, NRS | Depression | Pain | Nausea |
| | | Anxiety | Tiredness | Appetite |
| | | Well-being | Drowsiness | Dyspnoea |

ESAS Modifications:
[1]Replaced item 'activity' with 'fatigue'.
[2]Replaced items - 'activity' with 'tiredness', and 'shortness of breath' with 'dyspnoea', translated into Portuguese.
[3]Replaced item 'activity' with 'fatigue', added item 'sleep', different order (well-being at end), modified anchor (worst possible to worst imaginable), and specified symptom ratings over past 24 hours as distinct from at the time of assessment.
[4]Replaced items - 'activity' with 'fatigue', and 'shortness of breath' with 'dyspnoea'.

*Rasch analysis* is a scaling model that aims to reduce item variance (Kline, 2000). This modern approach to scaling items can be used when examining measures such as the ESAS (Bruera et al., 1991), where the responses to individual items can be tallied

to give a total score (Polit & Yang, 2016). Rasch analysis allows the researcher to examine to what extent the responses on a scale satisfy the construct being measured (Tennant & Conaghan, 2007), i.e. are the individual items on the ESAS (Bruera et al., 1991) contributing in a meaningful way to the measure of overall symptom distress. Furthermore, Rasch analysis investigates the structure of the responses. This provides an oppourtunity to confirm that the patients themselves report varying degrees of symptoms when they respond to items, rather than a measurement error introduced by factors that vary and are unintended to be measured, e.g. demographic data (Polit & Yang, 2016). This enables the researcher to identify items that may be affected by characteristics such as age, sex or diagnosis, i.e. biased items (Pallant & Tennant, 2007; Tennant & Conaghan, 2007).

The attractiveness of the Rasch model is its capacity to show exactly how respondents are using the rating scale. This process assists determining the optimal number of response categories for the scale, allowing for a re-scoring of the item-responses to ensure a high-quality valid measure (Bond & Fox, 2015). Allowing researchers to identify if the scale is unidimensional, and permits the development of an interval scale if the data fits the Rasch model, improving the precision of measurement (Rasch, 1960). An interval scale has measurements where the difference between points on the scale are measurable and exactly equal, e.g. the difference between moving from one number to the next on the scale is equal for each item (Rasch, 1960). This is important in clinical practice where these scales are used to make decisions as it makes the results easier to interpret when the level of difficultly or an increase in symptoms is equal from one-two as it is from nine-10.

Despite the substantial benefits to using a Rasch analysis on a clinical measure, only one previous study has undertaken Rasch analysis on the ESAS (Bruera et al., 1991); Cheifetz, Packham and Macdermid (2014) conducted a Rasch analysis of the 9-item version of the ESAS (Bruera et al., 1991), within an Oncology outpatient setting. They found all items apart from ‘wellbeing’ had an accurate fit in terms of measuring cancer symptoms, and confirmed that the ESAS (Bruera et al., 1991) measured one overall construct of disease burden (Cheifetz et al., 2014). The authors examined the rate of responses to each score of the individual symptoms. Although the ESAS provides responses ranging from zero (no symptom) to 10 (worst possible degree of

symptom), they found patients were not using the entire range of response options. Instead, patients reported their symptoms at the extreme low (0,1,2) or high end of the scale (7,8,9), i.e. they found it easier to rate mild and severe symptoms rather than moderate ones (Cheifetz et al., 2014). This polarity in scoring concerns clinicians as it suggests that moderate symptoms are not being effectively assessed using the ESAS (Bruera et al., 1991)

## 2.3 Summary of the Psychometric Properties of the ESAS

Previous research on the ESAS (Bruera et al., 1991) has been conducted using traditional statistical methods such as regression, correlation and descriptive statistics, with only a few using factor analysis and Rasch analysis. Despite several studies investigating the psychometric properties of the ESAS (Bruera et al., 1991), most have been limited methodologically. The ESAS (Bruera et al., 1991) has numerous different adaptations, including the response format, addition of symptoms, and translations into other languages that can impact on the psychometric properties. The establishment of the psychometric properties of the ESAS (Bruera et al., 1991) is therefore difficult, as there is no consistency in the structure of the measure itself. While this may account for some of the variability in reliability and factor structure observed between studies, we need to know more about the content validity of the ESAS (Bruera et al., 1991) in terms of measuring symptom distress in palliative patients with a cancer and non-cancer diagnosis.

A second issue is that the factor structure of the ESAS (Bruera et al., 1991) is inconsistent. The potential reasons for inconsistent factor structure include; how the analysis was undertaken, the variety of versions used for analysis, and the type of patients the analysis was completed on. We need to establish the factors within a New Zealand palliative population that includes both cancer and non-cancer diagnosis.

A third issue is that previous studies have predominantly focused on the use of the ESAS (Bruera et al., 1991) in patient groups with a cancer diagnosis. Despite this limitation in the validity evidence for the ESAS, clinical need has meant that the ESAS (Bruera et al., 1991) is routinely being used within New Zealand palliative care settings where patients typically have a mixture of both cancer and non-cancer health conditions. As it is not recommended to generalise the validity and reliability of a

scale from one population to the next (Cohen & Swerdlik, 2005), it is essential that research into use of the ESAS (Bruera et al., 1991) within a population of patients with mixed diagnoses (cancer and non-cancer) is undertaken promptly.

## 2.4 Research Aims

This study aims to evaluate the reliability and validity of the 12-item version of the ESAS (Bruera et al., 1991) within a New Zealand community palliative care setting that includes patients with both cancer and non-cancer diagnoses. Second, to investigate the structure and dimensionality of the 12-item version of the ESAS (Bruera et al., 1991) using exploratory factor analysis. Rasch analysis will be used to observe if the ESAS (Bruera et al., 1991) is unidimensional as it was designed, and to investigate how individual item scores are functioning on this measure. The study will also investigate if the responsiveness of the 12-item version of the ESAS (Bruera et al., 1991) can be improved by undertaking iterative rescoring. Finally, the present study will investigate the frequency of reassessment of patient symptoms using the 12-item version of the ESAS (Bruera et al., 1991) and determine the prevalence of proxy reporting in the research setting.

# Chapter 3 - Methods

## 3.1 Participants

All patients who were enrolled in a community palliative care/hospice service during the period of December 2015-May 2016 were assessed for inclusion in this study. Patients were seen in their own homes, or the residential care facility during last stages of their lives. They came from a variety of socioeconomic backgrounds, including immigrants without New Zealand residency status.. The facility is a charitable service which is free of charge to patients and families. The service is partially funded by the government; however, no funding is received for the care of those without New Zealand residency. As the ethos of this service is to provide equitable palliative care to all patients, those without New Zealand residency are not charged for these services and receive unrestricted access to the services provided. To evaluate frequency of retesting, patients who had been under this service for less than three months were excluded.

## 3.2 Measures

### 3.2.1 Eastern Cooperative Oncology Group (ECOG)

The ECOG (Oken et al., 1982) is a performance status measure that describes a patient's ability to undertake activities of daily living using a six-point scale (ECOG-ACRIN cancer research group, 2015). The ECOG (Oken et al., 1982) has three reported domains: ability to care for themselves, level of daily activity, and physical ability, and is scored from 0 = fully active, able to carry out all pre-disease performance without restriction, to 5 = dead (ECOG-ACRIN cancer research group, 2015).

The ECOG (Oken et al., 1982) was chosen as a measure of convergent validity with the ESAS (Bruera et al., 1991) as this was the performance status measure used in the research setting. The ECOG (Oken et al., 1982) is considered superior to similar performance status measures such as the Karnofsky Performance Status (KPS) (Karnofsky & Burchenal, 1949) due to its ease of use. Inter-rater reliability has been reported to have a stronger level of agreement when using the ECOG (Oken et al., 1982) compared with the KPS (Karnofsky & Burchenal, 1949). However, individual

correlation score where not reported (Kelly & Shahrokni, 2016; A. E. Taylor, Olver, Sivanthan, Chi & Purnell, 1999).

### 3.2.2 The Edmonton Symptom Assessment System (ESAS)

The ESAS (Bruera et al., 1991) was developed to measure symptom severity for inpatient palliative care patients with a cancer diagnosis. The original format of the ESAS (Bruera et al., 1991) measure has nine identified symptoms. However, many different adaptions are in current use throughout various clinical settings. The version of ESAS (Bruera et al., 1991) used in this clinical setting is a self-report, numeric rating scale where patients report their symptom intensity on a scale between 0 (least degree of symptom) to 10 (worst degree of symptom) at the time of assessment. It has 12 defined symptoms – pain, tiredness, nausea, depression, anxiety, drowsiness, appetite, wellbeing, shortness of breath, constipation, insomnia and complexity of care (staff completion), as well as an option for an 'open' rater-selected symptom. The version used in the present research has the following modifications: replaced the item 'activity' with 'tiredness' and includes three additional items; 'constipation', 'insomnia', and 'complexity of care'. (see Appendix B, i). Scoring can be assessed individually for each symptom or summed to get a total symptom distress score (ESAS-SDS).

## 3.3 Procedure

### 3.3.1 Ethics

Ethical approval was granted by the Auckland University of Technology Ethics Committee (AUTEC; 16/216), New Zealand (see Appendix A). Additional ethical approval was granted by the internal ethics board at the community hospice where the research was undertaken.

Patient anonymity was protected through the operations manager at the community hospice setting creating a report that detailed the names of all new admissions during December 2015-May 2016. Each record was checked by the primary researcher to ensure patients had signed their admission forms which gave consent for access to health information to "enable effective, safe and efficient care", including for "auditing, monitoring and evaluation purposes" (see Appendix B, ii). The completion

of these hospice admissions forms meant that AUTEC did not require the primary researcher to obtain specific consent for this research project.

The primary researcher who is employed full time in the research setting used the "new admissions" report mentioned above to access each individual patient record to obtain the ESAS (Bruera et al., 1991) and ECOG (Oken et al., 1982) scores. The researcher immediately de-identified data by removing the names and allocating each record an individualised number, transposing the de-identified data into an Excel spread sheet. Basic demographic data such as age, gender, ethnicity and diagnosis was also collected.

### 3.3.2 Data Collection

This study employed a positivist research approach in order to gain a better understanding of how the ESAS (Bruera et al., 1991) is utilised within the primary researcher's area of clinical practice, community hospice care in New Zealand.

As part of the usual admission to the hospice, a registered nurse undertook a thorough assessment of the patients' medical history and current concerns. The assessment includes completion of the ESAS (Bruera et al., 1991) alongside several other health measures available on PalCare (PalCare Pty Ltd, 2017); a computerised notes system. These include a general health assessment, oral assessment, pain assessment and a spiritual-emotional-social assessment, all of which are internal to PalCare (PalCare Pty Ltd, 2017) and not standardised. Furthermore, two standardised measures are also completed: the ECOG (Oken et al., 1982) performance status measure and the Waterlow (Waterlow, 1985) pressure risk assessment measure,. The nurse inputs the data collected during their assessment into PalCare (PalCare Pty Ltd, 2017) and a nursing management care-plan was created. Repeated testing was undertaken when deemed appropriate by their primary nurse. Local guidelines recommend the completion of routine assessment measures at admission, then when a change occurs, with a minimum of monthly re-testing (Hospice West Auckland, 2015). The present study specifically focused on the results from the ESAS (Bruera et al., 1991) and ECOG (Oken et al., 1982).

## 3.4 Statistical Analysis

*Descriptive statistics.* Data was initially analysed using SPSS version 24.0 (IBM Corp., Armonk, NY). Statistical procedures were undertaken to observe for frequencies and patterns in data, including missing data. Descriptive statistics were used to summarise these findings of the participants' demographic and clinical characteristics. In addition to the individual scores for items on the scale, the total scores for all items; ESAS-SDS (Bruera et al., 1991) and ECOG (Oken et al., 1982) scores were obtained.

*Item-level analysis*. Mean scores for individual items on admission were calculated and compared with repeated scores. Internal consistency was calculated using Cronbach's alpha.

*Convergent validity.* The ECOG (Oken et al., 1982) was chosen as a measure of convergent validity with the ESAS (Bruera et al., 1991) as this was the performance status measure used in the research setting. Item-item correlations were also assessed using Spearman's rho.

*Factor analysis.* Exploratory factor analysis was undertaken using SPSS (version 24.0 (IBM Corp., Armonk, NY). First the data was evaluated to ensure it was suitable for factor analysis. The principal component analysis was undertaken, followed by an inspection of the scree plot. Results were then rotated using oblimin rotation, to aid interpretation and identify the presence of distinct factors.

*Rasch analysis.* After establishing the suitability of the data, Rasch analysis was undertaken using RUMM2030 (RUMM Laboratory Pty Ltd, Duncraig, WA).

The Rasch analysis included the following sequential steps (Siegert, Tennant, & Turner-Stokes, 2010):

1. Initial fit statistics were completed to ensure the data set fitted the Rasch model.
2. Disordered thresholds were identified and rescoring was undertaken to obtain ordered thresholds.
3. Re-analysis for overall model and individual item fit.

4. Local dependency was examined and locally dependent items were combined into three subtests.
5. Final re-analysis for overall model and individual item fit.
6. Differential Item Functioning (DIF) was assessed for age, sex, ethnicity, and diagnosis. For the purpose of DIF analysis, data was re-allocated into fewer, equally-sized groups for comparison (see Table 4).
7. Test for unidimensionality.
8. Examination of the person-item threshold distribution.

Medvedev et al. (2016) state that Rasch is an iterative procedure whereby initial fit statistics are undertaken to ensure the overall fit of the data to the model, using the threshold map of the software. This identifies any items showing disordered thresholds and allows the researcher to examine the category structure (Tennant & Conaghan, 2007). A disordered threshold is when the patients capacity is higher than the ability of the construct, e.g. symptom distress (Medvedev et al., 2016). When disordered thresholds are present, the researcher has the option to collapse the score response points, until ordered thresholds are found (Tennant & Conaghan, 2007). After each step, the fit to the model is re-tested (Medvedev et al., 2016).

DIF was analysed as this can affect fit to the model; there should be no DIF based on demographics such as age, sex, diagnosis, or ethnicity (Tennant & Conaghan, 2007). The dimensionality of the ESAS (Bruera et al., 1991) was tested by assessing the principal component analysis, then examining local dependency (Tennant & Conaghan, 2007).

Table 4
*Reallocated groups for DIF analysis*

| Age Groups | Ethnicity Groups | Diagnosis Groups |
|---|---|---|
| < 65years (*n*=75) | Caucasian (*n*=154) | Lung cancer (*n*=43) |
| 66-80 years (*n*=88) | Maori (*n*=22) | Colorectal cancer (*n*=28) |
| 81 years + (*n*=66) | Pacifika (*n*=30) | Gastro-intestinal cancer (*n*=24) |
| | Asian (*n*=13) | Gynecological cancer (*n*=16) |
| | Other (*n*=10) | Breast cancer (*n*=14) |
| | | Other cancer (*n*=72) |
| | | Non-cancer diagnosis (*n*=32) |

# Chapter 4 - Results

## 4.1 Participant Characteristics

A total of 229 patients (133 Female, 96 Male, 25-99 years of age, M=71, SD = 14) with a terminal diagnosis under care of community hospice service in New Zealand, during the period of December 2015 to May 2016, provided baseline data at the time of admission to service. Participants came from six major ethnic groups, comprised of a total of 22 individual ethnicities. The majority of participants were of New Zealand European descent. All 229 patient files reviewed had been in service for at least three months and therefore were eligible for this research. There were 23 different diagnoses, with the majority of participants having a cancer diagnosis. Patient demographic data is outlined in Table 5 below.

Table 5
*Patient Demographic Data*

| | *N* | % |
|---|---|---|
| ***Age*** | | |
| 25-40 years | 8 | 3.49 |
| 41-60 years | 36 | 15.72 |
| 61-80 years | 113 | 49.34 |
| 81-99 years | 72 | 31.45 |
| ***Sex*** | | |
| Male | 96 | 41.92 |
| Female | 133 | 58.08 |
| ***Ethnicity*** | | |
| New Zealand European | 117 | 51.09 |
| Other European | 37 | 16.16 |
| Pacifika | 30 | 13.10 |
| Maori | 22 | 9.60 |
| Asian | 13 | 5.68 |
| Other | 10 | 4.37 |
| ***Diagnosis*** | | |
| Lung cancer | 43 | 18.78 |
| Colorectal cancer | 28 | 12.23 |
| Gastro-intestinal cancer | 24 | 10.48 |
| Gynaecological cancer | 16 | 6.99 |
| Breast cancer | 14 | 6.11 |
| Haematological cancer | 9 | 3.93 |
| Other cancer | 63 | 27.51 |
| Non-cancer diagnosis | 32 | 13.97 |

Only 28 of the 229 cases had a repeat ESAS (Bruera et al., 1991) completed. This group reported higher scores for all individual items (except insomnia and constipation) together with higher scores for symptom distress; ESAS-SDS (Bruera et al., 1991) and performance status; ECOG (Oken et al., 1982) (see Table 6).

Table 6
*Mean Scores for Individual Items, ESAS-SDS and ECOG*

| Admission Scores (*n*=177) | | Repeat Scores (*n*=27) | |
|---|---|---|---|
| Symptom | Mean | Symptom | Mean |
| Pain | 3.31 (SD = 2.94) | Pain | 4.72 (SD =3.21) |
| Tiredness | 4.16 (SD = 2.41) | Tiredness | 6.00 (SD = 2.96) |
| Nausea | 1.38 (SD = 2.17) | Nausea | 1.83 (SD = 3.02) |
| Depression | 1.16 (SD = 1.71) | Depression | 1.46 (SD = 2.25) |
| Anxiety | 1.82 (SD = 2.18) | Anxiety | 2.44 (SD = 2.89) |
| Drowsiness | 1.56 (SD = 2.46) | Drowsiness | 2.14 (SD = 3.34) |
| Appetite | 3.24 (SD = 2.46) | Appetite | 4.82 (SD = 3.57) |
| Wellbeing | 4.21 (SD = 2.21) | Wellbeing | 5.47 (SD = 2.44) |
| SOB | 2.05 (SD = 2.40) | SOB | 2.74 (SD = 3.09) |
| Complexity | 4.20 (SD = 2.46) | Complexity | 4.33 (SD = 2.40) |
| Constipation | 2.26 (SD = 2.80) | Constipation | 1.89 (SD = 2.27) |
| Insomnia | 1.50 (SD = 2.10) | Insomnia | 1.00 (SD = 1.97) |
| ESAS-SDS | 30.99 (SD = 15) | ESAS-SDS | 31.74 (SD = 20.39) |
| ECOG | 2.07 (SD = .92) | ECOG | 2.87 (SD = 1.01) |

The majority of ESAS (Bruera et al., 1991) scores were completed by the registered nurse (*n* =169, 95.48%) and the remainder were completed by the patient (*n* = 6, 3.39 %) or their carer (e.g. family member, *n* =2, 1.13%). Furthermore, only two raters out of the 177 cases that had an ESAS (Bruera et al., 1991) score completed on admission chose to add a rater-selected symptom; one rater selected ‘diarrhoea’, and the other selected ‘confusion’.

Missing data was common in this population. There were 134 cases where the reporter had completed all 12 items; meaning 41.48% of participants had not completed at least one item on the ESAS (Bruera et al., 1991). Fifty two cases (22.70%) had no response at all; i.e. they had not completed an ESAS (Bruera et al.,

1991) upon admission despite this being recommended best practice within this clinical setting. Table 7 identifies the number of missing responses by symptom. Missing data ranged from 23.14-36.24% per item. Insomnia was the most frequent source of missing data (36.24% incomplete) for respondents not completing this item.

Missing data was excluded pairwise for analysis; meaning that all data for individual items was retained for data analysis. However, to validate the total scores of this measure, all items must have been answered. Ninety-five cases were excluded because of missing data, the remaining 134 were retained for analysis related to the ESAS-SDS (Bruera et al., 1991).

Table 7
*ESAS Scores Including Missing Data, Ordered by Item*

| | Pain (Item 1) | Tiredness (Item 2) | Nausea (Item 3) | Depression (Item 4) | Anxiety (Item 5) | Drowsiness (Item 6) | Appetite (Item 7) | Well-being (Item 8) | Shortness of breath (Item 9) | Complexity (Item 10) | Constipation (Item 11) | Insomnia (Item 12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *N* Valid | 176 | 171 | 168 | 165 | 165 | 164 | 164 | 165 | 167 | 152 | 153 | 146 |
| Missing data | 53 | 58 | 61 | 64 | 64 | 65 | 65 | 64 | 62 | 77 | 76 | 83 |
| Mean | 3.31 | 4.16 | 1.38 | 1.16 | 1.82 | 1.56 | 3.24 | 4.21 | 2.05 | 4.20 | 2.26 | 1.50 |
| Standard Deviation | 2.94 | 2.41 | 2.17 | 1.71 | 2.18 | 2.46 | 2.46 | 2.21 | 2.40 | 2.46 | 2.76 | 2.10 |
| Minimum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Maximum | 10 | 10 | 8 | 7 | 10 | 10 | 10 | 10 | 9 | 10 | 10 | 10 |

## 4.2 Psychometric Properties

### 4.2.1 Reliability

The Cronbach's Alpha coefficient for the ESAS (Bruera et al., 1991) was .79. Furthermore, the Cronbach's alpha if the item was deleted marginally affected the internal consistency (see Table 8).

Table 8

*Item-Total Statistics for 12 ESAS Items*

| | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|
| Pain | .31 | .79 |
| Tiredness | .49 | .77 |
| Nausea | .29 | .78 |
| Depression | .31 | .78 |
| Anxiety | .38 | .78 |
| Drowsiness | .51 | .76 |
| Appetite | .52 | .76 |
| Wellbeing | .68 | .75 |
| Shortness of Breath | .26 | .79 |
| Complexity | .67 | .75 |
| Constipation | .36 | .78 |
| Insomnia | .43 | .78 |

### 4.2.2 Convergent Validity

The item-item correlation of the ESAS (Bruera et al., 1991) found that most items on this measure had a small to moderate positive correlation, see Table 9 below. Several items reported strong positive correlations; 'well-being and complexity' (.63), 'anxiety and depression' (.53), and 'drowsiness and tiredness' (.50). Four symptoms in particular were strongly correlated with the ESAS-SDS (Bruera et al., 1991) score; complexity (.75), well-being (.72), drowsiness (.62), and appetite (.57). Furthermore, there was a moderate correlation between ECOG (Oken et al., 1982) scores and the ESAS-SDS (Bruera et al., 1991) score (.48).

Table 9

*Item-Item Correlation Coefficients*

| Item | Pain | Tiredness | Nausea | Depression | Anxiety | Drowsiness | Appetite | Well-being | Shortness of breath | Complexity | Constipation | Insomnia | Total (ESAS-SDS) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pain | 1.00 | | | | | | | | | | | | |
| Tiredness | .24 | 1.00 | | | | | | | | | | | |
| Nausea | .12 | .17 | 1.00 | | | | | | | | | | |
| Depression | .11 | .16 | .22 | 1.0 | | | | | | | | | |
| Anxiety | .08 | .18 | .13 | .53 | 1.0 | | | | | | | | |
| Drowsiness | .30 | .50 | .13 | .31 | .30 | 1.00 | | | | | | | |
| Appetite | .11 | .39 | .28 | .26 | .20 | .37 | 1.00 | | | | | | |
| Well-being | .30 | .43 | .28 | .37 | .43 | .47 | .41 | 1.00 | | | | | |
| Shortness of breath | .06 | .16 | .07 | .14 | .29 | .10 | .17 | .22 | 1.00 | | | | |
| Complexity | .38 | .36 | .20 | .29 | .46 | .41 | .25 | .63 | .31 | 1.00 | | | |
| Constipation | .23 | .13 | .18 | .11 | .08 | .14 | .29 | .24 | .19 | .23 | 1.00 | | |
| Insomnia | .32 | .20 | .30 | .30 | .16 | .28 | .18 | .32 | .20 | .27 | .30 | 1.00 | |
| Total (ESAS-SDS) | .49 | .57 | .41 | .41 | .50 | .61 | .57 | .72 | .46 | .75 | .47 | .47 | 1.00 |

## 4.3 Factor Analysis

The 12 items of the ESAS (Bruera et al., 1991) were put forward for principal component analysis (PCA) using SPSS version 24.0. Before performing the PCA, the suitability of the data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .30 and above, results below .30 were excluded. The Kaiser-Meyer-Olkin measure was .81, exceeding the recommended value of .6 (Kaiser, 1970) and Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance, supporting the factorability of the correlation matrix. Factorability relates to the assumption that there are some correlations present in the items, allowing for the identification of coherent factors (Pallant, 2016).

Principal components analysis revealed four components with eigenvalues exceeding one, explaining 32.8%, 11.3%, 9.4% and 8.7% of the variance respectively. An inspection of the scree plot indicated a clear break after the second component. A three-factor solution was deemed appropriate following evaluation of the results from both the component and pattern matrix. The component matrix identified items loading strongly on three main components. The pattern matrix (shown in Table 10) identified three or more items loading on three components confirming that a three-factor solution was most appropriate.

The three-component solution explained 53.5% of the total variance: component one 32.8%; component two 11.3%; and component three 9.4%. To aid the interpretation of these three components, oblimin rotation was performed. The rotated solution revealed the presence of simple structure, with all components showing a number of strong loadings, and all variables loading substantially on only one component.

Table 10

*ESAS Principal Components Analysis with Oblimin Rotation Pattern Matrix*

| | 1 | 2 | 3 |
|---|---|---|---|
| Drowsiness | .90 | | |
| Tiredness | .82 | | |
| Appetite | .59 | | |
| Wellbeing | .56 | | |
| Complexity | .52 | | |
| Anxiety | | .85 | |
| Depression | | .73 | |
| Shortness of breath | | .51 | |
| Constipation | | | .76 |
| Nausea | | | .62 |
| Insomnia | | | .56 |
| Pain | | | .43 |

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalisation.
Rotation converged in six iterations.

Table 11

*ESAS Principal Components Analysis with Oblimin Rotation Structure Matrix*

| | 1 | 2 | 3 |
|---|---|---|---|
| Drowsiness | .84 | | |
| Tiredness | .79 | | |
| Wellbeing | .72 | .51 | .49 |
| Complexity | .62 | .54 | |
| Appetite | .64 | | |
| Anxiety | | .84 | |
| Depression | | .74 | |
| Shortness of breath | | .52 | |
| Constipation | | | .72 |
| Insomnia | | | .64 |
| Nausea | | | .63 |
| Pain | .47 | | .51 |

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalisation.

## 4.4 Rasch Analysis

### 4.4.1 Preliminary Test of the Overall Fit to the Rasch Model

The overall Rasch model fit statistics for the main stages of analysis are represented in Table 12. The person separation index (PSI) of .77 indicated strong reliability, and analysis of fit to the model was acceptable (Polit & Yang, 2016). However a significant Chi Square was found ($p$=.01); a score of less than .05 is deemed to misfit model expectation (Tennant & Conaghan, 2007). The individual item fit was reviewed and it was reported that none of the items showed a misfit to the Rasch model (Table 13). Furthermore, all items displayed disordered thresholds. See Figure 1 below for an example of a typical item category probability curve, displaying a disordered threshold. To improve thresholds, uniform rescoring of all ESAS (Bruera et al., 1991) items was undertaken before further analysis.

Table 12
*Summary of Overall Rasch Model Fit Statistics*

| | Item residual | | Person residual | | Goodness of fit | | PSI | Independent t-test | |
|---|---|---|---|---|---|---|---|---|---|
| Analysis | Mean | SD | Mean | SD | $X^2$ (df) | *P* | Value | % | 95% CI |
| 1 | .34 | 1.44 | -.20 | 1.16 | 155.41 (108) | .01 | .77 | 4.80 | 1.98 |
| 2 | .22 | 1.20 | -.32 | 1.21 | 142.28 (108) | .02 | .75 | 6.55 | 3.73 |
| 3 | .22 | 1.19 | -.29 | 1.08 | 29.56 (27) | .33 | .72 | 5.68 | 2.85 |

Note: PSI=Person separation index

Table 13
*Individual Item Fit*

| Item | Item-to-total correlation | Item loadings 1st PC | Item difficulty (location) | Item-fit residuals | Chi Square |
|---|---|---|---|---|---|
| 1. Pain | .17 | .42 | -.13 | 2.76 | 12.55 |
| 2. Tiredness | .39 | .62 | -.37 | .12 | 4.92 |
| 3. Nausea | .15 | .39 | .29 | .21 | 6.84 |
| 4. Depression | .18 | .42 | .59 | .65 | 9.63 |
| 5. Anxiety | .26 | .51 | .12 | .16 | 8.04 |
| 6. Drowsiness | .45 | .67 | .09 | -1.44 | 8.43 |
| 7. Appetite | .41 | .64 | -.17 | .46 | 4.63 |
| 8. Wellbeing | .64 | .80 | -.37 | -1.59 | 27.84 |
| 9. Shortness of Breath | .13 | .36 | .11 | 1.86 | 21.36 |
| 10. Complexity | .62 | .79 | -.40 | -1.21 | 20.21 |
| 11. Constipation | .20 | .44 | -.01 | 2.51 | 21.34 |
| 12. Insomnia | .27 | .52 | .25 | -.34 | 9.62 |

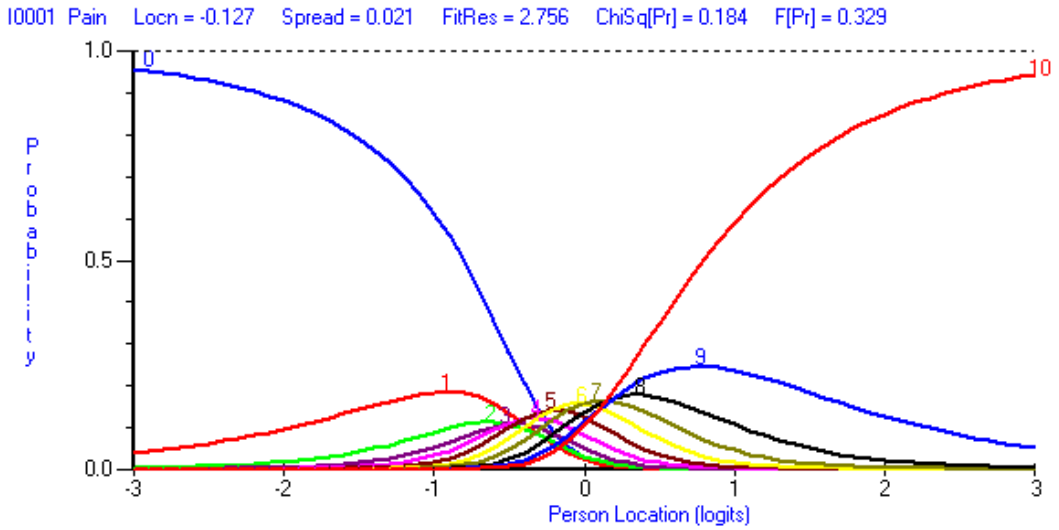



*Figure 1.* Item Category Probability Curves for Item 1 of the ESAS, before Re-scoring

### 4.4.2 Re-scoring of ESAS Items

Iterative rescoring of the ESAS items (Bruera et al., 1991) showed that optimal ordering of thresholds could be achieved using uniform rescoring of all the 11-point items by collapsing response categories. The first and the last response options were retained as they represented extreme conditions; for example, ‘no pain’ and ‘worst possible pain’.

The first response option was retained as zero, the response options one, two, three, four and five were rescored as one; the response options six, seven, eight, and nine were rescored as two, and the final response option as three, resulting in a four point scale (0-3). Figure 2 represents the item category probability curve for item one (pain) after rescoring, illustrating perfectly ordered thresholds.

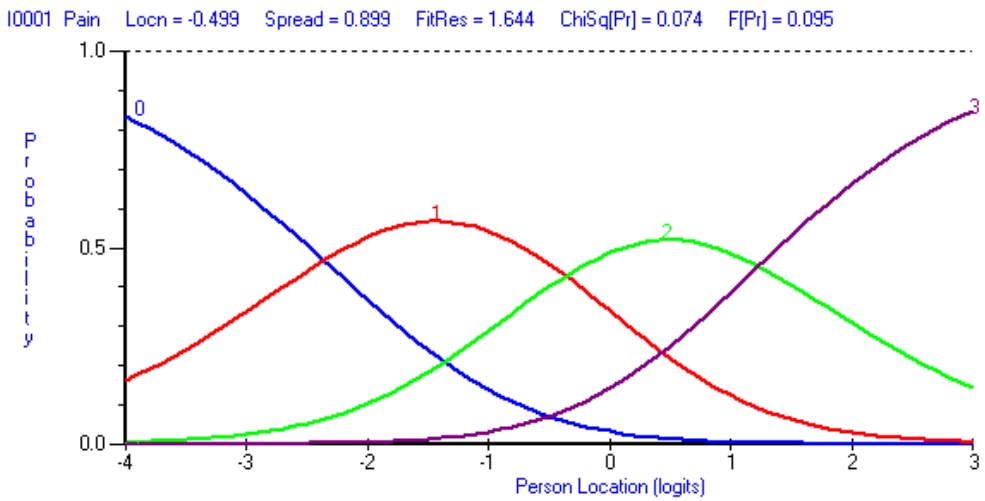



*Figure 2.* Item Category Probability Curves for Item 1 of the ESAS, after Re-scoring

Figure 3 shows the threshold map for all ESAS (Bruera et al., 1991) items after rescoring, ordered by degree of difficulty. It can be seen that there are no disordered thresholds, with 'complexity' and 'wellbeing' as the easiest to endorse items and 'drowsiness' and 'depression' as the most difficult to endorse items.

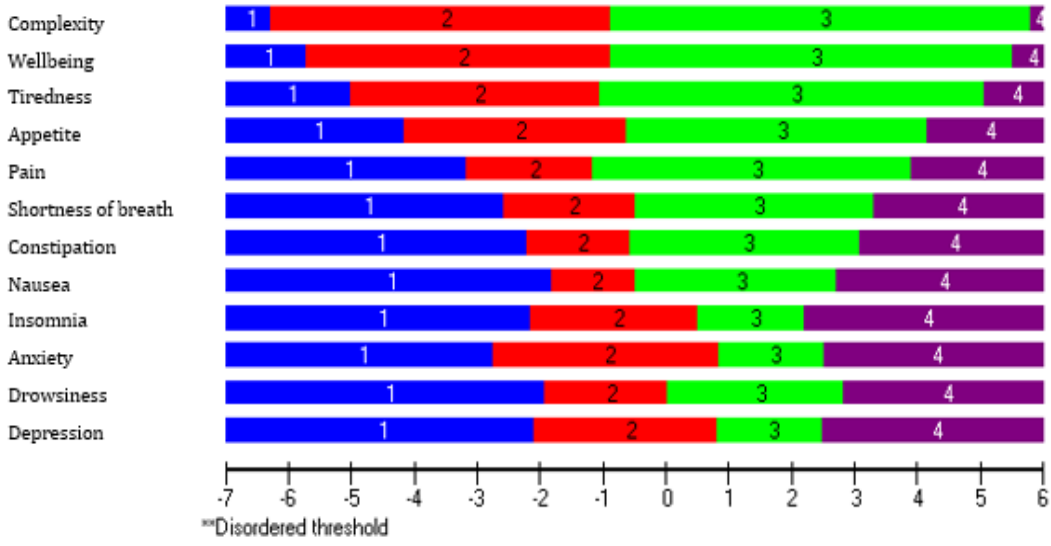



*Figure 3.* Threshold Map for the 12 ESAS Items after Uniform Re-Scoring, ordered by Difficulty

After rescoring the reliability of the scale remained strong and relatively unchanged (PSI=.75). Despite improvement in the overall model fit, it was still unsatisfactory with Chi Square (*p*=.02), see Table 12. At this stage, all items demonstrated satisfactory model fit, and local dependencies between individual items were examined because they can affect the overall fit to the Rasch model (Wright, 1996).

### 4.4.3 Local Dependency

The residual correlation matrix indicated local dependency between six items forming three independent pairs. All residual correlations exceeded the .2 cut off point above the mean of the residual correlations (Marais & Andrich, 2008). The highest correlation was found between 'depression and anxiety', followed by 'wellbeing and complexity' and 'drowsiness and appetite'. The researcher's clinical experience led to the expectation that these pairs of items would be highly correlated. Therefore, locally dependent items were combined into three subtests. This resulted in the best model fit indicated by the non-significant Chi Square and unidimensionality, see Table 12.

### 4.4.4 Differential Item Functioning

Differential Item Functioning (DIF) was analysed by controlling for gender, age, ethnicity and diagnosis. As a registered nurse was the major rater of this study (95.48%), DIF is in relation to a registered nurse's bias towards patients. There was no differential item functioning or item bias found, i.e. when nurses rated patients with the same underlying ability or attribute, in this case symptom load or severity, they had the same probability of recording the same score for each symptom, despite individual demographic factors.

### 4.4.5 Person-Item Threshold Distribution

Figure 4 shows the person-item threshold distribution where person ability and item difficulty are plotted on the same interval logit scale. The plot shows that 99% of individual symptom levels (abilities) are covered by item thresholds of the modified ESAS (Bruera et al., 1991). The number of item thresholds on the right hand side of the scale substantially exceeded the range of individuals in the upper level of sample distribution. This is an advantage for the clinical use of the measure scale, as it has potential to assess people experiencing more extreme symptoms.

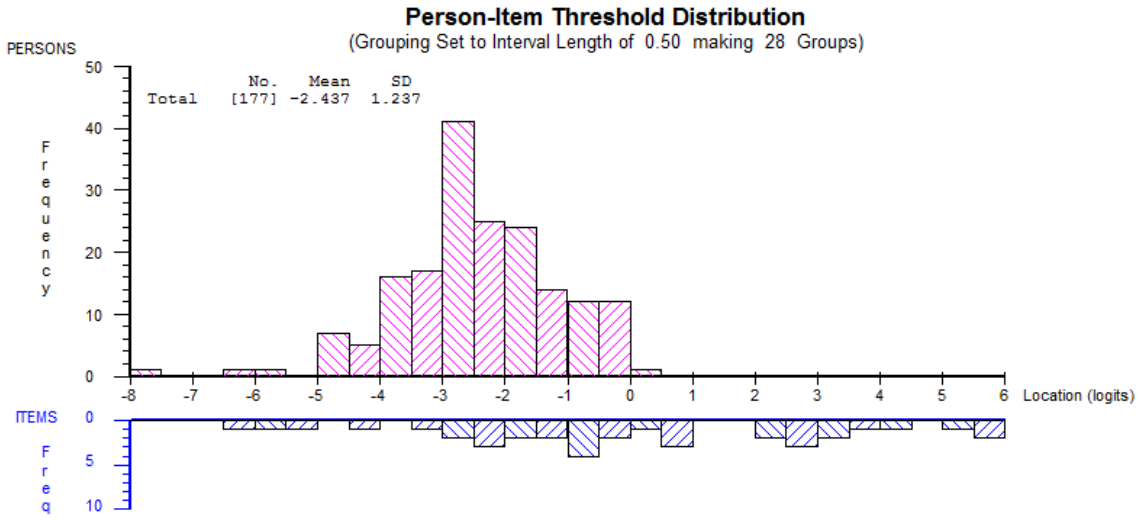



*Figure 4.* Person-Item Threshold Distribution

### 4.4.6 Ordinal-to-Interval Conversion

An algorithm was developed to convert ordinal scores to interval-level data to increase the accuracy of measurement (see Appendix C). Researchers and clinicians can transform ordinal scores without modifying the original response format of the scale by following the instructions provided with the table. This conversion table should only be used when the completed data for all items is available for the assessed individual.

# Chapter 5 – Discussion

## 5.1 Psychometric Properties

The present study met the aim of evaluating the psychometric properties of the 12-item version of the ESAS (Bruera et al., 1991). The study also observed that the usual practice was for registered nurses to complete the ESAS (Bruera et al., 1991) without direct patient input. Furthermore, re-testing was not regularly completed despite local guidelines instructing a minimum of monthly assessment. An evaluation of the structure and dimensionality of the 12-item version of the ESAS (Bruera et al., 1991) concluded the measure is unidimensional. Furthermore, a rescoring of item responses was undertaken resulting in a suggested modification to scoring options, thus improving the responsiveness of the scale.

*Internal consistency.* Strong internal consistency of .79 was found for the 12-item version of the ESAS (Bruera et al., 1991). For context values above .70 are considered acceptable for clinical use (DeVellis, 2012). Previous research reported internal consistency between .75 to.87, which provides further support for the reliability of the ESAS (Carvajal et al., 2011; Chang et al., 2000; Hannon et al., 2015). However, the clinicians' practice of adding items and changing the format of the ESAS (Bruera et al., 1991) mean that exact comparisons are not possible.

*Clinical utility.* The finding that the majority (95.48%) of the ESAS (Bruera et al., 1991) scores had been completed by the registered nurse has implications for the accuracy of measurement. A potential bias was introduced when proxy reporting was undertaken. Furthermore, concerns were raised about the validity of these scores as being a true representation of the patient's symptoms. Previous research found that nurses underestimated physical symptoms, and overestimated psychological symptoms (Nekolaichuk, Bruera, et al., 1999; Nekolaichuk, Maguire, et al., 1999; Pautex et al., 2003). The potential reasons for the high level proxy reporting include; completing the measure electronically when the nurse has returned to the main office after the visit, rather than alongside the patient during the visit. Additionally, within a palliative care setting where patients are presumably more fatigued, this type of reporting may be preferred by both patient and clinician.

Although the ESAS (Bruera et al., 1991) has an optional rater-selected item, this 'open' item was uncommonly utilised. Two clinicians added a rater-selected symptom; one

measure had the item ‘diarrhoea’ selected and the other had ‘confusion’ selected. This indicates that clinicians may consider that the ESAS (Bruera et al., 1991) covers the majority of symptoms experienced by this patient population. However, education about the appropriateness of an item to add may be required. Items such as confusion would be better suited to using an existing validated measure specific to that purpose.

*Frequency of assessment*. Re-assessment of patient symptoms using the ESAS (Bruera et al., 1991) was uncommon in this group, with 27 of the 229 (6.11%) cases completing a repeated ESAS (Bruera et al., 1991) following admission to service. This is contrary to recommended best practice as from *The Palliative Care Handbook, New Zealand* (MacLeod et al., 2016), which states alongside accurate and meticulous assessment, continuous reassessment is required to evaluate the effectiveness of interventions. Patients that had been retested during this study showed higher scores of symptom burden, raising concerns that clinicians were not accurately assessing and responding to symptoms as patient need increased throughout their disease trajectory.

*Convergent validity*. This study reported small to moderate positive correlations between all items on the ESAS (Bruera et al., 1991). Additionally, several items reported strong positive correlations; ‘well-being and complexity’ (.63), ‘anxiety and depression’ (.53), and ‘drowsiness and tiredness’ (.50). In the primary researcher’s clinical experience, it is expected that these items would be correlated. Furthermore, these pairings of items were also found in the EFA for this research (see section below for further discussion).

A strong positive correlation was found between the item (absence of) well-being and the sum of symptom scores; ESAS-SDS (Bruera et al., 1991) (.72). This suggests that the single item of well-being could be interpreted as a valid overall view of symptom distress. Furthermore, this is nearly identical to previous research that reported a correlation of .73 (Carvajal et al., 2011).

A moderate correlation (.48) was observed between ECOG (Oken et al., 1982) scores and the total ESAS (Bruera et al., 1991) score; ESAS-SDS. This finding indicates those reporting moderate levels of symptom distress are more likely to have poor physical performance. A potential reason for this level of correlation may be that the ECOG (Oken et al., 1982) is not a ‘gold standard’, this measure takes into consideration few activities regarding physical performance and may therefore be ‘construct

underrepresentation'. Further research is required as the ECOG (Oken et al., 1982) is beyond the scope of this research.

*Missing data.* Percentage of missing data reported in this research was higher (23-36% per item) than that reported by Moro et al. (2006), who reported 17-20% of missing data per item. The most common items not completed were insomnia, constipation, and complexity. Interestingly, these three symptoms are positioned at the bottom of the scale, i.e. last in the list of symptoms, and are those that have been added to supplement the 9-item version of the scale to create the 12-item version used in the present research. Potential reasons for this include; fatigue effect; clinicians perceiving the item less useful; and lack of education on the use of the measure.

## 5.1 Exploratory Factor Analysis

Three factors provided the best solution for the ESAS (Bruera et al., 1991); one: drowsiness, tiredness, appetite, wellbeing, and complexity, two: anxiety, depression, and shortness of breath, three: constipation, nausea, insomnia, and pain. Interestingly, three pairings of items found in factors one, two and three were reported as highly correlated in the item-item analysis; 'well-being and complexity', 'anxiety and depression', and 'drowsiness and tiredness'. In the primary researchers' clinical experience, some factors are logical: factor one is composed of symptoms that patients would rate poorly on at end of life when they are sleeping more and have lost their appetite or are no longer able to swallow. Factor two has mixed appeal, it is common for 'anxiety and shortness of breath' and 'anxiety and depression' to be seen together, however in clinical practice those with anxiety and shortness of breath do not necessary have depression. Factor three also has mixed appeal within the clinical setting; 'constipation and nausea' and 'insomnia and pain' are interrelated, however those with the first grouping, would not necessarily experience the second. Previous investigation into the overlap of symptoms has reported similarities to the factor structures identified above. Identifying the comorbidity of symptoms such as anxiety and depression (Clark & Watson, 1991), and insomnia and pain (D. Taylor et al., 2007) shared negative affect factor.

EFA previously undertaken on the ESAS (Bruera et al., 1991) also reported three factors. However, the items composing the factors are expressively different in each study. With the exception of depression and anxiety being paired in the same factor structure of each study (see Table 3). Haynes et al. (2011) states when undertaking

EFA it is important for clinicians not to mistake the results of factor analysis as a definitive and accurate description of the underlying dimensions in the phenomenon of interest, i.e. symptom burden in palliative care. There must be consideration of which variables (symptoms) were analysed and whether that set was an accurate, empirically informed representation of the full construct. The psychometric properties and the content of the ESAS (Bruera et al., 1991) has varied considerably from one study to the next, this has likely led to the variation in factor structures identified.

## 5.2 Rasch Analysis

The results of the Rasch analysis support the ESAS (Bruera et al., 1991) as a global unidimensional measure of symptom distress. Thus, the ESAS-SDS score is an accurate measure of overall symptom distress in the population researched for this study. This finding is equivalent to that of Cheifetz et al. (2014) who also established that the 9-item version of the ESAS (Bruera et al., 1991) fitted the requirements for the Rasch model and measured the same overall construct when used in a population of ambulatory cancer centre patients.

*Rescoring*. This study found that responses to the ESAS (Bruera et al., 1991) were clustered around the high and low ends. This is similar to the findings of Cheifetz et al. (2014), who that noted responses to individual items were clustered at both the low and the high response points of the scale. How raters use the scoring options for a measure is an important factor when considering the reliability of the score (Bond & Fox, 2015). The extreme scoring on the ESAS (Bruera et al., 1991) signified that patients are not easily able to discriminate between the 11 response points (0-10) of the ESAS (Bruera et al., 1991). No previous research that has undertaken iterative rescoring of the ESAS (Bruera et al., 1991).

Bond and Fox (2015) state that the iterative rescoring process evaluates the how raters are responding to the items and allows the researcher to adjust scoring options to achieve optimal results. Uniform rescoring of the 11 response points (0-10) to four response points (0 = no presence of symptom, 1 = mild presence of symptom, 2 = moderate presence of symptom, and 3 = severe presence of symptom) resulted in an improvement in the reliability, and overall fit to the Rasch model.

*Differential item functioning (DIF).* The evaluation of DIF revealed that nurses (as the predominant type of rater) were not biased in their scoring of participants based on

individual demographics such as age, sex, diagnosis, and ethnicity. This is aligned with the findings of Cheifetz et al. (2014) who reported no DIF for age, sex, and diagnosis, in a group where patients were the predominate rater. This indicates that there is no disadvantage for one group over the other, despite who is completing the scale.

# Chapter 6 - Conclusion and Implications

## 6.1 Limitations of the Present Research

As data was collected during admission to a community hospice service, and very few patients had a repeated score, the range of symptoms exhibited may be at a lower severity when compared with a group of scores taken at regular intervals throughout their admission. This means that the present study could not to investigate how the ESAS (Bruera et al., 1991) performed with scores at extreme ends of the scale. Having a data set with higher mean scores for each item could potentially affect the psychometric properties of the measure.

Secondly, there is an underrepresentation of the younger age group (M=70.86 years). Furthermore, as the study was undertaken within a small New Zealand population the ethnicities are likely to be notably different to those in larger cities overseas. Therefore a full evaluation of DIF by age and ethnicity may be limited, particularly in the context of the high uptake of proxy scoring in the present sample. This is a considerable limitation as the DIF results from this study only advise us if there are any differences in the way nurses are completing the ESAS (Bruera et al., 1991). While this study confirmed nurses are not biased in their scoring of patients based on their age, sex, diagnosis and ethnicity, it is vital similar analysis be conducted on data where patients had completed the ESAS (Bruera et al., 1991).

Finally, while some missing data is expected, the high prevalence of missing data in this research could mean that the overall statistical power of the research is reduced (Kang, 2013). However missing data were handled appropriately using statistical techniques such as pairwise exclusion to minimise the negative impact on the studies ability to draw conclusions.

## 6.2 Future Research

The present research has identified several key focus points for future research. There is an urgent need for research into the ability of the ESAS (Bruera et al., 1991) to accurately evaluate change in patient symptoms over time. This is particularly important for understanding change in a group of community palliative patients ESAS (Bruera et al., 1991) scores repeated at regular intervals throughout their admission. This would enable evaluation the performance of the ESAS (Bruera et al., 1991) with a wider range of symptom severity and allow researchers to establish the stability of the

measure over time. Further factor analysis should be undertaken across a variety of settings to determine whether a clear and replicable factor structure exists, and in which specific clinical contexts it might apply.

A thorough investigation into the clinical utility of the ESAS (Bruera et al., 1991) within a variety of settings is suggested. Firstly, to investigate why this measure is not being used in the intended way, as a patient self-report scale. Secondly, to explore barriers to reassessment using the ESAS (Bruera et al., 1991) and to understand why several items were commonly left unreported.

## 6.3 Conclusions

The results of this study has highlighted key concerns in terms of the factor structure and clinical utility of the 12-item version of the ESAS (Bruera et al., 1991) in the present research setting. The predominant practice of proxy reporting of the ESAS (Bruera et al., 1991) by the registered nurse is contrary to the intended use of the ESAS (Bruera et al., 1991) and may negatively affect the ability of the ESAS (Bruera et al., 1991) to accurately evaluate patient status. Furthermore, the practice of infrequent re-assessment of patients suggested that the ESAS (Bruera et al., 1991) was not being used to evaluate symptom management techniques that had been employed as a result of the initial scores. Future research should establish the barriers to using the ESAS (Bruera et al., 1991) as intended.

Psychometric evaluation of the 12-item version of the ESAS (Bruera et al., 1991) in the present research setting confirmed it was a reliable measure (.79). The three factors found using EFA were comprised of different items to previous research, which suggests a lack of consistency in factor structures.

Rasch analysis confirmed the unidimensionality of the12-item version of the ESAS (Bruera et al., 1991). While this study did not report any DIF, the majority of scores were completed by the registered nurse, therefore these results only observe for nursing bias when scoring the ESAS (Bruera et al., 1991). The highlight of this study is being first to undertake iterative rescoring using Rasch analysis on any version of the ESAS (Bruera et al., 1991). Future applications of the ESAS (Bruera et al., 1991) using collapsed scoring responses may therefore provide a measure that has increased reliability and is easier to use.

This study has provided vital information about the utility, reliability, and validity of the 12-item version of the ESAS (Bruera et al., 1991) within a New Zealand palliative care setting. To date very few studies have been undertaken within a setting that includes patients with both cancer and non-cancer diagnosis. This is important as in New Zealand and Australia it is common practice to care for both patient groups within the palliative care setting. It is therefore essential further investigation into the psychometric properties of the ESAS (Bruera et al., 1991) is undertaken across a variety of settings to observe if findings are similar. Furthermore, work focusing on the regulation of the numerous versions of the ESAS (Bruera et al., 1991) is highly recommended. The constant alterations to this measure makes it challenging for clinicians to accurately assess the reliability and validity of the ESAS (Bruera et al., 1991).

# References


Aktas, A., Walsh, D., & Kirkova, J. (2015). The psychometric properties of cancer multisymptom assessment instruments: a clinical review. *Supportive Care in Cancer, 23*(7), 2189-2202. doi:10.1007/s00520-015-2732-7

American Educational Research Association, A. P. A., National Council on Measurement in Education. (1999). Standards for Educational and Psychological Testing. Washington: American Educational Research Association.

Barata, P. C., Cardoso, A., Custodio, M. P., Alves, M., Papoila, A. L., António, B., & Lawlor, P. G. (2016). Symptom clusters and survival in Portuguese patients with advanced cancer. *Cancer Medicine, 5*(10), 2731-2739. doi:10.1002/cam4.860

Bartlett, M. S. (1954). A note on the multiplying factors for various Chi Square approximations. *Journal of the Royal Statistical Society*(16), 296-298.

Bond, T., & Fox, C. M. (2015). *Applying the Rasch Model* (Third ed.). London: Taylor and Francis Group. Retrieved from http://ebookcentral.proquest.com/lib/aut/detail.action?docID=3569314

Bruera, E., Kuehn, N., Miller, M., Selmser, P., & Macmillan, K. (1991). The Edmonton Symptom Assessment System (ESAS): A simple method for the assessment of palliative care patients. *Journal of Palliative Care, 7*(2), 6-9.

Carvajal, A., Centeno, C., Watson, R., & Bruera, E. (2011). A comprehensive study of psychometric properties of the Edmonton Symptom Assessment System (ESAS) in Spanish advanced cancer patients. *European Journal of Cancer, 47*(12), 1863-1872. doi:10.1016/j.ejca.2011.03.027

Chang, V. T., Hwang, S. S., & Feuerman, M. (2000). Validation of the Edmonton Symptom Assessment Scale. *Cancer, 88*(9), 2164-2171. doi:10.1002/(SICI)1097-0142(20000501)88:9<2164::AID-CNCR24>3.0.CO;2-5

Cheifetz, O., Packham, T. L., & Macdermid, J. C. (2014). Rasch analysis of the Edmonton Symptom Assessment System and research implications. *Current Oncology, 21*(2), e186-194. doi:10.3747/co.21.1735

Chinda, M., Jaturapatporn, D., Kirshen, A. J., & Udomsubpayakul, U. (2011). Reliability and validity of a Thai version of the edmonton symptom assessment scale (ESAS-Thai). *Journal of Pain and Symptom Management, 42*(6), 954-960. doi:10.1016/j.jpainsymman.2011.02.020

Chow, E., Fan, G., Hadi, S., & Filipczak, L. (2007). Symptom clusters in cancer patients with bone metastases. *Supportive Care in Cancer, 15*(9), 1035-1043. doi:10.1007/s00520-007-0241-z

Claessens, P., Menten, J., Schotsmans, P., & Broeckaert, B. (2011). Development and validation of a modified version of the Edmonton Symptom Assessment Scale in a Flemish palliative care population. *American Journal of Hospice and Palliative Care, 28*(7), 475-482. doi:10.1177/1049909111400724

Clark, L., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*, 316-336.

Cleeland, C., Mendoza, T. R., Wang, X. S., Chou, C., Harle, M. T., Morrissey, M., & Engstrom, M. C. (2000). Assessing symptom distress in cancer patients. *Cancer, 89*(7), 1634-1646. doi:10.1002/1097-0142(20001001)89:7<1634::AID-CNCR29>3.0.CO;2-V

Cohen, R., & Swerdlik, M. (2005). *Psychological testing and assessment* (sixth ed.). New York: McGraw Hill.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. doi:10.1007/BF02310555

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychology Bulletin, 52*(4), 281-302.

Davison, S. N., Jhangri, G. S., & Johnson, J. A. (2006). Cross-sectional validity of a modified Edmonton Symptom Assessment System in dialysis patients: A simple assessment of symptom burden. *Kidney International, 69*(9), 1621-1625. doi:10.1038/sj.ki.5000184

de Haes, J., Olschewski, M., Fayers, P., Visser, M., Cull, A., Hopwood, P., Sanderman, R. (2012). *The Rotterdam Symptom Checklist (RSCL)*. Netherlands: University of Groningen.

de Vet, H., Terwee, C., Mokkink, L., B.,, & Knol, D. (2011). *Measurement in Medicine*. United Kingdom: Cambridge University Press.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (third ed.). California: Sage.

ECOG-ACRIN cancer research group. (2015). *ECOG performance status*. Retrieved 01/02/2016, 2016, from http://ecog-acrin.org/resources/ecog-performance-status

Ferreira, K. A. S. L., Kimura, M., Teixeira, M. J., Mendoza, T. R., da Nóbrega, J. C. M., Graziani, S. R., & Takagaki, T. Y. (2008). Impact of cancer-related symptom synergisms on health-related quality of life and performance status. *Journal of Pain and Symptom Management, 35*(6), 604-616. doi:10.1016/j.jpainsymman.2007.07.010

Ferrero, A., Petri, D., Carbone, P., & Catelani, M. (2015). *Modern measurements*. Hoboken, USA: John Wiley & Sons, Incorporated. Retrieved from http://ebookcentral.proquest.com/lib/aut/detail.action?docID=4040589

Ganesh, V., Zhang, L., Chan, S., Wan, B. A., Drost, L., Tsao, M., . . . Chow, E. (2017). An update in symptom clusters using the Edmonton Symptom Assessment System in a palliative radiotherapy clinic. *Supportive Care in Cancer*. doi:10.1007/s00520-017-3749-x

Gapstur, R. L. (2007). Symptom burden: A concept analysis and implications for oncology nurses. *Oncology Nurses Forum, 34*(3), 673-680. doi:10.1188/07.onf.673-680

Ginty, A. T. (2013). Psychometric properties. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine* (pp. 1563-1564). New York, USA: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-1-4419-1005-9_480. doi:10.1007/978-1-4419-1005-9_480

Gorsuch, R. L. (2014). *Factor analysis : Classic edition*. Florence, USA: Taylor and Francis. Retrieved from http://ebookcentral.proquest.com/lib/aut/detail.action?docID=1873793

Hannon, B., Dyck, M., Pope, A., Swami, N., Banerjee, S., Mak, E., . . . Zimmermann, C. (2015). Modified Edmonton Symptom Assessment System including constipation and sleep: validation in outpatients with cancer. *Journal of Pain and Symptom Management, 49*(5), 945-952. doi:10.1016/j.jpainsymman.2014.10.013

Haynes, S., Richard, D., & Kubany, E. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238-247.

Haynes, S., Smith, G., & Hunsley, J. (2011). *Scientific foundations of clinical assessment*. New York, USA: Taylor and Francis Group.

Hearn, J., & Higginson, I. J. (1997). Outcome measures in palliative care for advanced cancer patients: A review. *Journal of Public Health, 19*(2), 193-199.

Homsi, J., Walsh, D., Rivera, N., Rybicki, L. A., Nelson, K. A., Legrand, S. B., . . . Pham, H. (2006). Symptom evaluation in palliative medicine: Patient report vs

systematic assessment. *Supportive Care in Cancer*, *14*(5), 444-453. doi:10.1007/s00520-005-0009-2

Hospice West Auckland. (2015). *Clinical assessment procedure* [Guideline]. (Available from the Hospice West Auckland)

Hui, D., Shamieh, O., Paiva, C. E., Khamash, O., Perez-Cruz, P. E., Kwon, J. H., . . . Bruera, E. (2016). Minimal clinically important difference in the physical, emotional, and total symptom distress scores of the Edmonton Symptom Assessment System. *Journal of Pain and Symptom Management, 51*(2), 262-269. doi:http://dx.doi.org/10.1016/j.jpainsymman.2015.10.004

Kaasa, S., & Loge, J. H. (2003). Quality of life in palliative care: Principles and practice. *Palliative Medicine*, *17*(1), 11-20. doi:10.1191/0269216303pm662ra

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401-415. doi:10.1007/bf02291817

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *65*(5), 402-406. doi:http://doi.org/10.4097/kjae.2013.64.5.402

Karnofsky, D. A., & Burchenal, J. H. (1949). *The clinical evaluation of chemotherapeutic agents in cancer in: C. M. MacLeod, Ed., Evaluation of chemotherapeutic agents*. New York: Columbia University Press.

Kelly, C. M., & Shahrokni, A. (2016). Moving beyond Karnofsky and ECOG performance status assessments with new technologies. *Journal of Oncology, 2016*, 13. doi:10.1155/2016/6186543

Kline, P. (2000). *Handbook of psychological testing* (second ed.). London: Routledge.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*(3), 151-160. doi:10.1007/BF02288391

Lopez, G., Liles, D. K., & Knupp, C. L. (2014). Edmonton Symptom Assessment System for outpatient symptom monitoring of sickle cell disease. *South Medical Journal*, *107*(12), 768-772. doi:10.14423/smj.0000000000000209

MacLeod, R., Vella-Brincat, J., & MacLeod, S. (2016). *The palliative care handbook* (Vol. eigth). Retrieved from http://www.hospice.org.nz/cms_show_download.php?id=1243

Marais, I., & Andrich, D. (2008). Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement*, *9*(3), 200-215.

McCorkle, R., & Young, K. (1978). Development of a symptom distress scale. *Cancer Nurs*, *1*(5), 373-378.

Medvedev, O., Siegert, R., Feng, X., Billington, D., Jang, J., & Krägeloh, C. (2016). Measuring trait mindfulness: How to improve the precision of the Mindful Attention Awareness Scale using a Rasch model. *Mindfulness*, *7*(2), 384-395. doi:10.1007/s12671-015-0454-z

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*, 5-11.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749. doi:10.1037/0003-066x.50.9.741

Moro, C., Brunelli, C., Miccinesi, G., Fallai, M., Morino, P., Piazza, M., . . . Ripamonti, C. (2006). Edmonton Symptom Assessment Scale: Italian validation in two palliative care settings. *Supportive Care in Cancer*, *14*(1), 30-37. doi:10.1007/s00520-005-0834-3

Nekolaichuk, C., Bruera, E., Spachynski, K., MacEachern, T., Hanson, J., & Maguire, T. O. (1999). A comparison of patient and proxy symptom assessments in advanced cancer patients. *Palliative Medicine, 13*(4), 311-323.

Nekolaichuk, C., Maguire, T. O., Suarez-Almazor, M., Rogers, W. T., & Bruera, E. (1999). Assessing the reliability of patient, nurse, and family caregiver symptom ratings in hospitalized advanced cancer patients. *Journal of Clinical Oncology, 17*(11), 3621-3630.

Nekolaichuk, C., Watanabe, S., & Beaumont, C. (2008). The Edmonton Symptom Assessment System: A 15-year retrospective review of validation studies (1991--2006). *Palliative Medicine, 22*(2), 111-122. doi:10.1177/0269216307087659

Oken, M. M., Creech, R. H., Tormey, D. C., Horton, J., Davis, T. E., McFadden, E. T., & Carbone, P. P. (1982). Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American Journal of Clinical Oncology, 5*(6), 649-655.

Paice, J. A. (2004). Assessment of symptom clusters in people with cancer. *Journal of the National Cancer Institute Monographs*(32), 98-102. doi:10.1093/jncimonographs/lgh009

Pallant, J. (2016). *SPSS survival manual* (sixth ed.). Sydney, Australia: Allen & Unwin.

Pallant, J., & Tennant, A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology, 46*(1), 1-18.

Pautex, S., Berger, A., Chatelain, C., Herrmann, F., & Zulian, G. B. (2003). Symptom assessment in elderly cancer patients receiving palliative care. *Critical Reviews in Oncology/Hematology, 47*(3), 281-286. doi:10.1016/S1040-8428(03)00043-X

Philip, J., Smith, W. B., Craft, P., & Lickiss, N. (1998). Concurrent validity of the modified Edmonton Symptom Assessment System with the Rotterdam Symptom Checklist and the Brief Pain Inventory. *Supportive Care in Cancer, 6*(6), 539-541. doi:10.1007/s005200050212

Polit, D. F., & Yang, F. M. (2016). *Measurement and the measurement of change.* Philadelphia, USA: Wolters Kluwer.

Portenoy, R. K., Thaler, H. T., Kornblith, A. B., Lepore, J. M., Friedlander-Klar, H., Kiyasu, E., . . . et al. (1994). The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *European Journal of Cancer, 30a*(9), 1326-1336.

Porter, M. E., Larsson, S., & Lee, T. H. (2016). Standardizing patient outcomes measurement. *New England Journal of Medicine, 374*(6), 504-506. doi:10.1056/NEJMp1511701

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*: Danmarks Paedagogiske Institut. Retrieved from https://books.google.co.nz/books?id=aB9qLgEACAAJ

Richardson, L. A., & Jones, G. W. (2009). *A review of the reliability and validity of the Edmonton Symptom Assessment System* (Vol. 16). Retrieved from http://www.current-oncology.com/index.php/oncology/article/view/261

Siegert, R., Tennant, A., & Turner-Stokes, L. (2010). Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disability and Rehabilitation, 32*(1), 8-17.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*(1), 72-101.

Strömgren, A. S., Groenvold, M., Pedersen, L., Olsen, A. K., & Sjogren, P. (2002). Symptomatology of cancer patients in palliative care. *European Journal of Cancer, 38*(6), 788-794. doi:10.1016/S0959-8049(01)00470-1

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. doi:10.5116/ijme.4dfb.8dfd

Taylor, A. E., Olver, I. N., Sivanthan, T., Chi, M., & Purnell, C. (1999). Observer error in grading performance status in cancer patients. *Supportive Care in Cancer, 7*(5), 332-335. doi:10.1007/s005200050271

Taylor, D., Mallory, L. J., Lichstein, K. L., Durrence, H. H., Riedel, B. W., & Bush, A. J. (2007). Comorbidity of chronic insomnia with medical problems. *Sleep, 30*(2), 213-218. doi:10.1093/sleep/30.2.213

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum, 57*(8), 1358-1362. doi:10.1002/art.23108

Valderas, J. M., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C. E., Halyard, M. Y., . . . Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: A systematic review of the literature. *Quality of Life Research, 17*(2), 179-193. doi:10.1007/s11136-007-9295-0

Watanabe, S., Nekolaichuk, C., Beaumont, C., & Mawani, A. (2009). The Edmonton symptom assessment system--what do patients think? *Supportive Care in Cancer, 17*(6), 675-683. doi:10.1007/s00520-008-0522-1

Waterlow, J. (1985). Pressure sores: a risk assessment card. *Nurs Times, 81*(48), 49-55.

Weisbord, S. D., Carmody, S. S., Bruns, F. J., Rotondi, A. J., Cohen, L. M., Zeidel, M. L., & Arnold, R. M. (2003). Symptom burden, quality of life, advance care planning and the potential value of palliative care in severely ill haemodialysis patients. *Nephrology Dialysis Transplantation, 18*(7), 1345-1352. doi:10.1093/ndt/gfg105

Wright, B. D. (1996). Local dependency, correlations and principal components. *Rasch Measurement Transactions*(10), 509-511.

Yokomichi, N., Morita, T., Nitto, A., Takahashi, N., Miyamoto, S., Nishie, H., . . . Ogawa, A. (2015). Validation of the Japanese version of the Edmonton Symptom Assessment System–revised. *Journal of Pain and Symptom Management, 50*(5), 718-723. doi:http://dx.doi.org/10.1016/j.jpainsymman.2015.05.014

# Appendix A: Ethics Approval

**AUTEC Secretariat**

Auckland University of Technology
D-88, WU406 Level 4 WU Building City Campus
T: +64 9 921 9999 ext. 8316
E: ethics@aut.ac.nz
www.aut.ac.nz/researchethics

16 June 2016

Margaret Roberts
Faculty of Health and Environmental Sciences

Dear Margaret

**Ethics Application:** **16/216 An evaluation of the validity of the Edmonton Symptom Assessment System when used in a community hospice setting.**

Thank you for submitting your application for ethical review. I am pleased to advise that the Auckland University of Technology Ethics Committee (AUTEC) approved your ethics application at their meeting on 13 June 2016 subject to the following condition:

1. Clarification about how and against what the comparison of the measurements will occur.

Please provide me with a response to the point raised in this condition, indicating either how you have satisfied the point or proposing an alternative approach. AUTEC also requires copies of any altered documents, such as Information Sheets, surveys etc. You are not required to resubmit the application form again. Any changes to responses in the form required by the committee in their conditions may be included in a supporting memorandum.

Please note that the Committee is always willing to discuss with applicants the points that have been made. There may be information that has not been made available to the Committee, or aspects of the research may not have been fully understood.

Once your response is received and confirmed as satisfying the Committee's points, you will be notified of the full approval of your ethics application. Full approval is not effective until all the conditions have been met. Data collection may not commence until full approval has been confirmed. If these conditions are not met within six months, your application may be closed and a new application will be required if you wish to continue with this research.

To enable us to provide you with efficient service, we ask that you use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at ethics@aut.ac.nz.

I look forward to hearing from you,

Yours sincerely

Kate O'Connor
Executive Secretary
**Auckland University of Technology Ethics Committee**

Cc: Emma Sprague emma_sprague@hotmail.com

Appendix B: Tools

i) Example of ESAS (Bruera et al., 1991) as it appears in PalCare (PalCare Pty Ltd, 2017).

Ms. Enquiries, Patient

ZZZ9999 | DOB: 01-Jan-00 | West Auckland

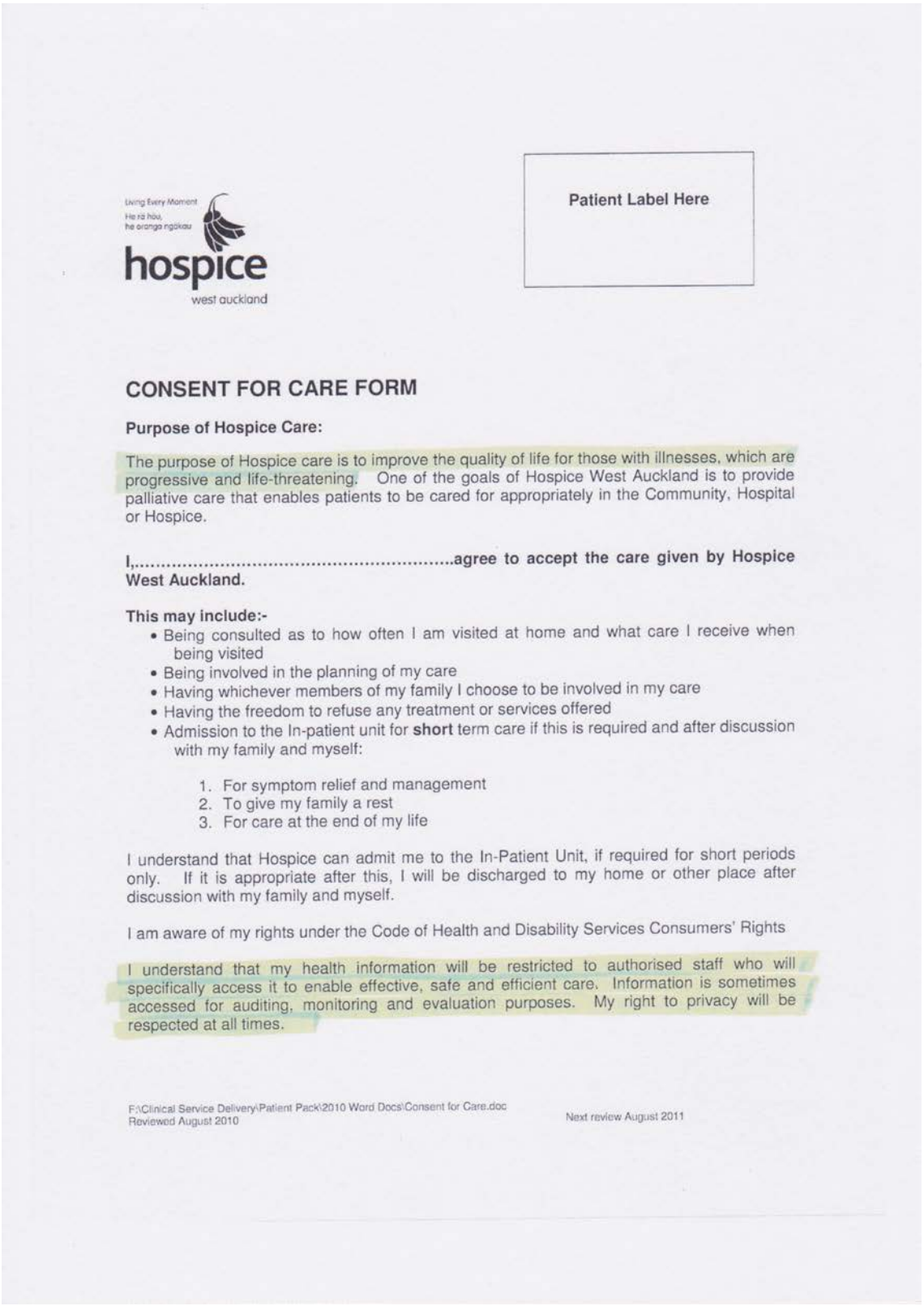
ESAS

Completed By: S
Phase: U
Date: 12 May 16

Pain
Tiredness
Nausea
Depression
Anxiety
Drowsiness
Appetite
Wellbeing
Shortness of Breath
Complexity of Care
Constipation
Insomnia

Other Symptoms

ii) Copy of consent for all patients at Hospice West Auckland, signed upon admission to service.



Living Every Moment
He rā hōu,
he oranga ngākau

hospice
west auckland

Patient Label Here

## CONSENT FOR CARE FORM

**Purpose of Hospice Care:**

The purpose of Hospice care is to improve the quality of life for those with illnesses, which are progressive and life-threatening. One of the goals of Hospice West Auckland is to provide palliative care that enables patients to be cared for appropriately in the Community, Hospital or Hospice.

**I,.....agree to accept the care given by Hospice West Auckland.**

**This may include:-**

- Being consulted as to how often I am visited at home and what care I receive when being visited
- Being involved in the planning of my care
- Having whichever members of my family I choose to be involved in my care
- Having the freedom to refuse any treatment or services offered
- Admission to the In-patient unit for **short** term care if this is required and after discussion with my family and myself:

    1. For symptom relief and management
    2. To give my family a rest
    3. For care at the end of my life

I understand that Hospice can admit me to the In-Patient Unit, if required for short periods only. If it is appropriate after this, I will be discharged to my home or other place after discussion with my family and myself.

I am aware of my rights under the Code of Health and Disability Services Consumers' Rights

I understand that my health information will be restricted to authorised staff who will specifically access it to enable effective, safe and efficient care. Information is sometimes accessed for auditing, monitoring and evaluation purposes. My right to privacy will be respected at all times.

F:\Clinical Service Delivery\Patient Pack\2010 Word Docs\Consent for Care.doc
Reviewed August 2010

Next review August 2011

I have read this Consent for Care Form and the information provided. I have had the opportunity for discussions with:

.....

I understand that if I have any concerns about my palliative care, I may contact the Community Palliative Care Team Manager, Hospice West Auckland Ph (09) 834 9755.

Signed: ..... Date: ...../...../.....

Witness:
(Hospice West Auckland Staff Member, including designation)

Signed: ..... Date: ...../...../.....

Designation: .....

Appendix C: Conversion of Ordinal Scores into Interval-Level Data, Logits and Scale Metric for the Modified ESAS.

| Ordinal | Interval (Logits) | Interval (Metric) |
|---|---|---|
| 0 | -8.64 | .00 |
| 1 | -5.69 | 7.45 |
| 2 | -4.21 | 11.18 |
| 3 | -3.47 | 13.05 |
| 4 | -2.92 | 14.43 |
| 5 | -2.49 | 15.53 |
| 6 | -2.13 | 16.42 |
| 7 | -1.84 | 17.18 |
| 8 | -1.58 | 17.83 |
| 9 | -1.35 | 18.42 |
| 10 | -1.13 | 18.95 |
| 11 | -.94 | 19.45 |
| 12 | -.75 | 19.91 |
| 13 | -.58 | 20.35 |
| 14 | -.42 | 20.75 |
| 15 | -.27 | 21.12 |
| 16 | -.14 | 21.47 |
| 17 | -.01 | 21.79 |
| 18 | .12 | 22.10 |
| 19 | .23 | 22.40 |
| 20 | .35 | 22.70 |
| 21 | .47 | 23.00 |
| 22 | .60 | 23.31 |
| 23 | .73 | 23.64 |
| 24 | .86 | 23.99 |
| 25 | 1.01 | 24.35 |
| 26 | 1.16 | 24.73 |
| 27 | 1.31 | 25.11 |
| 28 | 1.46 | 25.49 |
| 29 | 1.62 | 25.89 |
| 30 | 1.78 | 26.30 |
| 31 | 1.98 | 26.79 |
| 32 | 2.23 | 27.44 |
| 33 | 2.64 | 28.46 |
| 34 | 3.32 | 30.18 |
| 35 | 4.28 | 32.60 |
| 36 | 5.63 | 36.00 |

Note: The following rescoring is required before converting into interval scale: retain 0 unchanged, rescore 1 to 5 as 1, 6 to 9 as 2, and 10 as 3. Calculate the sum of rescored values and find the ordinal score and corresponding interval score in logit units and scale metric on the right side.