Measuring Frailty in Older Canadians: An Analysis of the Canadian Longitudinal Study on Aging

By

David M. Kanters, B.Sc.

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

TITLE: Measuring Frailty in Older Canadians: An Analysis of the Canadian Longitudinal Study on Aging (CLSA)

AUTHOR: David M. Kanters, B. Sc. (Hon), University of Waterloo

SUPERVISOR: Dr. Lauren Griffith

NUMBER OF PAGES: 109

ABSTRACT

Introduction:

Frailty is characterized by vulnerability to declining health and increased risk for adverse health outcomes. Measuring frailty would be beneficial for developing interventions and assessing healthcare resource needs. No standardized measurement tool for frailty has been established. The objective of this thesis was to evaluate the frailty of participants in the Canadian Longitudinal Study on Aging (CLSA).

Methods:

A Frailty Index (FI) was constructed for CLSA participants based on the cumulative deficit theory of frailty. Exploratory factor analysis was conducted to study the underlying constructs of frailty and identify key factors. A hypothesized measurement model for frailty was specified. The model was modified and tested using structural equation modelling (SEM) to improve goodness-of-fit. A new frailty measurement tool was created and the construct validity of the new tool and the Frailty Index were evaluated.

Results:

A FI was calculated for 20,874 CLSA participants (Mean 0.14 SD 0.07). The maximum FI value was 0.68. A model containing all hypothesized variables had good fit of the data, and all variables contributed significantly. A simplified model also showed good fit and included four domains: upper-body strength, lower-body strength, dexterity, and depressive symptoms. These results persisted in an independent dataset. A Simplified Frailty (SF) score was created based on this simplified model. The FI and SF scores showed significant agreement and associations with sociodemographic variables were as predicted.

Conclusions:

A FI was simple to construct in the CLSA, having good fit of the data and construct validity. These results are consistent with previous research on the cumulative deficit theory of frailty. A simplified frailty model revealed key domains of frailty and resulted in a potentially useful short screening tool. The FI is recommended as a valid and reproducible approach for measuring frailty in the CLSA and similar population datasets.

ACKNOWLEDGEMENTS

There are many people without whom this work could not have been accomplished. I could not ask for a better mentor and guide throughout this thesis than Dr. Lauren Griffith. Dr. Julie Richardson and Dr. Christopher Patterson have my sincere thanks for sitting on my committee and being so generous with their time, guidance, and feedback. I am also grateful to Dr. David Hogan, Dr. Parminder Raina, and Dr. Kenneth Rockwood for lending their considerable expertise and for challenging me to improve both myself and my work.

I would like to thank my family, friends, and colleagues who have helped and encouraged me along the way. A special thanks to Allie, whose love, patience, and support for me have yet to find their limits. Finally, I give thanks to Him who makes all things possible for me.

> Commit to the LORD whatever you do,
> and he will establish your plans.
>
> Proverbs 16:3

This research was made possible using data collected by the Canadian Longitudinal Study on Aging (CLSA) [Data set version 1.1]. Funding for the CLSA is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under grant reference: LSA 94473 and the Canada Foundation for Innovation. The CLSA is led by Drs. Parminder Raina, Christina Wolfson and Susan Kirkland.

LIST OF ABBREVIATIONS USED

ADL	Activities of Daily Living
AFT	Animal Fluency Test
AIC	Akaike Information Criterion
AMSTAR	A Measurement Tool to Assess Systematic Reviews
BMI	Body Mass Index
CCHS	Canadian Community Health Survey
CES-D 10	Centre for Epidemiological Studies Short Depression Scale
CFI	Comparative Fit Index
CLSA	Canadian Longitudinal Study on Aging
EFA	Exploratory Factor Analysis
FI	Frailty Index
IADL	Instrumental Activities of Daily Living
ICF	International Classification of Functioning, Disability, and Health
MOS	Medical Outcomes Study
MAT	Mental Alternation Test
NNFI	Bentler Non-Normed Fit Index
OARS	Older Adults Resources and Services Program
RAVLT	Rey Auditory Verbal Learning Test
SD	Standard Deviation
SEM	Structural Equation Modelling
SF	Simplified Frailty Score
SWLS	Satisfaction with Life Scale
WHO	World Health Organization

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background and Rationale

In Canada, people are living longer, and the population as a whole is aging.[1,2] With a growing population of older adults, understanding the changing healthcare needs of this population is important. Older adults aged 80+ have disproportionately high healthcare needs, making up 2% of the population but using up 20% of non-obstetrical hospital days, but not every older adult has a multitude of health problems.[2] In order to assess the healthcare needs of this population, it is necessary to evaluate vulnerability to declining health and adverse outcomes among people of the same age. This is commonly understood as frailty.[2] Frailty is defined as a "clinically recognizable state of increased vulnerability" across multiple physiologic systems resulting from age.[3] Frailty compromises one's ability to perform daily activities, take care of oneself, and deal with the stresses of day-to-day life[3] and is characterized by a state of heightened risk for adverse health outcomes, including mortality and institutionalization,[2–6] and functional decline that is related to but distinct from aging, disability, and multimorbidity.[7]

Currently there is no consensus regarding the etiology of frailty or how it should be detected in an aging persons.[4,5,8–11] Frailty is likely to result from the interplay of several factors and involve multiple systems. While age alone is a good predictor of the adverse health outcomes described above, studying frailty provides a better understanding of the heterogeneity of risk in populations who may be of a similar age. The ability to identify frail individuals could be helpful for researchers, clinicians, and health policy makers in a variety of ways.[11] Identifying populations who are frail will be helpful in planning healthcare resource needs to ensure adequate care can

be provided. With an improved understanding of frailty, preventative interventions can be developed to keep community-dwelling older adults healthy and in their own homes longer. With Canada's aging population, an improved understanding and management of frailty has the potential to make a significant impact on the healthcare system.

There are a number of competing approaches for measuring frailty, but attempts to operationalize frailty most commonly fall into one of two main theories:[2,12] The Fried phenotype of frailty model[4] and the cumulative deficit model proposed by Rockwood and Mitnitski.[13] The popular phenotype of frailty describes the degree of frailty as the presence of five criteria: exhaustion, weight loss, diminished activity level, slow walking speed, and weak grip strength.[4] The cumulative deficit model assesses the presence and severity of a list of health deficits, typically 30-70 items long.[13] The severity of frailty in this model is measured by the proportion of items on the list in which the person has deficits. Both the phenotype and cumulative deficit models have been shown to predict higher risks of adverse health outcomes.[14]

Several systematic reviews have been conducted with the purpose of identifying and comparing original frailty assessment tools.[12,15–20] Through a summary of these reviews, 58 original tools were identified, each identifying frailty using distinct criteria. The creation of so many tools shows continued uncertainty about the essential components and definition of frailty. The number and types of items included in the instruments vary widely, showing that the essential components of frailty, its mechanisms and determinants, are not well understood.[12,21,22] There is no consensus regarding the etiology of frailty or a standardized criterion reference (or gold standard) for how it should be detected in aging persons.[4,7,8,11] The published approaches report the ability to predict

higher risks of adverse health outcomes for those identified as frail compared to those classified as non-frail.[14] However, the differing methods for operationalizing frailty result in inconsistent prevalence rates of frailty among studies.[23] Substantial differences in feasibility and prognostic ability for all-cause mortality are present when multiple scales are used on a common dataset.[17]

Despite the extensive research on operational definitions of frailty, a greater understanding of the relationships between the proposed characteristics of frailty and measurable health outcomes associated with frailty is needed.[24,25] As part of this research, a conceptual framework for frailty will be used that includes biological, psychological, and social factors, organized into clearly-defined domains based on the International Classification of Functioning, Disability, and Health.[26] The use of this conceptual framework may help to make this and future frailty research more translatable and help move frailty researchers towards the eventual emergence of a successful consensus definition of frailty. An operational definition of frailty is necessary for its detection and measurement in the context of a large population dataset. An acceptable operational definition for frailty in the CLSA should foster collaboration, prevent duplication of effort, and enable synthesis of frailty research to drive forward the understanding of this syndrome to improve the care of those affected by frailty, and enhance support by their caregivers.

The Canadian Longitudinal Study on Aging (CLSA) is an excellent resource for the study of frailty, as it collects detailed data at multiple time points on a variety of determinants of health, including health conditions, physical and cognitive functioning, and environmental and social factors[27]. The CLSA includes a representative sample of Canadians from across the country. Previous research on frailty has often been restricted to populations ≥65 years. The detailed data collected in the CLSA provides an opportunity to study frailty in adults of a wider range of ages (45-85) than is

typically studied, which may provide new insights into the development and progression of frailty and pre-frailty, (the state of risk for becoming frail), in younger adults as well as study its epidemiology. The use of CLSA data to create a continuous frailty score will allow the use of this rich resource to explore potential predisposing and precipitating factors and determine the trajectory of frailty during the longitudinal component of the CLSA study.

1.2 Study Objectives and Hypotheses

The objective of this thesis is to evaluate the frailty of participants in the CLSA. This will be done by selecting variables within the CLSA database that have been shown in the literature to contribute to frailty, and using these variables to create a frailty score. This will enable the study of the underlying construct of frailty and identify factors included in the score that load on this construct. This project will lead to the inclusion of a frailty score into the CLSA as a new derived variable. This measurement of frailty can then be used by other researchers studying aging Canadians and emerging theories of frailty. The results of this study will assist other researchers to identify frail patients in similar large population datasets. This analysis is based on the tracking cohort of the CLSA. After this thesis is completed, the frailty score will be calculated for members of the comprehensive cohort, and compared with physical measures of frailty, such as gait speed, and biological markers of aging, which are available in the comprehensive but not tracking cohort.

The secondary objective is to evaluate the validity of the derived frailty score in the CLSA. This will be achieved first by exploring the latent constructs in the health indicators collected in the CLSA through exploratory factor analysis. Then, a measurement model will be developed based on the accumulated deficit theory of frailty and tested through structural equation modelling. Finally, the Frailty Index will be compared with a variety of physical and sociodemographic factors including: sex, age, education, number of chronic conditions, social participation, falls, and extent of care that individuals are receiving to determine if the correlations are valid as hypothesized. Since frailty cannot be measured directly, the frailty score resulting from this analysis must be tested against hypothesized associations with surrogate outcomes for frailty to evaluate its construct validity. If the relationships are as predicted, this study will provide convergent validity

evidence of our measure of frailty. This work will be beneficial for future work using the CLSA database to study emerging theories of frailty, as well as assisting other researchers to identify frail patients in similar large population datasets.

We hypothesize that frail and non-frail persons should be distinguishable in sociodemographic (e.g., sex, age, education), clinical (e.g., injuries, falls), and social characteristics (e.g., participation rates, extent of care received). Prior work on Canadian Community Health Survey, a database similar to the CLSA, indicates that frailty scores are higher in women and increase with age. Frail individuals are generally less autonomous and requiring and receiving more care. We predict that this will manifest in more frequent injuries and falls reported in those identified as frail. As a marker of frailty we plan to examine the amount of care received, including formal care, informal care, and the use of assistive devices. The amount and type of both paid and unpaid care received is reported in the CLSA, so this analysis will not be limited to those who can afford care. In addition, self-rated physical and mental health, and self-rated healthy aging are all predicted to be worse in frail participants. The most important criterion for a measure of frailty is the ability to predict adverse health outcomes, including mortality, development of new-onset or worsening disability, and nursing home admission. Although these outcomes cannot be tested in this dataset, the frailty score could be evaluated in future prospective studies by its ability to predict adverse health outcomes in CLSA participants.

1.3 Thesis Overview

This thesis will be presented in six chapters, the first of which is this introduction, which outlines the rationale for the study and research objectives.

Chapter two, *Models and Theories of Frailty*, provides a review of key theories of frailty that are used to guide this thesis.

Chapter three, *Systematic Reviews on Frailty Measurement*, presents a detailed review of the systematic reviews of primary studies developing measurements for frailty that were identified in a comprehensive literature review.

Chapter four, *Statistical Methods*, first presents the data source, the CLSA, including the sample collection method, variables of interest, outcomes assessed, and ethical considerations. The remainder of chapter four describes the development of the models used for factor analysis and Structural Equation Modelling (SEM).

Chapter five, *Results*, describes the dataset and the findings of the analyses conducted.

Chapter six, *Discussion and Conclusions*, details our interpretation of the results, a detailed discussion of the findings, the limitations of the study, and our recommendations moving forward in the research of frailty.

CHAPTER 2: MODELS AND THEORIES OF FRAILTY

This chapter reviews current frameworks for the concept of frailty, with emphasis on the two most prominent models currently in use: the phenotype of frailty developed by Fried[4] and the cumulative deficit model, or Frailty Index, developed by Rockwood and Mitnitski[13]. While the research examining the measurement of frailty is extensive, there is a lack of consensus regarding the conceptualization of frailty or its underlying mechanisms and determinants [21,22]. Researchers have suggested that a theoretical framework for frailty should include biological, psychological, and social factors to be useful[20,28]. Included in this chapter is a theoretical model of the organization of factors that make up frailty into clearly-defined domains based on the International Classification of Functioning (ICF) [29]. Using this established conceptual framework will help make this work translatable to that of other frailty researchers. We will consider the strengths and limitations of these models, as well as their potential for use in identifying frailty in CLSA participants.

2.1 Review of Key Theories of Frailty

Although the terms are sometimes used interchangeably to identify physically vulnerable older adults requiring enhanced care, frailty is related to, but separate from, disability and multimorbidity[7]. Multimorbidity is included in the set of risk factors theorized to contribute to frailty, while disability is often considered an outcome of frailty[4]. The measurement and clinical management of frailty is a distinct challenge. Two key models dominate the literature on frailty measurement: the phenotype of frailty model [4] and the cumulative deficit model [13]. Both the phenotype and cumulative deficit models have both been shown to predict higher risks of adverse health outcomes [14,30].

The phenotype of frailty model describes specific biological processes that together make up frailty. In this model, frailty is defined with following five criteria: self-reported exhaustion, unintentional weight loss, diminished activity level, slow walking speed, and weak grip strength.[4] Individuals are characterized as frail, pre-frail, or robust based on the number of criteria present; three or more defines frailty. Pre-frailty, or a state of at risk for becoming frail, is defined as one or two of the criteria for frailty. Pre-frailty describes individuals who are mildly frail and likely to become more frail, although this is not inevitable. It may be useful to identify pre-frail individuals, as they may be good candidates for preventative interventions to mitigate increasing frailty. This model forms the basis of much of the frailty literature. There are over 5000 citations of the original article and several measurement instruments for frailty developed building on this model.[32–34] There has also been a great deal of research establishing the predictive validity of the phenotype of frailty model in relation to cognitive functioning [35] and adverse health outcomes, including mortality.[4,36]

The cumulative deficit model assesses the presence and severity of items on a list of health deficits, typically 30-70 items long, and was initially developed to measure frailty with information that was typically collected in a comprehensive geriatric assessment.[13] Contrasting with the phenotypic model, in this model the severity of frailty (the Frailty Index) is measured by the proportion of items on the list in which the person has a deficit, and all deficits are weighted equally. Research has suggested that as long as the health deficits meet certain criteria, they can be selected at random, and still yield a valid measurement of an individual's frailty.[37] To meet these criteria, deficits included should represent a variety of organ systems and areas of health, including psychological, social, and environmental factors; deficits that are accumulated with age

should be included while those not sensitive to age should be avoided; no specific deficit is critical for the model and researchers can adapt the scale to fit the patient population they are studying and the data available. The Frailty Index provides a continuous variable showing where the person stands on a fit-to-frail spectrum. This conceptual model has also been extensively tested in validation studies and inspired several instruments designed to measure frailty through the accumulation of deficits.[38–41]

The main criticism of the phenotype model is that it may not account for all the relevant factors that contribute to frailty and may lack sensitivity, misclassifying some frail individuals as non-frail, simply because the factors contributing to their frailty are not included in the model. The cumulative deficit model incorporates many factors, but is limited in that it does not account for their relative importance or severity. When measuring frailty using the cumulative deficit model, each factor is considered independent in the Frailty Index, but many health deficits are likely to be interrelated. The inclusion of many factors may obscure frailty or pre-frailty in individuals affected by one or few key health attributes.

Both models are correlated with each other in their identification of frailty and prediction of adverse health outcomes,[42] with the cumulative deficit model proving at least equal if not superior to the phenotype of frailty for this purpose.[30] A direct comparison of the two models shows that a Frailty Index value of 0.25 or greater roughly corresponds to frailty in the phenotype model[42]. Individuals have been shown to move between states of frailty on both scales, with one third of all transitions from states of greater to less frailty.[43] This suggests that frailty is a dynamic process that might be either prevented or even remediated. Identification of those who are at risk of becoming frail or increasing in their severity of frailty may be useful in the future in determining

the need for interventions and predicting the burden on healthcare services. This study will compare different methods in order to make the first step towards measuring the frailty of adults in the CLSA database.

2.2 Model for Frailty and International Classification of Functioning, Disability and Health (ICF)

The previous chapter discussed the lack of a criterion reference for measuring frailty. One contributor to this ongoing problem is the continued lack of consensus among frailty researchers on a conceptual model or operational definition for frailty.[21,28,44,45] This study will use an operational definition for frailty that promotes understanding of the dynamic relationship between physical factors, changes in function, social, environmental, and psychological factors.[45]

Because of the lack of consensus in definitions of frailty, we plan to use the conceptual framework developed in the International Classification of Functioning, Disability and Health (ICF) to define frailty as a health state. The ICF is a conceptual basis for the definition and measurement of disability and health developed by the World Health Organization (WHO).[29] ICF has been shown to be an appropriate framework for an operational definition of frailty.[28,46,47] The ICF will inform the development and discussion of the model for frailty used in this thesis. The standardized language used in the ICF makes this work transparent to other researchers, facilitating the comparison of data from this study to data from the work of other researchers studying frailty, data from clinical settings, and population-based data.

The ICF conceptualizes a person's level of functioning as a dynamic interaction among the following areas: body structure and functions (anatomy and physiology), activity limitations (execution of tasks), participation (involvement in life situations), and environmental factors

which affect these experiences (facilitators and barriers). Each component of functioning and disability is made up of several domains, outlined in Table 2-1 below.[29]

Table 2-1: ICF components and domains/chapters

Body Function:	**Activities and Participation:**
Mental functions	Learning and applying knowledge
Sensory functions and pain	General tasks and demands
Voice and speech functions	Communication
Functions of the cardiovascular, haematological, immunological, and respiratory systems	Mobility
Functions of the digestive, metabolic, and endocrine systems	Self-care
Genitourinary and reproductive functions	Domestic life
Neuromuscular and movement-related functions	Interpersonal interactions and relationships
Functions of the skin and related structures	Major life areas
	Community, social, and civic life
Body Structure:	**Environmental Factors:**
Structure of the nervous system	Products and technology
The eye, ear, and related structures	Natural environment and human-made changes to environment
Structures involved in voice and speech	Support and relationships
Structures of the cardiovascular, haematological, immunological, and respiratory systems	Attitudes
Structures related to the digestive, metabolic, and endocrine systems	Services, systems, and policies
Structures related to the genitourinary and reproductive systems	
Structures related to movement	
Skin and related structures	

Source: WHO 2001[29]

The ICF model defines each of its components as interrelated, a concept that also applies to frailty. For example, one domain of frailty, such as declining social participation, may be indicative of declining function in other domains such as activities of daily living or mobility, which may in turn cause a more severe decline in social participation. The strong relationship among multiple domains is part of what makes the ICF model a good fit to operationalize frailty in this context.

We will examine frailty in detail, organizing the data collected in the CLSA into the domains relevant to frailty. We will be measuring participants' self-reported "performance" or what participants report they can do in their current environment. Ideally, we would have data on "capacity", or the performance of participants in a standardized environment collected through clinical assessment. The gap between performance and capacity would inform researchers on potential environmental factors affecting frailty, but is beyond the scope of this study. The model for frailty developed in this study may be incomplete, since we are restricted to data available in the CLSA tracking cohort, but an empirical cross-sectional study of this magnitude is an important starting point for describing frailty using the ICF core set for frailty. The core set of categories from the ICF describes functioning in clinical practice by providing lists of categories that are relevant for specific health conditions and healthcare contexts.[48] Core sets have been developed for other areas of health research, including geriatrics, but a core set specific to frailty would be a useful step forward for research in this area.[47] While the ICF provides guidance for measuring these domains, we will have data collected using the tools developed by the CLSA, and using the ICF to organize these results and create a conceptual model. We hypothesize that items within domains should be correlated more than items across domains. This hypothesis will useful when conducting SEM to assess the fit of this model and evaluating its construct validity.

The sections of questionnaire items in the CLSA, called modules, related to frailty are grouped according their corresponding ICF domains in Table 2-2. These groupings of modules were used to inform the base structural equation model for this analysis. Wherever possible, the ICF was used to guide the organization of modules into the most relevant domain[29]. A consensus was reached for the groupings of modules after review by experts in the fields of population health, geriatrics, and physiotherapy.

Table 2-2. Levels of the ICF and their measurement in the CLSA

Level of the ICF	Measurement in CLSA	Description
Health conditions	Number of chronic conditions Self-rated health Self-rated mental health	Multimorbidity and chronic conditions
Impairments to body structure and function	BMI Continence Sensory impairment	Physical impairments
	Depression Satisfaction with life Anxiety Mood disorders	Psychological impairments
	Cognition	Cognition impairments
Activity limitations	Functional status	Problems with simple tasks related to strength, mobility, or dexterity
	ADL IADL	Decreased self-care
Participation restrictions	Social participation	Lack of participation in life roles
Environmental or contextual factors	Social support – availability	Problems with services or support systems

Modified from Fairhall et al[47]

2.3 Surrogate Outcomes for Frailty

As this thesis is based on cross-sectional analysis, we are unable to assess prospective outcomes such as mortality, hospitalization and long-term care admission. In lieu of these more concrete outcomes, we have assembled a set of sociodemographic characteristics associated with frailty identified by review of the reported literature and consultation with experts in the area. Age, education, income, injuries, and falls are all expected to differ between populations who are robust and those who are frail. If these factors vary as expected between those indicated as frail and non-frail by the instruments for measuring frailty developed in this analysis, it would provide evidence for their construct validity.

In addition to sociodemographic variables, "care received" will be used as a surrogate outcome for health service use that is associated with frailty. Care receiving is an attractive option as a surrogate for frailty, since it is necessary to have a variable that reflects frailty, but is not considered a contributor. The relationship between frailty and care received in the form of healthcare utilization has previously been established, with frail individuals requiring more services than non-frail individuals.[40,50–52] The care needed for older community-dwelling adults includes a broad range of services including personal care, nursing care, and informal care such as home visits or meal delivery. Previous studies on frailty and the need for home care have shown links between frailty and use of home care services.[53–55] This evidence suggests that individuals who are frail are likely to need more care than those who are non-frail or mildly frail.

Care received is measured in great detail in the CLSA in the form of self-reported paid and unpaid care received by participants. We have also included self-reported use of assistive devices. We plan to use these data to evaluate the amount of care received as a function of the frailty score in

order to evaluate our measure. Testing the hypothesized relationship between care received and the instruments for measuring frailty developed in this study will be done to evaluate the construct validity of the instruments.

The results of this analysis should be interpreted cautiously, as factors that are related to the physical domain of frailty are likely to affect care receiving more than factors in the psychological or social domains, since frail elderly with severe physical deficits are in more apparent need of care than those with psychological or social impairments, who may still be able to care for themselves on a day-to-day basis (add a citation). However, the physical, psychological, and social domains of frailty have all been separately linked to indicators of healthcare utilization, including hospitalization, personal care, nursing care, and informal care.[55] We hypothesize that that all measured domains of frailty will be related to the level of care received by CLSA participants.

CHAPTER 3: SYSTEMATIC REVIEWS ON FRAILTY MEASUREMENT

This chapter presents the methodology and findings of an overview of the published systematic reviews on original measurement tools for evaluating frailty in adults using self-reported data. The objective of this overview was to identify health indicators for frailty that are measured in the CLSA Tracking cohort. Our frailty analysis will then be based upon items identified in this literature overview, and applied to the modules in the CLSA Tracking cohort.

3.1 Methodology

Several high-quality systematic reviews were identified which compared original frailty measurement tools. A review published by Sternberg et al (2011)[15] was found to be the first high-quality systematic review available that fit our research objectives. Included in the review are original studies developing measurement tools for frailty, which are compiled and compared based on their quality and content. The review summarizes identifying factors of frailty, clinical operational definitions, and tools for identifying the presence and severity of frailty. Rather than duplicating this work, we used it as a framework for this overview of current research on frailty measurement. We aimed to find all subsequent published reviews to ensure that our summary is comprehensive. Ovid MEDLINE, Embase, and Cochrane Database of Systematic Reviews were searched for systematic reviews of studies reporting original tools for measuring frailty. The medical subject heading "Frail Elderly" was used as the primary search term to identify relevant reviews, as it is broad enough to cover the literature related to the concept of frailty. This term was combined with a number of other related keywords and headings to focus the search and minimize the extraneous literature captured. These includes MeSH headings "aging", "aged", "vulnerable populations", "longevity," "health status," "geriatrics," "risk assessment," "risk

factors," "health status indicators," "disability evaluation," "forecasting," "patient care planning," "biomarkers," "health surveys," or "diagnosis," as well as the keywords "operationalization," "successful ag?ing," or "healthy ag?ing". The search was limited to "review articles" in Medline and articles containing the keyword "systematic review" in Embase to restrict the search results to published systematic reviews. The detailed search strategy is shown in Appendix A. Database searches covered the period of January 2010 to January 2015, subsequently updated to June 2016. The search was supplemented by consultation with experts in frailty measurement, as well as hand-searching of the references of identified reviews to minimize the risk of missing relevant reviews. We determined that a new systematic review was not necessary, as the reviews identified adequately summarized frailty measurement studies during the period of interest.

This literature overview focused on the domains of frailty measured by each tool, to select items in the CLSA that have the potential to measure frailty. Domains describe the area of functioning or health that that is included in each tool. Eligibility criteria included peer-reviewed, systematic reviews of primary studies that included original frailty measurement tools using self-reported data from community-dwelling adults. Due to time and funding constraints, the overview was restricted to English language publications. After duplicates were removed, full text copies of potentially eligible reviews were obtained. Reviews were excluded if they focused on institutionalized populations, used measurements of frailty unobtainable by self-report, or measured health statuses other than frailty, such as disability. Reviews that used a narrow definition of frailty were also excluded. For example, reviews based only on the Fried phenotype of frailty, were excluded, as the (five) frailty domains are already known. The screening and selection process for the literature overview is presented as a flow diagram in Figure 3-1. The first

author, DK, screened each of the reviews and determined eligibility. The quality of the included systematic reviews was assessed using the AMSTAR checklist, the first tool validated as a means to assess the methodological quality of systematic reviews[56,57].

3.2 Findings

Overall, 286 reviews were identified by the search strategy, and 5 were found to meet the inclusion criteria. The characteristics of these reviews are reported in Table 3-1, which includes the criteria for inclusion in each review, the number of original frailty measurement tools included, the characteristics of those tools that were reported, and the quality of each systematic review.

Figure 3-1: Literature Review Flow Diagram

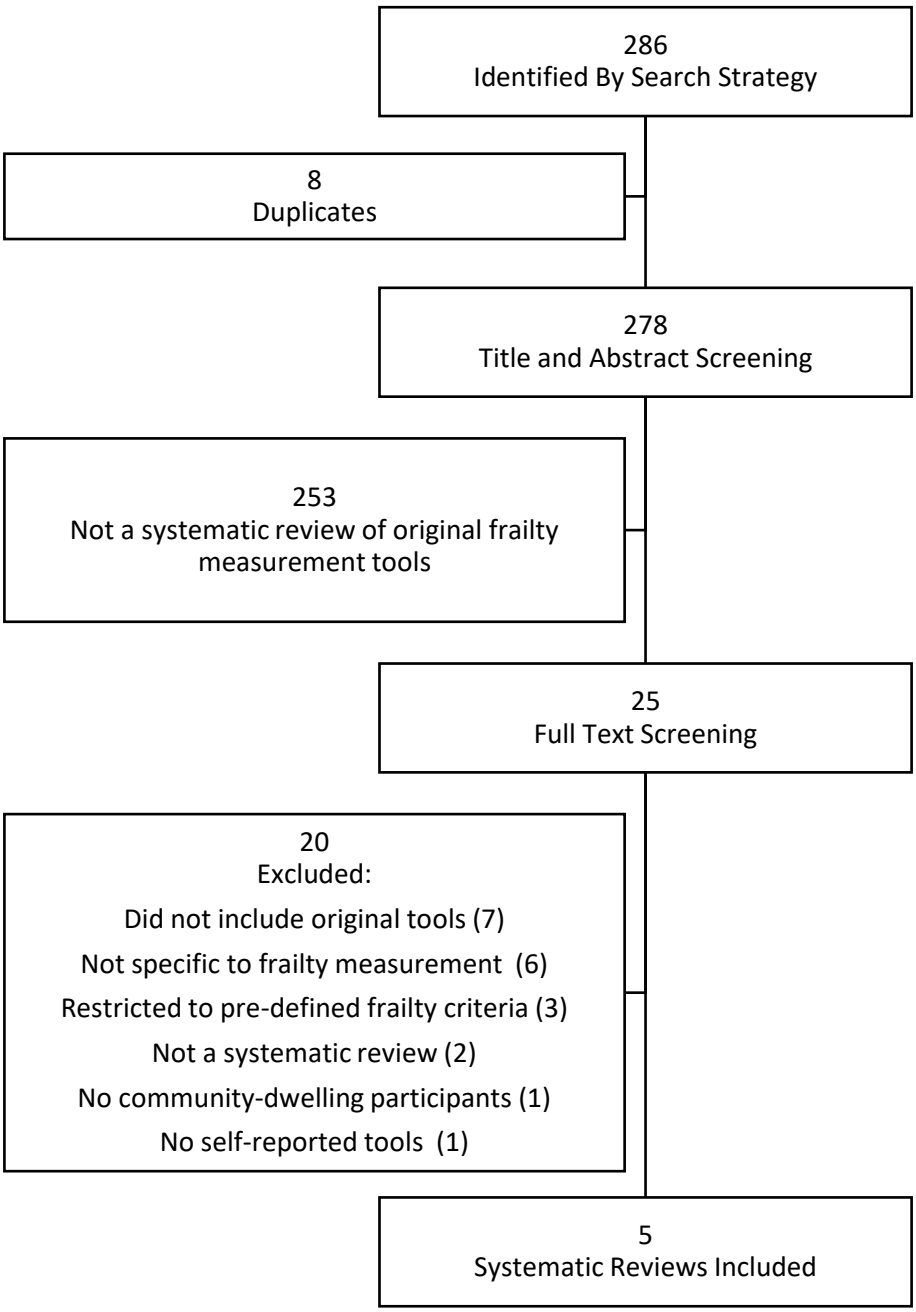

Table 3-1: Characteristics of included systematic reviews

First author (Year)	Inclusion Criteria				Number of Tools Included	Characteristics of included tools reported	AMSTAR Quality Score[56] (1-11)*
	Years	Language	Population	Frailty Measurement Tool			
Sternberg (2011)[15]	Published January 1997-December 2009	English or French	Community-dwelling adults aged 65 or older	Offered clinically-relevant outcomes, Described original tool	22	Quality, Recruitment period, Follow-up length, Sample size, Frailty domains assessed, Outcomes assessed	8
Bouillon (2013)[12]	Published 1948-May 2011	English, French, or Spanish	Participants aged 50 years and older	Measurement of frailty, Psychometric evaluation of frailty instrument	27	Setting, Sample size, Age, Frailty domains assessed, Criteria for frailty status, Reliability and validity	4
Hamaker (2012)[18]	Published before December 2011	None specified	Patients with cancer	Frailty screening method, comprehensive geriatric assessment (CGA)	7	Setting, Study population Sample size, Age, Sex, Frailty domains assessed, Criteria for frailty status	7
Pialoux (2012)[19]	Published before June 2011,	No restriction	Non-hospitalized patients with at least one frailty characteristic	Development/ psychometric evaluation of a screening tool, More than one item, Comparison to CGA, Tested in non-hospitalized population, Psychometric properties reported	10	Inclusion criteria, Number of items, Mode of administration, Language, Administration duration, Reference geriatric assessment, Frailty domains assessed, Psychometric properties	7
de Vries (2011)[20]	Published before February 2010	No restriction	No restriction	Development/ psychometric evaluation of a screening tool, Explicit and operational definition of tool	20	Frailty domains assessed, Scoring of instrument, Psychometric properties	5

*Two of the 11 AMSTAR items were not-applicable for this set of reviews as the results of the included primary studies were not pooled, making the 9/11 the highest achievable score.

These systematic reviews identified 57 original tools, many of which were included in more than one review. Specific items included in the tools varied widely, and in all tools taken together a total of 21 different domains were used to measure frailty. No domain was measured by all tools and no tool measured all domains. This literature overview supports the contention that there is no consensus on the definition of frailty or its measurement. There are likely a number of interrelated systems contributing to frailty, as represented in the breadth of domains included. Table 3-2 summarizes the frequency of various frailty domains assessed in the 57 tools and the relevant modules in the CLSA that measure these domains. The most common domains were physical functioning, which was measured in 40 of the 57 tools, and mobility and/or falls, measured in 37 tools. Each of the domains was assessed in a variety of ways in different tools. For example, items that measured grip strength, reaching, lifting objects, or chair stands were considered measure the physical function domain, while items measuring gait speed, difficulty walking, difficulty moving around the house, or falls were included in the mobility domain. Items in the depressive symptoms domain included feeling depressed, sadness, satisfaction with life and energy level (e.g. feel tired or have trouble get going). The cognition domain included measures of memory problems or executive functioning, as well as diagnosed dementia or cognitive impairment. The mood domain included reported or diagnosed mood disorders or anxiety. Nutrition was sometimes included in a common domain with weight loss or BMI, while other tools reported them as distinct. Social support included items on the extent of social support available (e.g. availability of others to provide help or conversation) social participation, and the desire for more social interaction.

Items for this analysis were selected from the CLSA dataset to correspond with domains identified through the literature search, in consultation with experts in geriatric medicine, physiotherapy, population health and frailty measurement. These items from the CLSA are summarized in Table 3-2 and discussed in greater detail in the analytical methods section in the next chapter.

Not all domains to evaluate frailty are included in the CLSA Tracking cohort dataset. Physical activity, health-related quality of life, nutrition, and medication data were not available in CLSA Tracking. Health service use is more often considered an outcome of frailty, and is strongly associated with disability. CLSA Tracking evaluates health service usage by the amount and type of care received, both formal (paid) and informal (unpaid) care, as well as the use of assistive devices for mobility. Sociodemographic variables of age, gender, and education were not included as potential components of frailty. These items were instead used to describe the population and to assess the construct validity of the frailty measurements developed in this study.

Table 3-2: Frailty domains included in original tools

Domain	Tools including each domain		Corresponding items in CLSA
	N	(%)	
Physical functioning	40	69%	Functional status (8 of 14 items)
Mobility	37	64%	Functional status (6 of 14 items) Falls causing injury (frequency)
Cognitive functioning	27	47%	Cognition (4 items)
Activities of daily living (ADL/IADL)	22	38%	ADL (7 items) OARS scale (7 items)
Depressive symptoms	21	36%	CES-D 10 scale (10 items) SWLS scale (5 items)
Weight loss/BMI	18	31%	BMI (1 item)
Sensory impairment (vision, hearing)	17	29%	Self-rated vision (1 item) Self-rated hearing (1 item)
Self-rated health	16	28%	Self-rated health (1 item) Self-rated mental health (1 item)
Comorbidities or health conditions	16	28%	Health conditions (27 items)
Nutrition	16	28%	None
Social support	14	24%	Social support availability (20 items) Social participation (frequency, type)
Physical activity	13	22%	None
Mood or anxiety disorders	12	21%	Self-reported mood disorder (1 item) Self-reported anxiety (1 item)
Health service use	11	19%	Informal care received (frequency, type) Formal care received (frequency, type) Assistive device use (type)
Medication/polypharmacy	10	17%	None
Age	9	16%	Age (years)
Continence	8	14%	Urinary incontinence (1 item) Bowel incontinence (1 item)
Subjective/general assessment	3	5%	None
Education	2	3%	Highest level of education
Sex	1	2%	Sex
Quality of life	1	2%	None

CHAPTER 4: STATISTICAL METHODS

This chapter presents a summary of the dataset used, as well as the methodology and analytical methods used to investigate the construct of frailty and its measurement in the Canadian Longitudinal Study on Aging (CLSA).

4.1 Data Source

The Canadian Longitudinal Study on Aging

The CLSA tracking cohort is the source of data for these analyses. Many studies on aging have been short-term and have focused on specific aspects of aging or disease. The CLSA helps address the need for large-scale population-based studies to examine the cumulative effect of many different factors on the health of Canada's aging population. CLSA examines the process of aging from mid-life to old age, enabling study of the development of frailty during its early stages. The wide extent of physical and psychological health indicators, social life circumstances, and disease outcomes collected for a large population makes the CLSA a resource well-suited for the study of the interrelationships among factors that affect frailty and healthy aging.[27]

4.2 Population and Recruitment

Population

The CLSA consists of a national stratified random sample of 50,000 Canadians aged 45-85 years at the time of recruitment[27]. This wide age range enables the study of factors in mid-life that may be associated with frailty later in life. This sample was selected to be representative of the Canadian population for provincial estimates of health determinants, health status, and health system utilization[27].

The CLSA population includes adults between the ages of 45 and 85 years at recruitment, who are able to complete the interview in English or French. Community-dwelling adults in households or transitional housing such as senor's residences with minimal care are included. Those in long-term care were excluded at baseline, while participants who become institutionalized during the study will continue to be followed. Also excluded were Canadian residents of the three territories, full-time members of the Canadian Armed Forces, and persons living on federal First Nations reserves or other First Nations settlements for logistical reasons and to ensure a sample representative of the population of Canadian adults as a whole. The presence of cognitive impairment at baseline is a criterion for exclusion, as this can compromise the capacity to give informed consent, and may affect the reliability and validity of interview responses.[27] These eligibility criteria were adapted from the Canadian Community Health Survey (CCHS).[58]

From the 50,000 CLSA participants, approximately 20,000 make up the CLSA Tracking cohort, a representative sample of Canadians designed for provincial-level estimates of health determinants and health system utilization.[27] The remaining 30,000 participants will be asked to provide more in-depth information through physical assessments and collection of biological specimens (blood and urine).[27] This analysis will use the responses from the CLSA tracking cohort, the only CLSA dataset currently available.

Recruitment

The first sampling frame for the CLSA was the Canadian Community Health Survey (CCHS). Contact information was obtained to approach participants for recruitment for the tracking cohort.[58] CLSA collaborated with Statistics Canada on the sampling strategy to assemble a representative sample of Canadians in the target age range.[58] Statistics Canada approached

participants in the CCHS and when consent was provided, they forwarded contact information to CLSA researchers.[27] , after which all further contact and follow-up became the responsibility of the CLSA research team. The consenting CCHS participants were sent an introductory letter containing information about the study and a study consent form.

The telephone interview consists of a 60-minute interview using Computer Assisted Telephone Interview (CATI) software, which assists trained interviewers to ask questions over the telephone. The interviewers enter responses directly into the system, improving transcription quality and security.[27] Interviewers at the CATI sites contacted potential participants by telephone within two weeks of the information package being sent. The CLSA interviewer explained the study in detail, answered questions, guided participants through the consent process, and if verbal consent was provided.,[27] they undertook the baseline interview. To participate in the Tracking cohort, a person had to complete the 60-minute interview and provide written consent to the CLSA National Coordinating Site. To reach the intended sample of 20,000 participants, additional Tracking participants were recruited through two additional sampling frames: Provincial Healthcare Registration Databases and random digit dialing.[58] A similar recruitment process was conducted for all sampling frames.

Ethical Considerations

All participants recruited through CCHS provided written consent to allow their contact information to be released and to be contacted by CLSA staff for recruitment in the study.[58] All CLSA participants have provided written informed consent. Extensive preparation has insured that province-specific regulations for the use of provincial healthcare registration databases has been met[59].

4.3 Measurement of Health Indicators

To minimize bias in the selection of variables for the analysis, a literature review of studies on the identification and measurement of frailty was conducted to identify all available variables associated with frailty. The final list of variables was selected through consultation with a team of experts, including geriatricians, physiotherapists, epidemiologists, population health researchers, and experts in the study of frailty.

The CLSA Tracking cohort includes measures selected by expert working groups. Measures were identified through literature reviews and chosen for their appropriateness, relevance to the population studied, availability and French and English, psychometric properties, and feasibility for telephone interviews[27]. The variables used for this analysis were selected from the data available in the CLSA specifically for their relevance to frailty.

4.3.1 Indicators of Frailty

The following health indicators available in the CLSA are hypothesized to associate with frailty. These will be used to develop the Frailty Index and be included in the EFA and SEM analyses.

Health Status Measures

The assessment of general health includes one five-point item on each of the following: self-rated general health, mental health, and healthy aging. Single-item scales have been shown to be reliable in the assessment of self-rated health.[60] To measure sensory impairment, participants were asked to rate their eyesight and hearing while using aids (corrective lenses, hearing aids, etc.), each on a five-point scale. Participants' best estimate of their own height in centimetres and weight in kilograms were also reported. From these measures, body mass index (BMI) was calculated as mass divided by square of body height and classified based on the international

standards set by the WHO.[61] Participants were categorized using their BMI as no risk (normal weight), mild risk (overweight), moderate risk (obese class 1), or high risk (underweight or obese class 2 and above). In aging populations, there is increased risk of mortality for those with low BMI, so this category was classified as high risk along with those classified as severely obese.[62]

Chronic Conditions

The chronic conditions assessed in the tracking cohort were selected for inclusion in the CLSA based on: relevance to adult populations, ability to be studied in a sample of this size, being understudied in existing population-based studies, and feasibility to ascertain without physical examination.[27] Participants were asked about the presence or absence of chronic health conditions items with the phrasing, "has a doctor ever told you that you have…". Chronic conditions assessed include arthritis (rheumatoid and osteoarthritis of the knee, hip, and hand), hypertension, cataracts, back problems, diabetes, cancer, mood disorders, migraine headaches, hyper- and hypothyroidism, asthma, heart disease, osteoporosis, bowel disorder, intestinal or stomach ulcers, anxiety disorder, peripheral vascular disease or poor circulation, emphysema, chronic bronchitis, chronic obstructive pulmonary disease, heart attack or myocardial infarction, angina, glaucoma, macular degeneration, mini-stroke or transient ischemic attack, kidney disease or failure, memory problems, stroke or cerebrovascular accident, epilepsy, Parkinsonism or Parkinson's disease, dementia, and Alzheimer's disease. Multiple sclerosis was excluded, as this disease was not thought to become more prevalent with age and would therefore not be indicative of frailty. Self-reported conditions diagnosed by a clinician have shown good test-retest reliability in population-based health surveys.[63]

Functional Status

Physical functioning is assessed using a set of fourteen 5-point items taken from the Framingham Disability Study, Established Populations for Epidemiologic Studies of the Elderly study, as well as the Nagi and Rosow-Breslau Scales.[64–67] Age-adjusted kappa statistic ranged from 0.41-0.95 for the included self-report items items.[68,69] Validity and reliability assessments of these instruments have shown good validity, test-retest reliability, responsiveness to change, and correlation with performance measures[68–72] and with lean and fat mass.[73] These items were grouped into three domains of physical functioning, based on the results of an exploratory factor analysis showing the constructs measured by the 14 functional status items, as well as our interpretation of the domains these items best measured. The lower body strength domain included the following items: difficulty walking 2-3 blocks, walking up and down stairs, standing for a long period, stooping, crouching, or kneeling, standing up after sitting, and sitting for a long period. The upper body strength domain included items on: difficulty making bed, pushing or pulling large objects, lifting ten pounds, taking a force or impact in the arms or hands, and washing one's back. The dexterity domain included items on difficulty using a knife, handling small objects, and extending arms above shoulders.

Activities of Daily Living

Basic Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) are measured using a modified version of the Older Americans Resources and Services (OARS) Multidimensional Assessment Questionnaire.[74] The total score is the number of the 13 daily activities that participants indicate they can do without help. The first six items pertain to basic or personal care such as eating and dressing, while the remaining seven evaluate higher level activities including

travelling, housework, shopping, and preparing meals. The use of these scales to evaluate self-care for the study healthy aging and frailty is widespread. The scales demonstrate high correlations with physical therapist measures of self-care capacity. Good validity was shown with intraclass correlation coefficients ranging from 0.66 on physical health to 0.87 on self-care capacity.[75]

Cognition

Two domains of cognition are assessed in the CLSA Tracking cohort, memory and executive function. Memory is assessed using the Rey Auditory Verbal Learning Test (RAVLT),[76] a 15-item word learning test that assesses learning and retention. The RAVLT is measured twice, once after the list is given, and once after the subsequent cognitive tests. The RAVLT has been shown to be reliable[77] and sensitive to the detection of early cognitive decline,[78] with an intraclass correlation coefficient of 0.71 for immediate and 0.76 for delayed recall tests showing good test-retest reliability.[79] Executive function is measured using the mental alternation test (MAT)[80] and animal naming test of verbal fluency. The MAT is a two-part test that first requires participants to count aloud from 1-20 and say the alphabet as quickly as they can. If they cannot do this, the test cannot be performed. If the participant proves capable of these tasks, they complete the second part where they are asked to alternate between numbers and letters (1-A, 2-B, 3-C, …) as quickly as possible for 30 seconds. Scores are based on the number of correct responses. The MAT is sensitive and specific for the detection of cognitive impairment as measured by the Mini-Mental State Examination in older adults[79] (Pearson correlation 0.84, sensitivity 0.91, specificity 1.00).[82] Verbal fluency is measured by asking participants to name as many animals as possible in 60 seconds. Animal naming has shown to be sensitive to cognitive decline and can discern normal

aging from early-stage dementia.[83] Category fluency tasks, including animal naming, have shown to be good at identifying participants with Alzheimer's disease[84] and strong predictors of incident Alzheimer's disease in longitudinal studies (sensitivity 1.00, specificity 0.93).[85,86] Scores from each of these tests were standardized to have a mean of 50 and standard deviation of 10, with no gender differences and unrelated to age or education, based on the validation of these cognitive function tests in the Canadian Community Health Survey.[87]

Depression

Depressive symptoms are measured using the Center for Epidemiologic Studies Short Depression Scale (CES-D 10).[88] This scale consists of ten 4-point items on feelings of depression, loneliness, hopefulness for the future, and restless sleep. The maximum score is 30, and a cut-off score of 10 or above indicates depression. This scale is among the best-known instruments for measuring depressive symptoms and has been used extensively in large population-based studies. The short form version shows close agreement with the full version[88] and comparable reliability, with a Cronbach's alpha of 0.80 compared to the 0.86 in the full version.[89]

Satisfaction with Life

The Satisfaction with Life Scale (SWLS) is included to measure positive mood state or life satisfaction, an important measure for self-assessment of health and well-being.[90] The SWLS consists of five 7-point items and evaluates several domains including social relationships, work or other role performance, and personal satisfaction with self, religious life, learning, and leisure. This scale is widely used and has shown good psychometric properties[91,92] with strong correlations with other measures of life satisfaction and Cronbach's alpha reliability of 0.83[92].

Social Support

Social Support is measured using the 19-item MOS Social Support Survey,[93] which is divided into the subscales that measure affection, emotional and informational support, social interaction, and tangible social support. The four separate social support subscales are scored by the difference between the observed and minimum possible scores, divided by the difference between the maximum and minimum possible scores. The scale was developed to measure social support for chronically ill community-dwelling adults, but is considered applicable to all adults.[94] The reliability of the scale and its validity for population-based studies have been studied and shown to be acceptable, with all Cronbach's alpha values greater than 0.91.[94]

Social Participation

CLSA Tracking includes eight 5-point items on the frequency of participation in a variety of social activities, including family or friend-based, faith community, sport or recreational, social club-based, educational, volunteer, and cultural. These items were developed as part of the CCHS-Healthy Aging.[95] Participants' responses were recorded on two axes:(?) the frequency of any type of community activity-related participation (never, yearly, monthly, weekly, daily) and their desire to participate more in social activities.

4.3.2 Construct Validation Variables

The following health indicators will be used to assess the construct validity of the frailty scores in this analysis. The direction of the frailty score correlation with each variable will be predicted *a priori* based on our theory of how a measure of frailty should perform.

Sociodemographic measures

Sociodemographic variables were self-reported for all participants, including age, province, sex, ethnicity, marital status, household income, and highest level of education achieved. Single marital status, low household income, low education, and older age are hypothesized to associate with increasing frailty. These variables will be compared to other available population datasets to ensure the population selected is representative of Canadian adults.

Care Receiving

The care receiving module collects information on both informal and formal care. Informal care describes performance of tasks by family, friends, or neighbors due to a health condition or limitation of the participant. Formal care describes the performance of these tasks by healthcare professionals, including nurses and personal support workers. This section also includes the use of assistive devices such as canes, wheelchairs, or personal alarms. These items were adapted from the General Social Survey (GSS)[96]. The number of weeks of care, as well as the average hours per week of care from formal and informal caregivers are recorded. We hypothesize that increasing frailty will associate positively with higher levels of care received.

Injury and Falls

Two items are included to screen for injuries in the past 12 months. If an injury is reported, the type, cause, location, and body part injured are assessed, as well as whether the injury resulted in fracture. If a fall is listed as the cause, a second battery of questions focusing on fall-related injuries and their consequences is also administered. This injury and falls module has been used numerous times as part of the CCHS and has undergone extensive validation. The number of injuries sustained from falls in the past 12 months will be included. Increasing injuries and falls are hypothesized to associate with higher levels of frailty.

Data Limitations

A limitation of secondary data analysis is the limited number of variables present in the dataset that has already been collected. Several key health indicators are absent from this dataset, including medications, nutrition, functional performance, muscle mass, and waist and hip circumference. However, these variables and others will be available in the CLSA Comprehensive dataset. This study of the CLSA Tracking cohort will provide an improved understanding of the health and frailty of aging Canadians and will act as important groundwork for future analyses of the Comprehensive data.

4.4 Statistical Methods

A Frailty Index will be constructed following the cumulative deficits model and evaluated for construct validity. Exploratory Factor Analysis and Structural Equation Modeling will be used to explore the relationships between these health indicators and the underlying construct of frailty.

4.4.1 Data Preparation and Descriptive Summary

The CLSA Tracking cohort population dataset exceeds the recommended sample size for the analyses performed, so that it can be split using simple random sampling, into two datasets: one for model development and one for validation. The subsamples will be tested using independent t-tests with Bonferroni correction (to minimize type 1 error) for continuous variables and chi-squared tests for categorical variables to ensure the two subsamples can be compared as two random samples of the same population.

Descriptive statistics for the variables will be calculated for the full sample and both subsamples. Complete case analysis with deletion of participants with missing data will be used for the exploratory and confirmatory factor analyses.

4.4.2 Frailty Index

A Frailty Index will be constructed from the variables present in the dataset, based on the recommendations of Searle et al (2008).[98] In this model, frailty is thought to be the result of the accumulation of health deficits as a person ages. With this model, a list of 30+ health deficits are chosen with the items weighted equally. All of the health deficits listed above will be transformed from 0 (no deficit) to 1 (most severe deficit) and included in the index. For example, a 4-point scale where 0 is the most severe option, such as "how often do you feel depressed" (all of the time, occasionally, some of the time, rarely or never) would transform to (1,0.66,0.33,0). The frailty of an individual is the proportion of deficits present divided by the total number considered. Frailty indices were not calculated for participants with more than 5% missing data. No specific deficit is necessary for a person to be considered frail, so this model is flexible for use in population datasets such as the CLSA. The Frailty Index is designed as a continuous variable and will be evaluated as such. A cut-off value of 0.25 has been shown to be suitable for the purposes of estimating the proportion of frail and non-frail participants.[42]

4.4.3 Exploratory Factor Analysis

Exploratory factor analysis (EFA) is a statistical method used to estimate factors, or latent variables, from a correlation matrix of measured variables that account for the maximum variance in the set of observed data. This method was chosen because it is commonly used to measure latent constructs behind inter-related variables and is relatively simple to conduct and interpret, facilitating the reproducibility of this research.

The assumption of EFA is that any variable may be associated with any factor, there is no *a priori* prediction for how the variables are related to the underlying factors[97].

The aim of this analysis is to evaluate frailty as an underlying construct that can be measured only by examining the relationship and covariance among observable health indicators. As a result, factor analysis is chosen as a method of analysis over principle component analysis, which does not account for the underlying structure of latent variables.[99] A matrix of polychoric correlations between ordinal variables and Tetrachoric correlations between dichotomous variables was used for the analysis, as these methods are more suited to exploratory and confirmatory factor analysis of categorical data than Pearson correlations.[100] The principal axes method of factor estimation will be used to extract factors. The Cattel scree test of graphed eigenvalues will be used to determine the number of factors retained by examining the graph to find the factors that explain the large amounts of the total variance.[101] Orthogonal varimax rotation will be used as a method for factor rotation to aid in factor interpretability, since we anticipate some correlation among frailty factors and this method is more appropriate when the factors are correlated.[99] When interpreting variable factor loadings, 0.30 will be considered the minimum cut-off to be considered part of a factor.[99] Research on appropriate sample size for factor analysis is conflicting, although a sample size greater than 150 is considered appropriate for an analysis of 10 variables or more, with very large samples making the analysis more generalizable or replicable.[99,102] Subject to item ratio of 10:1 is commonly used for determining a priori sample size required for factor analysis.[99] A population of 20,000 participants in CLSA Tracking far exceeds the necessary sample for this analysis.

4.4.4 Structural Equation Modelling

Structural equation modelling (SEM) is a method to evaluate a hypothesized measurement model explaining the relationships between the measured variables and underlying constructs of frailty.

This model is developed using both our theory of frailty, published literature on frailty, and the ICF[26], as discussed in chapter 2. The specified model is shown in Figure 5-2A.

SEM with maximum likelihood estimation will be used to test the fit of the model with the observed covariance structure of the observed health indicator variables.[97] Each latent variable will be considered exogenous and allowed to covary with every other latent variable.[97] To simplify the analysis, each indicator is assumed to measure only one construct. We hypothesize that each indicator will have high standardized loadings on that factor compared to others, suggesting convergent validity, and correlations between factors are not excessively high, showing discriminant validity.[97] Goodness-of-fit will be evaluated using the Comparative Fit Index (CFI), Bentler Non-Normed Fit Index (NNFI), Root Mean Squared Error of Approximation (RMSEA), and Akaike Information Criterion (AIC). CFI ranges from a 0-1, with larger values indicating better model fit. A CFI value greater than 0.90 is considered acceptable model fit.[103] NNFI values greater than 0.95 indicate good fit, whereas values less than 0.90 suggests that changes to the model are necessary.[103] RMSEA pertains to the residual in the model and has values ranging from 0-1 with smaller values indicating better model fit. A RMSEA value of 0.06 or less is considered acceptable model fit.[103] The AIC statistic is used for comparing candidate statistical models, and provides a comparison independent of model complexity.[104] Lower AIC values indicate better model fit. Modifications to the model will be made in a step-wise, exploratory process by removing poor-performing indicators until an acceptable model is reached. The model with the best face validity, or most consistent with the theory of frailty, and best fit of the data will be selected from the candidate models generated in the analysis. In SEM, it is necessary for there to be more than one variable in each factor, due to the modelling of error, so factors with only one remaining variable

were dropped or combined with other factors during model development.[97] The final model will be used to create a frailty score by equally weighting variables within each factor, then weighting each factor equally, based on recommendation of Streiner, Norman, and Cairney.[105] This score will be compared with the Frailty Index developed in the analysis describe below. The results from each will be reported and the best-performing measurement will be examined further for potential inclusion in the CLSA.

4.4.5 Model Validation

The frailty score created in this study must be able to measure frailty in a way that is clinically relevant and which demonstrates good validity and reliability. The sample size is sufficiently large that a randomly-split half of the data can be used to develop the frailty measurement models. The validation sample will be used to test the models by repeating the EFA and SEM analyses in this dataset and comparing the model fit. The score generated from the SEM analysis model will be compared to the Frailty Index by examining distribution (including skewness, kurtosis, floor and ceiling effects), correlation with one another, agreement on classification of participants as frail and non-frail, and correlation with relevant health indicators. Participants for whom the two scales disagree will be explored.

Since there is no way to measure frailty directly, the construct validity of the measures will be tested against hypothesized associations with sociodemographic and physical health variables. These hypotheses are based on the researchers' current theory of frailty, informed by the literature, as well as their research and clinical experience. We expect that frailty scores will be higher in women than in men, and will increase with age. Frail individuals will be less autonomous, requiring more formal and informal care and more likely to use assistive devices than the non-

frail. Frail persons are predicted to have more frequent falls and more severe injuries from falling. Highest level of education achieved and household income are both predicted to be lower in those identified as frail. If these hypotheses are met, the study will provide evidence for both the theory of frailty and the method of frailty measurement used.[105]

CHAPTER 5: RESULTS

This chapter presents a descriptive summary of the sample of CLSA participants in section 5.1. A summary of the Frailty Index in this population is shown in 5.2. The latent constructs emerging from the data in the exploratory factor analysis are presented in 5.3. The fit of the hypothesized measurement model for frailty in this dataset is evaluated through structural equation modelling in 5.4. The modifications to this model are described and the indices of model fit for the two subsamples are also presented in this section. The creation of a frailty score from the best performing measurement model is detailed in 5.5 and the construct validity of this score is compared to the Frailty Index in 5.6.

5.1 Sample Characteristics

Summary descriptive statistics of both the demographic factors and variables included in the frailty measurement models are displayed for the total sample and the subsamples included in the calibration and validation samples in Tables 1A and 1B. The calibration and validation subsamples had no significant difference in the distribution of sex, age, formal care received, assistive device use, or falls (Table 5-1A) and for age (treated as a continuous variable) and number of serious injuries (Table 5-1B). Significant differences were found between the calibration and validation samples in the income, education, and formal home care variables. Upon further examination, the absolute differences between the subsamples were very small and the very large sample size raises the likelihood of type 1 error for these tests. For the purpose of this analysis, the subsamples were considered to be similar for all factors.

5.2 Frailty Index

The Frailty Index was calculated using the 90 available indicators of frailty chosen from review of frailty measures and consultation with experts. Each variable was transformed to a decimal from 0 (no deficit) to 1 (highest deficit). A descriptive summary of these variables in the total population and in the calibration and validation subsamples are shown in Tables 5-2A and 5-2B. Table 5-2A includes the Frailty Index values assigned for each level of the categorical variables. For continuous variables in Table 5-2B, the Frailty Index value is calculated as the value divided by the maximum value for that category. In the total sample, 20874 (98%) participants had missing data for four or fewer variables and had a Frailty Index calculated and were thus included in the Frailty Index analysis. The Frailty Index had a mean of 0.14, standard deviation of 0.07. The distribution had a skewness of 1.55 and kurtosis of 3.8. The minimum Frailty Index for the full sample was 0.003, while the maximum was 0.68. Based on a cut-off of 0.25 for frailty, 1440 (6.92%) participants would be classified as frail. There were no floor or ceiling effects for the Frailty Index, as no participant had the lowest or highest possible score. The Frailty Index summary statistics are shown in Table 5-3.

5.3 Exploratory Factor Analysis

The polychoric bivariate correlations between the 90 observed variables in the calibration sample included in the exploratory factor analysis are shown in Table 5-S (supplementary file). The correlation between epilepsy and macular degeneration could not be calculated, as no person in the sample indicated having both. The correlations were generally low, with the highest correlations found between items in the same scales, including the SWLS (0.50-0.74), CES-D 10 scale (0.14-0.74), OARS scale for ADL and IADL (0.11-0.91), and MOS social support subscales

(0.69-0.83). Items within the same scale tended to load on the same factor, which will be discussed further. The minimum eigenvalue criterion was met by 16 factors. The Cattell scree plot (Figure 5-1) showed that 3 factors should be retained, with the first two explaining the majority of the variance in the sample. These three factors accounted for 33%, 17%, and 7% of the variance of the observations, for a total of 56% variance explained.

The factor loadings of the 3 retained factors after varimax rotation are shown in Table 5-4. The minimum cut-off was 0.30 for a variable to be considered a defining part of a factor. Factor 1 included very high (0.623-0.974) loadings from items in the functional status, OARS scale, and self-rated health, as well as moderate loadings for some chronic conditions, including urinary (0.611) and bowel incontinence (0.598), peripheral vascular disease (0.600), stroke (0.595), epilepsy (0.574), osteoporosis (0.560), memory problems (0.537), arthritis (0.487), back pain (0.389), diabetes (0.369), cataracts (0.349), Alzheimer's (0.336), and emphysema/bronchitis/COPD (0.319). Two items from the CES-D 10 depression scale loaded on factor 1: frequency "feel everything is an effort" (0.423) and frequency "have trouble get going" (0.326). Finally, the self-rated vision item also had a small loading on factor 1 (0.302).

Factor 2 included high loadings from the satisfaction with life scale (0.859-0.905). Some depression items also had high loadings on factor 2, including frequency feel happy (0.706), frequency feel hopeful for the future (0.723), frequency feel lonely (0.547), and frequency feel depressed (0.512). Self-rated mental health loaded on factor 2 (0.587), as did social support availability (0.380-0.480) and social participation items (0.350, 0.407). Minor loadings were found for anxiety (0.308) and some depression items, frequency "feel fearful" or " tearful" (0.315), frequency "feel everything is an effort" (0.300), and frequency "have trouble get going" (0.346).

Factor 3 had high loadings from depression items that did not load strongly on either of the previous factors. Frequency “fearful or tearful” (0.830), frequency “feel bothered” (0.819), frequency “have trouble concentrating” (0.803), “frequency feel depressed” (0.775), frequency “everything is an effort” (0.713), frequency “have trouble get going” (0.710) had high loadings. Moderate loadings were found for frequency “feel restless” (0.583), frequency “feel lonely” (0.555), frequency “feel happy” (0.501), and frequency feel “hopeful for the future” (0.347). Other mental health items also had moderate loadings, including mood disorder (0.510), anxiety (0.471), and self-rated mental health (0.469).

5.4 Structural Equation Modelling

In total, 2193 of 21241 participants (10%) in CLSA Tracking were excluded from the analysis due to missing data in at least one variable. The proportion of missing data for any one variable was low <5% and each was treated as missing at random. The highest proportion of missing data was found for the cognitive function tests, where 814 (3.8%) participants had no data available. When explored further, the data were most often reported as missing due to technical problems rather than participant refusal. When compared to those included, participants excluded due to missing data were found to have a higher mean age (67.0±11.1 vs. 62.5±10.5, $p<0.0001$) and greater mean number of chronic conditions reported (4.1±2.7 vs. 3.4±3.4, $p<0.0001$). This presents a limitation to this analysis, as the excluded participants may represent a more vulnerable population than the participants included.

After removing missing data, 9561 of 10621 (90%) participants in the calibration sample were included in the SEM analyses. Figure 5-2A and Table 5-5 show the parameter estimates for the hypothesized base measurement model using the calibration sample, including factor loadings,

squared multiple correlations, and factor covariances. All parameters were statistically significant contributors to the model ($p<0.001$), no parameter could be dropped based on the Wald tests. High covariances were found between health conditions and impairments to body structure and function (0.788), participation restrictions and impairments to body structure and function (1.026), and environmental factors and participation restrictions (0.796).

Modifications to the model were made in an exploratory fashion, one at a time. The modified model structure is shown in Figure 5-2B and the parameter estimates for the modified model are shown in Table 5-6. BMI classification and continence were moved from health conditions to the impairments to body structure and function domain, while self-rated mental health was moved to the health conditions domain. When two observed variables covary to a greater extent than would be expected by the model specification, an error covariance term was added to the model. These variables may contain similar items, measured in a similar fashion, or be closely related. Error covariance terms were added between self-rated health and self-rated mental health, BMI classification and lower body strength, continence and ADL, upper body strength and dexterity, and ADL and IADL. The modifications improved the model fit of the data. The strongest associations were found between the upper-body strength ($\beta=0.816$) and lower-body strength ($\beta=0.827$) scores and the activity limitations domain, and between the social support availability MOS Subscale scores and the environmental factors domain ($\beta=0.760$-0.905). The depression ($\beta=0.729$) and satisfaction with life ($\beta=0.712$) scores also had strong loadings on the impairments to body structure and function factors.

While all parameters in the model were statistically significant, the large sample size raises the chance of type 1 error for the Wald tests. The worst-performing parameters from the modified

model were removed in an exploratory fashion to achieve the best possible model fit. The best fit was found by a simplified model including only one factor. The structure of the model is shown in Figure 5-3 and the parameter estimates are detailed in Table 7. The strongest association with the frailty factor was found for the lower body strength (β=0.843) and upper body strength (β=0.787) scores. The fit statistics for the hypothesized base model, the modified model and the simplified model are shown in Table 8. Both the modified and simplified models had very good fit of the data. The simplified model had the best fit of the data, with very low RMSEA (0.015) and AIC (21.00), and CFI and NNFI values close to 1.

5.4.1 Cross-Validation

Each of the three measurement models used in the SEM analyses (base, modified, and simplified) were assessed using for goodness-of-fit in the validation sample. Complete cases were available for 9487 of 10620 (89%) participants in the validation sample to evaluate model fit. All of the goodness-of-fit indices assessed had similar values in the validation sample compared to the calibration sample. Both the modified model and simplified model had very good fit of the data. This suggests that the results obtained in the SEM analyses are unlikely to be due to chance. The fit statistics for the validation sample are reported in Table 8.

5.5 The Simplified Frailty Score

The close fit of the simplified model to the data prompted further investigation. A Simplified Frailty Score (SF) was constructed from the observations included in the model: CES-D 10 items measuring depression, and physical functioning items on upper body strength, lower body strength, and dexterity. Items within each of the four domains were weighted equally to create a subscale score. The subscale scores were then averaged to produce a SF score from 0 (robust) to

1 (maximum frailty possible). Complete data for the above items was available for 21105 (99%) participants, and a SF score was calculated for each. The mean SF score was 0.09 with a standard deviation of 0.10. The distribution was heavily skewed (2.53) and had high kurtosis (8.60). The minimum SF value was 0, and the maximum 0.96 for this population. There was a small floor effect seen for the SF score, as 6.7% of the population had the lowest possible score. Using a cut-off of 0.25 to classify participants as frail, the prevalence of frailty using the SF was 1303 (6.3%). The SF summary statistics are shown compared to the Frailty Index in Table 5-3.

5.6 Construct Validation

Both the Frailty Index and Simplified Frailty Scores were compared to population demographic variables thought to be related to frailty. These relationships were measured using Pearson correlation coefficients and are shown in Table 9. All correlations were statistically significant at the $p<0.0001$ level. All correlations were in the direction hypothesized for a measure of frailty. Both the FI and SF scores had positive correlations with age, fall status, injuries, and home care, and were negatively associated with income, education, and male gender. The measures were highly correlated with one another, with a Pearson correlation of 0.84 ($p<0.001$). The FI was more strongly associated with age, income, education, and fall status, while the SF was more strongly associated with formal home care. The distributions of scores by age category are shown in Figure 5-4. The FI has fewer outliers, and is more responsive to age than the SF. Using a cut-off value of 0.25, both score classified approximately 7% of the participants as frail. The two measures had moderate to high agreement beyond chance on frailty classification, with a Kappa of 0.67 (95%CI: 0.55-0.69), a statistically significant result at the $p<0.0001$ level. Of the 20800 participants for whom both scores could be calculated, the two agreed on 19949 (96%).

The 851 (4%) cases where the two scores disagreed in their classification of frailty status were examined, and the groups are compared in Table 5-10. Those classified as frail by the FI but not by the SF score ($FI\geq0.25$, $SF<0.25$) had more chronic conditions (8.4 vs. 4.5) and were more likely to have sustained a serious injury (22% vs. 18%) or injury from falling (12% vs. 8%) than those identified as frail by the SF but not the FI. The $SF\geq0.25$, $FI<0.25$ group had a lower mean age (64.8 years vs. 66.3) and were more likely to need formal home care (41% vs. 33%) than the $FI\geq0.25$, $SF<0.25$ group.

5.7 Summary

This project explored the underlying factor structure of 90 indicators of frailty available in the CLSA using exploratory factor analysis. The hypothesized measurement model organizing these factors based on the ICF framework was tested using structural equation modelling with maximum likelihood estimation. A model containing all available indicators showed good fit of the data and was theoretically meaningful. Many factors in a variety of health domains contributed to a measurement of frailty, a result consistent with the accumulated deficit model for measuring frailty. A simplified model also had good fit of the data, measuring only physical functioning and depressive symptoms. These results persisted when evaluated using an independent sample of the same population. A score was constructed from the simplified model and shown to be comparable to the Frailty Index. Both the Frailty Index and Simplified Frailty Scores showed good construct validity.

Table 5-1A: Descriptive Summary of Population Demographic Variables, Categorical

Variable	Total Sample		Calibration Sample		Validation Sample	
	N	%	N	%	N	%
Age						
45-54	5832	24.46	2870	27.02	2962	27.89
55-64	6563	30.90	3336	31.41	3227	30.39
65-74	4634	21.82	2326	21.90	2308	21.73
≥75	4212	19.83	2089	19.67	2123	19.99
Sex - Male	10405	48.99	5181	48.78	5224	49.19
Total Household Income*						
Less than $20 000	1347	6.78	686	6.92	661	6.64
$20 000-$50 000	5851	29.44	2875	28.99	2976	29.90
$50 000-$100 000	7218	36.32	3553	35.82	3665	36.82
$100 000-$150 000	3215	16.18	1699	17.13	1516	15.23
$150 000 or more	2240	11.27	1105	11.14	1135	11.40
Education*						
Less than secondary school graduation	1986	9.35	997	9.39	989	9.31
Secondary school graduation no post-secondary education	2882	13.57	1453	13.68	1429	13.46
Some post-secondary education	1623	7.64	790	7.44	833	7.84
Post-secondary degree/diploma	14667	69.05	7348	69.18	7319	68.92
Formal Home Care Services	1201	5.66	582	5.49	619	5.83
Informal Home Care Services*	2782	13.11	1444	13.62	1338	12.61
Assistive Devices	2729	12.86	1353	12.76	1376	12.96
Falls						
No falls reported	20184	95.03	10095	95.05	10089	95.02
No serious injury due to a fall or injury not receiving medical attention	367	1.73	185	1.74	182	1.71
Injury receiving medical attention without hospitalization	541	2.55	263	2.48	278	2.62
Injury receiving medical attention and hospitalization	147	0.69	78	0.73	69	0.65

*Indicates significant difference on chi-squared test

Table 5-1B: Descriptive Summary of Population Demographic Variables, Continuous

Variable	Entire Sample		Calibration Sample		Validation Sample	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Age (years)	21241	63.02 (10.67)	10621	63.07(10.67)	10620	62.96 (10.69)
Serious injuries reported	21217	0.17 (0.64)	10609	0.17 (0.57)	10608	0.18 (0.70)

No significant differences were found between continuous variables

Table 5-2A: Descriptive Summary of Analysis Variables, Categorical

Variable	Entire Sample		Calibration Sample		Validation Sample	
(Frailty Index Value)	N	%	N	%	N	%
Self-rated Health						
Excellent (0.00)	3977	18.74	2072	19.53	1905	17.96
Very good (0.25)	8123	38.28	4023	37.91	4100	38.65
Good (0.50)	6263	29.51	3105	29.26	3158	29.77
Fair (0.75)	2231	10.51	1115	10.51	1116	10.52
Poor (1.00)	626	2.95	297	2.80	329	3.10
Self-rated Mental Health						
Excellent (0.00)	6358	29.96	3216	30.31	3142	29.61
Very good (0.25)	8225	38.76	4120	38.83	4105	38.68
Good (0.50)	5514	25.98	2694	25.39	2820	26.57
Fair (0.75)	981	4.62	500	4.71	481	4.53
Poor (1.00)	145	0.68	80	0.75	65	0.61
Eyesight Rating						
Excellent (0.00)	4532	21.36	2324	21.90	2208	20.81
Very good (0.25)	7836	36.92	3887	36.64	3949	37.21
Good (0.50)	6952	32.76	3467	32.68	3485	32.84
Fair (0.75)	1532	7.22	739	6.97	793	7.47
Poor (1.00)	370	1.74	193	1.82	177	1.67
Hearing Rating						
Excellent (0.00)	4976	23.45	2517	23.72	2459	23.17
Very good (0.25)	6656	31.36	3268	30.80	3388	31.93
Good (0.50)	6929	32.65	3480	32.80	3449	32.50
Fair (0.75)	2238	10.55	1135	10.70	1103	10.39
Poor (1.00)	423	1.99	211	1.99	212	2.00
SWLS scale: Satisfied with life						
Strongly agree (1.00)	10464	49.37	5254	49.58	5210	49.16
Agree (0.83)	6191	29.21	3076	29.03	3115	29.40
Slightly agree (0.67)	1622	7.65	783	7.39	839	7.92
Neither agree nor disagree (0.50)	851	4.02	424	4.00	427	4.03
Slightly disagree (0.33)	909	4.29	457	4.31	452	4.27
Disagree (0.17)	654	3.09	334	3.15	320	3.02
Strongly disagree (0.00)	503	2.37	269	2.54	234	2.21
SWLS scale: Life close to ideal						
Strongly agree (1.00)	8357	39.53	4202	39.78	4155	39.29
Agree (0.83)	6138	29.04	3016	28.55	3122	29.52
Slightly agree (0.67)	1698	8.03	820	7.76	878	8.30
Neither agree nor disagree (0.50)	1764	8.34	922	8.73	842	7.96
Slightly disagree (0.33)	1425	6.74	697	6.60	728	6.88
Disagree (0.17)	1016	4.81	526	4.98	490	4.63
Strongly disagree (0.00)	741	3.51	381	3.61	360	3.40
SWLS scale: Have important things						

Strongly agree (1.00)	10527	49.78	5214	49.30	5313	50.26
Agree (0.83)	6375	30.15	3203	30.28	3172	30.01
Slightly agree (0.67)	1926	9.11	977	9.24	949	8.98
Neither agree nor disagree (0.50)	719	3.40	381	3.60	338	3.20
Slightly disagree (0.33)	737	3.49	367	3.47	370	3.50
Disagree (0.17)	522	2.47	256	2.42	266	2.52
Strongly disagree (0.00)	341	1.61	179	1.69	162	1.53
SWLS scale: Would change almost nothing						
Strongly agree (1.00)	6902	32.70	3505	33.22	3397	32.17
Agree (0.83)	5310	25.15	2644	25.06	2666	25.25
Slightly agree (0.67)	2188	10.36	1060	10.05	1128	10.68
Neither agree nor disagree (0.50)	955	4.52	487	4.62	468	4.43
Slightly disagree (0.33)	2243	10.63	1135	10.76	1108	10.49
Disagree (0.17)	1871	8.86	892	8.45	979	9.27
Strongly disagree (0.00)	1641	7.77	827	7.84	814	7.71
SWLS scale: Life conditions excellent						
Strongly agree (1.00)	8635	40.78	4355	41.14	4280	40.42
Agree (0.83)	5561	26.26	2808	26.53	2753	26.00
Slightly agree (0.67)	1898	8.96	946	8.94	952	8.99
Neither agree nor disagree (0.50)	1569	7.41	781	7.38	788	7.44
Slightly disagree (0.33)	1677	7.92	808	7.63	869	8.21
Disagree (0.17)	1118	5.28	520	4.91	598	5.65
Strongly disagree (0.00)	717	3.39	368	3.48	349	3.30
CES-D 10 scale: Frequency feel fearful or tearful						
Rarely or none of the time (0.00)	16189	76.38	8057	76.02	8132	76.74
Some or a little of the time (0.33)	3091	14.58	1544	14.57	1547	14.60
Occasionally (0.66)	1516	7.15	786	7.42	730	6.89
All of the time (1.00)	399	1.88	211	1.99	188	1.77
CES-D 10 scale: Frequency feel happy						
All of the time (0.00)	14097	66.64	7067	66.83	7030	66.45
Occasionally (0.33)	4756	22.48	2314	21.88	2442	23.08
Some or a little of the time (0.66)	1817	8.59	960	9.08	857	8.10
Rarely or none of the time (1.00)	484	2.29	234	2.21	250	2.36
CES-D 10 scale: Frequency feel could not get going						
Rarely or none of the time (0.00)	14526	68.52	7247	68.36	7279	68.68
Some or a little of the time (0.33)	3733	17.61	1875	17.69	1858	17.53
Occasionally (0.66)	2225	10.50	1110	10.47	1115	10.52
All of the time (1.00)	716	3.38	369	3.48	347	3.27
CES-D 10 scale: Frequency trouble concentrating						
Rarely or none of the time (0.00)	12858	60.74	6455	60.96	6403	60.52
Some or a little of the time (0.33)	4077	19.26	2028	19.15	2049	19.37

Occasionally (0.66)	3207	15.15	1585	14.97	1622	15.33
All of the time (1.00)	1027	4.85	521	4.92	506	4.78
CES-D 10 scale: Frequency feel depressed						
Rarely or none of the time (0.00)	16733	78.89	8365	78.89	8368	78.88
Some or a little of the time (0.33)	2587	12.20	1270	11.98	1317	12.42
Occasionally (0.66)	1437	6.77	738	6.96	699	6.59
All of the time (1.00)	454	2.14	230	2.17	224	2.11
CES-D 10 scale: Frequency feel everything is an effort						
Rarely or none of the time (0.00)	14175	66.90	7092	66.93	7083	66.86
Some or a little of the time (0.33)	3842	18.13	1931	18.22	1911	18.04
Occasionally (0.66)	2161	10.20	1072	10.12	1089	10.28
All of the time (1.00)	1011	4.77	501	4.73	510	4.81
CES-D 10 scale: Frequency sleep is restless						
Rarely or none of the time (0.00)	8258	38.97	4137	39.05	4121	38.90
Some or a little of the time (0.33)	5706	26.93	2851	26.91	2855	26.95
Occasionally (0.66)	4038	19.06	2042	19.28	1996	18.84
All of the time (1.00)	3187	15.04	1564	14.76	1623	15.32
CES-D 10 scale: Frequency easily bothered						
Rarely or none of the time (0.00)	13577	64.10	6744	63.70	6833	64.50
Some or a little of the time (0.33)	4124	19.47	2081	19.66	2043	19.28
Occasionally (0.66)	2791	13.18	1396	13.19	1395	13.17
All of the time (1.00)	689	3.25	366	3.46	323	3.05
CES-D 10 scale: Frequency feel hopeful about the future						
All of the time (0.00)	13062	62.51	6554	62.69	6508	62.33
Occasionally (0.33)	4112	19.68	2045	19.56	2067	19.80
Some or a little of the time (0.66)	2340	11.20	1158	11.08	1182	11.32
Rarely or none of the time (1.00)	1383	6.62	698	6.68	685	6.56
CES-D 10 scale: Frequency feel lonely						
Rarely or none of the time (0.00)	16389	77.28	8150	76.88	8239	77.68
Some or a little of the time (0.33)	2450	11.55	1215	11.46	1235	11.64
Occasionally (0.66)	1686	7.95	894	8.43	792	7.47
All of the time (1.00)	683	3.22	342	3.23	341	3.21
Difficulty walking up and down stairs						
No difficulty (0.00)	18735	88.26	9397	88.52	9338	88.00
A little difficult (0.25)	962	4.53	454	4.28	508	4.79
Somewhat difficult (0.50)	820	3.86	413	3.89	407	3.84
Very difficult (0.75)	525	2.47	262	2.47	263	2.48
Unable to do (1.00)	185	0.87	90	0.85	95	0.90
Difficulty walking 2-3 blocks						
No difficulty (0.00)	18894	89.11	9481	89.40	9413	88.83

A little difficult (0.25)	540	2.55	249	2.35	291	2.75
Somewhat difficult (0.50)	695	3.28	329	3.10	366	3.45
Very difficult (0.75)	657	3.10	330	3.11	327	3.09
Unable to do (1.00)	416	1.96	216	2.04	200	1.89
Difficulty taking force or impact in arms or hands						
No difficulty (0.00)	18325	86.76	9125	86.37	9200	87.15
A little difficult (0.25)	943	4.46	505	4.78	438	4.15
Somewhat difficult (0.50)	926	4.38	456	4.32	470	4.45
Very difficult (0.75)	541	2.56	279	2.64	262	2.48
Unable to do (1.00)	387	1.83	200	1.89	187	1.77
Difficulty stooping, crouching, or kneeling						
No difficulty (0.00)	14574	68.67	7239	68.21	7335	69.13
A little difficult (0.25)	2745	12.93	1395	13.14	1350	12.72
Somewhat difficult (0.50)	1972	9.29	989	9.32	983	9.26
Very difficult (0.75)	1352	6.37	694	6.54	658	6.20
Unable to do (1.00)	581	2.74	296	2.79	285	2.69
Difficulty pulling/ pushing large objects						
No difficulty (0.00)	18929	89.29	9461	89.27	9468	89.30
A little difficult (0.25)	658	3.10	324	3.06	334	3.15
Somewhat difficult (0.50)	715	3.37	355	3.35	360	3.40
Very difficult (0.75)	477	2.25	241	2.27	236	2.23
Unable to do (1.00)	421	1.99	217	2.05	204	1.92
Difficulty handling small objects						
No difficulty (0.00)	19742	92.96	9865	92.92	9877	93.01
A little difficult (0.25)	787	3.71	379	3.57	408	3.84
Somewhat difficult (0.50)	489	2.30	246	2.32	243	2.29
Very difficult (0.75)	195	0.92	112	1.05	83	0.78
Unable to do (1.00)	23	0.11	15	0.14	8	0.08
Difficulty making bed						
No difficulty (0.00)	20144	94.95	10062	94.85	10082	95.04
A little difficult (0.25)	358	1.69	190	1.79	168	1.58
Somewhat difficult (0.50)	398	1.88	193	1.82	205	1.93
Very difficult (0.75)	212	1.00	103	0.97	109	1.03
Unable to do (1.00)	104	0.49	60	0.57	44	0.41
Difficulty standing up after sitting						
No difficulty (0.00)	16627	78.32	8337	78.56	8290	78.08
A little difficult (0.25)	2852	13.43	1413	13.32	1439	13.55
Somewhat difficult (0.50)	1291	6.08	644	6.07	647	6.09
Very difficult (0.75)	409	1.93	194	1.83	215	2.03
Unable to do (1.00)	50	0.24	24	0.23	26	0.24
Difficulty washing back						
No difficulty (0.00)	19296	90.93	9675	91.17	9621	90.70

A little difficult (0.25)	768	3.62	357	3.36	411	3.87
Somewhat difficult (0.50)	553	2.61	263	2.48	290	2.73
Very difficult (0.75)	396	1.87	213	2.01	183	1.73
Unable to do (1.00)	207	0.98	104	0.98	103	0.97
Difficulty lifting 10 pounds						
No difficulty (0.00)	19594	92.31	9774	92.11	9820	92.51
A little difficult (0.25)	505	2.38	254	2.39	251	2.36
Somewhat difficult (0.50)	513	2.42	278	2.62	235	2.21
Very difficult (0.75)	350	1.65	176	1.66	174	1.64
Unable to do (1.00)	264	1.24	129	1.22	135	1.27
Difficulty standing for a long period						
No difficulty (0.00)	17998	84.84	9043	85.28	8955	84.39
A little difficult (0.25)	1063	5.01	517	4.88	546	5.15
Somewhat difficult (0.50)	1131	5.33	550	5.19	581	5.48
Very difficult (0.75)	822	3.87	409	3.86	413	3.89
Unable to do (1.00)	201	0.95	85	0.80	116	1.09
Difficulty sitting for a long period						
No difficulty (0.00)	18955	89.31	9473	89.28	9482	89.33
A little difficult (0.25)	1041	4.90	549	5.17	492	4.64
Somewhat difficult (0.50)	827	3.90	407	3.84	420	3.96
Very difficult (0.75)	348	1.64	154	1.45	194	1.83
Unable to do (1.00)	54	0.25	28	0.26	26	0.24
Difficulty extending arms above shoulders						
No difficulty (0.00)	18758	88.36	9378	88.35	9380	88.37
A little difficult (0.25)	1075	5.06	550	5.18	525	4.95
Somewhat difficult (0.50)	800	3.77	393	3.70	407	3.83
Very difficult (0.75)	507	2.39	249	2.35	258	2.43
Unable to do (1.00)	89	0.42	45	0.42	44	0.41
Difficulty using a knife						
No difficulty (0.00)	20725	97.60	10352	97.49	10373	97.71
A little difficult (0.25)	220	1.04	115	1.08	105	0.99
Somewhat difficult (0.50)	198	0.93	98	0.92	100	0.94
Very difficult (0.75)	71	0.33	39	0.37	32	0.30
Unable to do (1.00)	21	0.10	15	0.14	6	0.06
OARS scale: Able to travel (0)	20958	98.76	10478	98.78	10480	98.73
OARS scale: Able to do housework (0)	19899	93.96	9955	93.95	9944	93.96
OARS scale: Able to use telephone (0)	21148	99.58	10571	99.54	10577	99.62
OARS scale: Able to handle money (0)	21135	99.56	10569	99.58	10566	99.54
OARS scale: Able to go shopping (0)	20758	97.86	10385	97.89	10373	97.82
OARS scale: Able to prepare meals (0)	21020	99.08	10518	99.12	10502	99.04
OARS scale: Able to take medicine (0)	21142	99.63	10568	99.59	10574	99.66
OARS scale: Able to dress (0)	21006	98.98	10500	98.94	10506	99.01
OARS scale: Able to get out of bed (0)	21080	99.27	10542	99.27	10538	99.27
OARS scale: Able to feed (0)	21210	99.89	10542	99.27	10603	99.89

OARS scale: Able to take care of appearance (0)	21180	99.74	10596	99.77	10584	99.70
OARS scale: Able to walk (0)	20852	98.24	10434	98.31	10418	98.17
OARS scale: Able to take bath (0)	20944	98.64	10459	98.52	10485	98.77
OARS scale: Trouble getting to bathroom in time						
No trouble (0.00)	18733	88.37	9343	88.18	9390	88.55
Less than once a week (0.33)	1607	7.58	823	7.77	784	7.39
Once or twice a week (0.67)	455	2.15	220	2.08	235	2.22
Three times a week or more (1.00)	404	1.91	209	1.97	195	1.84
Frequency of Community-Related Activity Participation						
Once a day (0.00)	3276	15.58	1645	15.64	1631	15.52
Once a week (0.25)	14048	66.80	7008	66.63	7040	66.98
Once a month (0.50)	2946	14.01	1482	14.09	1464	13.93
Once a year (0.75)	617	2.93	308	2.93	309	2.94
Never (1.00)	142	0.68	75	0.71	67	0.64
Desire to participate in more activities (1)	8259	38.93	4188	39.51	4071	38.36
BMI Classification						
Normal (0.00)	7546	35.70	3807	36.03	3739	35.37
Overweight (0.33)	8233	38.95	4105	38.85	4128	39.05
Obese class 1 (0.67)	3517	16.64	1753	16.59	1764	16.69
Obese class 2+ or underweight (1.00)	1839	8.70	900	8.52	939	8.88
Chronic Conditions (1)						
Arthritis	8212	38.91	4105	38.89	4107	38.93
Asthma	2347	11.06	1168	11.01	1179	11.12
Stroke or CVA	390	1.84	200	1.89	190	1.79
Memory problem	449	2.11	229	2.16	220	2.07
Bowel disorder	1838	8.67	915	8.63	923	8.70
Glaucoma	959	4.52	465	4.39	494	4.66
Allergies	7853	37.07	3912	36.94	3941	37.21
Osteoporosis	2009	9.48	980	9.25	1029	9.71
High blood pressure	8104	38.20	4020	37.89	4084	38.51
Diabetes	3553	16.74	1769	16.67	1784	16.81
Heart attack	1317	6.21	644	6.07	673	6.35
Mini-stroke or TIA	749	3.54	365	3.45	384	3.63
Parkinson's Disease	78	0.37	39	0.37	39	0.37
Cataracts	5281	24.89	2631	24.81	2650	24.98
Back problems	5204	24.53	2582	24.34	2622	24.72
Heart disease	2191	10.34	1090	10.28	1101	10.39
Migraine headaches	2916	13.75	1461	13.78	1455	13.72
Intestinal or stomach ulcers	1637	7.72	828	7.80	809	7.63
Over-active thyroid gland	466	2.21	254	2.41	212	2.01

Emphysema, bronchitis, COPD	1436	6.77	730	6.88	706	6.65
Angina	1149	5.41	566	5.33	583	5.50
Peripheral vascular disease	1519	7.17	772	7.29	747	7.05
Dementia or Alzheimer's	43	0.20	21	0.20	22	0.21
Epilepsy	166	0.78	86	0.81	80	0.75
Urinary incontinence	1873	8.83	934	8.80	939	8.85
Macular degeneration	875	4.13	435	4.10	440	4.15
Under-active thyroid gland	2446	11.61	1235	11.73	1211	11.49
Kidney disease	593	2.80	300	2.83	293	2.76
Bowel incontinence	492	2.32	254	2.39	238	2.24
Cancer	3265	15.38	1623	15.29	1642	15.48
Mood disorder	3106	14.64	1587	14.95	1519	14.32
Anxiety disorder	1560	7.35	767	7.23	793	7.48

Table 5-2B: Descriptive Summary of Analysis Variables, Continuous

Variable	Entire Sample		Calibration Sample		Validation Sample	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Mental Alternation Test (MAT)*	18836	10.00 (3.00)	9405	9.99 (3.01)	9431	10.01 (2.99)
Rey Auditory Verbal Learning Test (First Recall)*	19584	10.00 (3.00)	9783	9.99 (3.01)	9801	10.01 (2.99)
Rey Auditory Verbal Learning Test (Delayed Recall)*	19474	10.00 (3.00)	9728	10.00 (3.02)	9746	10.00 (2.98)
Animal Fluency Test (AFT)*	20444	10.00 (3.00)	10216	9.98 (3.02)	10228	10.02 (2.98)
MOS Subscale: Affection	21241	11.09 (8.42)	10620	11.13 (8.61)	10621	11.06 (8.22)
MOS Subscale: Emotional and Informational Support	21241	28.43 (14.27)	10620	28.43 (14.32)	10621	28.43 (14.23)
MOS Subscale: Positive Social Interaction	21241	14.17 (10.28)	10620	14.14 (10.09)	10621	14.20 (10.46)
MOS Subscale: Tangible Social Support	21241	15.11 (13.52)	10620	15.19 (13.64)	10621	15.02 (13.39)

*Cognitive functioning variables were standardized to mean of 10 and standard deviation of 3

Table 5-3: Simple Statistics for Frailty Measures

Measure	Frailty Index	Simplified Frailty Scale
N	20874	21105
Mean (SD)	0.14 (0.07)	0.09 (0.10)
Skewness	1.55	2.53
Kurtosis	3.81	8.60
Range	0.67	0.96
Minimum	0.003	0.000
Maximum	0.677	0.956
% Minimum score	0	6.7
% Maximum score	0	0
N Missing (%)	367 (1.73)	136 (0.64)
Frailty prevalence at ≥0.25 cut-off (%)	1440 (6.9%)	1303 (6.3%)

Figure 5-1: Exploratory Factor Analysis, Cattel Scree Plot

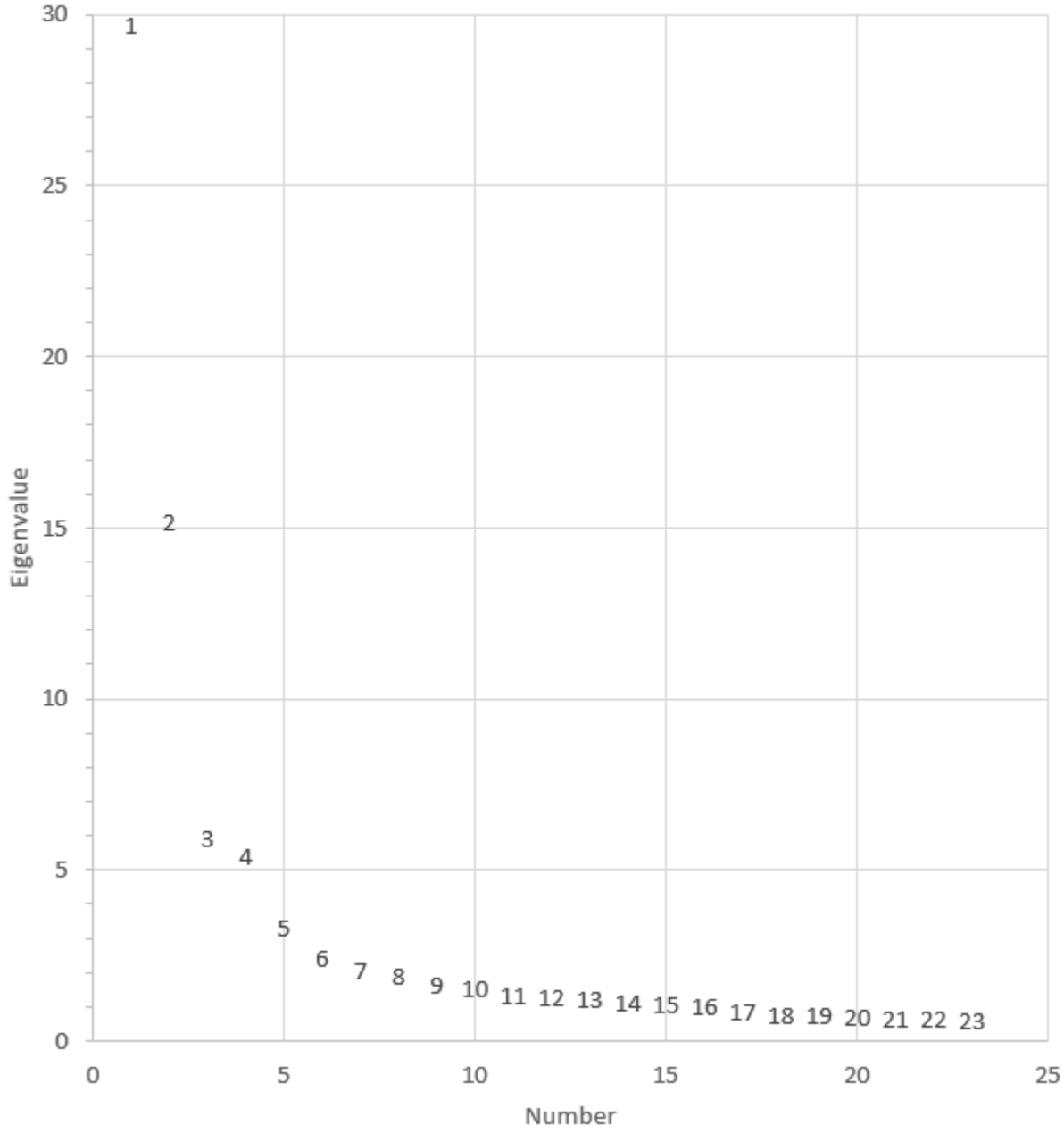

Table 5-4: Exploratory Factor Analysis, Rotated Factor Pattern

	Factor 1	Factor 2	Factor 3
Difficulty making bed	0.974	0.021	0.079
OARS scale: Able to get out of bed	0.964	0.078	0.046
OARS scale: Able to take bath	0.964	-0.006	-0.002
OARS scale: Able to do housework	0.960	0.007	0.052
OARS scale: Able to go shopping	0.958	0.044	0.049
OARS scale: Able to walk	0.957	0.029	0.030
OARS scale: Able to prepare meals	0.956	-0.005	0.022
Difficulty washing back	0.955	0.032	0.068
OARS scale: Able to dress	0.943	-0.034	-0.002
OARS scale: Able to travel	0.940	0.021	0.008
Difficulty pulling/ pushing large objects	0.933	0.035	0.109
Difficulty using a knife	0.933	0.042	0.049
Difficulty standing for a long period	0.930	0.022	0.076
Difficulty walking 2-3 blocks	0.928	0.015	0.052
Difficulty lifting 10 pounds	0.925	0.054	0.145
Difficulty walking up and down stairs	0.922	0.030	0.075
Difficulty standing up after sitting	0.914	-0.011	0.094
Difficulty taking force or impact in arms or hands	0.911	0.031	0.082
OARS scale: Able to take medicine	0.903	-0.052	-0.045
OARS scale: Able to take care of appearance	0.892	0.088	0.060
Difficulty extending arms above shoulders	0.888	-0.009	0.058
Difficulty handling small objects	0.882	-0.052	0.018
OARS scale: Able to handle money	0.861	0.049	0.096
Difficulty stooping, crouching, or kneeling	0.851	-0.004	0.059
OARS scale: Trouble getting to bathroom in time	0.776	-0.045	0.074
OARS scale: Able to use telephone	0.767	-0.105	-0.067
Parkinson's Disease	0.745	-0.073	-0.042
Self-rated Health	0.718	0.252	0.109
Difficulty sitting for a long period	0.697	0.097	0.268
OARS scale: Able to feed	0.623	0.143	0.116
Peripheral vascular disease	0.600	-0.108	-0.069
Stroke or CVA	0.595	-0.126	-0.112
Epilepsy	0.574	-0.001	0.026
Osteoporosis	0.560	-0.159	-0.024

Memory problem	0.537	0.173	0.221
Arthritis	0.487	-0.147	0.030
SWLS scale: Satisfied with life	0.078	0.905	0.292
SWLS scale: Life close to ideal	0.078	0.898	0.245
SWLS scale: Life conditions excellent	0.235	0.885	0.256
SWLS scale: Have important things	0.039	0.870	0.170
SWLS scale: Would change almost nothing	-0.086	0.859	0.191
CES-D 10 scale: Frequency feel hopeful about the future	-0.060	0.723	0.347
CES-D 10 scale: Frequency feel happy	-0.128	0.706	0.501
Self-rated Mental Health	0.091	0.587	0.469
Frequency of Community-Related Activity Participation	0.280	0.350	-0.067
CES-D 10 scale: Frequency feel fearful or tearful	0.021	0.315	0.830
CES-D 10 scale: Frequency easily bothered	0.095	0.284	0.819
CES-D 10 scale: Frequency trouble concentrating	0.085	0.213	0.803
CES-D 10 scale: Frequency feel depressed	0.032	0.512	0.775
CES-D 10 scale: Frequency feel everything is an effort	0.428	0.300	0.720
CES-D 10 scale: Frequency feel could not get going	0.326	0.346	0.710
CES-D 10 scale: Frequency sleep is restless	0.252	0.131	0.583
CES-D 10 scale: Frequency feel lonely	0.024	0.547	0.555
Mood disorder	0.033	0.402	0.510
Anxiety disorder	0.089	0.308	0.471
Macular degeneration	0.170	-0.161	-0.075
Mini-stroke or TIA	0.117	-0.101	0.001
Glaucoma	-0.084	-0.148	-0.062
Angina	0.185	-0.098	-0.090
Kidney disease	0.156	-0.119	-0.171
Heart attack	0.184	-0.072	-0.127
Dementia or Alzheimer's	0.336	0.064	-0.086
Cataracts	0.349	-0.338	-0.211
MOS Subscale: Emotional and Informational Support	-0.087	0.419	0.101
MOS Subscale: Tangible Social Support	-0.180	0.380	0.135
MOS Subscale: Positive Social Interaction	-0.044	0.469	0.146
MOS Subscale: Affection	-0.085	0.480	0.118
Heart disease	0.264	-0.166	-0.170
High blood pressure	0.241	-0.221	-0.158

Diabetes	0.369	-0.136	-0.125
Desire to participate in more activities	0.045	0.407	0.223
Asthma	0.089	-0.112	-0.028
Allergies	0.018	-0.107	0.026
Emphysema, bronchitis, COPD	0.319	0.021	-0.008
Migraine headaches	0.079	-0.054	0.239
Bowel disorder	0.175	0.078	0.156
Back problems	0.389	0.042	0.109
Intestinal or stomach ulcers	0.283	-0.030	-0.001
Rey Auditory Verbal Learning Test (Delayed Recall)	0.063	-0.041	-0.054
Rey Auditory Verbal Learning Test (First Recall)	0.131	-0.079	-0.088
Under-active thyroid gland	0.069	-0.082	0.029
BMI Classification	0.292	-0.006	-0.018
Urinary incontinence	0.611	-0.094	-0.009
Bowel incontinence	0.598	-0.001	0.030
Mental Alternation Test (MAT)	0.096	-0.029	-0.040
Animal Fluency Test (AFT)	0.197	-0.084	-0.058
Hearing Rating	0.020	0.013	-0.076
Eyesight Rating	0.302	0.107	0.048
Cancer	0.158	-0.226	-0.207
Over-active thyroid gland	-0.044	-0.066	-0.033

Figure 5-2A Hypothesized Base Measurement Model

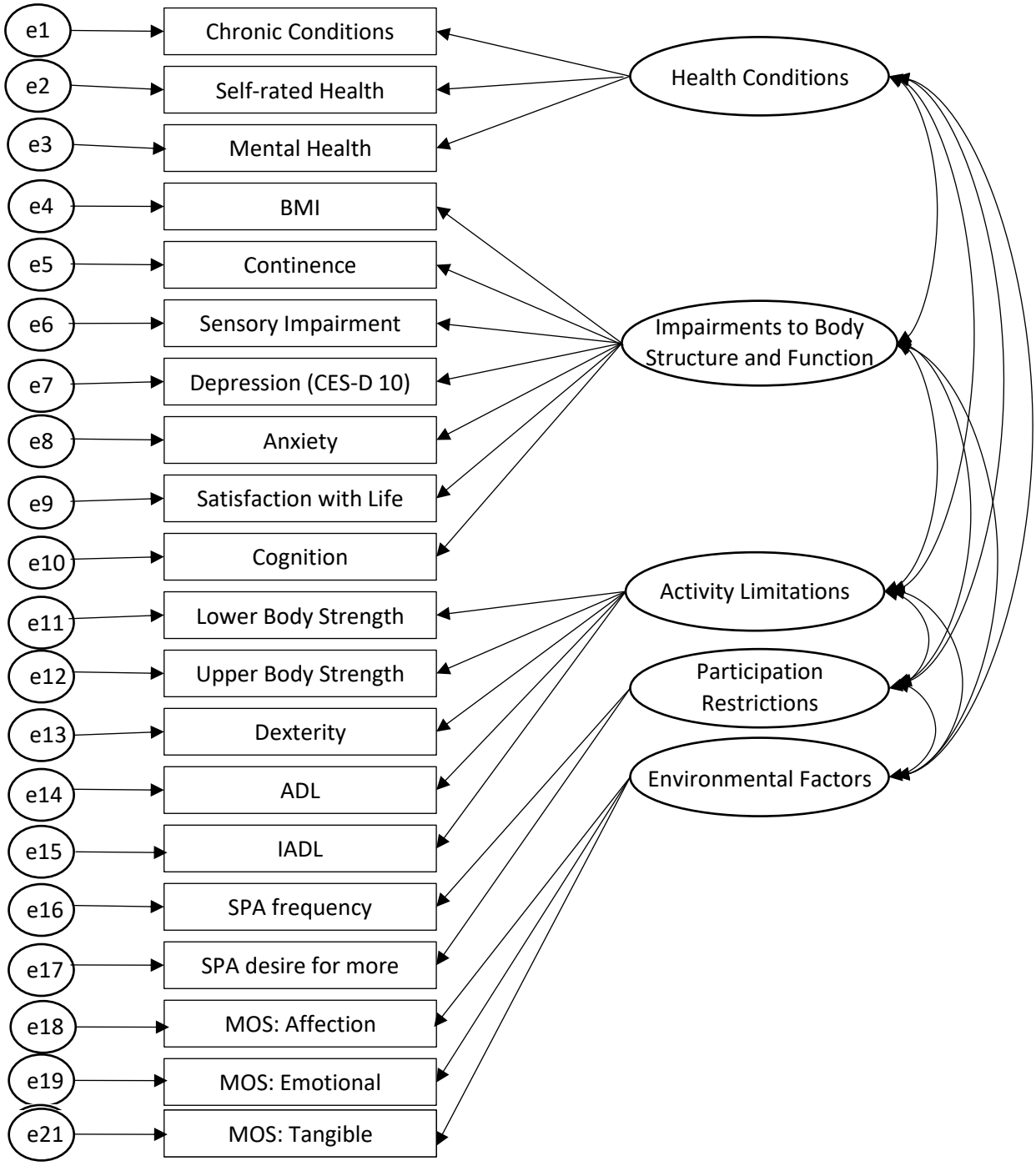

Figure 5-2B Modified Frailty Model

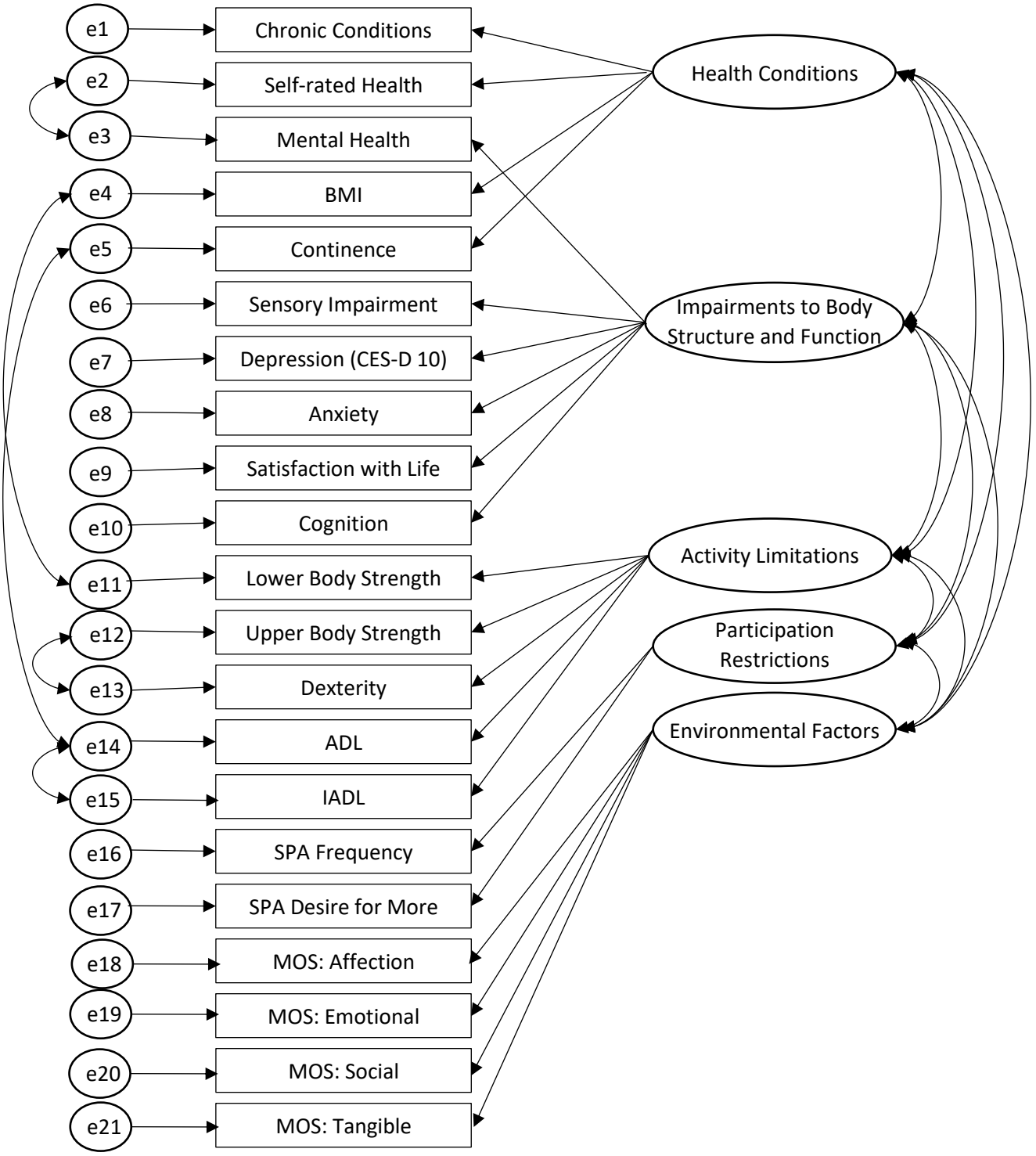

Table 5-5: Maximum Likelihood Parameter Estimates for Hypothesized Base Model

Parameter			Factor Loadings		Squared Correlations
			Estimate	Standard error	
Chronic conditions	←	Health conditions	0.525	0.011	0.276
Self-rated mental health	←	Health conditions	0.616	0.011	0.380
Self-rated health	←	Health conditions	0.751	0.010	0.563
BMI classification	←	Impairments to body structure and function	0.184	0.011	0.034
Continence	←	Impairments to body structure and function	0.227	0.011	0.052
Sensory impairment	←	Impairments to body structure and function	0.300	0.011	0.09
Depression	←	Impairments to body structure and function	0.697	0.010	0.485
Mood disorders	←	Impairments to body structure and function	0.389	0.011	0.152
Satisfaction with life scale	←	Impairments to body structure and function	0.700	0.010	0.490
Cognition	←	Impairments to body structure and function	0.160	0.011	0.026
Lower body strength	←	Activity limitations	0.782	0.009	0.611
Upper body strength	←	Activity limitations	0.837	0.009	0.701
Dexterity	←	Activity limitations	0.645	0.010	0.416
OARS scale: IADL	←	Activity limitations	0.684	0.010	0.468
OARS scale: ADL	←	Activity limitations	0.637	0.010	0.405
Social participation - frequency	←	Participation restrictions	0.258	0.023	0.066
Social participation - desire for more	←	Participation restrictions	0.232	0.021	0.054
MOS subscale: affection	←	Environmental factors	0.803	0.009	0.645
MOS subscale: emotional and informational support	←	Environmental factors	0.853	0.008	0.727
MOS subscale: positive social interaction	←	Environmental factors	0.904	0.008	0.818
MOS subscale: tangible social support	←	Environmental factors	0.761	0.009	0.579

			Factor Covariances	
			Estimate	Standard error
Impairments to body structure and function	↔	Health conditions	0.788	0.010
Activity limitations	↔	Health conditions	0.599	0.010
Activity limitations	↔	Impairments to body structure and function	0.493	0.011
Participation restrictions	↔	Health conditions	0.634	0.059
Participation restrictions	↔	Impairments to body structure and function	1.026	0.084
Participation restrictions	↔	Activity limitations	0.458	0.047
Environmental factors	↔	Health conditions	0.314	0.012
Environmental factors	↔	Impairments to body structure and function	0.574	0.010
Environmental factors	↔	Activity limitations	0.188	0.011
Environmental factors	↔	Participation restrictions	0.796	0.067

Table 5-6: Maximum Likelihood Parameter Estimates for Modified Model

Parameter			Factor Loadings		Squared Correlations
			Estimate	Standard error	
Chronic conditions	←	Health conditions	0.627	0.011	0.394
Self-rated mental health	←	Impairments to body structure and function	0.576	0.010	0.483
Self-rated health	←	Health conditions	0.695	0.011	0.338
BMI classification	←	Health conditions	0.252	0.012	0.064
Continence	←	Health conditions	0.303	0.012	0.091
Sensory impairment	←	Impairments to body structure and function	0.286	0.011	0.812
Depression	←	Impairments to body structure and function	0.729	0.010	0.531
Mood disorders	←	Impairments to body structure and function	0.411	0.011	0.169
Satisfaction with life scale	←	Impairments to body structure and function	0.712	0.010	0.507
Cognition	←	Impairments to body structure and function	0.153	0.011	0.024
Lower body strength	←	Activity limitations	0.827	0.009	0.681
Upper body strength	←	Activity limitations	0.816	0.009	0.666
Dexterity	←	Activity limitations	0.592	0.010	0.351
OARS scale: IADL	←	Activity limitations	0.644	0.010	0.415
OARS scale: ADL	←	Activity limitations	0.588	0.010	0.348
Social participation - frequency	←	Participation restrictions	0.259	0.023	0.067
Social participation - desire for more	←	Participation restrictions	0.231	0.021	0.054
MOS subscale: affection	←	Environmental factors	0.802	0.009	0.644
MOS subscale: emotional and informational support	←	Environmental factors	0.852	0.008	0.727
MOS subscale: positive social interaction	←	Environmental factors	0.905	0.008	0.819
MOS subscale: tangible social support	←	Environmental factors	0.760	0.009	0.578

			Factor Covariances	
			Estimate	Standard error
Impairments to body structure and function	↔	Health conditions	0.621	0.011
Activity limitations	↔	Health conditions	0.734	0.009
Activity limitations	↔	Impairments to body structure and function	0.468	0.011
Participation restrictions	↔	Health conditions	0.547	0.055
Participation restrictions	↔	Impairments to body structure and function	0.961	0.079
Participation restrictions	↔	Activity limitations	0.465	0.048
Environmental factors	↔	Health conditions	0.265	0.013
Environmental factors	↔	Impairments to body structure and function	0.548	0.009
Environmental factors	↔	Activity limitations	0.197	0.011
Environmental factors	↔	Participation restrictions	0.796	0.067

			Error Covariances	
			Estimate	Standard error
OARS scale: IADL	↔	OARS scale: ADL	0.194	0.008
Continence	↔	OARS scale: ADL	0.159	0.008
Self-rated health	↔	Self-rated mental health	0.218	0.008
Lower body strength	↔	BMI classification	0.111	0.007
Upper body strength	↔	Dexterity	0.105	0.007

Figure 5-3: Simplified Frailty Model

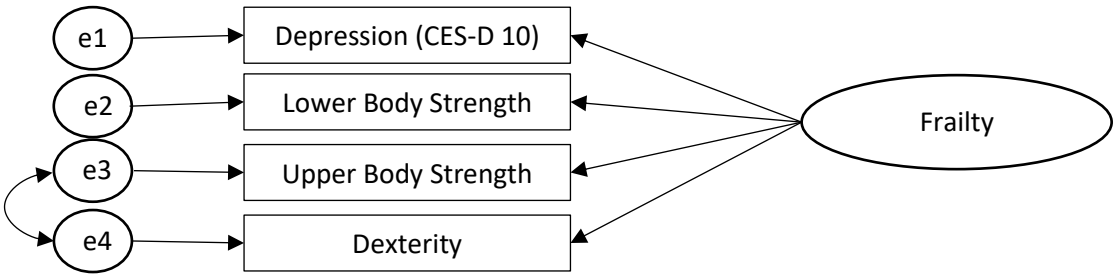

Table 5-7: Maximum Likelihood Parameter Estimates for Simplified Model

Parameter			Factor Loadings		Squared Correlations
			Estimate	Standard error	
Depression	←	Frailty	0.378	0.011	0.143
Lower body strength	←	Frailty	0.843	0.014	0.711
Upper body strength	←	Frailty	0.787	0.013	0.619
Dexterity	←	Frailty	0.550	0.013	0.302
			Error Covariances		
			Estimate	Standard error	
Upper body strength	↔	Dexterity	0.155	0.013	

Table 5-8: Comparison of Frailty Measurement Models, Goodness-of-fit Indices

	Model	Chi-square	Degrees of freedom	CFI	NNFI	RMSEA	AIC
Calibration Sample	Base	6865.25	179	0.900	0.882	0.063	6969.25
	Modified	2959.41	174	0.958	0.950	0.041	3073.40
	Simplified	2.99	1	0.999	0.999	0.015	20.00
Validation Sample	Base	6504.19	179	0.902	0.885	0.061	6608.19
	Modified	2764.88	174	0.960	0.952	0.040	2878.88
	Simplified	10.53	1	0.999	0.999	0.032	28.53

Table 5-9: Frailty Scale Correlations with Demographic Variables

	Age	Sex (M)	Income	Education	Injury from Fall	Serious Injuries	Informal Home care	Formal Home Care
Frailty Index	0.170	-0.119	-0.339	-0.174	0.122	0.122	0.324	0.300
Simplified Frailty	0.097	-0.118	-0.259	-0.155	0.107	0.123	0.329	0.315

Figure 5-4: Distribution of Frailty Index and Simplified Frailty Score by Age Category

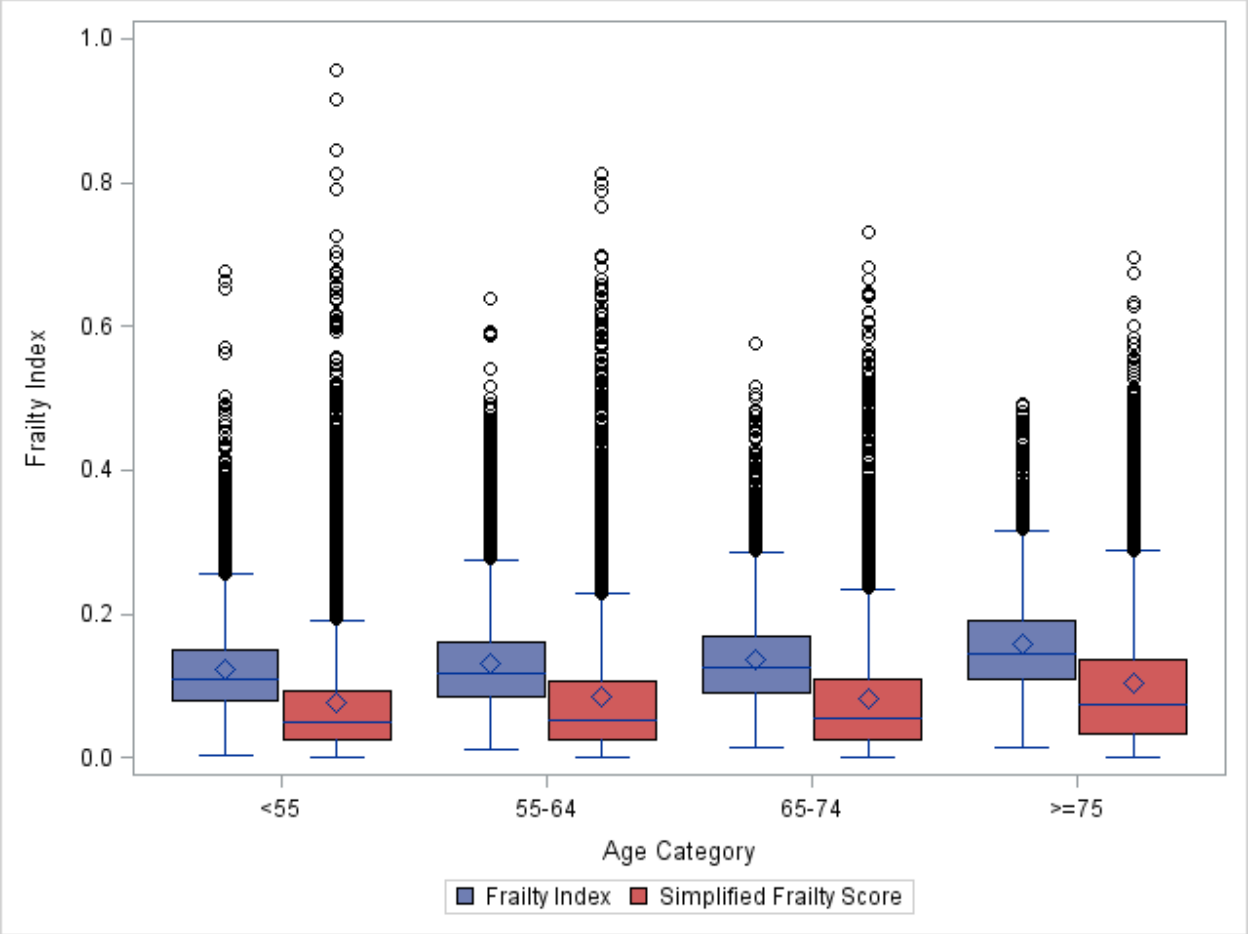

Table 5-10: Comparison of Participants where Frailty Index and Simplified Frailty Score Disagree on Frailty Classification

	Frailty Index ≥0.25, Simplified Frailty <0.25 N=494	Frailty Index <0.25, Simplified Frailty ≥0.25 N=357
	N (%)	N (%)
Age (category)		
45-54	102 (21)	73 (20)
55-64	128 (26)	122 (34)
65-74	106 (21)	74 (21)
≥75	158 (32)	88 (25)
Sex - Male	210 (42)	155 (43)
Informal Care	95 (19)	58 (16)
Formal Care	151 (31)	146 (41)
Injuries	107 (22)	63 (18)
Falls	61 (12)	27 (8)
Assistive Devices	173 (35)	160 (45)
	Mean (SD)	Mean SD)
Age (years)	66.3 (11.2)	64.8 (10.6)
Chronic Conditions	8.4 (2.7)	4.5 (2.2)

CHAPTER 7: DISCUSSION & CONCLUSIONS

This chapter includes the findings from the analyses conducted in this thesis project. The contribution of this project to current research on frailty, its limitations, and potential for future research will be discussed.

7.1 Discussion of Results

This thesis provides evidence for use of the cumulative deficit model for assessing frailty in participants in a large, population-based database. A Frailty Index can incorporate items that measure a number of physiological systems. The results from this analysis provide evidence to the theory that frailty involves multiple systems,[3] rather than just physical function, and the adaptability of the cumulative deficit model makes it better-suited to measure frailty than a measurement model that is restricted to certain domains.

7.1.1 Frailty Index

A Frailty Index score could be calculated for 98% of the population in the database, and the score had a good distribution with minimal skew and no floor or ceiling effects. The distribution of the Frailty Index in this population, including the maximum of 0.67, was consistent with results from the English Longitudinal Study on Ageing, a recent study that also used a Frailty Index to measure a similar population.[104] Further validation is needed to determine how closely the Frailty Index score reflects the true distribution of frailty in this population. The direction of correlations between the Frailty Index and sociodemographic variables, injuries, falls, and home care received, were consistent with those that would be expected for a measurement of frailty, providing evidence that frailty is the construct being measured by this instrument. The structural equation modelling results showed that a variety of health systems contributed to the frailty measurement

model, and included physical, psychological, social, and environmental factors. This is consistent with previous research, which suggests that deficits representing a variety of health systems should be included when developing a Frailty Index,[38] and provides evidence for the content validity of this instrument.

A Frailty Index was simple to construct from the health indicators available in the CLSA, using the method outlined by Searle et al for constructing a Frailty Index.[98] The selection of health deficits for inclusion in a Frailty Index can be done more or less at random,[37] which makes this procedure adaptable to any health database measuring older adults that collects enough information for an index to be constructed. A derived variable could easily be calculated and incorporated into the CLSA database or other similar population databases. A Frailty Index is currently underutilized in population-based studies compared to other methods of assessing frailty.[12] The continued development and validation of the Frailty Index for the CLSA and other similar population databases will facilitate the comparison of populations and promote better understanding of the determinants and outcomes of frailty.

7.1.2 Exploratory Factor Analysis

The exploratory analysis showed three latent constructs for the observed variables included in the model that may play an important role in defining frailty. The first and most important factor was labelled *Physical Frailty* and had very strong factor loadings from items in the three functional status scales (upper body strength, lower body strength, and dexterity) and basic and instrumental activities of daily living. Self-rated health and some chronic conditions had moderate loadings on this factor as well, including conditions that would be associated with lower physical functioning, such as osteoporosis, arthritis, and Parkinson's disease. Weak loadings were also

found on this primary factor from the two items from the CES-D 10 depression scale that are included in the Fried phenotype of frailty scale: "frequency have trouble get going", and frequency everything is an effort. This *Physical Frailty* domain explained the most variation in the observations included in the analysis, supporting current research stating that strength, mobility, and ability to care for oneself are the most important contributors to frailty.[7,105,106] The Fried phenotype of frailty model that measures primarily the *Physical Frailty* domain is commonly used in population-based studies.[12] However, variables that did not load strongly on this factor were still significant contributors to the factor model. The second and third factors identified, *Life Satisfaction* and *Depressive Symptoms*, suggest there is a strong psychological component to frailty. *Life Satisfaction* includes strong loadings from the Satisfaction with Life Scale, and related CES-D 10 depression items, including happiness, hopefulness for the future, and less feelings of depression or loneliness. Participants' perception of their mental health and social support also loaded on the *Life Satisfaction* factor. This combination of items suggests that a person's frailty is in part determined by their perception of their own health and quality of life. If they are happy, socially involved, and satisfied with their life as it is, participants may be less likely to be frail. The *Depressive Symptoms* factor had significant loadings from each item on the CES-D 10 scale, and the strongest loadings were from items on feeling fearful, depressed, or easily bothered, and having trouble concentrating. These results are consistent with recent research suggesting that psychological vulnerability may be an important component of frailty.[107] While physical functioning and self-care are the most important domains, items on life satisfaction and depression should not be overlooked when studying frailty.

7.1.3 Structural Equation Modelling

At the outset of the SEM analysis, the items were hypothesized to load on five latent constructs based on the domains of the framework of the International Classification of Functioning, Disability and Health (ICF).[26] Items from every ICF health domain were included in the model, based on the theory that every health domain would contribute to frailty. This model had acceptable fit of the data, but was improved by some small modifications. BMI and continence items had relatively low contributions to the *Impairments to Body Structure and Function* domain, and improved the model when moved to the *Health Conditions* domain, which included the other chronic conditions measured in the CLSA, and self-rated health. Self-rated mental health, initially included in the same domain as self-rated health and chronic conditions, was moved to *Impairments to Body Structure and Function* due to the association of this item with depression, anxiety, and life satisfaction. The addition of error covariance terms also improved the model substantially. Each of the covariance terms was between items that covary more than would be explained by the model. ADL and IADL are measured at the same time on the same scale, and contain similar item. The same reasons apply to upper-body strength and dexterity items, and self-rated health and self-rated mental health items. ADL and continence are related, as the ADL scale includes an item on continence. Finally, lower-body strength is expected to be related to BMI, as those with normal BMI are expected to be more active and mobile than those with abnormal BMI. The most prominent finding from the SEM results was the number and variety of items that made a contribution to the model. Every single item that was hypothesized to relate to frailty was a significant contributor and could not be removed based on the Wald tests. In the accumulated deficit model for measuring frailty, a list of health deficits from a variety of health systems and areas of function will be a reasonable measurement of frailty, so long as sufficient

health deficits (30-70) are included.[13] The modified model showed that a variety of health deficits, including physical, psychological, and social factors, all contributed to the measurement model, and had good fit of the data, supporting the accumulated deficit theory of frailty measurement. Each health deficit accumulated contributes to frailty in this model, so the more potential deficits included in the frailty measurement, the more accurate the measurement instrument will be. This makes the accumulated deficit model for measuring frailty appropriate for any setting where a sufficient number of health deficits can be collected, especially large population databases.

All things being equal, models that contain more parameters will generally have a better fit of the data.[108] For example, if every possible observable health attribute was included in a measurement model, we would anticipate that model to have a close to perfect fit of the data, as any variation between included participants would be explained. Based on this information, the improvement of the model fit upon removal of some variables was contrary to expectations. The most parsimonious model examined, the simplified frailty model, had the better fit of the data compared to the expanded model. This model included only the three physical functioning domains measured in the CLSA and depressive symptoms, loading on a single factor. Strong associations between subjective well-being and functional status variables were also found in the exploratory factor analysis, and these results are consistent with recent research on associations between latent constructs of frailty.[109,110]

7.1.4 Simplified Frailty Score

Items measuring physical functioning and depressive symptoms in the CLSA were the strongest contributors to both the EFA and SEM analyses of all of the variables measured. Based on these findings, we sought to examine if these items alone could be used to create a simplified frailty

scale. The score was constructed weighting each domain equally. This was done in order to avoid unintentional weighting due to the different number of items available for each domain.[105] Simplifying frailty measurement to fewer items may present some limitations. The Simplified Frailty Score had only 24 items, many of which had a very low endorsement rate. This lead to skew and floor effects, as most participants had a very low SF score. As a result, the SF score may be limited in its ability to detect pre-frailty, as participants who have very mild frailty may endorse very few or no items. When compared to the Frailty Index, the SF score had similar correlations with age, education, income, fall status, injuries, home care, and the FI and SF scores were significantly correlated with one another. The relationships between frailty scales and sociodemographic and healthcare utilization variables were as predicted. Frailty was associated with increasing age, female gender, less education, and more healthcare utilization. These correlations are consistent with those found in previous research on the association between socio-demographic factors associated with frailty.[111,112] These findings suggest that both scores are measuring the construct of frailty. We sought to examine the subpopulations of participants that each scale would identify as frail if a cut-off were used. If both scales are measurements of frailty, we would expect them to identify the same participants as frail. The two scores agreed on the majority of participants, but the kappa value of 0.67 suggested the agreement was less than would be appropriate for two scores that should measure same construct.

7.2 Contribution to Frailty Research

This thesis has calculated and evaluated two methods for measuring the level of frailty of participants included in the CLSA Tracking database. A Frailty Index was shown to be easy to calculate in population-based datasets and is the preferable of the two alternatives studied. The

distribution of the Frailty Index suggested good sensitivity to changing levels of frailty, as well as the ability to discern differing levels of frailty among individuals. The wide use of Frailty Indices helps make frailty research from population databases that incorporate this score more generalizable. The Frailty Index was showed good construct validity, each domain included was shown to make a significant contribution to the measurement model, which showed close fit of the data. The score correlated as hypothesized with sociodemographic and healthcare variables. The Simplified Frailty Score focused only on the domains from the structural equation model that best fit the data. These domains (upper body strength, lower body strength, dexterity, and depressive symptoms) also had the highest factor loadings in the exploratory factor analysis. The results suggested that these may be the most important contributors to frailty of those measured. The Simplified Frailty Score was found to be a reasonable approximation of the Frailty Index. The SF had many fewer variables, and may be useful as a frailty screening tool in contexts where data collection is limited, such as in studies where the burden on the patient is a problem. Further research is required to properly ascertain the reliability and construct validity of this measurement model using prospective data.

7.2.1 Recommendations for Frailty Indices

A Frailty Index would be simple to incorporate into the CLSA database as a derived variable available to any researchers requesting use of the dataset. The variables included are common health indicators that would be available in other population databases. It would be straightforward to construct a Frailty Index using the same methodology outlined in this report, as recommended by Searle et al (2008),[98] for any large database collecting a variety of health indicators. The SEM analysis results suggest that a Frailty Index should include items loading on

each of the domains included in the complete model: health conditions, body structure and function, activity limitations, participation restrictions, and environmental factors. Items on physical functioning, depressive symptoms, and life satisfaction were shown to be of particular importance in measuring frailty.

7.3 Limitations

The use of secondary data was ideal for this thesis, as it provided fast and cost-effective access to data for a population that would otherwise be well-beyond the scope of a student project. However, the dataset does present limitations to this analysis. Since the analysis was restricted to data that had already been collected, some hypothesized indicators of frailty were not assessed in this dataset, such as medication, nutrition, and physical activity levels. The subpopulations excluded (i.e. residents of the territories, full-time members of the Canadian Armed Forces, those who do not speak English or French) are likely to be small with little effect on the results. A cross-sectional dataset is restrictive, causing difficulty discerning predictors from outcomes. Some of the factors included in the model may be variables that are affected by a participant's frailty status but do not contribute to frailty. Additional validation is necessary for the measurements of frailty studied in this analysis, using longitudinal outcomes such as hospitalization, nursing home admission, and mortality. If a higher Frailty Index is shown to increase the risk for these outcomes, this would provide evidence that the score is a valid measurement to a participant's true frailty.

In the SEM analysis, there were a number of participants excluded due to missing data for one or more variables. While the proportion of data missing for any one variable was low (<5%), the participants excluded were significantly different from the population that was included in the analysis. Those excluded were older and had more chronic conditions, suggesting that this may

be a more vulnerable population. This presents a challenge for frailty assessment in large databases. Frailty has shown to be a complex condition involving multiple health domains, and as such will require a number of variables collected through self-report. If some vulnerable older adults have trouble completing these instruments, then assessments of frailty that include self-report items may be more likely miss frail adults in this population. For this reason, the number and percentage of participants for whom frailty could not be assessed should be reported for any population where frailty is measured. Although the participants excluded were not shown to be missing at random, missing data were not imputed, as the proportion with missing data was small and imputation was unlikely to affect the results.

SEM analysis is limited in secondary datasets, as the model specification is restricted to variables that are included in the dataset, which can lead to data affecting the theory and model.[97] The limitation of secondary data to this analysis is mitigated by two key factors. Firstly, the CLSA database was designed to be comprehensive, and to collect a wide array of health indicators for the study of healthy aging. This dataset is well-suited to study a complex condition affecting older adults that involves multiple health domains, such as frailty. Secondly, potential indicators for frailty were studied in the current literature, and this was used to guide the selection of health indicators from the CLSA for inclusion in this study.

Frailty and disability are distinct concepts, but difficult to distinguish from one another, as many of the same domains are included in both.[7] In the SEM analysis, a simplified model for measuring frailty had the best fit of the data, which included one latent construct and four observed variables: three functional status subscales measuring upper body strength, lower body strength and dexterity, as well as the CES-D 10 depression scale. These results suggest that the functional

limitations and depressive symptoms may be the most important contributors to frailty that were assessed in this study. Alternatively, it could be interpreted that this latent construct is measuring disability associated with frailty. Participants who have a disability would be expected to score poorly on the functional status items, and functional disability has previously shown strong association with depression.[113] Further validation on the simplified frailty scale using longitudinal data is necessary to ascertain the constructs it is measuring. These studies should include a separate measure of disability in order to further elucidate the measurement of these two related concepts. This is necessary, since the measurement instruments studied, especially the SF, and others focusing on physical functioning may be measuring disability as much as they are measuring frailty. Based on these results we would expect the Frailty Index and the simplified frailty scale to predict worsening outcomes such as hospital admission, institutionalization and mortality in a longitudinal study.

7.5 Conclusions

A number of different measurement models have been used to operationalize frailty, resulting in inconsistent reports of frailty prevalence.[23] The proportion of Canadians over the age of 65 is growing. For the first time ever, there are more persons aged 65 years or older in Canada than children under 15, nearly one in six Canadians (16.1%) are at least 65 years old.[1] This presents challenges to understanding the population of older adults who are frail, in both the assessment of need for healthcare resources and the design of interventions. This thesis presents a Frailty Index as a theoretically-meaningful model for frailty measurement that showed good validity in a large population-based sample of community-dwelling adults. This is an important precursor to future research that can further validate the frailty indices studied, develop standards for frailty

assessment, and ultimately promote better understanding of frailty. This thesis makes a valuable and necessary contribution to the study of frailty measurement.

REFERENCES

1. Government of Canada, S. C. The Daily — Canada's population estimates: Age and sex, July 1, 2015. (2015). Available at: http://www.statcan.gc.ca/daily-quotidien/150929/dq150929b-eng.htm. (Accessed: 4th August 2016)
2. Theou, O. & Rockwood, K. *Frailty in aging : biological, clinical, and social implications*. **41,** (Karger, 2015).
3. Xue, Q.-L. The Frailty Syndrome: Definition and Natural History. *Clin. Geriatr. Med.* **27,** 1–15 (2011).
4. Fried, L. P. *et al.* Frailty in older adults: evidence for a phenotype. *J. Gerontol. A. Biol. Sci. Med. Sci.* **56,** M146–156 (2001).
5. Morley, J. E., Perry, H. M. & Miller, D. K. Editorial: Something about frailty. *J. Gerontol. A. Biol. Sci. Med. Sci.* **57,** M698–704 (2002).
6. Van Kan, G. A. *et al.* The Assessment of Frailty in Older Adults. *Clin. Geriatr. Med.* **26,** 275–286 (2010).
7. Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D. & Anderson, G. Untangling the Concepts of Disability, Frailty, and Comorbidity: Implications for Improved Targeting and Care. *J. Gerontol. A. Biol. Sci. Med. Sci.* **59,** M255–M263 (2004).
8. Rockwood, K., Hogan, D. B. & MacKnight, C. Conceptualisation and measurement of frailty in elderly people. *Drugs Aging* **17,** 295–302 (2000).
9. Hamerman, D. Toward an understanding of frailty. *Ann. Intern. Med.* **130,** 945–950 (1999).
10. Gillick, M. Pinning down frailty. *J. Gerontol. A. Biol. Sci. Med. Sci.* **56,** M134–135 (2001).
11. Rockwood, K. What would make a definition of frailty successful? *Age Ageing* **34,** 432–434 (2005).

12. Bouillon, K. *et al.* Measures of frailty in population-based studies: an overview. *BMC Geriatr.* **13,** 64 (2013).
13. Rockwood, K. & Mitnitski, A. Frailty in relation to the accumulation of deficits. *J. Gerontol. A. Biol. Sci. Med. Sci.* **62,** 722–727 (2007).
14. Li, G. *et al.* Comparison between Frailty Index of Deficit Accumulation and Phenotypic Model to Predict Risk of Falls: Data from the Global Longitudinal Study of Osteoporosis in Women (GLOW) Hamilton Cohort. *PLoS ONE* **10,** e0120144 (2015).
15. Sternberg, S. A., Wershof Schwartz, A., Karunananthan, S., Bergman, H. & Mark Clarfield, A. The identification of frailty: a systematic literature review. *J. Am. Geriatr. Soc.* **59,** 2129–2138 (2011).
16. Clegg, A., Young, J., Iliffe, S., Rikkert, M. O. & Rockwood, K. Frailty in elderly people. *The Lancet* **381,** 752–762 (2013).
17. Theou, O., Brothers, T. D., Mitnitski, A. & Rockwood, K. Operationalization of frailty using eight commonly used scales and comparison of their ability to predict all-cause mortality. *J. Am. Geriatr. Soc.* **61,** 1537–1551 (2013).
18. Hamaker, M. E. *et al.* Frailty screening methods for predicting outcome of a comprehensive geriatric assessment in elderly patients with cancer: a systematic review. *Lancet Oncol.* **13,** e437–444 (2012).
19. Pialoux, T., Goyard, J. & Lesourd, B. Screening tools for frailty in primary health care: a systematic review. *Geriatr. Gerontol. Int.* **12,** 189–197 (2012).
20. De Vries, N. M. *et al.* Outcome instruments to measure frailty: a systematic review. *Ageing Res. Rev.* **10,** 104–114 (2011).

21. Hogan, D. B., MacKnight, C. & Bergman, H., Steering Committee, Canadian Initiative on Frailty and Aging. Models, definitions, and criteria of frailty. *Aging Clin. Exp. Res.* **15,** 1–29 (2003).
22. Levers, M.-J., Estabrooks, C. A. & Ross Kerr, J. C. Factors contributing to frailty: literature review. *J. Adv. Nurs.* **56,** 282–291 (2006).
23. Collard, R. M., Boter, H., Schoevers, R. A. & Oude Voshaar, R. C. Prevalence of frailty in community-dwelling older persons: a systematic review. *J. Am. Geriatr. Soc.* **60,** 1487–1492 (2012).
24. King-Kallimanis, B. L., Kenny, R. A. & Savva, G. M. Factor structure for the frailty syndrome was consistent across Europe. *J. Clin. Epidemiol.* **67,** 1008–1015 (2014).
25. Sarkisian, C. A., Gruenewald, T. L., John Boscardin, W. & Seeman, T. E. Preliminary Evidence for Subdimensions of Geriatric Frailty: The MacArthur Study of Successful Aging. *J. Am. Geriatr. Soc.* **56,** 2292–2297 (2008).
26. WHO | International Classification of Functioning, Disability and Health (ICF). *WHO* (2015). Available at: http://www.who.int/classifications/icf/en/. (Accessed: 11th August 2015)
27. Raina, P. S. *et al.* The Canadian Longitudinal Study on Aging (CLSA). *Can. J. Aging Rev. Can. Vieil.* **28,** 221–229 (2009).
28. Gobbens, R. J. J., Luijkx, K. G., Wijnen-Sponselee, M. T. & Schols, J. M. G. A. Towards an integral conceptual model of frailty. *J. Nutr. Health Aging* **14,** 175–181 (2010).
29. WHO | International Classification of Functioning, Disability and Health (ICF). *WHO* Available at: http://www.who.int/classifications/icf/en/. (Accessed: 11th August 2015)

30. Dent, E., Chapman, I., Howell, S., Piantadosi, C. & Visvanathan, R. Frailty and functional decline indices predict poor outcomes in hospitalised older people. *Age Ageing* **43,** 477–484 (2014).
31. Rothman, M. D., Leo-Summers, L. & Gill, T. M. Prognostic significance of potential frailty criteria. *J. Am. Geriatr. Soc.* **56,** 2211–2216 (2008).
32. Puts, M. T. E., Lips, P. & Deeg, D. J. H. Static and dynamic measures of frailty predicted decline in performance-based and self-reported physical functioning. *J. Clin. Epidemiol.* **58,** 1188–1198 (2005).
33. Ensrud, K. E. *et al.* Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. *Arch. Intern. Med.* **168,** 382–389 (2008).
34. Avila-Funes, J. A. *et al.* Cognitive impairment improves the predictive validity of the phenotype of frailty for adverse health outcomes: the three-city study. *J. Am. Geriatr. Soc.* **57,** 453–461 (2009).
35. Shamliyan, T., Talley, K. M. C., Ramakrishnan, R. & Kane, R. L. Association of frailty with survival: a systematic literature review. *Ageing Res. Rev.* **12,** 719–736 (2013).
36. Rockwood, K., Mitnitski, A., Song, X., Steen, B. & Skoog, I. Long-Term Risks of Death and Institutionalization of Elderly People in Relation to Deficit Accumulation at Age 70. *J. Am. Geriatr. Soc.* **54,** 975–979 (2006).
37. Rockwood, K. *et al.* A brief clinical instrument to classify frailty in elderly people. *Lancet Lond. Engl.* **353,** 205–206 (1999).
38. Rolfson, D. B., Majumdar, S. R., Tsuyuki, R. T., Tahir, A. & Rockwood, K. Validity and reliability of the Edmonton Frail Scale. *Age Ageing* **35,** 526–529 (2006).

39. Peters, L. L., Boter, H., Burgerhof, J. G. M., Slaets, J. P. J. & Buskens, E. Construct validity of the Groningen Frailty Indicator established in a large sample of home-dwelling elderly persons: Evidence of stability across age and gender. *Exp. Gerontol.* **69,** 129–141 (2015).
40. Aminzadeh, F., Dalziel, W. B. & Molnar, F. J. Targeting frail older adults for outpatient comprehensive geriatric assessment and management services: An overview of concepts and criteria. *Rev. Clin. Gerontol.* **12,** 82–92 (2002).
41. Rockwood, K., Andrew, M. & Mitnitski, A. A comparison of two approaches to measuring frailty in elderly people. *J. Gerontol. A. Biol. Sci. Med. Sci.* **62,** 738–743 (2007).
42. Gill, T. M., Gahbauer, E. A., Allore, H. G. & Han, L. Transitions between frailty states among community-living older persons. *Arch. Intern. Med.* **166,** 418–423 (2006).
43. Rockwood, K., Mitnitski, A. B. & MacKnight, C. Some mathematical models of frailty and their clinical implications. *Rev. Clin. Gerontol.* **12,** 109–117 (2002).
44. Bergman, H. *et al.* Frailty: an emerging research and clinical paradigm--issues and controversies. *J. Gerontol. A. Biol. Sci. Med. Sci.* **62,** 731–737 (2007).
45. Fairhall, N., Sherrington, C., Kurrle, S. E., Lord, S. R. & Cameron, I. D. ICF participation restriction is common in frail, community-dwelling older people: an observational cross-sectional study. *Physiotherapy* **97,** 26–32 (2011).
46. Fairhall, N. *et al.* Treating frailty-a practical guide. *BMC Med.* **9,** 83 (2011).
47. Selb, M. *et al.* A guide on how to develop an International Classification of Functioning, Disability and Health Core Set. *Eur. J. Phys. Rehabil. Med.* **51,** 105–117 (2015).
48. Gobbens, R. J. J., Assen, M. A. L. M. van, Luijkx, K. G. & Schols, J. M. G. A. The Predictive Validity of the Tilburg Frailty Indicator: Disability, Health Care Utilization, and Quality of Life in a Population at Risk. *The Gerontologist* **52,** 619–631 (2012).

49. Blodgett, J., Theou, O., Kirkland, S., Andreou, P. & Rockwood, K. Frailty in NHANES: Comparing the frailty index and phenotype. *Arch. Gerontol. Geriatr.* **60,** 464–470 (2015).

50. Ilinca, S. & Calciolari, S. The Patterns of Health Care Utilization by Elderly Europeans: Frailty and Its Implications for Health Systems. *Health Serv. Res.* **50,** 305–320 (2015).

51. Kato, G., Tamiya, N., Kashiwagi, M., Sato, M. & Takahashi, H. Relationship between home care service use and changes in the care needs level of Japanese elderly. *BMC Geriatr.* **9,** 58 (2009).

52. Mittmann, N. *et al.* Population-based home care services in breast cancer: utilization and costs. *Curr. Oncol.* **19,** e383–e391 (2012).

53. Gobbens, R. J. J. & van Assen, M. A. L. M. Frailty and its prediction of disability and health care utilization: The added value of interviews and physical measures following a self-report questionnaire. *Arch. Gerontol. Geriatr.* **55,** 369–379 (2012).

54. Shea, B. J. *et al.* Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med. Res. Methodol.* **7,** 10 (2007).

55. Smith, V., Devane, D., Begley, C. M. & Clarke, M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Med. Res. Methodol.* **11,** 15 (2011).

56. Wolfson, C. *et al.* The Canadian community health survey as a potential recruitment vehicle for the Canadian longitudinal study on aging. *Can. J. Aging Rev. Can. Vieil.* **28,** 243–249 (2009).

57. Raina, P. S. *et al.* Accessing Health Care Utilization Databases for Health Research: A Canadian Longitudinal Study on Aging Feasibility Study. *Can. J. Aging Rev. Can. Vieil.* **28,** 287–294 (2009).

58. Andrews, F. M. & Withey, S. B. *Social Indicators of Well-Being*. (Springer US, 1976).

59. WHO | Obesity: preventing and managing the global epidemic. *WHO* Available at: http://www.who.int/entity/nutrition/publications/obesity/WHO_TRS_894/en/index.html. (Accessed: 13th July 2016)

60. Winter, J. E., MacInnis, R. J., Wattanapenpaiboon, N. & Nowson, C. A. BMI and all-cause mortality in older adults: a meta-analysis. *Am. J. Clin. Nutr.* ajcn.068122 (2014). doi:10.3945/ajcn.113.068122

61. Raina, P., Bonnett, B., Waltner-Toews, D., Woodward, C. & Abernathy, T. How reliable are selected scales from population-based health surveys? An analysis among seniors. *Can. J. Public Health Rev. Can. Santé Publique* **90,** 60–64 (1999).

62. Branch, L. G. & Jette, A. M. The Framingham Disability Study: I. Social disability among the aging. *Am. J. Public Health* **71,** 1202–1210 (1981).

63. Nagi, S. Z. An epidemiology of disability among adults in the United States. *Milbank Mem. Fund Q. Health Soc.* **54,** 439–467 (1976).

64. Rosow, I. & Breslau, N. A Guttman health scale for the aged. *J. Gerontol.* **21,** 556–559 (1966).

65. Cornoni-Huntley, J. *et al.* Established populations for epidemiologic studies of the elderly: study design and methodology. *Aging Milan Italy* **5,** 27–37 (1993).

66. Ferrer, M., Lamarca, R., Orfila, F. & Alonso, J. Comparison of performance-based and self-rated functional capacity in Spanish elderly. *Am. J. Epidemiol.* **149,** 228–235 (1999).

67. Tager, I. B., Swanson, A. & Satariano, W. A. Reliability of physical performance and self-reported functional measures in an older population. *J. Gerontol. A. Biol. Sci. Med. Sci.* **53,** M295–300 (1998).

68. Kempen, G. I., Steverink, N., Ormel, J. & Deeg, D. J. The assessment of ADL among frail elderly in an interview survey: self-report versus performance-based tests and determinants of discrepancies. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* **51,** P254–260 (1996).

69. Simonsick, E. M. *et al.*, Health ABC Study Group. Measuring higher level physical function in well-functioning older adults: expanding familiar approaches in the Health ABC study. *J. Gerontol. A. Biol. Sci. Med. Sci.* **56,** M644–649 (2001).

70. Beaton, D. E. *et al.* Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J. Hand Ther. Off. J. Am. Soc. Hand Ther.* **14,** 128–146 (2001).

71. Sternfeld, B., Ngo, L., Satariano, W. A. & Tager, I. B. Associations of body composition with physical performance and self-reported functional limitation in elderly men and women. *Am. J. Epidemiol.* **156,** 110–121 (2002).

72. Fillenbaum, G. G. & Smyer, M. A. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. *J. Gerontol.* **36,** 428–434 (1981).

73. Fillenbaum, G. G. Screening the elderly. A brief instrumental activities of daily living measure. *J. Am. Geriatr. Soc.* **33,** 698–706 (1985).

74. Rey, A. *L'examen clinique en psychologie*. (Presses universitaires de France, 1964).

75. Lezak, M., Howieson, D. & Loring, D. *Neuropsychological Assessment*. (Oxford University Press, 2004).

76. Bäckman, L., Jones, S., Berger, A.-K., Laukka, E. J. & Small, B. J. Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. *Neuropsychology* **19,** 520–531 (2005).

77. Lemay, S., Bédard, M.-A., Rouleau, I. & Tremblay, P.-L. Practice Effect and Test-Retest Reliability of Attentional and Executive Tests in Middle-Aged to Elderly Subjects. *Clin. Neuropsychol.* **18,** 284–302 (2004).

78. Himmelfarb, S. & Murrell, S. A. Reliability and validity of five mental health scales in older persons. *J. Gerontol.* **38,** 333–339 (1983).

79. Folstein, M. F., Folstein, S. E. & McHugh, P. R. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12,** 189–198 (1975).

80. Billick, S. B., Siedenburg, E., Burgert, W. & Bruni-Solhkhah, S. M. Validation of the Mental Alternation Test with the Mini-Mental State Examination in geriatric psychiatric inpatients and normal controls. *Compr. Psychiatry* **42,** 202–205 (2001).

81. Crossley, M., D'Arcy, C. & Rawson, N. S. Letter and category fluency in community-dwelling Canadian seniors: a comparison of normal participants to those with dementia of the Alzheimer or vascular type. *J. Clin. Exp. Neuropsychol.* **19,** 52–62 (1997).

82. Fama, R. *et al.* Fluency Performance Patterns in Alzheimer's Disease and Parkinson's Disease. *Clin. Neuropsychol. Neuropsychol. Dev. Cogn. Sect. D* **12,** 487–499 (1998).

83. Tierney, M. C., Yao, C., Kiss, A. & McDowell, I. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology* **64,** 1853–1859 (2005).

84. Monsch, A. U. *et al.* Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Arch. Neurol.* **49,** 1253–1258 (1992).

85. Validation of cognitive functioning categories in the Canadian Community Health Survey - Healthy Aging. Available at: http://www.statcan.gc.ca/pub/82-003-x/2010004/article/11391-eng.htm. (Accessed: 14th July 2016)

86. Andresen, E. M., Malmgren, J. A., Carter, W. B. & Patrick, D. L. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am. J. Prev. Med.* **10,** 77–84 (1994).

87. Knight, R. G., Williams, S., McGee, R. & Olaman, S. Psychometric properties of the Centre for Epidemiologic Studies Depression Scale (CES-D) in a sample of women in middle life. *Behav. Res. Ther.* **35,** 373–380 (1997).

88. Diener, E., Emmons, R. A., Larsen, R. J. & Griffin, S. The Satisfaction With Life Scale. *J. Pers. Assess.* **49,** 71–75 (1985).

89. W. A. Arrindell, L. M. The Satisfaction With Life Scale (SWLS): Psychometric properties in a non-psychiatric medical outpatient sample. *Personal. Individ. Differ.* **12,** 117–123 (1991).

90. Pavot, W., Diener, E., Colvin, C. R. & Sandvik, E. Further Validation of the Satisfaction With Life Scale: Evidence for the Cross-Method Convergence of Well-Being Measures. *J. Pers. Assess.* **57,** 149–161 (1991).

91. Sherbourne, C. D. & Stewart, A. L. The MOS social support survey. *Soc. Sci. Med. 1982* **32,** 705–714 (1991).

92. McDowell, I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. (Oxford University Press, 2006).

93. Government of Canada, S. C. Canadian Community Health Survey - Healthy Aging (CCHS). (2008). Available at: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5146&lang=en&db=imdb&adm=8&dis=2. (Accessed: 2nd April 2016)

94. Government of Canada, S. C. The General Social Survey: An Overview. (2013). Available at: http://www.statcan.gc.ca/pub/89f0115x/89f0115x2013001-eng.htm. (Accessed: 2nd April 2016)

95. Searle, S. D., Mitnitski, A., Gahbauer, E. A., Gill, T. M. & Rockwood, K. A standard procedure for creating a frailty index. *BMC Geriatr.* **8,** 24 (2008).

96. Kline, R. B. *Principles and Practice of Structural Equation Modeling, Second Edition*. (The Guilford Press, 2004).

97. AB Costello, J. W. O. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Pract. Assess. Res. Amp Eval.* **10,** 1–9 (2005).

98. Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I. & Vila–Abad, E. Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Qual. Quant.* **44,** 153–166 (2008).

99. Cattell, R. B. The Scree Test For The Number Of Factors. *Multivar. Behav. Res.* **1,** 245–276 (1966).

100. Relation of Sample Size to the Stability of Component Patterns. **103,** 265–275

101. Hu, L. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* **6,** 1–55 (1999).

102. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19,** 716–723 (1974).

103. Streiner, D. L., Norman, G. R. & Cairney, J. *Health Measurement Scales: A practical guide to their development and use*. (Oxford University Press, 2014).

104. Marshall, A., Nazroo, J., Tampubolon, G. & Vanhoutte, B. Cohort differences in the levels and trajectories of frailty among older people in England. *J. Epidemiol. Community Health* jech–2014–204655 (2015). doi:10.1136/jech-2014-204655

105. Fried, L. P., Bandeen-Roche, K., Chaves, P. H. & Johnson, B. A. Preclinical mobility disability predicts incident mobility disability in older women. *J. Gerontol. A. Biol. Sci. Med. Sci.* **55,** M43–52 (2000).

106. Vermeulen, J., Neyens, J. C., van Rossum, E., Spreeuwenberg, M. D. & de Witte, L. P. Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: a systematic review. *BMC Geriatr.* **11,** 33 (2011).

107. Lohman, M., Dumenci, L. & Mezuk, B. Depression and Frailty in Late Life: Evidence for a Common Vulnerability. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* **71,** 630–640 (2016).

108. Forster, M. & Sober, E. How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions. *Br. J. Philos. Sci.* **45,** 1–35 (1994).

109. Mezuk, B., Lohman, M., Dumenci, L. & Lapane, K. L. Are depression and frailty overlapping syndromes in mid- and late-life? A latent variable analysis. *Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry* **21,** 560–569 (2013).

110. Simone, P. M. & Haas, A. L. Frailty, Leisure Activity and Functional Status in Older Adults: Relationship With Subjective Well Being. *Clin. Gerontol.* **36,** 275–293 (2013).

111. Salem, B. E. *et al.* Constructing and identifying predictors of frailty among homeless adults—a latent variable structural equations model approach. *Arch. Gerontol. Geriatr.* **58,** 248–256 (2014).

112. Mello, A. de C., Engstrom, E. M. & Alves, L. C. Health-related and socio-demographic factors associated with frailty in the elderly: a systematic literature review. *Cad. Saúde Pública* **30,** 1143–1168 (2014).

113. Von Korff M, Ormel J, Katon W, and Lin EB. Disability and depression among high utilizers of health care: A longitudinal analysis. *Arch. Gen. Psychiatry* **49,** 91–100 (1992).

APPENDIX A: LITERATURE SEARCH STRATEGY

Database: Ovid MEDLINE <1996 to June 2016>

Search Strategy:

#	Medical Subject Headings [Subcategories] and Keywords (mp)	Results
1	Frail Elderly/	7328
2	Aged/ph, px, sn [Physiology, Psychology, Statistics & Numerical Data]	4608
3	Aging/pa, ph, px [Pathology, Physiology, Psychology]	65756
4	Vulnerable Populations/cl, px, sn [Classification, Psychology, Statistics & Numerical Data]	1639
5	healthy ag?ing.mp.	2618
6	Longevity/ or Health Status/ or successful ag?ing.mp.	71448
7	Geriatrics/is, mt [Instrumentation, Methods]	818
8	Risk Assessment/	195794
9	Risk Factors/	560452
10	Health Status Indicators/	18452
11	Disability Evaluation/	28422
12	Forecasting/	53858
13	Patient Care Planning/	17482
14	Biomarkers/	183497
15	Health Surveys/	43429
16	Diagnosis/	2532
17	Operationalization.mp.	713
18	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17	1111488
19	1 and 18	2815
20	limit 19 to (yr="2010 - 2016")	1431
21	limit 20 to English language	1286
22	limit 21 to "review articles"	215

Database: EMBASE <1996 to June 2016>

Search Strategy:

#	Medical Subject Headings [Subcategories] and Keywords (mp)	Results
1	Frail Elderly/	7086
2	Aged/	1812656
3	Aging/	162533
4	Vulnerable Populations/cl, px, sn [Classification, Psychology, Statistics & Numerical Data]	9255
5	healthy ag?ing.mp.	4082
6	Longevity/ or Health Status/ or successful ag?ing.mp.	108061
7	Geriatrics/is, mt [Instrumentation, Methods]	18341
8	Risk Assessment/	369762
9	Risk Factors/	379195
10	Health Status Indicators/	1016
11	Disability Evaluation/	57397
12	Forecasting/	26701
13	Patient Care Planning/	14288
14	Biomarkers/	131006
15	Health Surveys/	132685
16	Diagnosis/	378025
17	Operationalization.mp.	971
18	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17	3159351
19	1 and 18	5977
20	limit 19 to yr="2010 - 2016"	3794
21	limit 20 to English language	3390
22	"systematic review"/	109864
23	20 and 21	64

Database: Cochrane Database of Systematic Reviews

Search Strategy:

#	Medical Subject Headings [Subcategories] and Keywords (mp)	Results
1	Frail Elderly/	7