Estimating the incidence of vertebral deformities in Canadian men and women

Philippe Carrière

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science in Epidemiology

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University, Montreal

November, 2004

Your file Votre référence
ISBN: 0-494-12410-5
Our file Notre référence
ISBN: 0-494-12410-5

ABSTRACT

BACKGROUND: Vertebral deformities are important sequelae of osteoporosis, but for feasibility and technical reasons their epidemiology has yet to be thoroughly described in Canada, especially in men.

OBJECTIVE: To estimate the incidence of osteoporotic vertebral deformities, from data collected by the Canadian Multicentre Osteoporosis Study (CaMos), a large cohort study of randomly selected Canadians radiographed at a five year interval.

METHODS: Sex- and age-specific incidence was estimated in men and women aged 55 years and older. Bayesian methods were employed, including adjustment for nonresponse and attrition biases using multiple imputation. Different assumptions for the missing data mechanism were used in a sensitivity analysis.

RESULTS: Weighted to the Canadian population, men aged 55+ have a crude incidence estimate of 17.7/1000 person-years (PY) (95% CrI: 13.5 - 22.1), whereas the corresponding estimate in women is 14.6/1000 PY (95% CrI: 12.2 - 17.1). Adjustment for bias due to attrition has only a slight effect on the estimates in women across all age groups and in men aged 65+ years, under the assumption that the missing data mechanism is ignorable. The rate estimates that are adjusted for both nonresponse and attrition biases variably diverge from the crude estimates both in magnitude and direction, depending on the assumptions made about the missing data mechanism.

CONCLUSIONS: A reasonable assumption for modeling the missing data mechanism is that the sex- and age-specific biases are at least as large, and in the same direction, as the differences between the respondent rates and the imputed rates for groups with missing deformity data. Therefore, in Canadians aged 55+ years, vertebral deformity rates that are adjusted for nonresponse and attrition biases are estimated as 14.4/1000 PY (95% CrI: 11.8 - 17.4) in women, and 23.8/1000 PY (95% CrI: 19.6 - 29.0) in men.

RESUMÉ

CONTEXTE: Les déformations vertébrales sont des séquelles importantes de l'ostéoporose, mais pour des raisons techniques et de faisabilité leur épidémiologie doit encore être décrite à fond au Canada, particulièrement chez les hommes.
OBJECTIF: Évaluer l'incidence des déformations vertébrales, en utilisant les données rassemblées par l'Étude canadienne multicentrique sur l'ostéoporose (CaMos), une grande étude de cohorte de Canadiens aléatoirement choisis et radiographiés à un intervalle de cinq ans.
MÉTHODOLOGIE: L'incidence spécifique à l'âge et au sexe a été évaluée chez les hommes et les femmes âgés de 55 ans et plus. Des méthodes bayésiennes ont été utilisées, y compris l'imputation multiple pour corriger le biais survenu des pertes de l'échantillon et de la non-réponse. Des suppositions différentes pour le mécanisme de constitution des données manquantes ont été utilisées dans une analyse de sensibilité.
RÉSULTATS: Pondérée à la population canadienne, l'estimation de l'incidence brute chez les hommes âgés de 55+ ans est de 17.7/1000 années-personnes (ICr à 95%: 13.5 - 22.1), tandis que l'évaluation femelle correspondante est de 14.6/1000 années-personnes (ICr à 95% : 12.2 - 17.1). Le réglage pour le biais d'attrition a peu d'effet chez les femmes dans toutes les tranches d'âge et chez les hommes âgés de 65+ ans, conformément à la supposition que les données manquantes sont ignorables. Les évaluations de taux, qui sont ajustées tant pour la nonréponse que pour l'attrition, divergent variablement des estimés brutes en ampleur et direction, selon les suppositions faites pour le mécanisme de constitution des données manquantes.
CONCLUSIONS: Une supposition raisonnable pour modéliser le mécanisme des données manquantes est que le biais spécifique à l'âge et au sexe est d'une ampleur aussi grande et dans la même direction que les différences entre les taux observés et les taux imputés pour les groupes n'ayant pas de données en ce qui concerne les déformations vertébrales. Ainsi, chez les Canadiens âgés de 55+ ans, les taux de déformations vertébrales qui sont ajustés pour le biais de nonréponse et d'attrition sont évalués comme étant 14.4/1000 années-personnes (ICr à 95%: 11.8 - 17.4) chez les femmes, et 23.8/1000 années-personnes (ICr à 95%: 19.6 - 29.0) chez les hommes.

ACKNOWLEDGMENTS

The completion of this thesis would not have been possible without the unwavering generosity of Lawrence Joseph. His genuine interest, infectious enthusiasm, and keen insights were, and remain, truly inspiring. It is my hope that Dr Joseph's intellectual and professional guidance will have left an indelible influence on my future endeavours.

I owe many thanks to Alan Tenenhouse, for sharing his vast and excellent knowledge of osteoporosis during my attempts to review the literature. Dr Tenenhouse was also instrumental in my receiving a Skeletal Health Training Program Award from the Canadian Institute of Health Research, funds which were well appreciated during the preparation of this thesis. I would also like to thank Stewart Jackson for his illuminating remarks on the problems in diagnosing vertebral deformities.

I am thankful to my family and friends for their encouraging words and kind deeds throughout this arduous but enriching task. I am particularly grateful to my wise and gentle Marie, who gave me the impetus to try, the courage to continue, and, most of all, the gift of her friendship. I dedicate this thesis to her.

All our knowledge is, ourselves to know.

Alexander Pope
An Essay on Man (1734)
Epistle 4, line 398

The aim of science is not to open the door to infinite wisdom, but to set a limit to infinite error.

Bertholt Brecht
The Life of Galileo (1939)
Scene 13

STATEMENT OF ORIGINALITY

This thesis contains elements that constitute original scholarship and an advancement of knowledge in Epidemiology. The first estimates of the incidence of vertebral deformities in the Canadian population are reported, as well as adjustments made for attrition and nonresponse biases. The estimates are based on data that were provided by the Canadian Multicentre Study of Osteoporosis (CaMos).

I certify that the original research within this thesis is the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with standard referencing practices. I acknowledge the helpful guidance and support of Professor Lawrence Joseph, who supervised me in the writing of this thesis within the Epidemiology stream of the Master of Science program, in the Department of Epidemiology and Biostatistics at McGill University.

Philippe Carrière

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION **10**

1.1 - RATIONALE 10
1.1.1 - OSTEOPOROSIS AS A SUBJECT OF INTEREST *10*
1.1.2 - VERTEBRAL DEFORMITIES AS A SUBJECT OF INTEREST *11*
1.1.3 - MULITPLE IMPUTATION AS A SUBJECT OF INTEREST *13*
1.2 - OBJECTIVES 14

CHAPTER 2 – LITERATURE REVIEW **15**

2.1 – EPIDEMIOLOGY OF OSTEOPOROSIS 15
2.1.1 - DISTRIBUTION *15*
2.1.1.i - Prevalence & Incidence 15
2.1.1.ii - Age 15
2.1.1.iii - Secular Trends 16
2.1.1.iv - Sex 16
2.1.1.v - Race between Nations 17
2.1.1.vi - Race within Nations 17
2.1.2 - MORBIDITY *18*
2.1.2.i - Hip Fractures 18
2.1.2.i - Vertebral Fractures 18
2.1.2.ii - Wrist Fractures 19
2.1.3 - MORTALITY *19*
2.1.4 - COST *20*
2.1.5 – PATHOGENESIS OF OSTEOPOROSIS *20*
2.1.5.i - Bone strength 20
2.1.5.ii - Modeling/Remodeling 21
2.1.5.iii - Secondary osteoporosis 23
2.1.6.i - Osteoporosis as an evolving construct 24
2.1.6.ii - The WHO criteria 25
2.1.6.iii - Other definitions 26
2.1.7 – RISK FACTORS *27*
2.1.7.i - Risk factors for low BMD 27
2.1.7.ii - Risk factors for fracture 28
2.1.8 – PREVENTION AND TREATMENT *29*
2.2 – EPIDEMIOLOGY OF VERTEBRAL DEFORMITIES 31
2.2.1 - INTRODUCTION *31*
2.2.1.i - Terminology 31
2.2.1.ii - Particularities of deformities as osteoporotic fractures 32
2.2.2 - PATHOGENESIS AND BIOMECHANICS OF VERTEBRAL DEFORMITIES *33*
2.2.3 - CLINICAL PRESENTATION AND RELEVANCE *35*
2.2.3.i - Acute presentation 35
2.2.3.ii - Chronic sequelae 36
2.2.3.iii - Mortality 37
2.2.3.iv - Prevalent deformities as risk factors for incident deformities 38
2.2.4 - DIAGNOSIS/ IMAGING *39*
2.2.4.i - Diagnosis without imaging techniques 39
2.2.4.ii - Conventional radiographs 40
2.2.4.iii - Detecting and defining vertebral deformities 41
2.2.4.iv - Visual approach (Qualitative) 41
2.2.4.v - Morphometric methods (Quantitative) 42
2.2.4.vi - Semiquantitative methods 46
2.2.4.vii - Combination of quantitative and visual approaches 47
2.2.5 - PREVALENCE *48*
2.2.5.i - Overall estimates of prevalence 48
2.2.5.ii - Age and sex 48
2.2.5.iii - Geographic area and ethnicity 49

2.2.6 - INCIDENCE ... *49*
2.2.6.i - Secular trends ... 49
2.2.6.ii - Population-based studies ... 49
2.3 - THE USE OF MULTIPLE IMPUTATION TO MANAGE MISSING DATA ... 51
2.3.1 - INTRODUCTION ... *51*
2.3.3 - MISSING DATA MECHANISMS ... *53*
2.3.4 - APPROACHES TO DATA WITH MISSINGNESS MECHANISMS THAT ARE IGNORABLE ... *55*
2.3.4.i - Case deletion ... 55
2.3.4.ii - Deterministic imputation ... 56
2.3.4.iii - Principled imputation methods ... 57
2.3.5 - MULTIPLE IMPUTATION ... *58*
2.3.5.i - Background ... 58
2.3.5.ii - Single random imputation ... 59
2.3.5.iii - Imputation models ... 61
2.3.5.iv - The Gibbs sampler ... 62
2.3.5.v - Multiple imputation algorithm using the Gibbs sampler ... 63
TABLE 2.1 - DIFFERENTIAL DIAGNOSES OF A RADIOLOGICAL VERTEBRAL DEFORMITY ... 65
FIGURE 2.1 - EXAMPLE OF SIX-POINT PLACEMENT FOR VERTEBRAL MORPHOMETRY ... 65
TABLE 2.2 - OVERALL DEFORMITY PREVALENCE ESTIMATES FROM VARIOUS STUDIES. ... 66
FIGURE 2.2 - INCIDENCE OF VERTEBRAL DEFORMITIES FROM THE EPOS ... 67
FIGURE 2.3 - THE INCIDENCE OF VERTEBRAL DEFORMITIES FROM THE ROTTERDAM STUDY ... 67

CHAPTER 3 – METHODS ... **68**

3.1 - INTRODUCTION ... 68
3.2 - CAMOS DESIGN ... 68
3.3 - VERTEBRAL DEFORMITY DEFINITION AND ASSESSMENT ... 71
3.3.1 - INTRODUCTION ... *71*
3.3.2 - RADIOGRAPHIC ASSESSMENT ... *71*
3.3.3 - ESTIMATION OF THE NORMAL POPULATION ... *72*
3.3.4 - VERTEBRAL DEFORMITY DEFINITION ... *73*
3.3.4.i - Original Definition ... 73
3.3.4.ii - Measurement errors in x-ray film digitization ... 74
3.4 - STATISTICAL ANALYSES ... 75
3.4.1 - DESCRIPTIVE STATISTICS ... *75*
3.4.2 - BAYESIAN ESTIMATES OF UNADJUSTED INCIDENCE RATES ... *76*
3.4.3 - IMPUTATION MODELS ... *77*
3.4.4 - INCIDENCE ESTIMATES ADJUSTED BY MULTIPLE IMPUTATION ... *80*
TABLE 3.1 - ESTIMATED AND ACTUAL SAMPLE SIZE BY AGE AND SEX GROUP FOR THE CAMOS. ... 82
TABLE 3.2 - DATA COLLECTED FOR THE NONRESPONDER GROUPS ... 82
FIGURE 3.1 - CAMOS DATA COLLECTION. ... 83

CHAPTER 4 – RESULTS ... **84**

4.1 - OVERVIEW ... 84
4.2 - DEMOGRAPHICS AND RISK FACTORS ... 84
4.3 - CRUDE ESTIMATES DERIVED DIRECTLY FROM OBSERVED DATA ... 87
4.4 - ESTIMATES DERIVED BY MULTIPLE IMPUTATION ... 88
4.5 - ESTIMATES ADJUSTED FOR ATTRITION BIAS ... 89
4.6 - ESTIMATES ADJUSTED FOR NONRESPONSE AND ATTRITION BIAS ... 90
TABLE 4.1 - COMPARISON OF VARIABLES AVAILABLE AT BASELINE FOR ALL **FEMALE** CAMOS CONTACTEES AGED ≥50 YEARS ... 92
TABLE 4.2 - COMPARISON OF VARIABLES AVAILABLE AT BASELINE FOR ALL **MALE** CAMOS CONTACTEES AGED ≥50 YEARS ... 93
TABLE 4.3 – BASELINE CHARACTERISTICS OF SUBJECTS WHO SUSTAINED NO INCIDENT DEFORMITY COMPARED TO THOSE WHO SUSTAINED AT LEAST ONE INCIDENT VERTEBRAL DEFORMITY OF **GRADES 1 AND 2 ONLY.** ... 94

TABLE 4.4 – BASELINE CHARACTERISTICS OF SUBJECTS WHO SUSTAINED NO INCIDENT DEFORMITY COMPARED TO THOSE WHO SUSTAINED AT LEAST ONE INCIDENT VERTEBRAL DEFORMITY OF **GRADES ½ AND HIGHER**. 95
TABLE 4.5 – NUMBER OF MEN AND WOMEN WITH INCIDENT VERTEBRAL DEFORMITIES, **GRADES 1 AND 2 ONLY**, STRATIFIED BY PRE-EXISTING DEFORMITY STATUS AT BASELINE. 96
TABLE 4.6 – NUMBER OF MEN AND WOMEN WITH INCIDENT VERTEBRAL DEFORMITIES OF **GRADES ½ AND HIGHER**, STRATIFIED BY PRE-EXISTING DEFORMITY STATUS AT BASELINE. 96
TABLE 4.7 – BAYESIAN ESTIMATES OF AGE-SPECIFIC INCIDENCE PER 1000 PY IN **WOMEN**, STRATIFIED BY THE PRESENCE OF VERTEBRAL DEFORMITIES AT BASELINE, ONLY **GRADES 1 AND 2** 97
TABLE 4.8 – BAYESIAN ESTIMATES OF AGE-SPECIFIC INCIDENCE PER 1000 PY IN **MEN**, STRATIFIED BY THE PRESENCE OF VERTEBRAL DEFORMITIES AT BASELINE, ONLY **GRADES 1 AND 2** 98
TABLE 4.9 – BAYESIAN ESTIMATES OF AGE-SPECIFIC INCIDENCE PER 1000 PY IN **WOMEN**, STRATIFIED BY THE PRESENCE OF VERTEBRAL DEFORMITIES AT BASELINE, **GRADES ½, 1, AND 2** 99
TABLE 4.10 - BAYESIAN ESTIMATES OF AGE-SPECIFIC INCIDENCE PER 1000 PY IN **MEN**, STRATIFIED BY THE PRESENCE OF VERTEBRAL DEFORMITIES AT BASELINE. NONINFORMATIVE PRIOR DISTRIBUTIONS WERE USED FOR ALL ESTIMATES, **GRADES ½, 1, AND 2** 100
TABLE 4.11 – FREQUENCY DISTRIBUTION OF DEFORMITIES BY VERTEBRAL LEVEL, **GRADES 1 AND 2**.... 101
FIGURE 4.1 - FREQUENCY DISTRIBUTION OF DEFORMITIES BY VERTEBRAL LEVEL IN MEN AND WOMEN, **GRADES 1 AND 2**. 101
TABLE 4.12 – FREQUENCY DISTRIBUTION OF INCIDENT DEFORMITIES BY VERTEBRAL LEVEL, **GRADES ½, 1, AND 2**. 102
FIGURE 4.2– FREQUENCY DISTRIBUTION OF INCIDENT DEFORMITIES BY VERTEBRAL LEVEL IN MEN AND WOMEN, **GRADES ½, 1, AND 2.** 102
TABLE 4.13 - IMPUTATION MODELS SELECTED FOR THE CENSORED GROUPS. 103
TABLE 4.14 - IMPUTATION MODELS SELCTED FOR THE REFUSAL QUESTIONNAIRE (RQ) GROUPS. 104
TABLE 4.15 - IMPUTATION MODELS SELECTED FOR THE TOTAL REFUSER GROUPS FOR WHICH DATA ON SEX, AGE, NUMBER OF HOUSEHOLD MEMBERS WERE AVAILABLE (**TR1 GROUP**). 105
TABLE 4.16 - IMPUTATION MODELS SLECTED FOR THE TOTAL REFUSER GROUPS FOR WHICH DATA ON SEX, AGE, NUMBER OF HOUSEHOLD MEMBERS WERE NOT AVAILABLE (**TR2 GROUP**). 105
TABLE 4.17 - OBSERVED AND IMPUTED ESTIMATES OF INCIDENCE PER 1000 PY IN **WOMEN**. 106
TABLE 4.18 – OBSERVED AND IMPUTED ESTIMATES OF INCIDENCE PER 1000 PY IN **MEN** 107
TABLE 4.19 - SENSITIVITY ANALYSIS OF ESTIMATES ADJUSTED FOR ATTRITION BIAS AMONG **WOMEN** LOST TO FOLLOW-UP (CENSORED GROUP) IN THE CAMOS. 108
TABLE 4.20 - SENSITIVITY ANALYSIS OF ESTIMATES ADJUSTED FOR ATTRITION BIAS AMONG **MEN** LOST TO FOLLOW-UP (CENSORED GROUP) IN THE CAMOS 109
TABLE 4.21 - SENSITIVITY ANALYSIS OF ESTIMATES ADJUSTED FOR NONRESPONSE AND ATTRITION BIASES AMONG **FEMALE** CAMOS CONTACTEES 110
TABLE 4.22 - SENSITIVITY ANALYSIS OF ESTIMATES ADJUSTED FOR NONRESPONSE AND ATTRITION BIASES AMONG **MALE** CAMOS CONTACTEES 111

CHAPTER 5 – SUMMARY AND CONCLUSIONS **112**

5.1 - ADJUSTED INCIDENCE ESTIMATES 112
5.3 - STRENGTHS AND WEAKNESSES 114
5.4 - FUTURE RESEARCH 115

REFERENCES **117**

APPENDIX 1 - WINBUGS PROGRAM TO ESTIMATE CRUDE RATES **131**

APPENDIX 2 - WINBUGS PROGRAM TO ESTIMATE RATES STRATIFIED BY PRE-EXISTING DEFORMITY STATUS **132**

APPENDIX 3 - S-PLUS PROGRAM TO IMPUTE SEX FOR THE TR2 GROUP 134

APPENDIX 4 - S-PLUS PROGRAM TO IMPUTE AGE FOR THE TR2 GROUP 135

APPENDIX 5 - WINBUGS PROGRAM TO ESTIMATE RATES ADJUSTED BY MULITPLE IMPUTATION 136

CHAPTER 1 – INTRODUCTION

This thesis provides the first Canada-wide estimates of the incidence of vertebral deformities, based on data collected as part of the Canadian Multicentre Study of Osteoporosis (CaMos). Aspects of osteoporosis and vertebral deformities that are relevant to this thesis are only briefly introduced in this chapter, since the literature review provided in Chapter 2 will expatiate on these themes. The objectives of the thesis are presented, followed by an overview of the contents of each chapter to follow.

1.1 - Rationale

1.1.1 - OSTEOPOROSIS AS A SUBJECT OF INTEREST

During the last fifty years, the control of infectious diseases in developed countries has led to increasing lifespan, which has allowed chronic diseases and their sequelae to emerge as the leading causes of mortality and morbidity[1]. Over the last decade, it has become increasingly evident that one of the most important of these diseases is osteoporosis; it has a worldwide distribution and it may be the most prevalent chronic disease in adults[2]. Its incidence is known to increase exponentially with age[3-11]. Although its epidemiology has been extensively reviewed in the medical literature in recent years, most of what is known about osteoporosis applies to white postmenopausal women.

The pathophysiology, epidemiology, diagnostics, consequences, and management of osteoporosis are all emergent and dynamic fields of research, and as such the apprentice reader of the literature may sometimes feel beleaguered by the various controversies and lack of consensus. As a case in point, there is even some uncertainty about how to best define the disease. Perhaps the most commonly used definition, at least currently, is that osteoporosis is a systemic disease of the skeleton which is characterized by reduced bone mass and microarchitectural deterioration of bone tissue. These changes, although

asymptomatic in themselves, lead to bone fragility and susceptibility to low-trauma or atraumatic fractures, which typically occur at the thoracolumbar spine, hip, and wrist.

Because osteoporosis can be silent for long periods of time, there are three approaches that are used to describe its epidemiology. Measurements of apparent bone mineral density (BMD) play such a dominant role in clinical diagnosis, that BMD has its own reported epidemiology. But low bone mass on its own is asymptomatic and causes no morbidity, whereas its associated sequela, the fragility fracture, can cause serious debility. Therefore the prevalence, incidence, and risks of fragility fractures (especially of the hip) are often reported as barometers of osteoporosis. Less common are studies that use osteoporosis as a clinical entity for reporting.

Fragility fractures, especially hip fractures, are recognized as major contributors to morbidity[6,12-20], escalating health care costs[18-25], and as markers of mortality[6,12,14,16,18,26-29]. There has thus been much interest in improving prevention and treatment modalities. The results of several large randomized controlled trials on the efficacy of new therapeutic agents, mostly funded by pharmaceutical companies, have served to promote the empiric treatment of osteoporosis[30]. Some clinicians have voiced concern that calls for more aggressive treatment are partly driven by the interests of pharmaceutical companies[31]. However, the suggestion that cost-effective prevention and management plans should be formalized seems well supported by the expanding literature on the epidemiology of osteoporosis, BMD, and fracture risk.

1.1.2 - VERTEBRAL DEFORMITIES AS A SUBJECT OF INTEREST

Osteoporotic vertebral fractures, also called vertebral deformities, are the most common sequelae of osteoporosis. They are associated with significant morbidity, although the health and economic consequences are more insidious than those for hip fractures. Valid estimates of incidence can be used to gauge the severity and progression of osteoporosis, especially since the occurrence of a deformity is typically not confounded by falling.

Additionally, vertebral deformities are important predictors of future fractures in the spine as well as at other sites[16,18,20]. Incidence estimates may thus serve as sentinel markers for the epidemiological study of osteoporosis, and can inform current treatment needs as well as future projections for public health policies.

The detection of vertebral deformities is more difficult than other types of osteoporotic fractures, because the majority of vertebral deformities are not preceded by trauma nor are they associated with symptoms severe enough to prompt a clinical consultation. Self-reports or clinical records are therefore unreliable to accurately ascertain their occurrence.

For biomechanical reasons, when a vertebra fractures because of underlying osteoporotic disease, its geometry is almost always altered along the cranio-caudal axis. A fractured vertebra therefore loses some of its 'height' to variable degrees, which may be evident on lateral x-ray radiographs. Epidemiologic studies can therefore make use of radiographic imaging, but there are major problems associated with radiographic surveys.

There are 12 thoracic vertebrae, numbered T1 - T12, and 5 lumbar vertebrae, numbered L1 - L5, and each one of these vertebral levels has its own height distribution within a population. Unless there is a severe deformity, it can be extremely difficult to distinguish a pathological loss of height from the normal variation of a level-specific vertebra. Adding to this problem is that there is currently no consensus on how to define a normal vertebral height, or how to quantify a height loss that characterizes a deformity. To make matters worse, once a definition for vertebral deformity has been arbitrarily chosen, there is controversy about whether thoracolumbar films should be read by radiologists, automated morphometry, or a combination of both.

The CaMos is a population-based longitudinal study that has so far accumulated 5 years of follow-up data. One of its a primary objectives is to accurately estimate the incidence of vertebral deformities in Canadian men and women aged 55 years and over, using standardized imaging and ascertainment techniques. The data analyzed in this thesis have been collected at baseline and at Year 5 of the CaMos. Note should be made that not all

of the Year 5 x-ray data had been entered into the CaMos database at the time that the analyses were conducted for this thesis, so that attrition is somewhat inflated in the results presented.

When the CaMos was designed a decade ago, the state-of-the-art in radiologic surveys of vertebral deformities was automated morphometry, and this was the methodology that was chosen. Since then, expert opinion in the radiological literature has shifted towards incorporating standardized visual assessments by radiologists into ascertainment algorithms. In view of this development, some researchers may question the accuracy of estimates of incidence derived solely by morphometry, such as those from the CaMos. Nevertheless, the results presented in this thesis should be of great interest, given the current state of knowledge in this area in Canada.

1.1.3 - MULITPLE IMPUTATION AS A SUBJECT OF INTEREST

Estimating the incidence of vertebral deformities requires repeated imaging over time, and this can lower response rates and increase the attrition of the original sample. Indeed, the CaMos had a response rate of 42%, and may have been able to radiograph only about half of its cohort at Year 5. As a result, there may be nonresponse and attrition biases in the estimates of incidence.

Multiple imputation is a Bayesian statistical technique that was first described 20 years ago, which 'fills in' missing data with plausible values, and accounts for the uncertainty in predicting the missing data. With the development of new computational methods and software, multiple imputation has recently become an accessible and practical approach for obtaining valid inferences from incomplete data sets. In this thesis, estimates adjusted for attrition and nonresponse biases will be derived from the combination of the observed and imputed data.

1.2 - Objectives

The main objectives of this thesis are:

1- To estimate sex- and age-specific incidence rates of vertebral deformities from the data that are currently available from the CaMos.

2- To derive plausible estimates that are adjusted for nonresponse and attrition biases by multiple imputation, under different assumptions for the missing data mechanism.

Chapter 2 first reviews the literature on the epidemiology of osteoporosis, and then proceeds to a more specific review of issues related to vertebral deformities. A brief discussion explaining the concepts and methods fundamental to understanding multiple imputation will close this section. Chapter 3 describes the CaMos design, focusing on the methods and criteria that the study used for the ascertainment of vertebral deformities. The statistical methods that were used to analyze the CaMos data specifically for this thesis are then described in detail. The results are reported in Chapter 4, and a summary and conclusions are presented in Chapter 5.

Tables and figures are located at the end of the chapter in which they are first cited.

CHAPTER 2 – LITERATURE REVIEW

In this chapter, a literature review of the main topics concerning osteoporosis provides the setting for vertebral deformities as a subject of investigation. The epidemiology of vertebral deformities is then presented as a motivation for the CaMos design and methods. A discussion of missing data and the concepts that underlie multiple imputation informs the rationale for the statistical analyses that were chosen for this thesis.

2.1 – Epidemiology of Osteoporosis

2.1.1 - DISTRIBUTION

2.1.1.i - Prevalence & Incidence

Osteoporosis is highly prevalent in all geographic areas that have been studied so far, which have mostly been in Europe and North America[12]. Comparisons across populations are difficult, as prevalence and incidence estimates vary according to the approach taken to normalize and assess bone mass, the skeletal site assessed, and the diagnostic criteria used[32]. It is estimated that 10 million individuals in the U.S. currently have osteoporosis[33]. The most recent Canadian population-based prevalence study estimates that 19% of women and 6% of men over the age of fifty have osteoporosis[34]. White women have a one-in-six lifetime risk of fracturing a hip, compared to a one-in-nine risk of developing breast cancer[35]. Based on results from the U.S., the annual incidence of vertebral fractures is even higher than that of hip fractures in white postmenopausal women[35]. In 2001, approximately 24,000 Canadians sustained a hip fracture related to osteoporosis[19].

2.1.1.ii - Age

The incidence of osteoporosis increases exponentially after age fifty[36], and most fractures in the elderly are related to osteoporosis[37]. In Western populations, nearly three quarters of osteoporotic fractures occur after the age of sixty-five[30,38,39]. Some investigators

classify osteoporotic fractures by the biphasic distribution of their occurrence during lifetime. So-called 'early' fractures occur in midlife (age 40-50), and 'late' fractures in the elderly (> 65 years). Early fractures occur mostly in the proportion of women who develop osteopenia in the years following the menopause. Coexisting osteopenia and low-impact falls increase the probability of fracture[40]. They are mostly sustained at the thoracolumbar vertebrae, wrist, and ankle. Late fractures are more common than early fractures, and tend to occur at a less advanced age in women compared to men. Voluntary physical activity can cause them. They typically affect the proxima of the extremity bones (humerus, tibia, and femur), the pelvis, and the thoracolumbar vertebrae[41].

As with other chronic age-related conditions like atherosclerosis and osteoarthritis[6], the incidence of osteoporosis is expected to rise exponentially in the next few decades due to the ageing populations of developed nations[23] and countries undergoing industrialization such as China[42]. Worldwide, the greatest increments in hip fracture prevalence will likely occur in Asia, Africa, and Latin America, in parallel with increments in their elderly populations[43]. It is projected that the number of hip fractures that occurred in Ontario in 1990 will double by the year 2010[44].

2.1.1.iii - Secular Trends

Ageing cannot entirely explain the increases in fragility fracture incidence that have been observed in developed and developing countries[6,43,45]. For example, from 1970 to the mid-80's there appears to have been an increase in age- and sex-specific hip fracture rates of 1-3 % per year. These "secular trends" seem to have levelled off in recent years in Australia, the U.K. and the U.S., but not in developing nations. The reasons for these secular changes are unknown. Declining daily physical activity levels, the increasing frailty of the elderly, and altered early environmental factors associated with modernisation are possible explanations[4,43].

2.1.1.iv - Sex

Fragility fractures are more common and occur earlier in women. Several explanations have been suggested from the current evidence. First, women have a higher life-

expectancy than men in most populations[6]. Second, as osteoporosis progresses with age, endosteal bone is gradually lost. The resulting decreases in bone strength are partly compensated by periosteal apposition in both sexes, but this process is known to be more effective in men than in women[46,47]. Third, as men on average have larger skeletons than women, they are able to accumulate more microarchitectural damage until a fracture threshold is reached[47]. Lastly, the acute estrogen deficiency that occurs in the perimenopause is associated with a period of accelerated bone loss that is superimposed on age-related bone loss, which predisposes women to developing osteoporosis at an earlier age than men[33].

2.1.1.v - Race between Nations

Differences in fracture risk between international populations have been reported. Studies done in Northern Europe[48] and North America[49] suggest that the incidence of hip fracture is less common in Asia[50]. The hip fracture rates in native Africans are lower than in white Africans, yet the formers' BMD values are lower[9]. Comparing developed nations with each other, the prevalence of osteoporosis at the femoral neck in postmenopausal women has been reported as 21% in Sweden[51], 17% in the U.S.[52], 12% in Japan[53], and 8% in Canada[54].

Factors other than race may explain some regional differences, although they are not yet fully understood. For example, vertebral fracture prevalence is greater in Japanese women compared to Japanese-Americans living in Hawaii[55]. Limb fractures rates vary widely within different regions of Europe, which has a relatively homogeneous Caucasian population. This phenomenon has been partly attributed to regional differences in fall rates[56].

2.1.1.vi - Race within Nations

Within the U.S., black women have a lower incidence of osteoporotic fractures than white women[33], in accordance with their higher bone density values in the axial and appendicular skeleton[57]. Age-adjusted osteoporosis prevalence estimates at the femoral neck are reported as 7%, 5%, and 3% in white, black, and Hispanic American men,

respectively[32]. Vertebral fracture rates are higher in white women than in non-Hispanic women[58]. Curiously, postmenopausal Japanese-American women seem to have lower hip fracture rates than Caucasian-American women, despite having lower peak BMD measurements at the hip[33,57]. These results may be explained by international comparisons: when adjusted for height, there may be no difference in mean BMD between Chinese, Indian, and European women of the same age[32]. A sedentary lifestyle has been suggested to explain the higher hip fracture rates observed in Canadian Natives compared to Caucasians[45].

2.1.2 - MORBIDITY

Sustaining an osteoporotic fracture can cause serious functional impairment in elderly women[18] and men[16]. Although relatively rare, some forms of osteoporosis such as osteogenesis imperfecta can also have devastating consequences in children[24].

2.1.2.i - Hip Fractures

A recent Canadian prospective cohort reported that only 59% of patients were residing back in the community one year following hip fracture[19]. In the U.S., an estimated 7% of white women with fractures of the hip, spine, or distal forearm become dependent, or experience increased dependence in the basic activities of daily living one year after the event[13]. Of these, hip fracture is the most important cause of disability, and almost all require acute hospitalization[6]. Forty to 79% of patients sustaining a hip fracture only regain their prior ambulation function after 12 months or more, with older people having a worse prognosis[28]. Co-morbidity is strongly associated with hip fractures. The odds ratio for hospitalization due to non-orthopaedic diseases is higher in the years preceding a first hip fracture[15].

2.1.2.i - Vertebral Fractures

The majority of vertebral fractures may be asymptomatic[59], or at least their symptomology does not prompt clinical consultation or the ordering of radiographs of the

spine. Only approximately $^1/_3$ of vertebral deformities are severe enough to be recognized clinically[28]. New vertebral fractures may nevertheless cause severe back pain for several weeks, kyphosis, height loss, and may evolve into chronic back pain[18].

2.1.2.ii - Wrist Fractures

The acute and chronic consequences of wrist fractures on activities of daily living are more specific and less disabling than hip or severe vertebral fractures. Surprisingly, hand pain and weakness are present in roughly one $^1/_3$ of elderly patients in the 6-10 years following a wrist fracture. Algodystrophy (hand pain, limited finger movement, and vasomotor instability) may be an important sequela of wrist fracture, although accurate prevalence estimates have been impeded by different definitions of this syndrome[28].

2.1.3 - MORTALITY

Prospective studies done in Canada and the U.S. have shown that proximal femur[29] and vertebral fractures[14] are associated with decreased survival rates in postmenopausal women. As many as 22% of Canadians die in the first year following a hip fracture[19]. Although hip fractures have the highest excess mortality of all fracture types, mortality during the acute hospitalization for hip fracture is uncommon[28]. A prospective cohort study done in Australia found that other major osteoporotic fractures (pelvis, distal femur, proximal tibia, multiple ribs, proximal humerus) sustained at, or after age 60, are also associated with increased mortality[27]. It is doubtful, however, that any of these fractures are major direct contributors to excess mortality. They may be markers for comorbid conditions that increase the risk of death[28], such as cardiovascular disease and cancer[14]. It has been shown that low bone density predicts higher mortality independent of fracture, supporting the concept that the association between mortality and fragility fractures reflects the underlying health status of an individual[26]. Fractures of the wrist and ankle are not associated with altered survival, at least not in women[28].

2.1.4 - COST

Concomitant with the rising prevalence of osteoporosis, the economic costs associated with diagnosis and treatment are projected to increase at alarming rates. In Canada, the total cost of acute care related to osteoporosis, which includes admission to hospital, outpatient care, and drug therapy, was estimated at over 1.3 billion Canadian dollars (CAD) in 1993[20]. Cost projections made in other regions, such as Europe[21], the United Kingdom[18] and the U.S.[22], have also estimated enormous annual costs associated with the treatment of osteoporosis, expenditures that could eventually outstrip resources.

Hip fractures account for the majority of annual osteoporosis-related expenditures, although all types of fractures have the potential to cause disability and incur substantial costs[18]. The Canadian expenditures associated with hip fractures alone are currently estimated to be 650 million CAD per year. A conservative estimate of the mean cost of a hip fracture over one year is 26 500 CAD. If current trends continue, these annual costs are expected to reach 2.2 billion CAD by the year 2041[19] .

2.1.5 – PATHOGENESIS OF OSTEOPOROSIS

Research of the pathogenesis of osteoporosis is a developing field that is both vast and complex. The structural and biomechanical components responsible for bone fragility are not fully understood. Understanding the literature on bone fragility is additionally hampered by variable definitions, terminology, and approaches to concepts. For purposes of clarity and brevity, the approach used by Frost[41] to explain the physical component of fragility will be described here. The role of remodeling is only briefly summarized, and regulation at the cellular and molecular levels is omitted from this review.

2.1.5.i - Bone strength

The major determinants of the increased fracture risk associated with osteoporosis are bone strength and propensity to fall. This review will cover only the characteristics of

bone believed to be determinants of bone strength. There are four physical determinants: (1) the amount and kind of bone present (bone mass); (2) the shape and size of the bone; (3) the mechanical properties of the bone tissue, such as trabecular connectivity; and (4) the amount of microfatigue in a bone[41]. Some investigators prefer to group the last three factors together as "bone quality"[33], a term that is increasingly seen in the literature.

Frost proposes that a healthy bone is one that has enough strength to keep voluntary loads (*i.e.* intentional) from causing nontraumatic fractures. Accordingly, an interplay between these four physical determinants is thought to maintain whole-bone strength[41].

The exact pathogenesis of bone fragility leading to increased fracture risk is still being researched. *In vivo*, it seems likely that a compromise of a nonspecific combination of the physical factors just described, rather than a single one such as bone mass, is necessary to reach the threshold for nontraumatic fractures[33,46,60,61]. An ideal clinical evaluation of whole-bone strength would therefore, in theory, measure both bone mass and bone quality.

2.1.5.ii - Modeling/Remodeling

Osteoporosis results from the progressive loss of bone strength that occurs with increasing age[3,62,63], and from extrinsic and intrinsic factors that may exaggerate this process. These changes may be superimposed on a low peak bone mass[3]. Age-related loss of bone mass and quality is linked with the impairment of a physiological process called bone *remodeling*. The brief description normal of bone physiology that follows may clarify the subsequent discussion on the pathophysiology of osteoporosis.

The initial formation of the skeleton is achieved by the direct apposition of new bone, and is called *modeling*. Another process, remodeling, is carried out by temporary anatomic structures called basic multicellular units (BMUs) throughout life. The BMUs continuously excavate and replace cortical and trabecular bone, which repairs microdamage, maintains skeletal strength, and plays a part in calcium homeostasis[41,63]. It is hypothesized that a 'mechanostat' feedback system ensures that the greater a bone is

loaded over time, the stronger it becomes by the remodeling process (up to a limit), and *vice versa*[41] .

Remodeling of the skeleton begins in early fetal life, and becomes the dominant metabolic activity of the skeleton by the end of puberty. The gain in bone mass that is achieved during childhood and adolescence is associated with large and fluctuating rates of bone remodeling. Once peak bone mass is achieved in young adulthood, BMUs resorb as much as they make bone until mid-life (age 40-50)[62,63].

For reasons as yet unclear, BMU-based remodeling at midlife goes into a negative balance mode, by the 'uncoupling' of bone resorption and formation. Bone mass loss may occur because of accelerated bone resorption and/or because bone formation fails to keep pace with accelerated resorption[47,63]. A slow, continuous loss of endosteal and trabecular bone thus results[62] at an average rate of about 1% per year in women, and 0.5% per year in men[64,65]. This slow phase continues through the remainder of life, but at age- and site-specific rates. The abnormal remodeling reduces trabecular number and connectivity and increases porosity[47,66]. In addition, older bone is replaced with younger, less mineralized bone which is less resistant to mechanical stress. Microdamage thus accumulates and is less efficiently repaired[47]. Studies of postmenopausal women suggest that calcium deficiency and the resulting secondary hyperparathyroidism are the predominant causes. The decrease in renal calcium conservation and calcium absorption that occur in the elderly (and perhaps vitamin D deficiency), increase parathyroid hormone levels which may drive most of the bone resorption. The slow phase of bone loss may be responsible for losses of 20-30% in both trabecular and cortical bone. Not surprisingly, the 'late' osteoporotic fracture sites (proximal long bones, pelvis, vertebrae) are rich in both types of bones, except the vertebrae which are mostly comprised of trabecular bone[62-64].

In women, there is also an accelerated and transient phase of bone loss that is superimposed on the slow phase. It is caused by the estrogen deficiency that originates during the perimenopause and lasts about 10-15 years subsequently, after which the slow continuous rate resumes[62]. There is, however, considerable individual variation as to the

rate, timing, and duration of bone loss[67]. On average, this accelerated phase accounts for further trabecular bone losses of 20-30%, and cortical bone losses of 5-10%. In about 10-20% of women, this accelerated phase of bone loss clinically manifests as 'early' fractures at sites rich in trabecular bone, such as the vertebrae, distal forearm, and ankle[62-64]. Partly because of the absence of an accelerated phase in men, age-adjusted overall bone loss is only about $^2/_3$ of that seen in women[62].

Even though osteoporosis is a systemic disease, patterns of bone loss differ by skeletal site[32]. It can be asymptomatic for several decades, until bone strength is sufficiently compromised. Indeed, osteoporosis tends to be recognized only late in its clinical course by the occurrence of fragility fractures[30]. Over a lifetime, fractures occur most commonly in the thoracolumbar vertebrae, the proximal femur, and the distal forearm. Curiously, the cervical spine is rarely fractured[4,7].

The mechanics of falls and their incidence aside, fragility fractures may be the consequence of a positive feedback between microdamage (reduced bone quality) and the reduced bone density caused by increased remodeling[66]. Fractures may result from minimal trauma such as a fall from a standing position or less, or even activities of daily living[30]. It is estimated that over the age of 65, about 90-98% of hip fractures result from a fall of less than or equal to standing height[68].

2.1.5.iii - Secondary osteoporosis

Apart from the uncoupled remodeling processes that are associated with menopause and ageing, the reduction of bone strength can result from specific medical and iatrogenic disorders. It may be useful in the clinical setting to view this type of osteoporotic disease as 'secondary osteoporosis', because the underlying causes or comorbidity can sometimes be treated[61]. According to the Osteoporosis Society of Canada (OSC), the most important determinants are: an exposure to systemic glucocorticosteroids longer than 3 months, propensity of falling (neuromuscular disorders, decreased visual depth and acuity), primary hyperparathyroidism, malabsorption syndromes (*e.g.* celiac disease, cystic fibrosis, inflammatory bowel disease), and hypogonadism[30].

2.1.6 – DEFINING OSTEOPOROSIS

2.1.6.i - Osteoporosis as an evolving construct

In the 1990's, the definition of osteoporosis underwent revision with the emerging knowledge of its pathophysiology indicating that osteoporosis is a systemic disease, as opposed to a disease of bone tissue. Even so, definitions remain vague and descriptive because the cause of osteoporosis has not been found at the molecular level[69]. Clinicians, who are mainly interested in preventing fractures, generally use diagnostic criteria that try to predict fracture risk, whereas the main emphasis of definitions from academic sources is on aspects of the bone. There seems, however, to be uniform consensus in the literature that a definition of osteoporosis should include reference to generalized skeletal fragility, in which bone strength is sufficiently weak that fractures may occur with minimal trauma[30,33,61,69]. Skeletal fragility is considered the consequence of multiple genetic, physical, hormonal, and nutritional factors acting alone or in concert to diminish skeletal integrity[61].

There is no direct measure of whole-bone strength *in vivo* currently available; neither bone mass nor bone quality can directly be measured. The only clinical index of bone quality at present is a patient's history of fragility fracture[30]. Quantitative ultrasound may measure aspects of bone quality, but its usefulness is still under investigation[67]. Bone mass can be approximated by estimating bone density with non-invasive imaging techniques, such as dual-energy X-ray absorptiometry (DXA). However, BMD values derived from DXA are themselves only approximations of density, as they measure *apparent* bone mineral density: BMD values represent the bone mineral content that is contained not only in bone, but also in the marrow and other surrounding tissues, and values are derived per *area* scanned rather than per volume. Routine assessments therefore do not consider confounders such as bone size and geometry. Because of these problems, opinions differ as to what the diagnostic criteria for osteoporosis should be.

2.1.6.ii - The WHO criteria

In 1994, a WHO expert panel operationally defined osteoporosis based exclusively on BMD values as measured by DXA. The number of standard deviations (SD) away from the mean of a normal young adult population of the same sex and race is called a "T-score". A cut-off T-score of -2.5 was suggested as diagnostic of osteoporosis. Osteopenia was defined as T-scores between -1 to -2.5 SD below the young adult mean. Severe osteoporosis was described patients with a T-score below -2.5 in the presence of a fragility fracture[70].

The rationale for using BMD values as surrogates for bone strength is that decreased BMD is roughly indicative of decreased bone strength and increasing susceptibility to fracture[64]. The age-specific risk of fracture continuously and progressively increases as BMD declines[71]. A reduction in BMD of 1 standard deviation from the mean value for an age-specific population corresponds to a two- to three-fold increase in long-term fracture risk, depending on the site measured[61,71,72]. One comprehensive review of the literature has concluded that a BMD value measured at a particular site is the best predictor of fracture risk for that site[71]. Some determinants of bone quality are also strongly correlated with BMD estimates[61,73]. BMD measurements thus serve as diagnostic indices in terms of fracture risk, much as blood pressure predicts risk for stroke[36].

There are several limitations to the WHO approach. First, there is no threshold of BMD that reduces the fracture risk to zero[74]. There is substantial overlap in the distribution of BMD values in patients with and without osteoporotic fractures, so that any T-score cut-off does not predict the presence of osteoporotic fractures[61]. Whenever a fragility fracture is present, osteoporosis should be diagnosed, regardless of BMD measurement values[75]. Second, age must always be considered along with BMD measurement in the determination of fracture risk. A low BMD in a 65 year old is predictive of a high 10-year fracture risk; but a 25 year old with a low BMD has a very low 10-year risk of fracture, that is not substantially greater than a 25 year old with a high BMD[71]. Third, there is no evidence yet that a -2.5 cut-off is appropriate for predicting fracture risk in men or different ethnic groups. Fourth, a low bone density offers no information regarding peak

bone mass attained, the amount of bone that has been lost, or the quality of bone that remains in an individual. Finally, BMD values vary with skeletal sites, so that different sites may yield different conclusions about whether an individual has osteoporosis[61].

The intended use for the WHO classification was to enable the regional comparison of epidemiologic data concerning osteoporosis. The reference ranges were not specifically developed for clinical decision-making in the treatment of individual patients. Nevertheless, the WHO classification system is commonly used as diagnostic criteria. The OSC recommends its conditional use with established norms for peak BMD values of different populations (sex- and race-matched). The OSC also recommends that certain risk factors, especially a history of fragility fracture, be considered in addition to BMD values to establish a diagnosis[30].

2.1.6.iii - Other definitions

A formal definition of osteoporosis which is often cited was elucidated at a 1993 Consensus Development Conference[12]: "Osteoporosis is a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, with a resultant increase in fragility and risk of fracture". To emphasize the central role of bone strength (as opposed to bone mass), this definition was modified by a National Institute of Health consensus in 2000 as: "a skeletal disorder characterized by compromised bone strength predisposing a person to an increased risk of fracture. Bone strength reflects the integration of 2 main factors: bone density and bone quality"[33].

Some experts still recommend that osteoporosis be diagnosed only when a fragility fracture has occurred. Frost, for example, emphasizes that it is the interdependence between whole bone strength and loading that prevents nontraumatic fractures. As long as these fractures are avoided, he claims, then the bone is considered non-osteoporotic, even if severe osteopenia is present[40]. The obvious disadvantage with this definition is that diagnosis could be delayed, when prevention is the optimum treatment.

2.1.7 – RISK FACTORS

Given that there are no symptoms prior to an osteoporotic fracture, the recognition of risk factors plays an important role in identifying those at risk, if used in conjunction with estimations of bone mass[74]. In the literature, the data on the clinical risk factors that are associated with osteoporosis are reported either as risk factors for low BMD, or risk factors for fracture. The risk factors in these two categories are not mutually exclusive. A review of this topic is complicated by some contradictory evidence emerging from epidemiologic data. A great deal more research is needed before clinical risk factors alone can identify patients at high risk for osteoporosis[32]. What follows is a brief summary of the risk factors most commonly mentioned as important by different reviewers.

2.1.7.i - Risk factors for low BMD

Studies of twins indicate that 50-80% of the variance of peak BMD measured at the femoral neck and lumbar spine in white women may be genetically determined[69,76]. However, these may be overestimates due to the phenotypic similarity of monozygotic twins that are typically used in genetic studies[69]. Moreover, the influence of heredity in age-related bone loss is perhaps diluted by environmental influences[77]. Programming (the interplay of genetic and environmental factors during development) of several metabolic and endocrine systems may explain an observed correlation between weight in infancy and adult BMD measurements of the hip and spine[2,60]. One expert has said that "osteoporosis is a paediatric disease with geriatric consequences"[78], in reference to the importance of genetics and nutrition during childhood and adolescence towards the achievement of peak bone mass[33, 60].

As mentioned earlier, the role of estrogen deficiency is critical in the transient accelerated phase of bone loss, but it also plays a role in the slow continuous phase. Women who are 20-35 years postmenopausal still have a substantial antiresorptive response to estrogen[67].

Maternal history of hip fracture has been found to be a consistent risk factor for low BMD values at the distal radius, lumbar spine, and femoral neck. Positive associations between

BMD and later age at menopause, calcium intake, grip strength, and greater height and weight have also been observed at different sites. Many other risk factors have been postulated, but the evidence is not consistent[32].

2.1.7.ii - Risk factors for fracture

Fracture risk is an extremely broad topic with numerous identified predictors, but few genetic ones[77]. After a review of the literature, the OSC in 2002 identified four key predictors of future fracture in men and women: age > 65, low BMD, prior fragility fracture, and a family history of osteoporotic fracture in first degree relatives of either sex (especially maternal hip fracture)[30]. Sex is not overtly stated as a risk factor by the OSC, but its importance is implicit in that risk factors and treatment guidelines are reported by sex.

Smoking history and caffeine intake moderately increase the risk of fracture. Physical activity is generally protective against fractures. Fragility fractures are also associated with the numerous risk factors for falling. These include lean body weight, dementia, reduced visual acuity and depth perception, and the use of long-acting benzodiazepines[32]. There is no strong relationship between calcium intake and fractures beyond adolescence[74], although calcium and vitamin D deficiency may increase fracture rates in elderly women[67]. Body weight does not independently affect fracture risk from its effect on bone mass[74].

The risk factors for fracture have a cumulative effect if concomitantly present. For example, a combined low BMD and a history of a prior fragility fracture may increase the risk of subsequent factor by more than 7-fold, compared to having only one of these risk factors[79].

Of all these risk factors, a history of an osteoporotic fracture is the single most important factor in evaluating fracture risk in postmenopausal women[80]. Even fractures that are premenopausal but not related to motor vehicle accidents, although they have low incidence, are predictive of future postmenopausal fractures in white women[80,81]. A

positive history of a fracture since age 40 increases future fracture risk twofold, similar to a 1 standard deviation further decline in BMD[74]. Peri- and postmenopausal women with prior nonvertebral fractures have at least twice the risk of sustaining an incident fracture, compared with those without prior fractures[80]. The magnitude of the increased risk for subsequent fracture varies depending on the number of prior fractures, the site of the incident fracture, and the age at assessment[30,37]. Vertebral fractures have been best studied in this regard. A vertebral fracture increases the risk of a subsequent vertebral fracture at least 4-fold, and there is also an increased risk of fracture at other sites[30].

2.1.8 – PREVENTION AND TREATMENT

Osteoporosis is reportedly underdiagnosed and undertreated in Canada[82-84] and in other countries, such as Israel[85], the United Kingdom[86,87], and the U.S.[88]. Increasing public and clinician awareness has shifted the care of the osteoporotic patient from specialists to family practitioners, who may lack experience in managing osteoporosis[11]. The numerous new agents that are now available have also led to the need for up-to-date clinical guidelines. There have been several recent summaries of evidenced-based treatment[65,67,89-91], and working guidelines for Canadian clinicians using current evidence have been published by the OSC[30]. A brief summary of these guidelines is presented.

Medically, the prevention and treatment of osteoporosis are essentially the same. The clinician's aim is early detection of those at risk for osteoporotic fractures, and it ideally involves the retardation of bone loss before fractures occur[30,65,67]. Except for the acute orthopaedic management of fractures, the main modality for the treatment of osteoporosis is chemotherapeutic. It is necessary to assess the metabolic status of a patient before deciding whether conservative measures suffice (calcium and vitamin D_3 supplementation, exercise and fall prevention), or if there is indication for adjunct pharmacologic therapy[31,67].

There are several classes of drugs that are currently available in Canada for the prevention and treatment of osteoporosis, but only a modest reduction in the risk of fractures generally occurs within the first year of treatment[30,93,94]. All of these drugs have antiresorptive effects on bone, meaning that they increase the density of bone by inhibiting bone remodelling. Partly because these agents do not promote either the formation of new bone or the repair of osteoporotic microarchitectural damage, there remains an important residual risk for fracture in patients undergoing treatment.

The OSC[30] currently recommends the use of bisphosphonates in the prevention and first line treatment of osteoporosis in men and postmenopausal women. Raloxifene, a selective estrogen-receptor modulator, is also recommended as first line preventive and therapeutic agent, but only in postmenopausal women. Bisphosphonates and raloxifene have both been found to increase BMD at the femoral neck and spine as well as decrease osteoporotic fracture rates. Nasal calcitonin is suggested only as a second-line treatment in postmenopausal women, as the evidence is weaker regarding its ability to prevent fragility fractures. It is advocated as a first-line treatment of men with acute vertebral fractures and may also be useful in preventing fractures in premenopausal women. Hormone replacement therapy is well known to increase BMD and reduce fracture risk, but its role has recently been put into question with reported associations with breast cancer, coronary heart disease, deep venous thrombosis, and stroke[92]. Although not yet approved in Canada, a synthetic analogue of parathyroid hormone, hPTH(1-34), is promising as another first-line treatment. It is the only agent known to increase bone volume as well as bone density.

2.2 – Epidemiology of Vertebral Deformities

2.2.1 - INTRODUCTION

2.2.1.i - Terminology

Before discussing the epidemiology of vertebral deformities, a note on the nomenclature used in this thesis is necessary. The term vertebral *fracture* will denote an osteoporotic deformation that is detected after a patient presents to a clinician for symptoms related to the fracture. A vertebral *deformity* will be used as a more general term that refers to both symptomatic and clinically silent vertebral fractures, as in the context of a radiographic survey. In the medical literature the term 'vertebral fracture' tends to be used indiscriminately. Even more confusing is that radiologists prefer to use 'vertebral fracture' to designate radiologic evidence of a traumatic vertebral event of any cause, and the term 'deformity' to indicate radiologic changes to a vertebra that are caused by several different metabolic pathologies; either term is used by radiologists to describe an osteoporotic vertebral event[95].

Various morphologic descriptive terms are also used to describe vertebral deformities: *compression* (or interchangeably *wedge*), *biconcave* (or *endplate*), and *burst* (or *crush*) fractures. These terms are sometimes used arbitrarily in the literature. Results from the CaMos have shown that the wedge deformity is the most common type in men and women[96]. However, morphologic terms have been criticized as being unstandardized and of uncertain clinical relevance[95,97,98]. Most deformities contain a combination of morphologic features from all three types, and these features are influenced by the local biomechanics of the spinal level involved[95]. Moreover, deformities are often rapidly progressive: a compression fracture may evolve into a burst fracture within days[99]. As a result, modern and validated radiographic methods evaluating the spine do not classify vertebral deformities based on morphological types[100].

2.2.1.ii - Particularities of deformities as osteoporotic fractures

Vertebral deformities are the most prevalent type of osteoporotic fracture[101]. They are so common among older women that some clinicians may accept them as an inevitable consequence of ageing that does not necessarily require treatment. Even when a vertebral deformity is noted incidentally on an x-ray, radiologists infrequently note it in their reports[102].

Vertebral deformities have several distinct features that set them apart from other osteoporotic fractures. Most osteoporotic fractures at sites other than the spine are events that are readily observed and symptomatic, occur after some form of trauma, and should always be orthopaedically treated to restore any abnormal geometry of bone. In contrast, vertebral deformities may be difficult to detect (especially at their onset), are frequently progressive, are often asymptomatic, usually occur without trauma, and orthopaedic management is usually unnecessary even though bone geometry can be considerably altered[103].

The epidemiology of thoracolumbar deformities has not yet been thoroughly described, especially in men, and there are several reasons for this. First, the majority of deformities are not clinically diagnosed. Only approximately ¼ to ⅓ are associated with symptoms that prompt a clinical consultation[14,105]. Some studies include only clinically diagnosed vertebral fractures, whether reporting estimates of prevalence and incidence, risk factors, or clinical outcomes[20]. However, radiographic surveys of the general population are required to generate valid estimates of prevalence and incidence, although the exact time of occurrence of morphometric vertebral deformities is usually unknown[37]. One of the goals of the CaMos is to conduct such a survey on an essentially random sample of the Canadian population. Second, minor degrees of collapse of the vertebral endplate can be difficult to detect, although these deformities can rapidly progress within days or weeks[99]. Third, and most importantly, there is a lack of a universal consensus as to the definition of vertebral deformities from lateral thoracolumbar radiographs, and the radiologic techniques that should be used to assess them[100,105]. Past studies have used different definitions as well as different radiographic ascertainment methods, making

epidemiologic comparisons difficult. However, recent large population studies in Europe and Canada, such as the CaMos, have used definitions and standardized assessment techniques that are more similar across studies, and this has permitted more accurate comparisons of prevalence and incidence estimates[4].

Valid estimates of prevalence and incidence are greatly needed, as they can be used to gauge the severity and progression of osteoporosis. Vertebral deformities have been synonymous with osteoporosis since its earliest description as a metabolic disorder[106]. Knowledge of their epidemiology can also be useful to understand trends in the prevalence and incidence of osteoporosis, since their occurrence is typically not confounded by falling[4]. Vertebral deformities may therefore serve as sentinel markers for the epidemiological study of osteoporosis[12,107].

Vertebral deformities predict future vertebral and nonvertebral fractures more than any other type of osteoporotic fracture[37,108]. This may be because most spine deformities are thought to occur spontaneously and are attributed to bone fragility, whereas nonspinal fragility fractures are mostly associated with falls and other minor trauma[37]. Therefore patients with vertebral deformities experience the greatest reduction in the risk of incident vertebral and hip fractures from pharmacologic therapies[93,94,109,110]. They are often the benchmark by which the efficacy of new osteoporotic treatments is established, since prevalent vertebral deformities are relatively common and increase the likelihood of substantial numbers of incident vertebral deformities[93]. The presence of vertebral deformities can also guide the clinician on how aggressive treatment should be[30].

2.2.2 - PATHOGENESIS AND BIOMECHANICS OF VERTEBRAL DEFORMITIES

There is a wide variation in the presentation and progression of vertebral deformities. In general, the earliest vertebral deformities occur in the upper thoracic spine, and many of these are asymptomatic[99]. Progressive collapse of multiple vertebrae in this area can lead to substantial dorsal kyphosis, sometimes and regrettably referred to as a 'dowager's

hump'. As a whole, vertebral deformities occur most frequently at areas of the spine with the greatest biomechanical vulnerabilities: where the normal dorsal kyphosis is most prominent (T6-T8), and at the thoracolumbar junction (T12-L1), where the rigid thoracic spine meets the more flexible lumbar segment[96,99,14]. Prevalence estimates from the CaMos have shown that women tend to have more severe deformities than men[96]. Vertebral deformities can evolve rapidly, where an initial radiograph is barely abnormal, and progress within days or weeks to a vertebral collapse[99]. They are thought to be irreversible[111].

The normal vertebral body is characterized by a central core of trabecular bone surrounded by a thin covering of condensed trabecular bone (the 'cortical shell'). The cortical shell accounts for only 10-30% of vertebral strength in healthy individuals[99]. As mentioned in Section 2.1.5, the uncoupled remodeling associated with age decreases bone strength by reducing bone mass and compromising bone quality[47,64]. Trabecular number, thickness, and connectivity all decline in the central core along with decreasing density, whereas trabecular separation and anisotropy increases[112]. Due to the reduced bone mass and quality of the central core, the osteoporotic cortical shell bears 50-90% of weight loads[99], thus further increasing bone fragility.

It is likely that both bone fragility and skeletal loading play important roles in the etiology of vertebral deformities, although this has not been as clearly demonstrated as in hip fractures[112]. As in the other osteoporotic fractures, vertebral deformities are ultimately due to the structural failure of bone, which occurs when the load-bearing capacity of a vertebra is exceeded. Unlike most other skeletal sites, loads are applied to the spine during nearly all daily physical activities[112], and most vertebral fractures probably occur under axial compression[99]. A U.S. population-based study found that approximately 50% of clinically diagnosed vertebral fractures in people aged 60 and over were associated with a loading activity, such as lifting a heavy object (9%) or falling (33%); the remainder were either spontaneous (33%) or diagnosed incidentally (16%)[14]. A 65 year old woman whose spine BMD is one SD below the mean for her age (the WHO definition of osteopenia) is at high risk for fracture by lifting a 15 kg object[113]. In the WHO

osteoporotic range for BMD, vertebral deformities may result during simple daily activities such as tying a shoe, or opening a window[112].

Once a vertebral deformity has occurred, the risk of future deformities greatly increases[37]. It is likely that part of this risk results from local changes in load bearing due to the initial fracture[74]. The development of biomechanical models to estimate the load on the spine during controlled activities could help in eliciting prevention strategies directed at reducing loads in the elderly[112].

2.2.3 - CLINICAL PRESENTATION AND RELEVANCE

2.2.3.i - Acute presentation

The acute presentation of osteoporotic vertebral fractures varies widely. The most common symptom is back pain at the level of the involved vertebra, which may radiate in a radicular distribution[114]. The pain severity can be minimal, or completely debilitating for several weeks or months[115]. In other cases, individuals have a reduced ability to perform daily household and self-care tasks, such as cooking, vacuuming, bathing, and dressing[116]. Most of the vertebral fractures that are associated with back pain and disability occur at the lower thoracolumbar levels[95,117]. The reasons for the wide variability in symptoms are unknown.

About ½ of women with vertebral deformities do not report having had any back pain, and about $^{2}/_{3}$ do not seek medical attention[37]. Approximately 25% of vertebral deformities cause back pain that is too mild to prompt a search for a fracture by x-ray. A decade ago in the U.S., only about $^{1}/_{3}$ of vertebral deformities were diagnosed at the time that they occurred, and less than 10% require admission to hospital[14,118]. Between 2% and 10% of all patients with vertebral fractures require hospitalization in Europe and the U.S.[115,119]. One U.S. population-based study estimated that 23% of the elderly who sustain vertebral fractures need to be hospitalized[14]. A Canadian case-control study found that those

patients hospitalized for a vertebral fracture require a mean length of stay in hospital of at least 5 days, after adjusting for comorbid conditions[120].

On average, women with vertebral fractures have an increased number of days spent in bed and days of limited activity, as well as lengthier hospital admissions than women without fractures[120,121]. These associations hold for radiographic vertebral deformities, although to a lesser extent[121]. The relative impact of vertebral fractures increases with age, so that functional impairment becomes similar to that seen with hip fracture in elderly women[122]. Prevalent vertebral deformities predict decreased pulmonary capacity and increased risk of pulmonary death[17,123,124]. Severely painful vertebral fractures may lead to a cascade of impaired mobilization, muscle weakness, accelerated bone loss, and further frailty[101].

2.2.3.ii - Chronic sequelae

Long-term consequences also vary widely and can persist for at least several years[115]. After adjusting for comorbid conditions, chronic pain and disability among elderly patients with vertebral fractures remain substantially greater on average than among people without osteoporotic fractures[115,116]. On average, physical function is impaired among people with vertebral deformities, whether or not they currently report back pain. Functional impairment of varying degrees is present in 30-50% of women who sustain clinically diagnosed vertebral fractures during their lifetime[115]. The severity and number of vertebral deformities both correlate with decreased general functioning, but only severe degrees of deformity are strongly associated with substantial functional impairment[126,127]. Other chronic sequelae of vertebral deformities are postural deformity, abdominal crowding, altered body image, decreased quality of life, social withdrawal, and depression[116].

Longitudinal studies suggest that most vertebral fractures are associated with pain and functional impairment only in the first 4 years after their occurrence, and are unrelated to functional limitation thereafter[116,125]. The number of deformities in this time period predicts the extent of functional impairment[116]. Overall, cross-sectional radiographic

surveys have also found that back pain and disabilities are more severe when the degree of the deformity is high[20]. However, the number and severity of vertebral deformities may not entirely predict the clinical outcome. In a study of hospitalized patients diagnosed with vertebral fractures, only moderate associations were found between the number and severity of vertebral deformities, pain intensity, and functional impairments[128].

Elderly patients with prevalent vertebral deformities may experience reduced pulmonary function, due to thoracic spinal deformation[17,124]. Of 132 women referred to a Canadian osteoporosis clinic, forced lung vital capacity declined by 10% for each prevalent thoracic anterior wedge deformity[17]. These small physiologic changes may be of clinical importance when cardiopulmonary comorbidity is present[124].

A decrease in health-related quality of life seems to be only moderately associated with the presence of vertebral deformities[129]. One recent study found that a clinically identified vertebral fracture alone is not associated with quality-adjusted life years, a measure of cost-effectiveness[130]. One investigator speculated that the cumulative impact of vertebral deformities on quality of life may surpass that of hip fractures, because hip fractures are less frequent and occur later in life[115]. However, a recent study that used the CaMos data was unable to find a cumulative impact on SF-36 scores in women with multiple deformities, perhaps due to a lack of precision in the estimates[129]. It may also be that more specific measurement tools, such as the The Osteoporosis Quality of Life Questionnaire, are needed to detect the impact of vertebral deformities more accurately.

2.2.3.iii - Mortality

Incident vertebral deformities are rarely directly fatal. A recent prospective cohort reported that incident vertebral deformities do not predict mortality independent of frailty (*e.g.* weight loss, decreased physical function) in older women[131]. However, the prevalence of both fractures and deformities are associated with an increased mortality of 1.5- to 2.5-fold in men and women[14,59,111,123], despite controlling for various demographic variables and comorbid conditions. Men seem to be slightly more at risk than women[59], although women tend to have a higher prevalence of severe deformities than men[96].

Moreover, mortality rises with increasing numbers of fractures[123]. In some rare cases, severe vertebral deformities and kyphosis explain the excess mortality from pulmonary causes[159,123]. Otherwise the mechanism by which vertebral deformities influence death rates remains unclear[106]. Interestingly, vertebral deformities in women have been found to be associated with subsequent death from cancer, after adjustment for confounders such as age and comorbidity[59,123]. Kado *et al* have suggested that the underlying pathogenesis of vertebral deformities and neoplastic disease share common factors; bisphosphonates, well known to reduce the incidence of vertebral deformities, also decrease the number of bone metastases and mortality in women with breast cancer[123]. Future death from cardiovascular disease is also associated with prevalent vertebral deformities in men[59].

2.2.3.iv - Prevalent deformities as risk factors for incident deformities

Osteoporotic vertebral deformities are clinically relevant even if they are asymptomatic. Their presence (as with other fracture types) serves as markers for severe osteoporosis, and their direct effect on the spine increases the risk of incident vertebral deformities[112]. Several population studies have reported that women are 2-5 times more likely to experience subsequent vertebral deformities for each prevalent vertebral deformity[37,132]. The risk for incident deformities remains strong even after adjusting for BMD[106,133]. The combination of a prevalent deformity and low BMD yields a stronger prediction of fracture risk than having either risk factor alone. An individual with a normal BMD and a vertebral fracture is at slightly greater risk than an individual with low BMD but no fractures[134]. Given a history of a prior osteoporotic fracture, the type of fracture that is the strongest risk for an incident fracture is a vertebral deformity[37]. A prior vertebral deformity is a strong risk factor for hip fracture as well as for other osteoporotic fractures in men as well as pre- and postmenopausal women[37,59]. Larger deformities (*i.e.* those with a more pronounced loss of vertebral height) convey a greater increase in risk of future fracture[106]. Despite the importance of the vertebral deformity as a risk factor, it is relatively uncommon that plain x-ray radiographs are taken in the course of evaluating a patient for osteoporosis[95].

2.2.4 - DIAGNOSIS/ IMAGING

2.2.4.i - Diagnosis without imaging techniques

The clinical diagnosis of a vertebral fracture without resource to imaging techniques is very difficult, even for experienced clinicians[95,132]. Less than 1% of back pain episodes are related to vertebral fractures[135]. There are clinical signs that are indicative of vertebral deformities, but their measurements are difficult to standardize and have low predictive values.

Height loss can be a sign of vertebral deformity, but it is nonspecific even when spinal osteoporosis has been established. Height loss is expected with ageing, due to postural changes and the compression of intervertebral discs[95]. If less than 4 cm, a height loss is an unreliable indicator of deformity status[105], and results depend on the measurement technique used and variability in posture[114]. When secondary to osteoporosis, a height loss of 4 cm or more is usually an index of quite severe disease, and is therefore not useful in identifying osteoporosis in early stages[95].

Kyphosis can be used as an index for vertebral deformities in individuals with advanced osteoporosis, but valid measures are also difficult to obtain[114]. Besides, the majority of patients with prevalent vertebral deformities do not experience substantial kyphosis. Like substantial losses of height, kyphosis is usually found only in severe and advanced cases.

Relying on patient recall tends to yield inaccurate results with regards to vertebral fracture history. Self-reports are unreliable in the estimate of prevalent and incident deformities[14,136,137] since most vertebral deformities are either asymptomatic or associated with acute back pain that is usually mild and transient. Studies that have investigated the validity of self-report of vertebral fractures reveal that the true incidence of deformities can be over-reported by as much as 20%, or under-reported by 40%[136,137].

2.2.4.ii - Conventional radiographs

Lateral radiographs of the thoracolumbar spine currently remain the best method for assessing the presence of vertebral deformities[95,105]. An initial assessment of the spine for osteoporotic vertebral deformities usually includes both AP and lateral views of the thoracolumbar spine from T4 to L4 levels[100]. The main use of the AP view is to verify that all the relevant vertebrae are present and are clearly visible on a radiograph, and it is the lateral view that is examined for vertebral deformities. Only lateral views are subsequently needed to monitor incident deformities[95]. Assessments of vertebrae at higher levels than T4 are usually not made, since their view is obstructed by the overlap of scapulae and the shoulder, and vertebral deformities at these levels rarely occur[100]. Viewing the L5 vertebra can also be difficult because of the thickness of the pelvis and a parallax effect[95].

As with other types of radiologic assessments, the accuracy of a reading depends on the experience and training of the observer. There are several differential diagnoses to be considered when assessing for the presence of an osteoporotic vertebral deformity (**Table 2.1**), and this can only be achieved by a visual inspection and an expert interpretation[132]. The presence and progression of vertebral deformities due to osteoporosis can be difficult to assess if an individual has certain pathologies that either block visualization (*e.g.* severe scoliosis), or mimic osteoporotic changes (*e.g.* osteoarthritis). There also remain a number of other questions to resolve, such as the meaning of mild degrees of anterior wedging. The latter do not seem to be associated with low bone mass or subsequent vertebral deformities[67]. In a clinical setting, once a vertebral fracture is diagnosed by imaging, or a deformity is incidentally detected by a radiologist, secondary causes of osteoporosis and other metabolic bone diseases should be investigated by laboratory investigations[114].

The method that is used to take the radiograph also needs to be standardized. For example, many studies have shown that correct positioning of the patient for the radiographic assessment of vertebral deformities can greatly affect the results[100]. Well-defined protocols for taking the radiographs (e.g. tube-to-film distance, correct positioning of the patient) are needed to ensure comparability of films[105].

High-resolution fan-beam DXA, using technology similar to computed tomography, could eventually prove to be a low-radiation alternative to lateral x-ray films. The radiation dose is 100 times lower than conventional radiographs[95]. It is also less vulnerable to technical artefacts of positioning and can image the lateral spine in about 10 seconds. BMD measurements could be obtained at the same time. Further development and validation are however required for its clinical implementation, as the superior image resolution of radiographs currently outweighs the advantages of DXA imaging[138].

2.2.4.iii - Detecting and defining vertebral deformities

There is no consensus on how to best detect the presence of vertebral deformities. The most efficient method to assess radiographs is in contention, as is the threshold to categorize normal and abnormal vertebrae. The choice of using one method over another, or a particular definition for deformity, depends on the goal. For most epidemiologic and clinical situations, high specificity is desirable as this produces stronger associations between vertebral deformities and their risk factors and consequences[105].

Methods of assessment can rely on purely visual readings done by radiologists (qualitative), purely morphometric evaluations by automated instruments (quantitative), or standardized visual (semiquantitative) techniques. Some population surveys, like the European Prospective Osteoporosis Study (EPOS) and the Rotterdam Study, use a quantitative approach that is supplemented by expert visual confirmation. Within each radiologic ascertainment approach, several different criteria have been used to define an osteoporotic vertebral deformity. Some methods use percentage cut-offs, while others rely on a certain number of standard deviations away from the mean of a predefined population. There is also no uniform approach to establishing normative values.

2.2.4.iv - Visual approach (Qualitative)

There is considerable intra- and interobserver variation in the identification of deformities with purely qualitative readings[100,105]. Unstandardized visual inspection of lateral radiographs of the thoracolumbar spine is said to be inadequate for the diagnosis of

vertebral fractures even in routine clinical practice[100]. When attention is not specifically focused on the issue of a vertebral deformity, this approach to radiologic reporting has been shown to be unreliable[95]. The accuracy of spinal radiographic diagnosis of vertebral deformities was recently examined by the worldwide IMPACT prospective study[132]. The results from local radiographic reports were re-analyzed at a single central site using Genant's semiquantitative method[140]. Using over 2 000 radiographic reports from five continents, the study found that vertebral deformities were considerably under-diagnosed, despite a clear vertebral deformity definition and a protocol that minimized inadequate film quality. These results illustrate the need for a standardized assessment protocol. It can also be inferred that a central reading site may help with validity and reliability, when a large number of radiographs are to be assessed.

2.2.4.v - Morphometric methods (Quantitative)

From 1988 to 1993, computer-assisted methods that rely solely on the quantitative assessment of vertebral deformities were developed to serve the needs of epidemiologic studies and clinical trials. These studies require definable, reproducible, and objective methods to detect vertebral deformities in large numbers of radiographs, when there may not be a clinical indication to guide the observer[100]. Essentially, several different points are placed on the laterally projected vertebral body with a cursor (typically 6 points), and the dimensions of the anterior (Ha), middle (Hm), and posterior (Hp) heights are measured using electronic digitizing procedures (**Figure 2.1**)[132]. Most quantitative methods use ratios of different heights that are measured on individual vertebrae, as part of their algorithm. The calculated parameters that are most commonly used for deformity detection are the wedge ratio (Ha/Hp), the biconcave ratio (Hm/Hp), and the compression ratio ($Ha/Hp_{\pm 1}$), where "± 1" or some other subscript indicates the vertebral body above or below[100].

The ratio measurements are compared to an explicitly defined reference for normal values. As there are variations in absolute dimensions between individuals and populations, defining vertebral deformities in terms of ratios of vertebral heights within vertebrae is one way to minimize this normal variation in size[141]. The National Osteoporosis Foundation (NOF) has recommended that each vertebral level should have

established mean and SD values, specific for the population under study[105]. Using the CaMos data, Jackson *et al* also demonstrated that normative values are required for each sex-specific and vertebral-level-specific ratio[96].

The definition of deformities also needs to be put forth, and various empirical and statistical approaches are commonly used. Melton[142] utilized a 15% reduction in any of the Ha/Hp, Hm/Hp, or $Hp/H_{\pm 1}$ ratios to define a deformity, and then adjusted the percentage for normal variability in vertebral shape. Eastell[97] modified this method by defining a deformity as a reduction in height over 3 SD from the assumed norm, instead of fixed arbitrary percentages. Minne[143] proposed that vertebral heights be compared to normalized values for body size of healthy young women, by dividing all values by the corresponding values of the T4 vertebra.

McCloskey[98] increased the specificity of the Eastell/Melton standard criteria by first predicting posterior heights (H_{pred}) from the Hp of up to four adjacent vertebrae. The McCloskey definition for deformity then requires two criteria: decreases in either Ha/Hp, Hm/Hp, or Ha/H_{pred} accompanied by decreases of over 3 SD in Ha/H_{pred}, Hm/H_{pred}, Hp/H_{pred}, respectively. This method has been utilized by the EPOS and the Rotterdam Study. In contrast to other methods, the McCloskey method predicts normal values for the individual patient rather than comparing height ratios exclusively to a reference population, which results in a lower false positive rate.

The NOF recommended in 1995 that a reduction of 3 SD or more from normal mean ratios of dimensions for a particular vertebral level be used as the definition for a vertebral deformity. Furthermore, it is suggested that the definition should include at least two criteria (*e.g.* the reduction of one height ratio must accompany another's) in order to reduce the number of false positives[105]. Representative of the lack of consensus around this issue, the CaMos is currently using only one height reduction criteria to define a vertebral deformity.

Quantitative methods are objective and reliable measurements if point placement and normative references are rigorously defined, and if there are well-defined algorithms for a deformity's definition[95,105]. Despite efforts to remove subjectivity from the data, there remain several disadvantages in using solely quantitative methods.

Different quantitative methods tend to yield different estimates of prevalence and incidence[141]. Most of these differences are due to the way that the normative vertebral dimensions are derived and the way the vertebral deformity is defined[100]. Sensitivity and specificity vary according to how and which criteria are used to define abnormal values[105,132]. More specific criteria yield less false positives, but more deformities are classified as normal. For example, the prevalence estimates from the European Vertebral Osteoporosis Study (EVOS) differed by as much as 4-fold, depending on whether Eastell's or McCloskey's method was used[107]. The Eastell method reported the larger prevalence estimates, as it is less specific than the McCloskey method which utilizes variation within an individual's spine and two criteria to define a deformity. Yet another example of the divergence in opinions is that the CaMos opted to use Eastell's method.

Some investigators claim that definitions based on assumed statistical distributions are more efficient than percentage cut-offs[96,97,144]. The distributions of height ratios of non-deformed vertebrae within each spinal level are approximately Gaussian[96,144], and cut-offs such as 3 or 4 SD away from the mean can be chosen to define deformities. These definitions are however still arbitrary. Level-specific height ratio values that dichotomize vertebrae into normal and abnormal have not yet been clearly defined. Moreover, variability in the shape of the distribution of normative data may change where the cut-off value falls. Still another difficulty is that there are a large number of measurements made per individual, and therefore definitions based on statistical distributions will classify some normal vertebrae as deformed by chance alone[105]. This problem is compounded when large samples are examined, as in population-based studies.

According to some experts, substantial numbers of mild deformities that could be detected by a visual reading can be missed by morphometry. End plate deformities, a lack

of parallelism of end plates, and the general altered appearance of a vertebra compared to its neighbours are all visual cues of mild deformities that only a visual assessment can consider. Another way that mild deformities may be missed by morphometry is that six-point placement is not able to capture certain visual characteristics of a vertebral body[132]. There is an inevitable variation on serial radiographs in positioning and parallax, which may result in different point placement[95]. Thus an alteration in projection can be mistaken for an incident deformity by morphometry, an artefact that can usually be identified by visualization.

Digitizing techniques require experience and training for consistent point placement. The protocol used for the location of point placement will also yield results with varying sensitivity and specificity[95,132]. Valid quantitative assessments also depend on proper positioning of the patient parallel to the x-ray table, so that patients with mobility difficulties, such as in scoliosis, may not be assessed[100].

The most important limitation of purely quantitative methods is that large numbers of false-positives are usually obtained. Automated methods are unable to distinguish anatomic variants, differential causes, and technical artefacts from osteoporotic vertebral deformities. The phenomenon of "fracture rebound", which is not uncommon with morphometry, illustrates this: vertebrae labeled as deformities are re-classified as normal over time, which clinically is highly unlikely. The CaMos has consequently modified Eastell's approach to correct for random measurement error *post hoc*, and this will be detailed in the Methods' chapter.

Only a visual inspection by an expert in osteoporosis can confirm that the anomalies captured by quantitative technique are indeed osteoporotic deformities[132]. Genant recommends that a quantitative method should not be used without a visual assessment to confirm that prevalent and incident deformities are secondary to osteoporosis. Most of the large clinical trials investigating drugs for osteoporosis in the last decade, as well as the EPOS and the Rotterdam Study have used a combination of quantitative and semiquantitative methods[100].

2.2.4.vi - Semiquantitative methods

Concurrent with the development of quantitative methods, standardized visual assessments (or semiquantitative) were also advanced in the last ten years. These methods use a visual approach, but with extensive training and clear protocols for assessment. Direct measurements of vertebral dimensions are not done with the semiquantitative approach, but approximated by visually comparing adjacent vertebral heights with each other, or with same vertebrae seen on previous films. The shape of vertebrae is explicitly considered as part of the diagnostic algorithm. A numeric score is then assigned to vertebral deformities according to their severity and type, or they are classified into distinct categories[95].

Standardized visual grading schemes have been found to be more reproducible than inspection of radiographs without specific criteria for deformity diagnosis[105,132]. They have also been found to be more specific than quantitative methods in detecting mild deformities. The advantage of semiquantitative methods is that the entire spectrum of visible features is used to identify deformities. Assessments done by trained and experienced readers can distinguish between normal vertebrae and anomalies that simulate mild deformities (*e.g.* Schmorl's nodes), and differential diagnoses can be considered. Semiquantitative assessments have been used in several randomized clinical trials[100].

The method of Genant is perhaps the most validated semiquantitative method, as it emphasizes the differences between osteoporotic deformities and those caused by other pathologies. It combines an assessment of vertebral body height with a qualitative inspection of its shape and configuration relative to adjacent vertebrae. Deformities are graded (normal, mild, moderate, and severe) by visual inspection according to the percentage reduction in expected vertebral height. Unlike other semiquantitative methods, the type of deformity (wedge, biconcavity, compression) is not associated with the grading of a deformity. The severity of a fracture is assessed solely by the extent of vertebral height reduction and morphological change. The approximate degree of height reduction determines the grade of severity assigned. Because morphologies that are

typical of osteoporotic deformities are used, the positive predictive value increases. Refracturing of vertebrae can also be identified in serial radiographs, as the progressive nature of deformities is considered[95].

Genant's semiquantitative method has been shown to be valid, and has high interobserver and intraobserver reliability for the diagnosis of both prevalent and incident vertebral deformities[140,145]. Sensitivity and specificity have been shown to be relatively high if the readings are done by expert radiologists or clinicians. However, reminiscent of qualitative and morphometric approaches, Genant's semiquantitative method has limitations. The differentiation of mild deformities from normal anatomic variations of vertebrae relies on the expertise and training of the observer. Normative data for the quotient from anterior and posterior vertebral heights have revealed that there is considerable variation dependent on the vertebral level: smaller quotients are found in the middle thoracic spine (T6-T8) and at the thoracolumbar junction than other areas of the spine. Prevalence values are therefore particularly susceptible to overestimation, especially when large numbers of films are assessed[95]. Incidence estimates may be less susceptible, as radiographs of individuals are examined in series.

2.2.4.vii - Combination of quantitative and visual approaches

In large epidemiological studies, a rigorously objective approach is needed, given the enormous number of vertebrae assessed, the lack of a specific clinical indication to guide the assessor, and the stringent requirements for reproducibility. Contrary to clinical situations, radiologists usually do not have access to additional imaging techniques, such as bone scintigraphy or magnetic resonance imaging. The basic goals are however the same as in the clinic: vertebral height reductions must be assessed in a non-biased, reliable manner, and differential diagnoses must be part of the algorithm to minimize false positives.

Several studies[145-147] have shown that there is only moderate agreement between semi- and quantitative methods, with substantially better agreement between observers occurring when Genant's semiquantitative method is used. There may be higher

concordance between the two approaches when mild deformities are excluded from the analysis[146]. In the current literature, Genant's method is advocated for both clinical and research applications[95,132]. When morphometric methods are necessary as in large surveys, it is recommended that a semiquantitative inspection of the positive films should be done as confirmation[105]. Using sensitive morphometric criteria to identify potential deformities in combination with a qualitative assessment for increased specificity is one way to obtain valid prevalence and incidence estimates from population studies[100,105,145].

2.2.5 - PREVALENCE

2.2.5.i - Overall estimates of prevalence

Since falls are only involved in approximately one third of vertebral deformities[14], their prevalence may serve as a rough guide to the epidemiology of severe osteoporosis in a population. From the previous discussion, it is expected that prevalence estimates will vary according to the method used to define normative values and the definition for deformity chosen . When comparable methods and definitions are used, the results of different surveys can be evaluated for potential underlying risk factors. Overall prevalence estimates are similar among Canadian, European, and white women in the northern U.S. (**Table 2.2**). The overall prevalence of deformities in Japanese women is much lower than in Dutch women, as estimated by Genant's semiquantitative method. Among different centres within Europe, the overall prevalence of vertebral deformities can vary between 10-24%[107].

2.2.5.ii - Age and sex

The estimated prevalence of vertebral deformity increases with age, as was reported by the CaMos[96]. Women tend to have more severe deformities than men. Men in the younger age-group (50-59 years) have a higher prevalence of deformities than women in the same age group. Starting at age 70, women have a markedly higher prevalence than men. This trend also holds if only severe deformities are reported (> 4 SD). Trauma may play an etiologic role among young and middle-aged men; the risk of deformity in men increases

with high levels of physical activity. In contrast, physical activity seems protective in women[4].

The CaMos findings for sex- and age-adjusted overall prevalence concur with those of American[148,149] and European studies[59,107,126].

2.2.5.iii - Geographic area and ethnicity

When examined by age-stratification across the world, prevalence estimates are somewhat more dissimilar compared to overall estimates[6]. For example, across all age groups Chinese women have fewer deformities than either Canadian men and women, or white women from the North-Western U.S.. However, factors other than race may explain some geographical differences; some of the observed variation across countries can be accounted for by study methodology, including differing age distributions. The age-stratified estimates from the EVOS are all noticeably greater than ones from Rochester, Minnesota, which are both homogeneously Caucasian. Within the EVOS, there was considerable variation between the 36 centres' overall prevalence estimates[107]. The highest rates were observed in the Scandinavian countries. This variability is perhaps partly related to differences in geographic female-to-male ratios.

2.2.6 - INCIDENCE

2.2.6.i - Secular trends

Incidence rates for osteoporotic fractures at various skeletal sites, including the spine, have risen during the last fifty years. The incidence of vertebral fractures among postmenopausal women in the northern U.S. and Sweden increased until the 1960s, and then reached a plateau[4].

2.2.6.ii - Population-based studies

Only two population-based prospective cohort studies were found after a Medline search for studies investigating the incidence of vertebral deformities: the prospective arm of the EVOS (called EPOS)[151], and the Rotterdam Study[152]. Several other studies have

investigated incidence in population samples of men and women, but using only clinically diagnosed fractures[153,154,155]. Among them is the seminal study done in Rochester, Minnesota, which used only vertebral fractures identified retrospectively from medical records from 1985-1989[153], so that the Rochester study relied on qualitative readings for outcome ascertainment. Only a subsample was analyzed with morphometry. Epidemiologic comparisons between the European and American estimates are therefore difficult to interpret.

The incidence of vertebral deformities increases strongly with age in both men and women. However, the age-related increases in incidence have been observed to be more pronounced in women than in men. These increases occur mainly in subjects with prevalent deformities at baseline. **Figures 2.2** and **2.3** show that a greater incidence of vertebral deformities was reported in European women compared to men across all age-groups, with the exception of the 55-59 age group where rates are similar. After about the age of 60, women in Europe have a 2-fold to 3-fold greater incidence than men. As expected, the Rochester estimates of absolute incidence are about 3 times lower than to the European estimates, confirming that only about $^1/_3$ of vertebral deformities prompt clinical evaluation. Interestingly, the Rochester results revealed a similar disparity to the other two studies in the gradient of incidence between men and women across all age groups.

In all three studies, the rate of increase is constant for women until the age of 70 years, after which there is a steep rise. Men also experience a steep rate of change after a period of a constant rate of increase, but this occurs later than in women. In the EPOS, the rate increase was found after age 80 in men, but it occurred after age 75 in the Rochester and Rotterdam studies. At older ages in the Rotterdam Study, the incidence rate ratio between subjects with and without prevalent deformities is similar in men and women.

The Rotterdam study also found 2 peaks of incidence in both sexes, in the midthoracic spine and at the thoracolumbar junction, with vertebral levels T12 and L1 being the most frequently affected. Although the absolute incidence of vertebral deformities was lower in

men than in women, the risk for an incident deformity was similar at any given level of lumbar BMD in men and women, after adjusting for age and prevalent deformities. This suggests that the difference in absolute incidence in men and women may be explained by sex-related BMD differences, as men tend to have higher peak BMD values and lower bone loss rates than women.

2.3 - The Use of Multiple Imputation to Manage Missing Data

2.3.1 - INTRODUCTION

Missing data commonly occur in medical research, and they are an almost inevitable occurrence in the setting of the multivariate and longitudinal studies utilized in epidemiology[156,157]. Missingness may occur on only some of the variables (*item* nonresponse) or entire cases may be missing (*unit* nonresponse). Selected subjects may not agree to participate, or be lost to follow-up before the end of a study. For example, low-income individuals tend to refuse to participate in surveys, whereas high-income individuals tend to refuse to respond only to certain items, especially questions realting to income[158].

Unit nonresponse can be substantial in large population surveys. For example, the CaMos obtained consent from 42% of subjects asked to participate[34], and the EVOS response rate was 49%[107]. In contrast, the Rotterdam study was able to recruit 78% of invited participants, perhaps in part because all subjects were from a single, confined district in the Netherlands[126]. Even so, only 35% of those initially recruited attended a follow-up visit that was necessary for providing estimates of vertebral deformity incidence.

This section will briefly review different approaches for handling missing data, focusing in particular on multiple imputation. The rationale for using this procedure is demonstrated by first describing missing data mechanisms and the constraints they

impose on the options available to deal with missing data, and then describing the principles of multiple imputation.

2.3.2 - MISSING DATA AND BIAS

In surveys, it is reasonable to suspect that nonrespondents and subjects lost to follow-up are systematically different from full respondents. If the expected value of parameter estimates differs between respondents and those individuals with missing data, then nonresponse and/or attrition biases are present. Given that inference to the general population is the main goal of population surveys, the presence of bias can affect the external validity of a study, as well as its internal validity. A common way to assess whether estimates are biased is to measure those characteristics of the nonparticipants and the censored that are predictive of the study's outcome of interest[158] .

Methods for dealing with missing data include both qualitative and quantitative approaches. After presenting the differences between certain characteristics of participants and nonparticipants, most investigators simply discuss the possible impact on the results of the incomplete data set. For example, the Rotterdam investigators postulated that they had underestimated the incidence of vertebral deformities, as nonparticipants were found to be older, in poorer health, and to have higher mortality than participants[126]. With this approach, there is no quantification of the error introduced by selection bias, so that bias-adjusted estimates cannot be reported.

One quantitative method to adjust estimates for bias is to use a regression model of the related characteristics to predict unobserved future outcomes, as was done by the EPOS[151]. A Poisson regression model was developed from the baseline characteristics and outcomes of the participants without missing data. The model was then used to predict the expected vertebral deformity incidence of nonparticipants by age groups, and these results were compared to observed estimates. The assumptions underlying this approach are that the relationship between baseline risk factors and subsequent fracture

was the same in both participants and nonparticipants, and that sufficient risk factor data were available for good predictions. In this way, it was found that nonresponse bias was present in the oldest age group (75-79 years), where the incidence was higher in both male and female nonparticipants. The investigators were thus able to adjust an estimate of the expected age-standardized incidence for those aged 50-79 years. A major problem with this approach is that the missing incidence data that is predicted by the regression model is treated as observed data.

An alternative approach to adjusting for nonresponse bias is multiple imputation, which was recently illustrated by Kmetic *et al* in estimating the prevalence of osteoporosis from the CaMos[34]. Multiple imputation is a general technique that uses a model to 'fill in', or impute, all missing data items, thereby creating "complete" data sets. Anywhere from two to several thousand imputed values are typically generated, so that the uncertainty in the values of the missing data is integrated into the inferences. Estimates that are derived by multiple imputation can assume different mechanisms for the missingness of the missing data. These mechanisms are now discussed.

2.3.3 - MISSING DATA MECHANISMS

Before deciding on an approach to analyzing a data set with missing values, the underlying mechanism for the missingness must be considered. The ideal situation with incomplete data sets is that the missing values be no different, on average, from the observed values. If this is the case, any analyses can simply ignore the missing data with the only loss being one of statistical efficiency (true variances are at least as small as the variances obtained from any other consistent estimators). Otherwise, the assumption made about the mechanism for missingness needs to be described, as different techniques for dealing with missing data may be invalid under some assumptions. For this reason, Rubin[159] formally defined the different assumptions that can be made about missing data mechanisms.

Data are *missing completely at random* (MCAR), if the probability of a missing data item on a variable is unrelated to the value of that item, or to the values of any other variables in the data set[160]. For example, a missing value for gender that is no more likely in men than in women and is unrelated to the other variables in the dataset, is MCAR. It is important to emphasize that the MCAR definition allows for the possibility that the missingness of a variable be related to the missingness of some other variable. For example, if subjects who tend to refuse to report their age also tend to refuse to report their smoking status, the data can still be MCAR if, say, older subjects are not on average smokers. If missing data are MCAR, most statistical approaches will give valid inferences[156]. Nonetheless, MCAR is a very strong assumption that is not often met in clinical research settings.

A more realistic assumption is that the data are *missing at random* (MAR). This means that the probability of missing data on a variable is unrelated to the value of that variable, after controlling for the other non-missing variables in the analysis[160]. In other words, the probability that an observation is missing may depend on the observed values, but not on the missing values[161]. Therefore, the pattern of missing data given the observed data does not supply any information about the parameters of interest[34]. For example, if the probability of missing data on self-reported vertebral fractures depends on age, but within each age category, the probability of a missing fracture history is unrelated to vertebral fracture status itself, then the data are MAR. In some circumstances, the MAR condition is known to hold exactly. One such situation is when a subset of nonrespondents are randomly re-sampled in large surveys; if all resampled original nonrespondents now participate, then it has been shown that the remaining nonrespondents are MAR[161].

A further constraint is found when missing data on the parameter of interest are said to be *ignorable*. In this case, the data need to be **(i)** at least MAR, and **(ii)** the parameters that define the missing data mechanism are independent of the parameters used to model the observed data[34,158]. An advantage of assuming that the data are ignorable is that the missing data mechanism does not need to be modeled as part of the estimation process. Some authors such as Allison[160], point out that condition **(ii)** is almost always

encountered in data that is MAR, and therefore they use the terms MAR and ignorable interchangeably. Moreover, if the data are MAR but the distinctness condition **(ii)** does not hold, inference based on the assumption of ignorable data are still valid, if not fully efficient[162]. Nonresponse in surveys is rarely known to be ignorable[158].

Nonresponse mechanisms are *nonignorable* when there are systematic differences between missing and observed values, even after accounting for all observed data[160]. More specifically, nonresponse bias exists if a respondent and a nonrespondent with exactly the same values for the observed variables have systematically different values for variables missing for the nonrespondent[158]. In this case, the missing data mechanism needs to be explicitly modeled in order to get valid estimates of the parameters of interest[156].

Even though in any real data situation it is impossible to verify whether missing data are ignorable or nonignorable[158,162], virtually all missing-data procedures that are used, even so-called principled ones, rely at least implicitly on an assumption of ignorability[161]. Multiple imputation techniques are known to have good properties if the missing data are ignorable. However, multiple imputation techniques can also permit the consideration of different assumptions for the response mechanism; imputations can be generated with models either under an assumption of ignorable missing data, or of nonignorable missing data. A comparison of the separate inferences under different models displays the range of bias that may be present in any analysis[34,158].

2.3.4 - APPROACHES TO DATA WITH MISSINGNESS MECHANISMS THAT ARE IGNORABLE

2.3.4.i - Case deletion

There are several statistical techniques that have been developed to analyze incomplete data sets, but many require specialized knowledge and techniques not readily accessible to most epidemiologic researchers[162]. Most statistical analytic techniques used by

researchers work on the assumption that a data set is complete[158,160]. Excluding missing data and then applying standard methods of analysis, called case or listwise deletion, is perhaps the easiest and the most commonly used approach to handling missing data. Most statistical software programs by default exclude all cases that have missing data on any of the variables included in an analysis. Case deletion relies on the assumption that the missing data are MCAR, which is unlikely in the context of large surveys. Even so, when the incomplete cases comprise less than 5% of all cases, then case deletion may be a reasonable solution, as the effects of bias and statistical inefficiency are minimized by the small amount of missing data[162,163].

Analyzing only the complete cases may however create several problems. In large studies, the selection of optimal covariates for modeling, already an onerous task given the many correlated variables that are usually collected, is further complicated by the presence of missing data. There may be a tendency to exclude from consideration those variables with many missing cases[164]. When the amount of missing data is substantial (> 5%), case deletion has a strong propensity to introduce bias into the estimates and to overestimate the standard errors (SE)[158,162,165,166]. It will also cause a loss of statistical precision in estimation, especially in a large study with several covariates, where the probability of missing data on any one covariate becomes high[164]. In general, because the MCAR assumption is usually not known to hold, and because of the loss of precision, case deletion is not a recommended procedure[160].

2.3.4.ii - Deterministic imputation

Imputation is a generic term that is defined as the "practice of filling-in missing data with plausible values"[162]. The main advantages of single imputation techniques are that they are relatively easy to use and allow a "complete data set" to be analyzed with standard methods. Imputation is said to be deterministic if missing data are filled in using a non-random mechanism.

One of the simplest ways to impute is to substitute the missing values of a variable with its mean or median. Although easy, such substitutions underestimate the SE's of the

estimates and do not take advantage of relationships between variables[156]. As such, mean and median substitution are well known to bias point estimates and to lead to confidence intervals that are too narrow[162,163].

Apart from mean substitution, several other deterministic imputation methods have been described (*e.g.* "hot deck" single imputation). These methods select likely values for each missing item from some sort of model or "nearest neighbor" technique[161], and then analyze the data as if the missing values had not occurred. Because these approaches treat imputed data as real data, they can underestimate the SE's[162,163]. Treating missing values as known quantities rather than as random variables is therefore a poor general strategy[165].

2.3.4.iii - Principled imputation methods

Although no technique for dealing with missing data can replace the ideal of having a 100% response rate[167], there exist several accessible and optimal techniques that use existing data to impute values, while accounting for their uncertainty. Principled imputation procedures are based on either an observed-data likelihood (*e.g.* the expectation-maximization (EM) algorithm), or an observed-data posterior distribution (*e.g.* multiple imputation). These procedures rely on the ignorability assumption and attempt to remove bias from the analyses by using the assumed relationship between the observed and unobserved data[161].

A brief mention of approaches using maximum likelihood (*e.g.* EM algorithm) is warranted, as they have been shown to yield results that are as valid as multiple imputation in fully parametric models[160,162]. The SE estimates may be somewhat more efficient than multiple imputation, since they involve no simulation[163]. The disadvantage of a maximum likelihood approach is that it can only be used for certain types of linear models, and it requires a model for the joint distribution of the variables with missing data[160]. Multiple imputation is much more flexible, and can accommodate virtually any type of model, and may work reasonably well even with nonignorable missing data[162], although this is difficult to verify in practice.

2.3.5 - MULTIPLE IMPUTATION

2.3.5.i - Background

In 1977, Rubin first proposed a stochastic imputation paradigm for the problem of missing data, referred to as multiple imputation[168]. Multiple imputation is a general approach that replaces each missing value with two or more values representing a distribution of possibilities[158]. Originally developed to deal with non-response in large sample surveys and censuses[167], it has since been used in several types of health-care databases and study designs[166].

Its basic objective is to obtain valid results for incomplete data sets, without having to use specialized analyses of the incomplete data. This does not imply that the inferences obtained from standard analytic tools on complete data are expected to be the same as those resulting from imputed data sets. As Rubin points out, this would be analogous to having as an objective that a survey's answers be identical to a complete census, which is an unachievable goal[162]. The goal is to obtain reasonably valid inferences, even though the data set has missing values[168].

Frequentist inference via multiple imputation usually proceeds in three stages. First, data are imputed under an appropriate model and *m* imputed data sets are created. In principle, the *m* imputations of the missing values are *m* random draws from the posterior predictive distribution of the missing values[169]. Second, the *m* versions of the imputed data set are each analyzed separately by complete-data techniques to obtain the desired parameter estimates and SE's. Third, the results of the *m* analyses are combined by computing the mean of the *m* parameter estimates, and a variance estimate that incorporates both within-imputation and between-imputation variances[161].

It has been shown that estimates derived from data sets that are "completed" by multiple imputation are statistically consistent (estimates are approximately unbiased in large samples), asymptotically efficient, and asymptotically normal, assuming the data are MAR[158,162]. Multiple imputation makes use of all available data and preserves sample

size. Rubin's method is said to be "without serious competition" with regards to solving incomplete-data problems[170].

The techniques involved in multiple imputation require algorithms that can be computationally intensive. The use of multiple imputation was therefore limited by the software and memory capacities available at the time they were first proposed[163]. For example, in 1995 Greenland and Finkle[157] reviewed several methods of handling missing covariates in regression analysis. They concluded that the more complex methods of multiple imputation were preferable, but their implementation was impractical because of a lack of packaged software. Multiple imputation was therefore rarely used as a practical tool in research, even though it was acknowledged by statisticians as a rigorously sound methodology, far surpassing the most commonly used methods in many situations[162,163].

With the advent of increased hard disc space, computer speed, and new computational methods and software, multiple imputation is now more accessible. Multiple imputation has since been utilized in a variety of study design types, and its inherent flexibility avails it to the most complicated modeling analyses[168]. Although it is not yet widely used in medical research settings[156], it is receiving notice in epidemiologic journals[34,169].

Even though implementing multiple imputation is becoming easier[168], in depth understanding of multiple imputation requires more than a passing acquaintance with statistics. Researchers without quantitative backgrounds may need to rely on the interpretation of Rubin's works by more accessible reviewers, like Allison[160] or Schafer[163]. Various multiple imputation techniques are now part of standard analytic software packages[160], but their blind use has the potential of yielding worse results than simply ignoring the missing data[162,168], for example if the missing data mechanism is nonignorable. The investment of the extra effort needed to understand multiple imputation procedures could reward the researcher seeking valid results.

2.3.5.ii - Single random imputation

In order to understand the principles of multiple imputation, one of its basic components, single random imputation, will first be described. By deliberately introducing a random

component to an imputation process, the problem of underestimating the variance for variables with missing data can be largely solved[160,161].

A simple example taken from Allison[160] will be used to illustrate the procedure. Suppose the goal of an experiment is to estimate the correlation between variables X and Y, but there exist 50% missing cases on X. First, each missing value for X is imputed by regressing X on Y. The resulting least-squares regression equation is then used to generate predicted values for the missing X value. The new imputed data set, which comprises both observed and imputed X values, can now be analyzed by conventional methods.

However, as Allison demonstrates by simulation, if the X values are filled in by the mean predicted value, the inherent randomness of a predicted value isn't considered and the correlation coefficient estimate will be substantially overestimated compared to the true parameter. The reason for this biased result is that all imputed values lie exactly on the estimated regression line. The standard deviation of X (including the imputed values) is underestimated, and yields an overestimate of the correlation.

One way towards solving this problem is to take random draws from the residual distribution of X. In this example the residual distribution of X is normal, or $\epsilon \sim N(0, \sigma)$, where σ can be estimated from the regression of X on Y. The imputed values for X are therefore 'corrected' by adding the predicted value from the regression of X on Y, to the product of σ times a randomly chosen N (0,1) residual value. In this way the standard deviation for X is increased, and this results in point estimates with less potential for bias[160].

Despite the insertion of randomness to the imputed value, there remains a serious problem with single random imputation. Complete-data analytic methods will use these imputed data as if it were observed data, which is generally not correct. A single imputed value cannot reflect the sampling variability about the actual value when only one model for nonresponse is being considered. The additional uncertainty that occurs when more than

one imputation model is considered is also ignored[158,166]. This means that estimates of the SE are still too small[160-162,169].

The solution is to repeat the imputation procedure $m > 1$ times, thus creating variability in any value of X that has been imputed. Note that the variability in X arises from two different sources. First the parameters from the model that produces predicted values of X are estimated with uncertainty, and second, given the model parameters, the predictions are uncertain. This suggests a two-stage imputation procedure, where parameter values are first selected from their joint posterior distribution, and given this parameter selection, random values of X from the model are imputed. In this way, all inherent uncertainty is included in the imputations. This results in *m* "complete" data sets, and if within- and between-variability are included, accounts for the fact that the data are imputed with uncertainty, as well as parameter uncertainty in the imputation model.

2.3.5.iii - Imputation models

The validity of multiple imputation methods depends on how the imputations are generated[165]. The imputation model must preserve all important associations among the variables that will be included in the final analysis[156,161,163,165]. For example, say a group of correlated, categorical covariates containing missing values is identified in a longitudinal study, and the imputation model chosen specifies a log-linear distribution. The missing values for each variable should then be predicted from all other variables in the correlated group. If variables that are predictive of missing data are not included in the imputation model, but subsequently used in the analyses, the imputed subset of each variable will not be related to the other variables[164]. The assumptions or models used to create the imputed data sets must be compatible with those used to analyze them. Statistical inconsistency may arise if the imputation models do not reasonably preserve those distributional features (*e.g.* associations) that will be the subject of future analyses[162,163].

Multiple imputation methods do not necessarily require that missing data are ignorable. Current theory allows for the construction of imputation models tailor-made for each

unique data set. In principle, imputations may be created under any kind of assumptions or model for the missing-data mechanism, and the resulting inferences will be valid, assuming that the mechanism is correct and the imputation model 'proper'[163]. In practice, however, the missing data mechanism remains unknown, so that one can never be certain if the inferences are valid or not.

The essential goal of a multiple imputation model is to have the imputed values derived from a predictive distribution of the missing values given the data observed, while preserving the associations between the variables[171]. There are "off-the-shelf" models that can be used as good approximations for a wide range of multivariate data sets, and they have the advantage of being incorporated into accessible software. These models include linear regression for continuous outcomes, and logistic or polytomous regression for categorical variables. More complex models can be created for better point estimate prediction, but the small gain in the accuracy of prediction may not be worth the effort[160].

2.3.5.iv - The Gibbs sampler

Often the most complicated component of the multiple imputation procedure is the elucidation of posterior distributions required for the random draws of the imputation model's parameters. Even with simple probability models, posterior distributions often cannot be described in closed form, as they are usually complex multidimensional functions[161].

Recently several new simulation methods have been applied to the problem of Bayesian estimation in complicated parametric models. These methods, collectively known as Markov chain Monte Carlo (MCMC), are often able to get around the problem of intractable probability distributions by utilizing simulation. MCMC methods, such as the Gibbs sampler, simulate random draws from nonstandard distributions via Markov chains[163]. A Markov chain is a sequence of generated variates of a random variable, in which the distribution of each variate depends on the value of the previous one. If the simulating steps of the chain are repeated a large number of times, the MCMC algorithms are designed such that random draws converge to random samples from the target

density[161]. MCMC methods can thus simulate the joint posterior distribution of an unknown parameter, and simulation-based estimates of virtually any features of the posterior can be obtained.

Since imputations are created under Bayesian arguments, it seems natural to use a Bayesian approach to analyze data sets. All estimates are given a probability distribution that represents the uncertainty of their values, and missing data are considered simply as additional unknown parameters to be estimated[34]. While a frequentist analysis of the *m* imputed data sets can be done, the imputations themselves are usually drawn using a Bayesian procedure, since unknown parameters are given a distribution.

2.3.5.v - Multiple imputation algorithm using the Gibbs sampler

Before running a model-based multiple imputation algorithm, an investigator needs to first select the set of variables that will be used to model the missing data. The distributional assumptions of the model selected need to be considered, so that some variables may need to be transformed[160,161]. Once the model has been constructed, imputations are created through a Bayesian process. A parametric model for the complete data is specified (and, if necessary, a model for the mechanism by which data become missing), a prior distribution to the unknown model parameters is applied, and *m* independent draws from the conditional distribution of the missing data given the observed data by Bayes' Theorem are simulated. Although the creation of imputations can be performed explicitly through formulas in simple problems, special computational techniques such as MCMC usually must be applied for more complex models[161]. Since the missing data are simply considered as additional unknown parameters to be estimated, using multiple imputation via the Gibbs sampler has the additional advantage of estimating all posterior distributions at once, including the posterior distributions of parameters of main interest.

A concern that is sometimes raised by the researchers is that multiple imputation seems to be "creating data out of nothing". This protestation has merit if single imputation is utilized, since in this case imputed values are treated as observed ones[166]. However, as

Schafer states, "multiple imputation is nothing more than a device for representing missing-data uncertainty" [163]. By averaging over a predictive distribution of the missing data by simulation, it accurately represents the observed data set while preserving randomness of the imputed values[163,168]. Although multiple imputation may not be necessarily the best method for any given missing-data problem[163,172], it is a flexible approach to imputing missing data that, when used properly, yields valid estimates of parameters and SE's[168].

Table 2.1 - Differential diagnoses of a radiological vertebral deformity. From Guermazi *et al*[132].

Osteoporosis
Trauma
Degenerative disease
Scheuermann's disease
Congenital anomaly
Neoplastic disease and haematopoietic disorders
Infectious disease
Paget's disease

Figure 2.1 - Example of six-point placement for vertebral morphometry. From Guermazi *et al*[132].

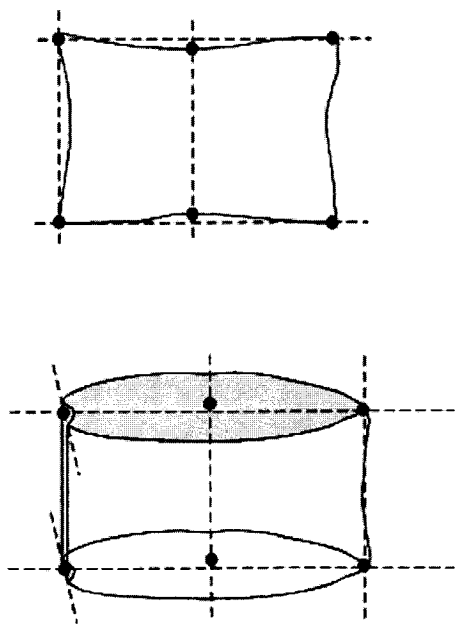


Table 2.2 - Overall deformity prevalence estimates from various studies.

Reference	Method	Prevalence	Sample	Population
Melton *et al*[142]	≥15% (adjusted)	26%	White women ≥50 years	Rochester, Minn.
Eastell *et al*[97]	3 SD 4 SD	21% 11%	195 white postmenopausal Women > 47 years	Rochester, Minn.
Black *et al*[144] SOF	≥20% 3 SD	12.5% 20%	2992 women 65-70 years	Northern U.S.
Kado *et al*[123] SOF	3 SD 4 SD	20.0% 13.1%	9575 women ≥65 years	Northern U.S.
O'Neill *et al*[107] EVOS	3 SD	20.2% 20.2%	8331 women 7239 men ≥50 years	Europe
Jackson *et al*[96] CaMos	3 SD 4 SD	23.5% 9.5%	4613 women > 50 years	Canada
	3 SD 4 SD	21.5% 7.3%	1820 men > 50 years	
Pluijm *et al*[127]	Genant's semiquantitative method	39% 39%	267 women 260 men > 65 years	Netherlands
Burger *et al*[126] Rotterdam Study	3 SD 4 SD	17% 8%	750 women ≥55 years	Netherlands
	3 SD 4 SD	12% 4%	750 men ≥55 years	
Kitazawa *et al*[150]	Genant's semiquantitative method	4.7%	1092 women 45-69 years	Japan

Figure 2.2 - Incidence of vertebral deformities in women and men in Europe by age group, from the EPOS[151]. Vertebral deformities were defined as 20% and at least 4 mm height reduction using the McCloskey method.

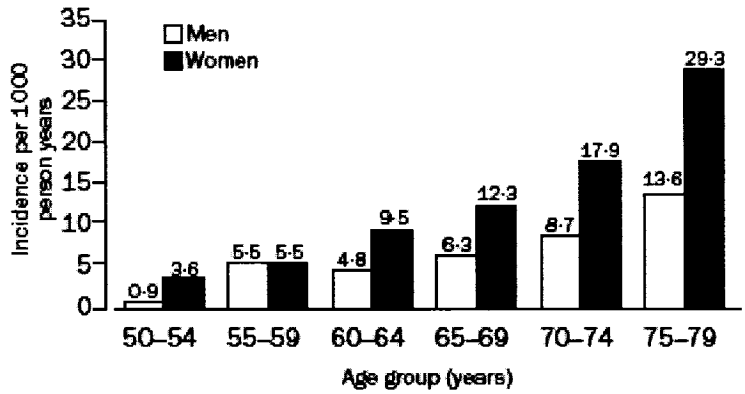
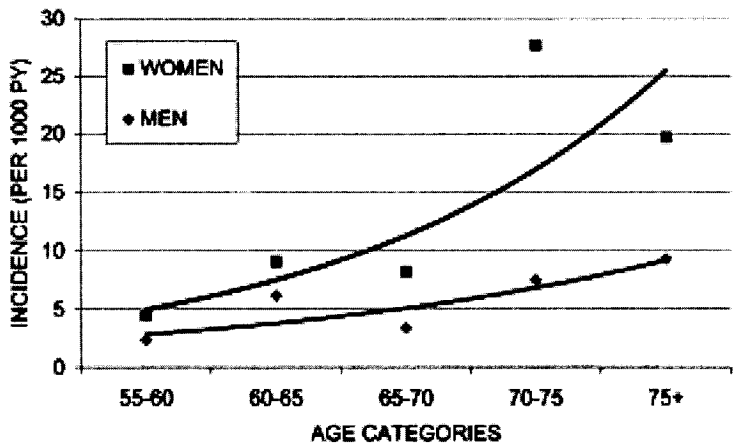


Figure 2.3 - The incidence of vertebral deformities per 5-year age strata in men and women, from the Rotterdam Study[152]. Vertebral deformities were defined as a 15% and at least 4.6 mm height reduction using the McCloskey method.



Chapter 3 – METHODS

3.1 - Introduction

This thesis proposes to estimate the incidence of vertebral deformities in the Canadian population, using data collected by the CaMos. This chapter therefore describes the design and conduct of the CaMos, and details the manner in which vertebral deformities were defined and radiologically assessed. The statistical methods that were chosen by the author of this thesis to analyze the CaMos data are also described.

3.2 - CaMos Design

The CaMos is an ongoing population-based longitudinal study that has so far accumulated 5 years of follow-up data. The study's basic goals are the assessment of: **1)** the prevalence and incidence of osteoporotic fractures in Canadians aged 50 and older; **2)** the cost of osteoporosis and related fractures; **3)** the effect of osteoporosis on health status and quality of life; **4)** the relationship between osteoporosis and sociodemographic and other characteristics; and **5)** the estimation of the distribution of bone density in Canadians aged 25 and over. More specifically, and related to this thesis, a primary objective is the derivation of accurate estimates of the incidence of vertebral deformities in men and women aged 50 years and over.

The ideal of randomly sampling all Canadians over the age of 25 was not possible for logistic and financial reasons. The source population therefore consisted of noninstitutionalized Canadians aged 25 years of age and older, residing within 50 km of the following nine centres: Vancouver, Calgary, Saskatoon, Toronto, Hamilton, Kingston, Quebec City, Halifax, and St John's. With this geographical inclusion criterion, it is estimated that the study population was sampled from 37% of the Canadian population, from both urban and rural areas. Based on the nine study centres, seven geographic regions were identified as follows: British Columbia (Vancouver), Alberta (Calgary),

Prairies (Saskatoon), Southern Ontario (Toronto, Hamilton), Eastern Ontario (Kingston), Quebec (Quebec City), and the Atlantic (Halifax, St. John's).

The required sample sizes were largely driven by rough prevalence estimates within age- and sex- strata, based on confidence interval widths (**Table 3.1**). Since incidence measures will be numerically less than the corresponding prevalences, and since all prevalences are below 50%, more than sufficient numbers were ensured for incidence estimates.

The target study sample was stratified by sex, 8 age groups, the 9 study centres, and blocked by three-month calendar periods. The sampling frame consisted of all residential telephone subscribers within specified postal codes. Random sample blocks were purchased from provincial telephone companies, except in Nova Scotia and Newfoundland where the random selection of residential subscribers was done manually. All participants were recruited between February 1996 and September 1997 by randomly selecting from the telephone listings. In each household contacted, a census was taken of all persons in the household and based on the stratification scheme, one randomly selected member 25 years of age or older was invited to participate.

In all, 80 163 households were contacted, and 22 436 were eligible to participate based on the sampling criteria (**Figure 3.1**). Forty-two percent (9 423) of eligible persons contacted consented to fully participate, of which 69% were female and 31% were male (the respondent group). Recruitment ranged from 900 to 1 133 subjects within the nine study centres.

The CaMos attempted to collect the major risk factors for osteoporosis in a brief Refusal Questionnaire (RQ) that was administered to those eligible subjects who declined to participate (the RQ group, n = 5 519 for those ≥50 years). The RQ gathered information on age, gender, race, fracture history (yes/no and if yes before or after age 50), family history of osteoporosis (yes/no, osteoporosis in family members including the individual responding), and current cigarette smoking status (yes/no).

There were 6 107 subjects who declined to answer the RQ that were either aged ≥50 years or their age was unknown, and henceforth they are referred to as the Total Refuser (TR) group. Some of these subjects provided their age and gender (the TR1 group), and most of these contactees also provided the number of household members older than 25 years of age. For the remainder of the Total Refusers, the only information available was the number of telephone calls needed to reach the individual, the urban centre, and either **(i)** age or gender (but not both), or **(ii)** neither age nor gender (the TR2 group). The data collected on the RQ and TR groups (**Table 3.2**) will be used to estimate the probability that these contactees would have sustained incident vertebral deformities, thereby allowing for the adjustment of nonresponse biases, under different assumptions for the mechanism of missingness.

At baseline, a Full Questionnaire was administered by interviewers to all consenting respondents, which collected socio-demographic information, medical and fracture history, height and weight, dietary intake, physical activity, tobacco smoking and secondary exposure, family history. The interviewers also administered the Rand SF-36 Questionnaire[173] and the Health Utilities Index[174]. Bone density was measured using DXA at the lumbar spine (L1-L4) and the femoral neck, as well as ultrasound at the calcaneus. Baseline radiographs of the thoracolumbar spine of all Respondents aged 50 years and over (n = 6 233) were taken in order to determine the prevalence of vertebral deformities.

Annual Follow-up Questionnaires collected information on nonvertebral incident fractures. At Year 3, subjects who were 40-60 years at baseline once again completed the Full Questionnaire and BMD at the lumbar spine and hip were re-measured. At Year 5, all subjects were administered the Full Questionnaire and BMD was once again measured. Respondents aged ≥50 years were asked to undergo a spine x-ray at Year 5, so that the incidence of vertebral deformities could be estimated. Note that a small number of subjects turned 50 during the 5-year follow-up period, so that these respondents did not have the baseline radiographs that were needed for comparison with the Year 5 radiographs.

Not all the data collected from respondents who were radiographed both at baseline and at Year 5 had been entered into the database at the time that the analyses were conducted. It will however be assumed, for the exercise of analysing these data within the context of this thesis, that all of the 3 469 subjects who had no data available from x-rays taken at Year 5 had been lost to follow-up (the censored group). Therefore 56% of the original cohort over age 50 (n = 3 479) will be considered lost to follow-up, and 2 754 radiographs (1 982 female, 772 male), which represented 44% of the baseline radiographs, were available for the analyses pertaining to this thesis. Because of logistic reasons, it is not known exactly how many subjects had morphometric data yet to be entered into the CaMos database, so that the true attrition rate is unknown at this time.

More detailed descriptions of the CaMos protocol are presented in Kreiger *et al*[175], and Tenenhouse *et al*[176].

3.3 - Vertebral Deformity Definition and Assessment

3.3.1 - INTRODUCTION

As discussed in Chapter 2, the choice made for the definition criteria of a deformity, as well as the methods with which the radiographs are analyzed, will profoundly influence the interpretation of the results. This section describes the radiographic approach that was used for ascertainment, how normal vertebrae for the population were estimated, and the criteria that were used to define osteoporotic deformities.

3.3.2 - RADIOGRAPHIC ASSESSMENT

All respondents aged 50 and over were radiographed at baseline and invited to be x-rayed again at year five. Spinal radiography proceeded with a defined, detailed, and standardized protocol (tube-to-film distance 100 cm; thoracic films centered on T8, lumbar films centered on L3), which ensured sufficient quality to allow for accurate

assessments by morphometry. All radiographic films were first examined for quality by a qualified technologist, before being transported to a central radiographic analysis centre for digital morphometry. The films were then analyzed by a single trained morphometry technologist using a backlit transparent digitizing tablet (Scriptel) and a custom-designed software program. The technologist was trained in conjunction with an established expert group at the University of California in San Francisco, and continuing collaborative analysis of test samples have ensured a uniform and consistent analysis. Individual vertebrae with pathology which made point placement difficult were identified either by the morphometry technologist or flagged by the software, and excluded from the morphometric analysis. The vertebrae at levels T4-L4 were analyzed by measuring the posterior, mid, and anterior heights. These measurements were used to calculate 4 vertebral height ratios for each vertebra: H_a/H_p, H_m/H_p, H_p/H_{p+1}, H_p/H_{p-1}.

3.3.3 - ESTIMATION OF THE NORMAL POPULATION

Sex- and vertebral-level-specific reference norms specific to the CaMos population were developed by using a sub-sample of the actual study group. The normal data were extracted from the sub-sample data, which contained both normal and abnormal vertebrae, in a method similar to the one described by Black *et al*[144]. A separate healthy group of individuals therefore did not need to be radiographed to establish population norms.

A random subset of 3 971 radiographs were first selected from the CaMos subjects (2 827 female, 1 144 male) that had a baseline spine radiograph done. Based on a qualitative visual inspection, any deformity that was diagnosed by each study centre radiologist was excluded. All films were then sent to the CaMos central radiographic centre, and read by the morphology technologist. Any vertebra identified as an obvious deformity was excluded in this second qualitative reading. The remaining radiographs were then analyzed by morphometry and the means for the 4 height ratios of each vertebral level were computed. Any vertebral height ratio that was reduced more than

20% compared to the sex-specific mean for that vertebral level was excluded. In this way, 3 056 subjects (2 225 female, 801 male) were deemed to have normal vertebrae out of the original subsample. It was assumed that each sex- and vertebral level-specific height ratio was normally distributed, based on the inspection of histograms and normal-probability plots. The means and SD of each height ratio were then utilized as reference norms for the CaMos, from which deformities could be defined. Jackson *et al* have described the derivation of normal vertebral height ratios for the CaMos in detail[96].

3.3.4 - VERTEBRAL DEFORMITY DEFINITION

3.3.4.i - Original Definition

The choice of morphometric parameters that were used to define abnormal values for vertebral body height ratios were guided by recommendations made by the National Osteoporosis Foundation Working Group on Fractures[105]. As was first described by Eastell *et al*[97], any vertebral height ratio greater than 3 SD below the mean norm and equal to or above 4 SD defines a vertebra as having a Grade 1 osteoporotic deformity. A Grade 2 deformity is defined as any height ratio greater than 4 SD below the mean. If a vertebra has more than one abnormal height ratio, then it is classified by the highest grade of deformity present in its measured height ratios.

The estimated incidence of vertebral deformities will not be stratified by grade severity in these analyses. The operational definition of a deformity for this thesis is thus any vertebra which has at least one height ratio that is more than 3 SD below the sex- and vertebral level-specific reference norm that was derived specifically for the CaMos population. A vertebra was identified as having sustained an incident deformity when the same vertebra at baseline had normal vertebral height ratios, but satisfied the deformity criteria at Year 5. This approach was used to reduce the fracture rebound phenomenon as described in Chapter 2, Section 2.2.4.v, thus potentially lowering the false-positive rate.

Quality control assessments have shown that the methods used in the CaMos to ascertain vertebral deformity occurrence are highly reliable. Interobserver reliability between a study-designated expert radiologist and the morphometrically assessed radiographs at one centre was 99%. Similarly, intraobserver reliability for morphometry measurements varied between 98-99% over a period of 3 years[96].

3.3.4.ii - Measurement errors in x-ray film digitization

Once all the baseline x-ray films had been analyzed by morphometry, repeated measurements of a sub-sample of films revealed a random measurement standard error of approximately 3% for each vertebral height ratio. This measurement error was thought to be partly responsible for the fracture rebound that is known to occur with morphometry. The threshold ratios for deformity, previously determined from the population data by Jackson *et al*[96], were consequently corrected by subtracting 3% of their values for the analyses presented in this thesis. Vertebral height ratios were then re-classified into the usual normal, Grade 1, and Grade 2 categories based on their variance away from the corrected population norms. In addition, a new category was created in attempt to capture 'borderline' deformities. A "Grade ½" category is therefore defined as a height ratio whose value is more than 3 SD away from the uncorrected mean norm, but less than 3 SD away from the norm mean that was corrected for the 3% measurement error. It may be that so-called Grade ½ deformities represent mostly measurement error, instead of subtle deformities. Age-specific estimates of incidence were therefore reported with and without the inclusion of the Grade ½ deformities, in order to assess the impact on incidence estimates.

3.4 - Statistical Analyses

3.4.1 - DESCRIPTIVE STATISTICS

Throughout, separate analyses were carried out for men and women. All of the descriptive analyses were performed with and without the inclusion of the Grade ½ deformities, in order to assess the impact of the 'borderline' grade on the estimates.

Within the sex-specific groups, CaMos contactees were initially compared across their participation status (respondents, and the censored, RQ, and TR groups). Some of the variables included in these comparisons are known to be associated with vertebral deformities: age, race, history of fracture and osteoporosis, smoking status, pre-existing vertebral deformities, spine and hip BMD, history of treatment for osteoporosis, BMI, and ambulatory difficulties. Other variables included were the number of telephone call required to establish contact, the number of household members ≥25 years, and the city of residence. These latter variables were especially important in groups where few other variables were available because of nonparticipation in the rest of the study. They were included in the hope that they were also correlated with deformities; for example, because subjects in poorer health may be home more often, they hence may be more likely to answer the phone in fewer attempts.

The baseline characteristics of subjects who sustained no incident deformity during the 5 years of follow-up were compared to those who sustained at least one incident deformity. The effect of including the Grade ½ vertebral deformities in the analyses was also investigated by comparing the frequencies of subjects with and without incident deformities, stratified by the presence of prevalent deformities.

The extraction and transformation of data from the CaMos database, as well as all descriptive statistics, were performed with SAS software (Release 8.02) for Windows.

3.4.2 - BAYESIAN ESTIMATES OF UNADJUSTED INCIDENCE RATES

Incidence was defined as the number of subjects who sustained at least one deformity during the follow-up period, per 1000 Person Years (PY). Since x-rays were taken only at baseline and Year 5, the exact time at which new vertebral deformities occurred is unknown. Therefore, the PY contributed by each case was taken at the half-way point of the follow-up period, which should be close to the correct value, at least on average.

A Poisson likelihood function was utilized for each of the sex- and age-specific incidence rates, when inferences were made without the consideration of prevalent deformities. The Gamma distribution, which is the conjugate distribution for the Poisson likelihood, was used as a prior distribution. It was parameterized with $\alpha = 0.001$ and $\beta = 0.001$, where $\mu = \alpha\beta^{-1}$, so that the posterior distributions were almost exclusively driven by the data. When stratified by the presence and absence of prevalent deformities, the incidence estimates were computed by using a Poisson regression with prevalence (yes/no) as the sole independent variable. In this instance, diffuse prior distributions were specified as Normal for both the intercept and the independent variable, with $\mu = 0$ and $\tau = 0.0001$, where $\sigma^2 = 1/\tau$.

All rate estimates were stratified by sex and by 4 age groups: 55-64, 65-74, 75 years and older (75+), 55 years and older (55+). Strictly speaking, all of the estimates are therefore adjusted rates for sex and age. Even so, for the sake of clarity the incidence rates that have not been adjusted for attrition or nonresponse biases will be referred to as 'crude' or 'unadjusted' rates in this thesis.

The rate estimates reported for subjects aged 55+ were weighted to the 1996 Canadian population census by direct standardization. For the estimates in the 55+ age group that were stratified by pre-existing deformity status, the proportion of Canadians without and with prevalent deformities for each sex- and age-specific category was approximated using the proportions that were observed in the CaMos study sample, and the weights were then adjusted accordingly.

Point estimates and 95% credible intervals (Cr I) were calculated with WinBUGS, version 1.4 software. The mean of the posterior distribution for each incidence rate was taken as the point estimate. The convergence of the simulations was assessed by running two or more iterative Gibbs sampler chains simultaneously, so that the stability of the simulations could be verified with history plots of the sample values of each parameter versus the iterations. The width of the central 80%-interval of the pooled runs, as well as the average width of the 80%-intervals within the individual runs, were also examined for their convergence to stability. The Gelman-Rubin convergence statistic[177], which is the ratio of the pooled and within 80%-interval widths, was examined for its convergence to 1 for all parameters.

Once convergence had been evaluated as to have likely occurred, the number of iterations required to obtain accurate inferences was determined by running the algorithms until the Monte Carlo error for each parameter of interest was less than 5% of the sample standard deviation, and the posterior density plots appeared 'smooth'. A minimum burn-in of 1 000 iterations was followed by 5 000 further iterations before the posterior means and 95% credible intervals were taken as summaries of the Gibbs sampler outputs. When convergence and accuracy were of concern, 10 000 iterations were performed after a burn-in of 5 000 iterations. Templates of the WinBUGS programs that were used to estimate the crude incidence rates, with and without stratification by pre-existing deformities, are provided in **Appendices 1** and **2**, respectively.

The sex-specific distributions of deformities across the thoracolumbar levels of the spine were inspected with and without the inclusion of Grade ½ deformities.

3.4.3 - IMPUTATION MODELS

Incidence estimates for the censored and nonrespondent (RQ, TR1, and TR2) groups were obtained by first constructing sex- and age group-specific imputation models for each of these 4 groups. The incident and covariate data that were available for respondents at

baseline were used to build unconditional multivariate logistic regression models, using the independent variables available for each group whose incidence was to be imputed. The coefficients for the logistic models were estimated by iterative reweighted least squares using S-Plus software (version 4.5). With each successive level of participation, there were progressively fewer independent variables available for the candidate imputation models (see **Table 3.2**). Censored individuals had of course the same covariate data collected at baseline as the respondents. Model selection was then guided by the Bayesian Information Criterion (BIC)[178], which approximates Bayes Factors, by using the bic.glm function developed for S-Plus software by Raftery[180]. The BIC has optimality properties in predicting the dependent variable for future subjects, among all possible candidate models[179].

Although the BIC tends to select parsimonious models, imputation models should be as large as reasonably possible, and include variables associated with the parameter of interest as well as the probability of missingness[168]. Therefore, when independent variables judged to be important predictors were missing from the BIC-selected models, the variables were forced back into the models, but retained only if the resulting coefficient estimates were in the expected direction. The independent variables that were deemed as essential for models (when data were available for the different groups) were: age, prevalent deformities, history of fracture $\geq$50 years, family/personal history of osteoporosis, BMD of the lumbar spine, and history of having been pharmaceutically treated for osteoporosis.

As previously mentioned, the TR2 group was comprised of contactees on whom age and/or gender were missing. As these variables are perhaps two of the most important predictors of vertebral deformities, these missing data were imputed for the TR2 group. The probability of being female given the data available was estimated for each TR2 contactee using logistic regression. The respective probabilities were then used to parameterize the Bernoulli distribution, from which random draws were made to singly impute gender (S-Plus program in **Appendix 3**). Age was imputed using linear regression, where the mean predicted age and the SE for an individual prediction were

estimated for each TR2 contactee, and subsequent random draws made from the Normal distribution accordingly parameterized (S-Plus program in **Appendix 4**). Since only respondents aged ≥50 years at baseline were radiographed for deformities, only the TR2 individuals whose imputed age was ≥50 years were retained in the analyses.

Although these procedures only impute age and gender randomly according to the age/gender proportions already seen in the CaMos, it is a way to incorporate the additional uncertainty due to these missing variables into the model. Because of the paucity of demographic information available for the TR2 group, executing a single imputation was considered sufficient, there being little variation from imputation to imputation given the simplicity of the model.

Besides age and gender, only a few variables were available for selection by the BIC in the TR1 and TR2 groups. The number of household members ≥25 years and the number of telephone calls made to establish contact variables were therefore forced into the imputation models, and only retained if the estimated coefficients were in the expected direction. It was assumed that as incidence increases, the number of household members and the number of phone calls needed for contact both tended towards one; the implication is that the elderly, who are more prone to develop vertebral deformities, tend to live alone and be at home more often compared to younger individuals.

The fit of each imputation model was verified by stratifying the respondents into 8-10 mutually exclusive groups according to their status on the covariates, and then comparing the observed rates to the expected predicted rates by the models. Once the imputation models were chosen, each group with missing data on the dependent variable could be combined with the respondent group, so that the multiple imputation program embedded in WinBUGS software (version 1.4.1) could estimate the missing incidence data.

3.4.4 - INCIDENCE ESTIMATES ADJUSTED BY MULTIPLE IMPUTATION

First, only the respondent and censored groups were combined, so that estimates could be adjusted solely for attrition bias. The missing incidence data in the age-specific censored groups were then imputed. Using the proportion of respondents and censored subjects in each of the age-specific groups as weights ($\sum$weights = 1 for each age group), the rate estimate for each age group was adjusted for attrition bias under the ignorable assumption by taking the weighted averages of the observed and imputed estimates.

Let Δ_i be the absolute value of the difference between the observed incidence of vertebral fractures within a sex- and age-specific group *i*, and the incidence imputed for the corresponding censored group *i*, where *i* = either age group 55-64, 65-74, or 75+. In a method similar to that described by Kmetic *et al*[34], a sensitivity analysis of the ignorability assumption was conducted, where the missing data mechanism was assumed to be nonignorable following one of the three possibilities listed below:

(i) The attrition bias is at least as large as Δ_i, and in the same direction as the adjustment made by multiple imputation. Therefore Δ_i is either added to, or subtracted from the censored $rate_i$, depending if the censored $rate_i$ is larger, or smaller than the respondent $rate_i$, respectively. The censored $rate_i$, already adjusted by imputation based on the observed covariates, is thus further adjusted for unobserved factors that may be associated with incidence as well as missingness. By taking the weighted average rate of the respondent $rate_i$ and the "censored $rate_i \pm \Delta_i$" (weights being the $proportion_i$ of subjects, where $\sum_i$ $weights_i = 1$), the rate estimate for group *i* is thus adjusted for attrition bias (referred to subsequently as the "$+\Delta_i$" model, where the sign '+' refers to the direction of the adjustment being the same as that made by multiple imputation);

(ii) The attrition bias is only half as large as Δ_i, and in the same direction (the "$+ \frac{1}{2} \Delta_i$" model);

(iii) The attrition bias is half as large as Δ_i, but in the opposite direction (the "$- \frac{1}{2} \Delta_i$" model, where the sign '-' refers to the direction of the adjustment being opposite to the one made by imputation).

While there is no strict justification for these assumptions, they seem to be an intuitively reasonable way to find bounds for a sensitivity analysis.

In order to adjust incidence estimates for nonresponse bias as well as attrition bias, all contactees were combined within sex- and age-specific groups, *i.e.* the respondents, censored, and the nonrespondents (RQ, TR1, and TR2 groups). Under the assumption that the missing data are ignorable, weighted averages of the observed and multiple imputation-derived estimates are thus adjusted for both attrition and nonresponse biases, assuming that the imputation models are proper.

Let Δ_{ij} be the absolute value of the difference between the observed incidence within a sex- and age-specific group *i*, and the imputed $rate_{ij}$ for a missing data group, where *j* = either censored, RQ, TR1, or TR2 groups. Under a range of nonignorability assumptions concerning the subjects with missing data, a sensitivity analysis to the ignorability assumption was carried out, similar to the one described above, again utilizing three models: **(i)** The combined effect of the nonresponse and attrition biases is at least as large as Δ_{ij}, and in the same direction as the adjustment made by imputation. Therefore Δ_{ij} is either added to, or subtracted from the imputed $rate_{ij}$ depending if the imputed $rate_{ij}$ is larger, or smaller than the respondent $rate_i$, respectively. A weighted average of the observed $rate_i$ and the "imputed $rates_{ij} \pm \Delta ij$" thus represents an estimate adjusted for both nonresponse and attrition biases under the nonignorable assumption, if the imputation models are proper (referred to subsequently as the "$+\Delta_{ij}$" model, the sign '+' refers to the direction of the adjustment being the same as made by multiple imputation); **(ii)** The combined effect of the nonresponse and attrition biases is only half as large as Δ_{ij}, and in the same direction (the "$+ \frac{1}{2} \Delta_{ij}$" model); **(iii)** The combined effect of the nonresponse and attrition biases is only half as large as Δ_{ij}, but in the opposite direction (the "$- \frac{1}{2} \Delta_{ij}$" model, where the sign '-' refers to the direction of the adjustment being opposite to the one made by imputation).

Appendix 5 contains the WinBUGS template program that was used for the estimation of all adjusted estimates.

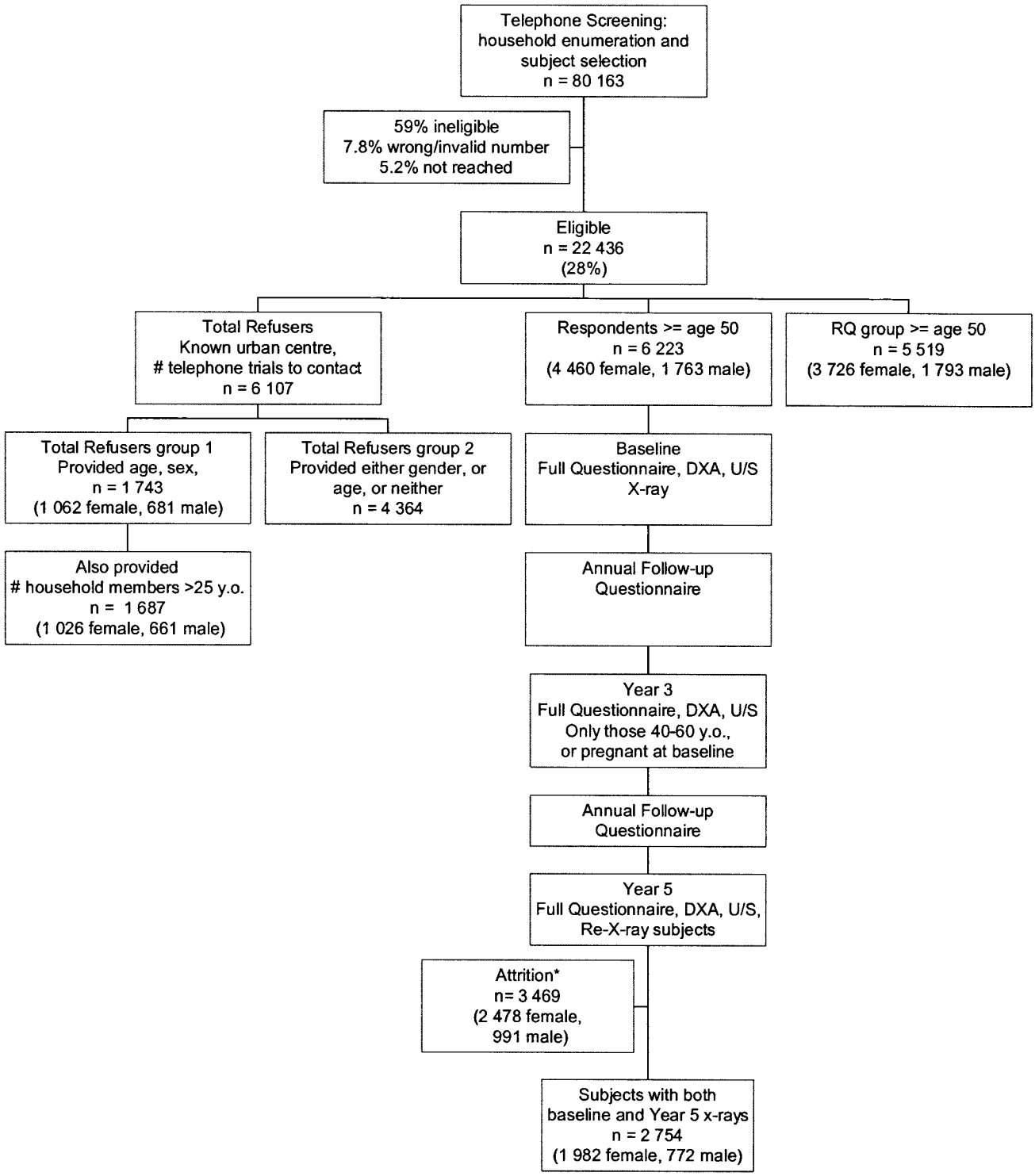
Table 3.1 - Estimated and actual sample size by age and sex group for the CaMos.

	Women		Men	
Age stratum	*Estimate*	*Actual*	*Estimate*	*Actual*
25-45	278	534	278	443
46-50	482	534	278	311
51-55	599	641	278	300
56-60	715	748	278	304
61-65	944	964	323	351
66-70	1168	1157	403	424
71-80	1014	1518	369	587
81+	983	443	382	164
Total	6 183	6 539	2 589	2 884

Table 3.2 - Data collected for the nonresponder groups.

Data Element	**Nonresponder Group**		
	Refusal Questionnaire (RQ)	***Total Refusers Subgroup 1 (TR1)***	***Total Refusers Subgroup 2 (TR2)***
Age	✓	✓	
Gender	✓	✓	
Race	✓		
Fracture history (yes/no and if yes before or after age 50)	✓		
Family history of osteoporosis (yes/no including individual responding)	✓		
Current cigarette smoking status (yes/no)	✓		
Study center	✓	✓	✓
Number of calls required to establish contact	✓	✓	✓
Number of household members 25 or older	✓	✓	

Figure 3.1 - CaMos data collection. Not all data collected were available at the time of analyses, so that attrition numbers are inflated.



* Subjects whose x-rays at Year 5 were not yet entered into the CaMos database were considered lost to follow-up for these analyses.

CHAPTER 4 – RESULTS

4.1 - Overview

The reader is reminded that at the time these analyses were performed, not all of the morphometric data that were extracted from the Year-5 x-rays had been entered into the CaMos database, so that attrition is somewhat exaggerated in these analyses. Once these missing data are included in the analyses, incidence rates can be estimated with more accuracy.

The baseline characteristics of the individuals within the different participation groups are first presented, followed by the distributions of these characteristics when subjects are stratified according to their incidence status. The proportions of men and women who did and did not sustain an incident deformity are then stratified by pre-existing deformity status at baseline. The crude estimates of incidence are reported, followed by the effect of including the Grade ½ deformities on the distribution of incident deformities along the spine. The imputation models and the incidence estimates that they yielded for the different participation groups are then presented. Finally, the rate estimates adjusted for attrition bias only, and those adjusted for both nonresponse and attrition biases are reported.

4.2 - Demographics and Risk Factors

Some of the risk factors for vertebral deformities are distributed differently in the censored group compared to the full respondents (**Tables 4.1** and **4.2**). Although women comprise approximately the same proportion of subjects in the censored and full respondents (about 70%), the female subjects that were lost to follow-up are on average older, had more skeletal fractures reported after age 50, more vertebral deformities present at baseline, and were less ambulatory than full respondents. Prevalent deformities, smoking, and difficulties with walking were more common in the censored men, who also

tended to be older than the male respondents were. The sex- and age-specific estimates based on the full respondents may therefore be biased towards underestimation of incidence, if only attrition bias is considered.

Tables 4.1 and **4.2** show that nonresponse bias may also affect the incidence estimates derived from the respondent data only. The RQ and TR nonrespondents were on average older, and the RQ contactees were more likely to be smokers than the respondents, which are risk factors for osteoporosis. However, female RQ nonrespondents were substantially less likely to report a personal/family history of osteoporosis and a history of skeletal fractures occurring over age fifty. Although men in the RQ group also reported, on average, less osteoporosis in their personal/family history, they were more likely to have a history of fractures. Nonrespondents within the RQ and TR groups were differently distributed on gender, with 68% of the RQ contactees and 61% of the TR1 contactees being female, perhaps indicating that these two groups are sufficiently different on at least one variable that is associated with deformity incidence. These results demonstrate that a univariate analysis of the baseline characteristics of participants *vs.* non- or partial participants, can be limited in its usefulness in forecasting the direction of bias, quite apart from the magnitude.

The baseline characteristics of respondents, stratified by the occurrence of incident deformities (**Table 4.3**), suggest that factors associated with increasing incidence are increasing age, pre-existing deformities, slightly lower BMD measurements at both the lumbar spine and femoral neck, a personal or family history of osteoporosis, a history of skeletal fractures occurring after age 50, having been treated for osteoporosis, and ambulatory difficulties. Remarkably, the proportion of individuals with incident deformities who are female (66.5%) is not greater than the proportion of subjects who are female and did not fracture (72.4%). Calculating the unadjusted relative risk from the data available in **Table 4.3**, men had 1.3 times the risk of developing a vertebral deformity compared to women. Female sex has been consistently reported as a risk factor in the literature, even when unadjusted for other covariates, so that this finding is unexpected.

There were 212 Respondents who sustained an incident vertebral deformity (7.7%), if only Grades 1 and 2 are included in the analyses. The addition of Grade ½ deformities almost doubles the number of cases to 370 (13% of Respondents) (**Table 4.4**).

The frequency distributions of male and female subjects with and without incident deformities, stratified by the presence of prevalent deformities at baseline, are shown in **Table 4.5**. Being male continues to be associated with an increased probability of sustaining an incident vertebral deformity, despite adjustment for pre-existing deformity status. Although 141 of the 212 subjects (66.5%) who fractured were female, there was a slightly greater proportion of men who had no pre-existing deformities who went on to sustain one or more incident deformities (5.5%), compared to women (3.2%). When pre-existing deformities are present at baseline, the proportion of men with incident deformities is 23.9%, *vs.* 22.8% in women. Prevalent deformities have therefore a relatively more pronounced effect (relative risk = 6.0) on incidence rate compared to sex in the CaMos respondents. If the relative risk is stratified by sex, the presence of pre-existing vertebral deformities increases the risk of deformity approximately 7 times in women, compared to 4 times in men (Incidence rate ratio of 1.6). A Poisson regression model with incident deformity as outcome, and sex (female = 1, male = 0), age, and previous deformity as covariates, estimates a negative coefficient for sex, albeit with a confidence interval (CI) that crosses zero (β = -0.26; 95% CI -0.54, 0.02).

The inclusion of Grade ½ deformities seems to re-establish female sex as a risk factor for incident deformity in the presence of pre-existing deformities (**Table 4.6**). Of the now 370 subjects with incident deformities, 80 (13.0%) and 40 (25.8%) were men without and with pre-existing deformities, respectively, compared to 128 (8.7%) and 112 (28.3%) for the respective proportions in women.

4.3 - Crude Estimates Derived Directly from Observed Data

Incidence rates per 1000 PY, stratified according to sex and 10-year age groups, are shown in **Tables 4.7** and **4.8**. The overall incidence estimate in men aged ≥55 years (weighted to the 1996 Canadian census) is greater than in women (17.7/1000 PY, 95% CrI 13.5 - 22.1; and 14.6/1000 PY, 95% CrI 12.2 - 17.1, respectively). Larger rate estimates in men compared to women are also observed in the 55-64 and 65-74 age groups. A larger point estimate in women is observed only in the elderly (75+ years), where the female incidence estimate is 35.7/1000 PY (95% CrI 28.5 - 43.2), and the male estimate is 31.1/1000 PY (95% CrI 21.0 - 42.1). When prevalent deformities are present at baseline, rate estimates are larger in men than in women across all age groups, although the credible intervals are quite large and there is considerable overlap across age groups as well as between the sexes. The effect of pre-existing deformities on the age-related increases of the estimates is especially remarkable in younger males. Whereas the point estimate increases from 9.2/1000 PY to 49.5/1000 PY in men aged 55-64 when the presence of pre-existing deformities is considered, the estimated rate increases from 3.0/1000 PY to 8.2/1000 PY in women of the same age group.

Including the Grade ½ deformities (**Tables 4.9** and **4.10**) considerably amplifies the rate estimates, except in the 55-64, 75+, and 55+ age groups of men who had pre-existing deformities at baseline; furthermore, the rates now increase across the 3 age groups of these men. The female point estimates for those who had prevalent deformities in the age groups 65-74, 75+, and 55+ years are now greater than the male rate counterparts (54.5, 94.6, and 67.1/1000 PY *vs.* 53.5, 84.5, and 55.8/1000 PY, respectively).

The bimodal distribution of vertebral deformities within the spine reflects the two areas that are most vulnerable to biomechanical stresses (**Tables 4.11** and **4.12**, **Figures 4.1** and **4.2**), and this finding has been corroborated in other studies. In women, incident deformities are more frequent in the T5-T8 and T10-L2 areas, but this pattern is attenuated when Grade ½ deformities are included, as their inclusion increases the number of deformities throughout the spine in a non-uniform manner. If the biphasic

distribution is held to be true, then the inclusion of the Grade ½ deformities may capture more false-positive morphometric readings in women, especially in vertebrae where the ascertainment of occurrence may be more subtle. In men, the bimodal distribution is retained with the inclusion of Grade ½ deformities, perhaps indicating that male vertebrae are less prone to subtle confounding pathologies, at least in the case of morphometric readings.

At this point in the analyses, it is worth mentioning that based on the results presented so far, the Grade ½ definition criterion, although perhaps capturing some of the more subtle height ratio reductions, probably inflates the estimates with substantial numbers of false-positives. Its inclusion incurs a loss of specificity that may be too substantial to be of practical value. It was therefore decided to discontinue the comparison of estimates with and without Grade ½ deformities in the remainder of the analyses. The subsequent analyses, which attempt to adjust the estimates for attrition and nonresponse bias, were therefore performed with the inclusion of only the Grade 1 and Grade 2 deformities.

4.4 - Estimates Derived by Multiple Imputation

The imputation models that were ultimately selected for each missing data group (censored, RQ, TR1, and TR2 groups), stratified by sex and age group are shown in **Tables 4.13** to **4.16**. Many of the independent variables were ultimately forced into the models, as the BIC criterion tends to choose small models, and these did not necessarily include variables known to be predictive of vertebral deformities. In a number of cases, the highest posterior probability estimated for the model by the BIC was one in which none of the candidate independent variables were present, meaning that the data may not be highly predictive of the measured outcome. Indeed, the majority of the coefficient estimates within the imputation models have 95% confidence intervals that include zero. Nevertheless, these variables can be useful in carrying out imputations, because if the direction of the coefficients is correct, the model will still adjust the possibly biased rates

in the correct direction, even if adjustments in the opposite direction occur in some iterations of the multiple imputation algorithm.

Estimates of incidence derived by multiple imputation compared with the observed rates in women and men are presented in **Table 4.17** and **4.18,** respectively. Across the different participation groups within each age group, the 95% credible intervals for incidence substantially overlap with each other. The age-specific, imputed estimates for women that differ most from the corresponding observed rates are those in the censored and RQ groups (**Table 4.17**). The estimated incidence rates in the censored females are substantially larger than the respondent estimates, whereas the female RQ group's estimates are smaller, suggesting that attrition and nonresponse biases may be in opposite directions in the female observed estimates, at least if the nonresponse is ignorable and if the imputation model is correct.

In all of the male age groups, the imputed incidence estimates are larger than in the respondents (**Table 4.18**). Especially for the 55-64 age group, the difference between the imputed estimates in the censored males (16.7/1000 PY, 95% CrI 6.3 - 30.5) and the observed rates (12.7/1000 PY, 95% CrI 6.8 - 19.5) is conspicuously large. For the RQ group, the imputed rate for men aged 55-64 (14.7/1000 PY, 95% CrI 5.0 - 28.3) is again the largest difference seen across the RQ age groups. The greater uncertainty in the imputed estimates is demonstrated by the narrower credible intervals seen within the male respondent age groups, compared to those, for example, of the RQ groups which all have larger sample sizes.

4.5 - Estimates Adjusted for Attrition Bias

If the missingness of the censored data is considered ignorable, a weighted average of the observed estimates pooled with those imputed in the censored can adjust for attrition bias. Under the ignorable assumption, the female adjusted estimates do not substantially differ from the crude incidence rates across age groups, the adjustments ranging between 3%

and 9% from the corresponding crude estimates (**Table 4.19**). Within the nonignorable models proposed, the female adjusted estimates are somewhat more divergent from the corresponding crude rates, especially in the 65-74 age group; the nonignorable + Δ_i model estimate (12.6/1000 PY, 95% CrI 6.8 - 19.3) and + $\frac{1}{2}\Delta_i$ model estimate (12.1/1000 PY, 95% CrI 7.6 - 17.2) differ substantially from the crude estimate (10.6/1000 PY, 95% CrI 7.5 - 13.9). These adjustments represent increases of 19% and 14%, respectively, from the crude rate. The larger adjustments in the female 65-74 age group suggest that the crude estimate is sufficiently biased to question its validity, if the nonignorable assumption holds and only attrition bias is considered. However, the ignorable assumption is reasonable, given the many covariates that were available for the imputation procedure. Moreover, the effect of attrition bias is small under all of the different models proposed for missingness, when the female age groups are collapsed into the 55+ group (3% to 12% increase from the crude rate).

Incidence estimates in men are affected more by attrition bias than in women, especially in the youngest age group (**Table 4.20**). The adjusted incidence rates in the 55-64 age group under the ignorable assumption (14.8/1000 PY, 95% CrI 8.5 - 22.6), the nonignorable + Δ_i model (16.9/1000 PY, 95% CrI 5.9 - 31.1), and the + $\frac{1}{2}\Delta_i$ model (15.8/1000 PY, 95% CrI 7.5 - 26.6) represent adjustment increments of 17%, 33%, and 24%, respectively, relative to the crude rate (12.7/1000 PY, 95% CrI 6.8 - 19.5). The observed incidence in elderly males (31.1/1000 PY, 95% CrI 21.0 - 42.1) may also be importantly biased by attrition, as the estimate for this group under the nonignorable + Δ_i model is adjusted to 37.1/1000 PY (95% CrI 9.9 - 69.7).

4.6 - Estimates Adjusted for Nonresponse and Attrition Bias

The available data on all contactees (respondent, censored, RQ, and TR groups) were pooled to provide overall estimates that are potentially adjusted for both nonresponse and attrition biases. These adjusted sex- and age-specific estimates are now presented under

different assumptions for the mechanism that is assumed to underlie the missingness of missing data in women (**Table 4.21**), and in men (**Table 4.22**).

Within the female age groups under the ignorable assumption, the point estimates of incidence are very similar to the crude rates. If the missing data are indeed ignorable, this suggests that the underestimation due to nonresponse bias in the crude female estimates is effectively cancelled out by the effect of attrition bias in the opposite direction. In contrast, only the estimates for the 65-74 age group are robust to the different nonignorable models proposed. In the + Δ_{ij} model, the adjustment increment is 16% for the 55-64 age group, and -5% for the elderly group. The largest adjustments made to the estimates in the + $\frac{1}{2}\Delta_{ij}$ and - $\frac{1}{2}\Delta_{ij}$ models are within the 55-64 age group: 9% and -6%, respectively.

In men, the variability in magnitude and direction of the adjusted rate estimates also demonstrate a dependence on the assumptions for missingness (**Table 4.22**). The adjusted rate estimates that differ the least from the corresponding crude age-specific rates are found under the ignorable assumption, while the + Δ_{ij} model for an nonignorable assumption yields rate estimates that most diverge from the crude rate estimates.

Although there were some important predictors of osteoporosis collected for the RQ group, it is unclear whether these subjects can be assumed to resemble the TR group, about whom very little is known, so that an assumption of ignorability is more uncertain than one of nonignorability. Hence, except for the female 65-74 age group, the magnitude and direction of the bias adjustments required depend on the mechanism of missingness that is proposed.

Table 4.1 - Comparison of variables available at baseline for all **female** CaMos contactees aged ≥50 years. Values are presented as numbers with percentages, or mean ± standard deviation.

Factors	Respondents n = 1 982 No.	%	Censored n = 2 478 No.	%	RQ group n = 3 726 No.	%	TR group* n = 1 062 No.	%
Age (mean ± SD)	64.3 ± 8.4		67.1 ± 9.0		70.6 ± 9.8		71.0 ± 10.8	
Race	1 885	95.1	2 402	97.2	3 504	94.3	NA	
History of fracture ≥age 50	453	22.9	697	28.3	581	18.4	NA	
History of osteoporosis in subject or family	474	24.2	569 n = 2 399	23.7	262	7.6	NA	
Current Smoker	246	12.4	340	13.8	507	16.9	NA	
≥ 1 Deformity at Baseline	222	11.2	386 n = 2 441	15.8	NA			
BMD spine	0.92 ± 0.17		0.90 ± 0.17		NA		NA	
BMD femoral neck	0.70 ± 0.12		0.68 ± 0.12		NA		NA	
Treatment for osteoporosis	140	7.1	213	8.6	NA		NA	
BMI kg/m^2	26.7 ± 5.0		27.3 ± 5.1		NA		NA	
Walks with difficulty/aid	156	7.9	428	17.3	NA		NA	
Number of telephone calls required to establish contact (mean ± SD)	2.1 ± 1.9		2.0 ± 1.9		1.9 ± 1.8		2.0 ± 1.9	
Number of household members ≥25 years of age (mean ± SD)	1.7 ± 0.6		1.7 ± 0.7		1.7 ± 0.8		1.7 ± 0.8 n = 1 026	
City								
Calgary	144	7.3	479	19.3	407	10.9	98	9.2
Hamilton	197	9.9	317	12.8	391	10.5	82	7.7
Halifax	2	0.1	536	21.6	474	12.7	125	11.8
Kingston	60	3.0	212	8.6	303	8.1	123	11.6
Quebec	370	18.7	121	4.9	277	7.4	44	4.1
St John's	12	0.6	457	18.4	435	11.7	155	14.6
Saskatoon	499	25.2	107	4.3	271	7.3	145	13.7
Toronto	270	13.6	114	4.6	520	14.0	223	21.0
Vancouver	428	21.6	135	5.5	648	17.4	67	6.3

NA = not applicable RQ = Refusal questionnaire TR = Total refuser

* Only the observed age and sex in the TR1 group within the overall TR are reported.

Table 4.2 - Comparison of variables available at baseline for all **male** CaMos contactees aged ≥50 years. Values are presented as numbers with percentages, or mean ± standard deviation.

Factors	Respondents n = 772 No.	%	Censored n = 991 No.	%	RQ group n = 1 793 No.	%	TR group* n = 681 No.	%
Age (mean ± SD)	64.3 ± 8.4		66.6 ± 9.5		68.3 ± 10.4		68.0 ± 11.0	
Race	719	93.1	930	94.6	1 630	91.2	NA	
History of fracture ≥age 50	117	15.2	146	14.9	428	25.8	NA	
History of osteoporosis in subject or family	70	9.1	92 n = 974	9.5	41	2.4	NA	
Current Smoker	95	12.3	176	17.9	296	20.3	NA	
≥ 1 Deformity at Baseline	94	12.2	142 n = 977	14.5	NA		NA	
BMD spine	1.03 ± 0.18		1.04 ± 0.17 n = 967		NA		NA	
BMD femoral neck	0.80 ± 0.12		0.78 ± 0.13 n = 962		NA		NA	
Treatment for osteoporosis	5	0.7	10	1.0	NA		NA	
BMI kg/m^2	27.0 ± 3.8		27.3 ± 3.9		NA		NA	
Walks with difficulty/aid	49	6.4	126	12.8	NA		NA	
Number of telephone calls required to establish contact (mean ± SD)	2.2 ± 1.9		2.0 ± 1.9		1.9 ± 1.8		2.1 ± 2.0	
Household members ≥age 25 (mean ± SD)	1.8 ± 0.6		1.9 ± 0.6		1.9 ± 0.7		1.8 ± 0.7 n = 661	
City								
Calgary	52	6.7	185	18.7	196	10.9	56	8.2
Hamilton	92	11.9	125	12.6	167	9.3	51	7.5
Halifax	0	0	220	22.2	224	12.5	89	13.1
Kingston	17	2.2	84	8.5	131	7.3	63	9.3
Quebec	140	18.3	48	4.8	114	6.4	18	2.6
St John's	2	0.3	162	16.4	217	12.1	113	16.6
Saskatoon	185	23.9	46	4.6	114	6.4	87	12.8
Toronto	120	15.5	55	5.6	324	18.1	158	23.2
Vancouver	164	21.2	66	6.7	306	17.1	46	6.8

NA = not applicable RQ = Refusal questionnaire TR = Total refuser

* Only the observed age and sex in the TR1 group within the overall TR are reported.

Table 4.3 – Baseline characteristics of subjects who sustained no incident deformity compared to those who sustained at least one incident vertebral deformity of **Grades 1 and 2 only**. Values are means ± SD, or numbers with percentages.

	All respondents	*No incident vertebral deformity*	*At least one incident vertebral deformity*
Number of subjects	2 754	2 542 (92.3)	212 (7.7)
Age	64.3 ± 8.4	63.9 ± 8.3	69.7 ± 7.8
Number of women	1 982 (72.0)	1 841 (72.4)	141 (66.5)
≥ 1 prevalent vertebral deformity at Year 0	316 (11.5)	240 (9.4)	76 (35.8)
BMD femoral neck	0.73 ± 0.13	0.73 ± 0.13	0.69 ± 0.13
BMD lumbar spine	0.95 ± 0.18	0.95 ± 0.18	0.91 ± 0.19
History of fracture ≥ age 50	570 (20.1)	485 (19.1)	85 (40.1)
Personal/Family History of Osteoporosis	544 (19.9) n = 2 728	481 (19.1) n = 2 517	63 (29.9) n = 211
History of treatment for osteoporosis	145 (5.3) n = 2 751	113 (4.5) n = 2 539	32 (15.1)
BMI kg/m^2	26.8 ± 4.7	26.7 ± 4.7	27.1 ± 5.0
Current smoker	341 (12.4)	321 (11.6)	20 (9.4)
Walking with difficulty or need a walking aid	205 (7.4)	171 (6.7)	34 (16.2)
Centre			
Vancouver	592 (21.5)	544 (21.4)	48 (22.6)
Calgary	196 (7.1)	181 (7.1)	15 (7.1)
Saskatoon	684 (24.8)	621 (24.4)	63 (29.7)
Hamilton	289 (10.5)	263 (10.3)	26 (12.3)
Toronto	390 (14.2)	371 (14.6)	19 (9.0)
Kingston	77 (2.8)	71 (2.8)	6 (2.8)
Quebec	510 (18.5)	477 (18.8)	33 (15.6)
Halifax	2 (0.1)	2 (0.1)	0
St John's	14 (0.5)	12 (0.5)	2 (0.9)

Table 4.4 – Baseline characteristics of subjects who sustained no incident deformity compared to those who sustained at least one incident vertebral deformity of **Grades ½ and higher**. Values are means ± SD, or numbers with percentages.

	All respondents	*No incident vertebral deformity*	*At least one incident vertebral deformity*
Number of subjects	2 754	2 384 (86.6)	370 (13.4)
Age	69.3 ± 8.4	63.7 ± 8.3	68.2 ± 8.3
Number of women	1 982 (72.0)	1 732 (72.7)	250 (67.6)
≥1 prevalent vertebral deformities at Year 0	550 (20.0)	398 (16.7)	152 (41.1)
BMD femoral neck	0.73 ± 0.13	0.73 ± 0.13	0.70 ± 0.13
BMD lumbar spine	0.95 ± 0.18	0.96 ± 0.17	0.91 ± 0.19
History of fracture ≥ age 50	570 (20.7)	454 (19.0)	116 (31.4)
Personal/Family History of Osteoporosis	544 (19.9) n = 2 728	454 (19.0) n =2 362	90 (24.6) n = 366
History of treatment for osteoporosis	145 (5.3) n = 2 751	105 (4.4) n = 2 381	40 (10.8) n = 370
BMI kg/m^2	26.8 ± 4.7	26.8 ± 4.7	26.8 ± 4.4
Current smoker	341 (12.4)	302 (12.7)	39 (10.5)
Walking with difficulty or need a walking aid	205 (7.4)	161 (6.8)	44 (11.9)
Centre			
Vancouver	592 (21.5)	211 (21.4)	81 (21.9)
Calgary	196 (7.1)	176 (7.4)	20 (5.4)
Saskatoon	684 (24.8)	575 (24.1)	109 (29.5)
Hamilton	289 (10.5)	236 (9.9)	53 (14.3)
Toronto	390 (14.2)	349 (14.6)	41 (11.1)
Kingston	77 (2.8)	69 (2.9)	8 (2.2)
Quebec	510 (18.5)	454 (19.1)	56 (15.1)
Halifax	2 (0.1)	2 (0.1)	0
St John's	14 (0.5)	12 (0.5)	2 (0.5)

Table 4.5 – Number of men and women with incident vertebral deformities, **Grades 1 and 2 only**, stratified by pre-existing deformity status at baseline. Values are numbers with percentages.

	Men	*Women*
No prevalent deformity		
No incident deformity	583 (94.5)	1 536 (96.8)
One incident deformity	32 (5.2)	44 (2.8)
More than one incident deformity	2 (0.3)	7 (0.4)
At least one prevalent deformity		
No incident deformity	118 (76.1)	305 (77.2)
One incident deformity	32 (20.7)	72 (18.2)
More than one incident deformity	5 (3.2)	18 (4.6)

Table 4.6 – Number of men and women with incident vertebral deformities of **Grades ½ and higher**, stratified by pre-existing deformity status at baseline. Values are numbers with percentages.

	Men	*Women*
No prevalent deformity		
No incident deformity	537 (87.0)	1 449 (91.3)
One incident deformity	71 (11.5)	120 (7.6)
More than one incident deformity	9 (1.5)	18 (1.1)
At least one prevalent deformity		
No incident deformity	115 (74.2)	283 (71.7)
One incident deformity	34 (21.9)	87 (22.0)
More than one incident deformity	6 (3.9)	25 (6.3)

Table 4.7 – Bayesian estimates of age-specific incidence per 1000 PY in **women**, stratified by the presence of vertebral deformities at baseline. Noninformative prior distributions were used for all estimates. Vertebral deformities of only **Grades 1 and 2 were included in these analyses.**

	No. of women	No. of new deformities[1]	No. of women with ≥ 1 new deformities	No. of PY	Incidence point estimate	95% CrI[2]
Overall						
55-64 years	634	11	10	3 129.7	3.2	1.3 - 5.4
65-74 years	798	47	41	3 863.8	10.6	7.5 - 13.9
75 +	550	117	90	2 514.7	35.7	28.5 - 43.2
55+[3]	1 982	175	141	9 508.2	14.6	12.2 - 17.1
No prevalent deformity						
55-64 years	609	10	9	3 008.1	3.0	1.4 - 5.3
65-74 years	712	29	26	3 472.8	7.5	4.9 - 10.6
75+	439	63	55	2 048.0	26.8	20.2 - 34.4
55+[3]	1 760	102	90	8 528.9	10.4	8.3 - 12.6
≥1 Prevalent deformity						
55-64 years	25	1	1	121.6	8.2	0.2 - 30.3
65-74 years	86	18	15	391.0	38.4	21.3 - 60.1
75+	111	54	35	466.7	75.2	52.1 - 102.1
55+[3]	222	73	51	979.3	53.0	39.5 - 68.5

[1]Represents the total number of incident deformities of Grades 1 and 2 in each group *i.e.* those subjects who had ≥ 1 incident deformities contributed ≥ 1 times to the count.
[2]95% credible intervals.
[3]Weighted to the 1996 Canadian population census.

NOTE: As the time of occurrence for deformities is unknown, the time for follow-up for cases was taken as the halfway mark between the dates of baseline and Year-5 radiographs.

Table 4.8 – Bayesian estimates of age-specific incidence per 1000 PY in **men**, stratified by the presence of vertebral deformities at baseline. Noninformative prior distributions were used for all estimates. Vertebral deformities of only **Grades 1 and 2 were included in these analyses**.

	No. of men	No. of new deformities[1]	No. of men with ≥ 1 new deformities	No. of PY	Incidence point estimate	95% CrI[2]
Overall						
55-64 years	244	20	15	1 179.7	12.7	6.8 - 19.5
65-74 years	313	26	25	1 494.8	16.6	10.7 - 23.3
75 +	215	37	31	990.6	31.1	21.0 - 42.1
55+[3]	772	83	71	3 665.0	17.7	13.5 - 22.1
No prevalent deformity						
55-64 years	221	14	10	1 078.1	9.2	4.4 - 15.7
65-74 years	277	19	18	1 332.8	13.5	8.0 - 20.4
75+	180	23	18	850.5	21.2	12.6 - 32.2
55+[3]	678	56	46	3 260.4	13.0	9.4 - 17.2
≥1 Prevalent deformity						
55-64 years	23	6	5	101.6	49.5	16.1 - 101.9
65-74 years	36	7	7	162.0	43.3	17.3 - 80.8
75+	35	22	13	141.1	92.2	49.1 - 148.6
55+[3]	94	27	25	404.7	59.2	37.7 - 85.9

[1]Represents the total number of incident deformities of Grades 1 and 2 in each group *i.e.* those subjects who had ≥ 1 incident deformities contributed ≥ 1 times to the count.
[2]95% credible intervals.
[3]Weighted to the 1996 Canadian population census.

NOTE: As the time of occurrence for deformities is unknown, the time for follow-up for cases was taken as the half-way mark between the dates of baseline and Year-5 radiographs.

Table 4.9 – Bayesian estimates of age-specific incidence per 1000 PY in **women**, stratified by the presence of vertebral deformities at baseline. Noninformative prior distributions were used for all estimates. Vertebral deformities of **Grades ½, 1, and 2 were included in these analyses**.

	No. of women	No. of new deformities[1]	No. of women with ≥1 new deformities	No. of PY	Incidence point estimate	95% CrI[2]
Overall						
55-64 years	634	41	38	3 059.4	12.4	8.8 - 16.7
65-74 years	798	98	82	3 761.5	21.8	17.3 - 26.8
75 +	550	169	130	2 414.2	53.8	44.9 - 63.4
55+[3]	1 982	308	250	9 235.1	26.9	23.6 - 30.4
No prevalent deformity						
55-64 years	570	32	30	2 762.8	10.9	7.3 - 15.1
65-74 years	643	51	45	3 082.4	14.6	10.7 - 19.2
75+	374	77	63	1 704.4	36.9	28.3 - 46.7
55+[3]	1 587	160	138	7 549.5	18.2	15.2 - 21.4
≥1 Prevalent deformity						
55-64 years	64	9	8	296.7	27.0	11.7 - 48.8
65-74 years	155	47	37	679.0	54.5	38.2 - 73.3
75+	176	92	67	709.8	94.6	73.3 - 118.4
55+[3]	395	148	112	1 685.5	67.1	55.1 - 80.2

[1]Represents the total number of incident deformities of Grades ½ and higher in each group *i.e.* those subjects who had ≥ 1 incident deformities contributed ≥1 times to the count.
[2]95% credible intervals.
[3]Weighted to the 1996 Canadian population census.

NOTE: As the time of occurrence for deformities is unknown, the time for follow-up for cases was taken as the half-way mark between the dates of baseline and Year-5 radiographs.

Table 4.10 - Bayesian estimates of age-specific incidence per 1000 PY in **men**, stratified by the presence of vertebral deformities at baseline. Noninformative prior distributions were used for all estimates. Vertebral deformities of **Grades ½, 1, and 2 were included in these analyses**.

	No. of men	No. of new deformities[1]	No. of men with ≥1 new deformities	No. of PY	Incidence point estimate	95% CrI[2]
Overall						
55-64 years	244	36	29	1 144.8	25.3	16.9 - 35.3
65-74 years	313	51	44	1 447.3	30.4	22.0 - 39.9
75 +	215	60	47	950.7	49.4	36.2 - 64.4
55+[3]	772	147	120	3 542.8	31.8	26.2 - 38.0
No prevalent deformity						
55-64 years	205	29	22	968.0	22.7	14.2 - 33.3
65-74 years	249	31	29	1 166.4	24.9	16.7 - 34.6
75+	163	38	29	737.3	39.3	26.4 - 54.6
55+[3]	617	98	80	2 871.7	26.6	20.9 - 33.0
≥1 Prevalent deformity						
55-64 years	39	7	7	176.8	39.4	16.0 - 73.3
65-74 years	64	20	15	281.0	53.5	29.9 - 83.8
75+	52	22	18	213.4	84.5	50.3 - 127.4
55+[3]	155	49	40	671.2	55.8	39.5 - 75.2

[1]Represents the total number of incident deformities of Grades ½ and higher in each group *i.e.* those subjects who had ≥ 1 incident deformities contributed ≥1 times to the count.
[2]95% credible intervals with uninformative priors.
[3]Weighted to the 1996 Canadian population census.

NOTE: As the time of occurrence for deformities is unknown, the time for follow-up for cases was taken as the half-way mark between the dates of baseline and Year-5 radiographs.

Table 4.11 – Frequency distribution of deformities by vertebral level in men and women, **Grades 1 and 2 only**.

	T4	*T5*	*T6*	*T7*	*T8*	*T9*	*T10*	*T11*	*T12*	*L1*	*L2*	*L3*	*L4*	*Total*
Male	3	4	7	13	5	5	1	5	7	17	9	4	3	83
Female	5	17	12	18	22	9	15	16	17	19	12	8	5	175

Figure 4.1

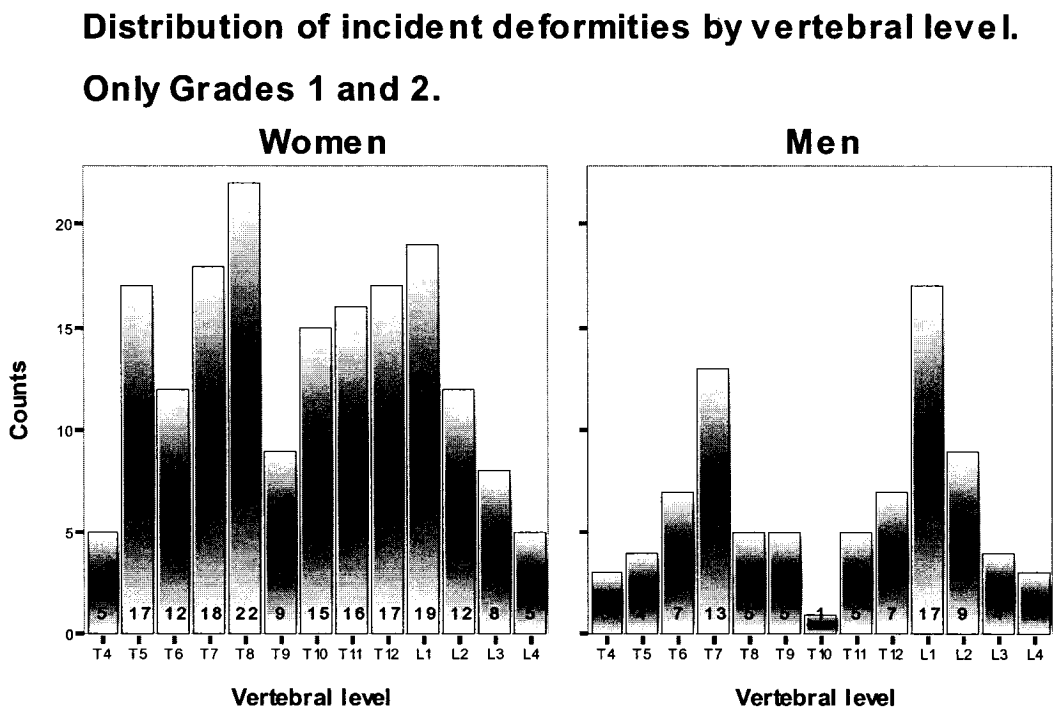


Table 4.12 – Frequency distribution of incident deformities by vertebral level in men and women, including **Grades ½, 1, and 2**.

	T4	*T5*	*T6*	*T7*	*T8*	*T9*	*T10*	*T11*	*T12*	*L1*	*L2*	*L3*	*L4*	*Total*
Male	7	10	9	23	14	11	3	9	15	22	15	5	4	147
Female	18	26	22	34	31	22	22	25	25	27	24	24	8	308

Figure 4.2

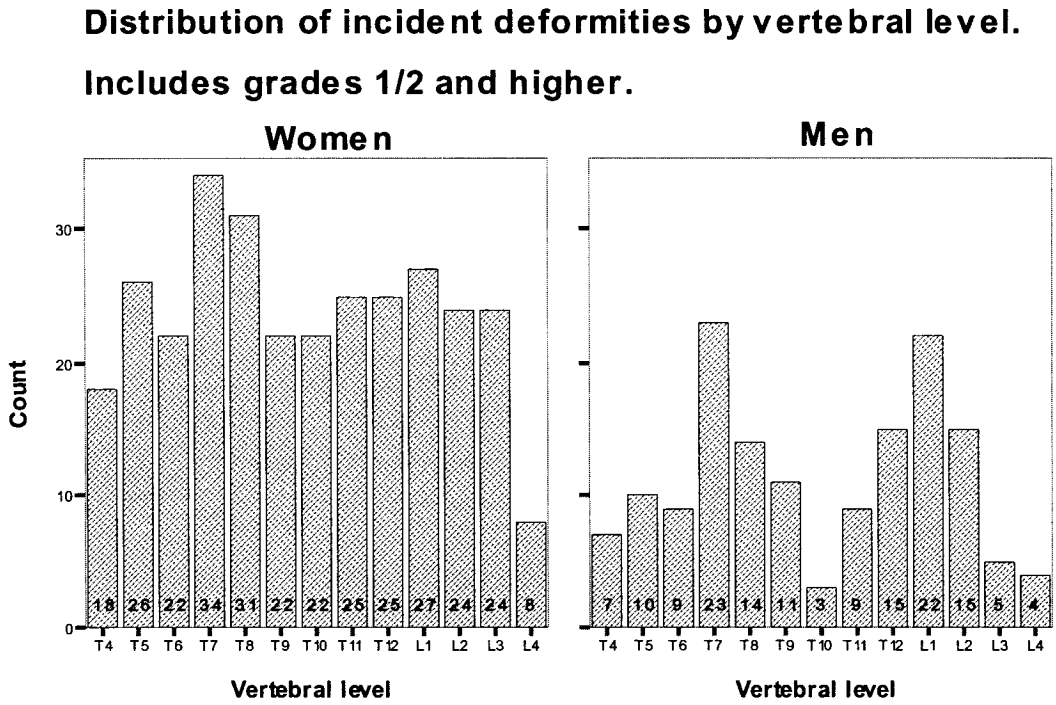


Table 4.13 - Coefficient estimates and Wald 95% confidence intervals of the unconditional logistic regression models that were selected as imputation models for the censored groups. Independent variables that were forced into the models have their coefficients highlighted in grey, whereas other covariates were selected by the BIC. Candidate variables for the models that were ultimately not selected are also shown.

Independent variables available for model selection	Female age groups (years)			Male age groups (years)		
	55-64	65-74	75+	55-64	65-74	75+
Intercept	-2.20 (-18.0, 13.6)	-11.57 (-20.0, -2.9)	-0.84 (-5.7, 4.0)	-8.34 (-22.5, 5.8)	-5.39 (-15.8, 5.0)	0.65 (-7.2, 8.5)
Previous deformity	0.86 (-1.4, 3.1)	1.56 (0.9, 2.3)	0.84 (0.3, 1.4)	1.86 (0.5, 3.2)	1.08 (0.08, 2.1)	1.71 (0.8, 2.6)
Age	-0.05 (-0.3, 0.2)	0.11 (-0.01, 0.2)	0.001 (-0.06, 0.06)	0.02 (-0.2, 0.2)	0.04 (-0.1, 0.2)	-0.04 (-0.1, 0.06)
History of fracture ≥50 years	0.91 (-0.8, 2.7)	0.47 (-0.2, 1.6)	0.49 (0.01, 1.0)	0.67 (-0.7, 2.0)	0.77 (-0.2, 1.7)	0.30 (-0.6, 1.2)
Family/Personal history of osteoporosis	1.03 (-0.3, 2.4)	0.42 (-0.3, 1.1)		1.28 (-0.1, 2.7)	0.18 (-1.1, 1.5)	
BMD lumbar spine	0.30 (-4.6, 5.2)		-1.66 (-3.1, -0.2)	-1.0 (-5.4, 3.3)		
Treatment for osteoporosis			0.90 (0.3, 1.5)			
BMI				0.16 (0.03, 0.3)		
BMD femoral neck						
Walking with aid/difficulty						
Caucasian						
Smoker						
Household members ≥25 years						
Telephone calls to contact						
Study centre						

Table 4.14 - Coefficient estimates and Wald 95% confidence intervals of the unconditional logistic regression models that were selected as imputation models for the Refusal Questionnaire (RQ) groups. Independent variables that were forced into the models have their coefficients highlighted in grey, whereas other covariates were selected by the BIC. Candidate variables for the models that were ultimately not selected are also shown.

Independent variables available for model selection	**Female agegroups (years)**			**Male agegroups (years)**		
	55-64	**65-74**	**75+**	**55-64**	**65-74**	**75+**
Intercept	-4.63 (-18.1, 8.8)	-12.28 (-20.9, -3.7)	-2.66 (-7.3, 2.0)	-6.20 (-18.0, 5.6)	-7.06 (-17.2, 3.1)	-2.43 (-9.9, 5.1)
Age	-0.001 (-0.2, 0.2)	0.13 (0.01, 0.3)	0.01 (-0.05, 0.06)	0.05 (-0.1, 0.2)	0.06 (-0.1, 0.2)	.01 (-0.1, 0.1)
History of fracture ≥50 years	0.77 (-0.9, 2.5)	0.60 (-0.07, 1.3)	0.70 (0.3, 1.2)	1.0 (-0.3, 2.2)	0.93 (-0.01, 1.9)	0.39 (-0.5, 1.3)
Family/Personal history of osteoporosis	1.3 (0.03, 2.5)	0.63 (-0.03, 1.3)	0.59 (0.1, 1.1)	1.1 (-0.3, 2.4)	0.17 (-1.1, 1.4)	
Caucasian						
Smoker						
Household members ≥25 years						
Telephone calls to contact						
Study centre						

Table 4.15 - Coefficient estimates and Wald 95% confidence intervals of the unconditional logistic regression models that were selected as imputation models for the Total Refuser groups for which data on sex, age, number of household members were available (**TR1 group**).

Independent variables available for model selection	**Female agegroups (years)**			**Male agegroups (years)**		
	55-64	**65-74**	**75+**	**55-64**	**65-74**	**75+**
Intercept	7.59 (-20.7, 5.5)	-11.1 (-19.9, -2.2)	-2.70 (-7.4, 2.0)	-6.45 (-17.8, 4.9)	-7.18 (-17.6, 3.2)	-1.91 (-9.6, 5.8)
Age	0.03 (-0.2, 0.2)	0.13 (0.01, 0.3)	0.01 (-0.04, 0.1)	0.06 (-0.1, 0.3)	0.07 (-0.1, 0.2)	0.01 (-0.9, 0.1)
Household members ≥25 years	-0.21 (-1.3, 0.9)	-0.47 (-1.0, 0.1)	-0.10 (-0.5, 0.3)	-0.05 (-1.0, 0.9)	-0.02 (-0.7, 0.7)	-0.24 (-1.0, 0.5)
Telephone calls to contact						
Study centre						

Table 4.16 - Coefficient estimates and Wald 95% confidence intervals of the unconditional logistic regression models that were selected as imputation models for the Total Refuser groups for which data on sex, age, number of household members were not available (**TR2 group**). Single imputation was utilized to infer data on sex and gender.

Independent variables available for model selection	**Female agegroups (years)**			**Male agegroups (years)**		
	55-64	**65-74**	**75+**	**55-64**	**65-74**	**75+**
Intercept	-5.49 (-18.4, 7.4)	-12.70 (-21.4, -4.1)	-3.03 (-7.5, 1.5)	-6.52 (-17.8, 4.7)	-7.10 (-17.2, 3.0)	-2.51 (-10.0, 8.0)
Age	0.02 (-0.2, 0.2)	0.14 (0.02, 0.3)	0.02 (-0.04, 0.1)	0.06 (-0.1, 0.3)	0.07 (-0.1, 0.2)	0.01 (-0.1, 0.1)
Telephone calls to contact	-0.02 (-0.4, 0.3)	-0.07 (-0.3, 0.1)	-0.01 (-0.2, 0.1)			
Household members ≥25 years						
Study centre						

Table 4.17 - Observed and imputed estimates of incidence per 1000 PY in **women** contacted by the CaMos. Shown are estimates based on observed vertebral deformities (respondent group), and estimates derived by multiple imputation in the different participation groups (censored, RQ, and TR groups).

	Age Range (years)			
	55-64	**65-74**	**75+**	**55+***
Respondent group				
N	634	798	550	1 982
Incidence	3.2	10.6	35.7	14.6
95% CrI	1.3 - 5.4	7.5 - 13.9	28.5 - 43.2	12.2 - 17.1
Censored group				
N	657	962	494	2 113
Incidence	3.5	12.3	39.2	16.3
95% CrI	0.6 - 8.0	7.0 - 18.5	26.4 - 54.3	12.1 - 21.1
RQ group				
N	647	1 115	982	2 744
Incidence	2.9	8.9	26.2	11.3
95% CrI	0.3 - 6.9	4.7 - 14.1	17.0 - 36.8	8.2 - 14.9
TR group 1				
N	203	357	392	952
Incidence	3.3	10.4	37.4	15.0
95% CrI	0.0 - 10.0	4.0 - 18.6	23.1 - 54.4	10.0 - 20.8
TR group 2**				
N	517	532	571	1 620
Incidence	3.2	11.2	38.3	15.5
95% CrI	0.4 - 7.5	5.4 - 18.6	24.6 - 54.6	11.0 - 20.7

* Weighted to the 1996 Canadian census.

** Total refuser group 2 = subjects with imputed age and sex.

95% CrI = 95% credible interval.

Table 4.18 – Observed and imputed estimates of incidence per 1000 PY in **men** contacted by the CaMos. Shown are estimates based on observed vertebral deformities (respondent group), and estimates derived by multiple imputation in the different participation groups (censored, RQ, and TR groups).

	Age Range (years)			
	55-64	**65-74**	**75+**	**55+***
Respondent group				
N	244	313	215	772
Incidence	12.7	16.6	31.1	17.7
95% CrI	6.8 - 19.5	10.7 - 23.3	21.0 - 42.1	13.5 - 22.1
Censored group				
N	262	358	209	829
Incidence	16.7	17.8	34.5	20.6
95% CrI	6.3 - 30.5	8.0 - 30.2	17.0 - 56.1	13.3 - 28.8
RQ group				
N	369	546	449	1 364
Incidence	14.7	18.2	31.9	19.3
95% CrI	5.0 - 28.3	9.0 - 29.9	18.0 - 48.8	12.8 - 27.2
TR group 1				
N	164	207	201	572
Incidence	13.0	16.8	32.1	18.1
95% CrI	2.5 - 28.0	5.9 - 31.6	14.6 - 54.1	10.4 - 27.3
TR group 2**				
N	439	393	400	1 232
Incidence	13.0	16.9	32.9	18.3
95% CrI	5.1 - 23.8	7.8 - 28.7	16.2 - 55.8	12.1 - 25.8

* Weighted to the 1996 Canadian census.

** Total refuser group 2 = subjects with imputed age and sex.

95% CrI = 95% credible interval.

Table 4.19 - Sensitivity analysis of estimates adjusted for attrition bias among **women** lost to follow-up (censored group) in the CaMos. Multiple imputation-adjusted estimates of incidence per 1000 PY are compared under different assumptions for missingness. Bias-unadjusted, age-specific estimates for the respondents are included as a reference.

	Female age groups			
	55-64	**65-74**	**75+**	**55+***
Respondents (crude)				
N	634	798	550	1 982
Incidence	3.2	10.6	35.7	14.6
95% CrI	1.3 - 5.4	7.5 - 13.9	28.5 - 43.2	12.2 - 17.1
Respondents + Censored (ignorable)				
N	1 283	1 746	1 041	4 070
Incidence	3.3	11.6	37.4	15.4
95% CrI	1.6 - 5.8	8.3 - 15.3	30.0 - 45.3	13.0 - 18.1
Respondents + Censored (nonignorable + Δ_i)				
N	1 283	1 746	1 041	4 070
Incidence	3.5	12.6	39.0	16.3
95% CrI	0.5 - 8.1	6.8 - 19.3	26.8 - 53.4	12.2 - 21.0
Respondents + Censored (nonignorable + $\frac{1}{2}\Delta_i$)				
N	1 283	1 746	1 041	4 070
Incidence	3.4	12.1	38.2	15.9
95% CrI	1.1 - 6.8	7.6 - 17.2	28.7 - 49.1	12.7 - 19.5
Respondents + Censored (nonignorable - $\frac{1}{2}\Delta_i$)				
N	1 283	1 746	1 041	4 070
Incidence	3.3	11.1	36.6	15.0
95% CrI	1.7 - 5.1	8.3 - 14.0	30.2 - 43.3	12.9 - 17.2

* Weighted to the 1996 Canadian census.
95% CrI = 95% credible interval.
Δ_i = the absolute value of the mean difference between the observed$_i$ and the imputed rate$_i$.

Table 4.20 - Sensitivity analysis of estimates adjusted for attrition bias among **men** lost to follow-up (censored group) in the CaMos. Multiple imputation-adjusted estimates of incidence per 1000 PY are compared under different assumptions for missingness. Bias-unadjusted, age-specific estimates for the respondents are included as a reference.

	Male age groups			
	55-64	**65-74**	**75+**	**55+***
Respondents (crude)				
N	244	313	215	772
Incidence	12.7	16.6	31.1	17.7
95% CrI	6.8 - 19.5	10.7 - 23.3	21.0 - 42.1	13.5 - 22.1
Respondents + Censored (ignorable)				
N	503	669	209	1 381
Incidence	14.8	17.3	32.8	19.2
95% CrI	8.5 - 22.6	11.1 - 24.5	22.4 - 44.7	14.9 - 24.0
Respondents + Censored (nonignorable + Δ_i)				
N	503	669	209	1 381
Incidence	16.9	17.6	37.1	21.1
95% CrI	5.9 - 31.3	7.1 - 30.8	9.9 - 69.7	12.4 - 31.0
Respondents + Censored (nonignorable + $\frac{1}{2}\Delta_i$)				
N	503	669	209	1 381
Incidence	15.8	17.3	34.1	20.0
95% CrI	7.4 - 26.7	9.4 - 27.3	20.3 - 50.5	14.2 - 26.5
Respondents + Censored (nonignorable - $\frac{1}{2}\Delta_i$)				
N	503	669	209	1 381
Incidence	13.8	16.9	32.1	18.4
95% CrI	8.3 - 19.8	11.5 - 22.7	22.9 - 42.2	14.8 - 22.4

* Weighted to the 1996 Canadian census.
95% CrI = 95% credible interval.
Δ_i = the absolute value of the mean difference between the observed$_i$ and the imputed rate$_i$.

Table 4.21 - Sensitivity analysis of estimates adjusted for nonresponse and attrition biases among **female** CaMos contactees. Multiple imputation-adjusted estimates of incidence per 1000 PY are compared under different assumptions for missingness. Bias-unadjusted, age-specific estimates for the respondents are included as a reference.

	Age range (years)			
	55-64	**65-74**	**75+**	**55+***
Respondents (crude)				
N	634	798	550	1 982
Incidence	3.2	10.6	35.7	14.6
95% CrI	1.3 - 5.4	7.5 - 13.9	28.5 - 43.2	12.2 - 17.1
All contactees (ignorable)				
N	2 658	3 764	2 989	9 411
Incidence	3.2	10.6	33.9	14.1
95% CrI	1.8 - 4.9	8.3 - 13.2	28.7 - 39.5	12.4 - 16.0
All contactees (nonignorable + Δ_{ij})				
N	2 658	3 764	2 989	9 411
Incidence	3.7	10.7	33.9	14.4
95% CrI	2.0 - 6.1	7.6 - 14.5	26.1 - 43.4	11.8 - 17.4
All contactees (nonignorable + $\frac{1}{2}\Delta_{ij}$)				
N	2 658	3 764	2 989	9 411
Incidence	3.5	10.7	33.8	14.2
95% CrI	1.9 - 5.5	8.0 - 13.8	27.5 - 41.2	12.1 - 16.6
All contactees (nonignorable - $\frac{1}{2}\Delta_{ij}$)				
N	2 658	3 764	2 989	9 411
Incidence	3.0	10.6	33.7	14.0
95% CrI	1.6 - 4.7	8.3 - 13.1	28.8 - 39.0	12.3 - 15.8

* Weighted to the 1996 Canadian census.
95% CrI = 95% credible interval.
Δ_{ij} = the absolute value of the mean difference between the observed$_i$ and the imputed rate$_{ij}$ of a particular missing data group.

Table 4.22 - Sensitivity analysis of estimates adjusted for nonresponse and attrition biases among **male** CaMos contactees. Multiple imputation-adjusted estimates of incidence per 1000 PY are compared under different assumptions for missingness. Bias-unadjusted, age-specific estimates for the respondents are included as a reference.

	Age range (years)			
	55-64	**65-74**	**75+**	**55+***
Respondents (crude)				
N	244	313	215	772
Incidence	12.7	16.6	31.1	17.7
95% CrI	6.8 - 19.5	10.7 - 23.3	21.0 - 42.1	13.5 - 22.1
All contactees (ignorable)				
N	1 478	1 817	1 474	4 769
Incidence	14.0	17.4	32.5	18.8
95% CrI	9.5 - 19.3	12.8 - 22.6	24.7 - 41.3	15.7 - 22.2
All contactees (nonignorable + Δ_{ij})				
N	1 478	1 817	1 474	4 769
Incidence	18.5	21.6	40.0	23.8
95% CrI	12.4 - 27.1	15.7 - 29.6	30.1 - 54.0	19.6 - 29.0
All contactees (nonignorable + $\frac{1}{2}\Delta_{ij}$)				
N	1 478	1 817	1 474	4 769
Incidence	16.3	19.5	36.3	21.3
95% CrI	11.1 - 23.0	14.4 - 25.7	27.7 - 47.3	17.8 - 25.5
All contactees (nonignorable - $\frac{1}{2}\Delta_{ij}$)				
N	1 478	1 817	1 474	4 769
Incidence	11.7	15.1	28.8	16.3
95% CrI	7.2 - 16.4	10.4 - 19.8	20.7 - 36.7	13.2 - 19.4

* Weighted to the 1996 Canadian census.
95% CrI = 95% credible interval.
Δ_{ij} = the absolute value of the mean difference between the observed$_i$ and the imputed rate$_{ij}$ of a particular missing data group.

CHAPTER 5 – SUMMARY AND CONCLUSIONS

5.1 - Adjusted Incidence Estimates

The male incidence estimates, whether crude or bias-adjusted, are larger than the female estimates in the 55-64 and 65-74 age groups. Somewhat congruent with these results is that the CaMos has previously reported a higher prevalence of vertebral deformities in Canadian men aged 50-59 compared to women[96]. However, the contributions made by confounding and ascertainment methods in the association of sex with incident vertebral deformities needs to be clarified in future research.

There are noticeable, and in some cases, large effects of nonresponse and attrition biases on the crude incidence estimates, but in most cases the magnitude and direction of required adjustments depend on the model put forward for the missing data mechanism. Although the validity of the latter cannot ever be verified, one can make a heuristic choice of one model over another, in order to report "best guess" summary rates.

In the context of this study, an assumption of nonignorability is more conservative than one of ignorable missing data. Although several covariates known to be predictive of deformities were available for the construction of imputation models in the censored and RQ groups, there is greater uncertainty in the imputed estimates for the TR groups, about whom little is known. The possibility of a nonignorable missing data mechanism is thus salient. Another reasonable conjecture is that the direction of bias is in the same direction as the adjustment derived by multiple imputation, and in this case the $+\Delta_{ij}$ model is the more conservative choice. This study therefore "best estimates" that the incidence rates of vertebral deformities in Canadians aged 55+ years, adjusted for nonresponse and attrition biases, are 14.4/1000 PY (95% CrI 11.8 - 17.4) in women and 23.8/1000 PY (95% CrI 19.6 - 29.0) in men. Although this suggests that the male respondent estimate is underestimated by 34% in Canadian males aged 55+, the overall adjusted estimate proposed for women is in effect identical to the corresponding female respondent rate.

5.2 - Comparisons with Other Studies

It is difficult to compare the findings of this study with those of similar longitudinal studies, as there are several obstacles that arise quite apart from the differences that are inherent to the investigated populations. First, the Rotterdam study[152] reported incidence as the number of vertebral deformities per 1000 PY, as opposed to the EPOS and this study, which defined incidence as the number of individuals with ≥1 new deformities. Second, differences in the distribution of PY for the different age groups across studies can also affect the rate estimates; the rate of change in incidence is not linear over time, as once a vertebral deformity is present, the spine becomes increasingly vulnerable to re-fracturing with each subsequent new deformity. Third, the EPOS[151] did not stratify their sex- and age-specific estimates by pre-existing deformity status, when the latter is at least, if not more predictive of incidence than age or sex. Fourth, although both European studies tried to account for the effect of nonresponse bias, the Rotterdam Study only compared baseline characteristics of cases and non-cases in a univariate analysis, and the EPOS adjusted their estimates using deterministic imputation without considering missing data mechanism. Neither study attempted to assess the effect of loss to follow-up, despite similar attrition to the CaMos. Last, and perhaps most important, the CaMos methods used to define and assess the outcome differed in important aspects from the methods chosen by the EPOS and Rotterdam studies.

Nevertheless, all three studies have observed some commonalities. In both sexes there is steep rise in the rate of increase in incidence in elderly subjects, and this trend endures when subjects have pre-existing deformities. The presence of vertebral deformities is a strong risk factor for re-fracture across all age groups and in both sexes.

5.3 - Strengths and weaknesses

One of the strengths of the CaMos is that the normative values for vertebral height ratios for each sex- and level-specific vertebra were derived from a sub-sample of the CaMos participants, and that a standardized and reliable protocol was used to define and ascertain the outcome. Given the lack of consensus on these issues, the CaMos used methods that were considered state-of-the-art during the time of its conception. However, the enormous number of automated measurements made (3 heights per vertebra, 14 vertebrae per person, measured at baseline and at Year 5 in a sample of 2 754 subjects), makes it inevitable that some normal vertebrae were classified as abnormal by measurement error or chance alone.

This thesis investigated the effect of including into the analyses deformities of Grade ½, which are those vertebrae with height ratios that fall within the estimated measurement standard error. Their inclusion distorts the typical frequency distribution of incident deformities that is usually observed along the spine and unduly inflates the estimates. In some cases including the Grade ½ deformities alters the relationship among age-specific rates between the two sexes, as well as the direction and magnitude of the rate of change in incidence across age groups. It is therefore likely that the CaMos' attempt to correct the original normative values for measurement error shifted the deformity definition towards greater specificity.

On the other hand, the Eastell method, which was utilized for morphometric assessment in the CaMos, is known to be less specific than other techniques, such as the McCloskey method. The CaMos definition criteria for a deformity requires the reduction of any one vertebral height ratio, whereas the NOF has since recommended that a morphometric definition include at least two reduced ratios within a vertebra. Additionally, both the EPOS and the Rotterdam Studies confirmed all deformities that were initially identified by morphometry with a visual inspection by a radiologist, thus further improving the positive predictive value of their assessments. Roughly 30% of the CaMos respondents analyzed for this thesis were aged 75+ years, and it is recognized that certain common

age-related pathologies, like osteoarthritis, can mimic osteoporotic changes on x-ray assessed by morphometry.

Another major strength of the CaMos is that an attempt was made to collect data predictive of osteoporosis on nonrespondents, in view of adjusting estimated parameters for nonresponse bias. This permitted the use of principled imputation techniques, so that a range of plausible estimates adjusted for both nonresponse and attrition biases could be reported.

5.4 - Future Research

Although the analyses within this thesis are fairly extensive, there remains potential for more research utilizing these data. For example, to account for important covariates other than age and sex, such as pre-existing deformities and BMD, the incidence estimates could be further adjusted using a regression model; the effect of sex on the estimates could then be more accurately reported. The distributions of deformities by grade of severity could also be explored, as well as their associations with functional impairment. Most of the x-ray data that were slower to be entered into the database originate from 3 study centres, and although unlikely, future analyses that utilize all of the available data and account for random effects may yield different estimates than the ones reported.

The extent of bias due to outcome ascertainment could also be more thoroughly assessed in a future study. The CaMos vertebral height ratio data could be re-analyzed by altering the protocol for defining and/or radiologically assessing vertebral deformities. For example, an algorithm could be used to redefine a deformity with more than one height ratio reduction criteria. Alternatively, the McCloskey method could be utilized, as this method predicts normative values from the adjacent vertebrae within individuals, and definition criteria for deformity are more stringent. With either option, specificity could thus be improved, albeit at the expense of sensitivity. A visual confirmation by an expert radiologist of all the positive vertebrae flagged by morphometry, although costly, could

also ensure that differential and artefactual causes are excluded. It would be impractical in an epidemiologic survey to apply semi-quantitative techniques to the entire sample, but reading a sub-sample of the radiographs with this approach could enable a comparison with morphometric results. A semi-quantitative reading of a sub-sample could also be used to scrutinize the frequency and the radiological attributes of fracture rebound (morphometrically detected deformities that disappear on a subsequent reading).

REFERENCES

1. Hennekens CH, Buring JE (1987) *Epidemiology in Medicine*. Philadelphia: Lippicott Williams & Wilkins, 383 pages.

2. Eisman JA (1999) *Genetics of osteoporosis*. Endocrine Reviews 20: 788-804.

3. Black DM, Cooper C (2000) *Epidemiology of fractures and assessment of fracture risk*. Clinics in Laboratory Medicine 20(3): 439-453.

4. Cooper C (1999) *Epidemiology of osteoporosis*, Osteoporosis International, Suppl 2: S2-S8.

5. Christodoulou C, Cooper C (2003) *What is osteoporosis*. Postgraduate Medical Journal 79: 133-138.

6. Cummings SR, Melton LJ III (2002) *Epidemiology and outcomes of osteoporotic fractures*. Lancet 359: 1761-1767.

7. Dennison E, Cooper C (2000) *Epidemiology of osteoporotic fractures*. Hormone Research 54(Suppl 1): 58-63.

8. Karsh J (2001) *Diagnostic challenges in osteoporosis*. Canadian Family Physician 47: 1244-1250.

9. Melton LJ III (2003) *Epidemiology worldwide*. Endocrinology and Metabolism Clinics of North America 32(1): 1-13.

10. Raisz LG (1999) *Osteoporosis: Current approaches in diagnosis, pathogenesis, and management*. Journal of Bone and Mineral Metabolism 17(2): 79-89.

11. Wade JP (2001) *Clinical basics of rheumatology: 15. Osteoporosis*. Canadian Medical Association Journal 165(1): 45-50.

12. Consensus Development Conference (1993) *Diagnosis, prophylaxis, and treatment of osteoporosis*. American Journal of Medicine 94: 646-650.

13. Chrischilles EA, Butler CD, Davis CS, Wallace RB (1991) *A model of lifetime osteoporosis impact*. Archives of Internal Medicine 151: 2026-2032.

14. Cooper C, Atkinson EJ, et al (1993) *Population-based study of survival following osteoporotic fractures*. American Journal of Epidemiology 137: 1001-1005.

15. Johnell O, Haglund B (1999) *Co-morbidity and mortality in hip fractures: A population based study*. Journal of Bone and Mineral Research 14(Suppl 1): S160.

16. Johnell O, Kanis J, Gullberg G (2001) *Mortality, morbidity, and assessment of fracture risk in male osteoporosis.* Calcified Tissue International 69(4): 182-184.

17. Leech JA, Dulberg C, Kellie S, et al (1990) *Relationship of lung function to severity of osteoporosis in women.* The American Review of Respiratory Disease 141: 68-71.

18. Melton LJ III (2003) *Adverse outcomes of osteoporotic fractures in the general population.* Journal of Bone and Mineral Research 18(6): 1139-1141.

19. Wiktorowicz ME, Goeree R, et al (2001) *Economic implications of hip fracture: Health service use, institutional care and cost in Canada.* Osteoporosis International 12(4): 271-278.

20. Greendale GA, Barrett-Connor E (2001) *Outcomes of osteoporotic fractures*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J, San Diego: Academic Press, pages 819-829.

21. Goeree ROB, Pertitt DB, Fernaz M, Adachi J (1996) *An assessment of the burden of illness due to osteoporosis in Canada.* Journal of the Society of Obstetricians and Gynaecologists of Canada 18(Suppl July): 15-24.

22. De Laet C, Van Hout B, et al (1999) *Incremental cost of medical care after hip fracture and first vertebral fracture: The Rotterdam Study.* Osteoporosis International 10: 66-72.

23. Ray NF, Chan JK, Thamer M, Melton LJ III (1997) *Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: Report from the National Osteoporosis Foundation.* Journal of Bone and Mineral Research 12(1): 24-35.

24. Papaidmitropoulos EA, Coyte PC, Josse RG, Greenwood CE (1997) *Current and projected rates of hip fracture in Canada.* Canadian Medical Association Journal, 157(10): 1357-1363.

25. Rauch F, Glorieux FH (2000) *Osteoporosis in children*, in *The Osteoporosis Primer.* Edited by Henderson JE and Goltzman D. Cambridge: Cambridge University Press, pages 186-196.

26. Browner WS, Pressman AR, Nevitt MC, Cummings SR (1996) *Mortality following fractures in older women. The Study of Osteoporotic fractures.* Archives of Internal Medicine 156: 1521-1525.

27. Browner WS, Seeley DG, Vogt TM, et al (1991) *Non-trauma mortality in elderly women with low bone density.* Lancet 338: 355-358.

28. Center JR, Nguyen TV, Schneider D, et al (1999) *Mortality after all major types of osteoporotic fracture in men and women: An observational study.* Lancet 353: 878-882.

29. Iacovino JR (2001) *Mortality outcomes after osteoporotic fractures in men and women*. The Journal of Insurance Medicine 33(4):316-320.

30. Brown JP, Josse RG, for the Scientific Advisory Council of the Osteoporosis Society of Canada (2002) *2002 Clinical practice guidelines for the diagnosis and management of osteoporosis in Canada*. Canadian Medical Association Journal 167(10 Suppl): S1-S34.

31. Nordin BEC (2003) *Should the treatment of osteoporosis be more selective?* Osteoporosis International 14: 99-102.

32. Melton LJ III, Wahner HW (1989) *Defining osteoporosis*. Calcified Tissue International 45: 263-264.

33. Anonymous (2000) *Osteoporosis prevention, diagnosis and therapy*. NIH Consensus Statement March 27-29, 17(1): 1-45.

34. Kmetic A, Joseph L, Berger C, Tenenhouse A (2002) *Multiple imputation to account for missing data in a survey: Estimating the prevalence of osteoporosis*. Epidemiology 13(4): 434-444.

35. Melton LJ, Lane AW, Cooper C, et al (1993) *Prevalence and incidence of vertebral deformities*. Osteoporosis International 3: 113-119.

36. Kanis JA, Glüer CC, for the Committee of Scientific Advisors, International Osteoporosis Foundation (2000) *An update on the diagnosis and assessment of osteoporosis with densitometry*. Osteoporosis International 11(3): 192-202.

37. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, Berger M (2000) *Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis*. Journal of Bone and Mineral Research 15: 721–739.

38. Melton LJ III, Crowson CS, O'Fallon WM (1999) *Fracture incidence in Olmsted County, Minnesota: Comparison of urban with rural rates and changes in urban rates over time*. Osteoporosis International 9:29-37.

39. National Osteoporosis Foundation (1998) *Osteoporosis: Review of the evidence for prevention, diagnosis, and treatment and cost-effectiveness analysis*. Osteoporosis International 8(Suppl 4): S7-S80.

40. Frost HM (2002) *Emerging views about "osteoporosis", bone health, strength, fragility, and their determinants*. Journal of Bone and Mineral Research 20: 319-325.

41. Doran PM, Khosla S (2000) *Senile osteoporosis*, in *The Osteoporosis Primer*. Edited by Henderson JE and Goltzman D. Cambridge: Cambridge University Press, pages 225-235.

42. Zhonghou L, Junhong P, et al (2002) *The diagnostic criteria for primary osteoporosis and the incidence of osteoporosis in China*. Journal of Bone and Mineral Metabolism 20: 181-189.

43. Gullberg G, Johnell O, Kanis JA (1997) *World-wide projections for hip fracture*. Osteoporosis International 7(5): 407-413.

44. Jaglal SB, Sherry PG, Scatzker J (1996) *The impact and consequences of hip fractures in Ontario*. Canadian Journal of Surgery 39(2): 105-111.

45. Fujita T (2001) *Global assessment for risk factors for osteoporosis*. Journal of Bone and Mineral Metabolism 19:131-132.

46. Ahlborg HG, Johnell O, et al (2003) *Bone loss and bone size after the menopause*. New England Journal of Medicine 349(4): 327-342.

47. Seeman E (2003) *The structural and biomechanical basis of the gain and loss of bone strength in women and men*. Endocrinology and Metabolism Clinics of North America 32(1): 25-38.

48. Luthje P (1991) *The incidence of hip fracture in Finland in the year 2000*. Acta Orthopaedica Scandinavica 62(Suppl 241): 35-37.

49. Cummings SR, Rubin SM, Black D (1990) *The future of hip fracture in the United States*. Clinical Orthopaedics and Related Research 252: 163-166.

50. Lau EMC (1997) *Epidemiology of osteoporosis in urbanized Asian populations*. Osteoporosis International (Suppl 3): S91-S95.

51. Kanis JA, Johnell O, Oden A, et al (2000) *Risk of hip fracture according to the World Health Organization criteria for osteopenia and osteoporosis*. Bone 27: 585-90.

52. Looker AC, Orwoll ES, Johnston Jr. CC, et al (1997) *Prevalence of low femoral bone density in older U.S. adults from NHANES III*. Journal of Bone and Mineral Research 12: 1761-8.

53. Iki M, Kagamimori S, Kagawa Y, et al (2001) *Bone mineral density of the spine, hip and distal forearm in representative samples of the Japanese female population: Japanese Population-Based Osteoporosis (JPOS) Study*. Osteoporosis International 12: 529-37.

54. Tenenhouse A, Joseph L, Kreiger N, et al (2000) *Estimation of the prevalence of low bone density in Canadian women and men using a population-specific DXA reference standard: the Canadian Multicentre Osteoporosis Study (CaMos)*. Osteoporosis International 11: 897-904.

55. Ross PD, Fujiwara S (1995) *Vertebral fracture prevalence in women in Hiroshima compared to Caucasians or Japanese in the US.* International Journal of Epidemiology 24: 1171-1177.

56. Roy DK, Pye SR, et al (2002) *Falls explain between-centre differences in the incidence of limb fracture across Europe.* Bone 31(6): 712-717.

57. Finkelstein JS, Lee M-LT, et al (2002) *Ethnic variation in premenopausal and early perimenopausal women: Effects of anthropomorphometric and lifestyle factors.* Journal of Clinical Endocrinology & Metabolism 87(7): 3057-3067.

58. Looker AC (2002) *Editorial: The skeleton, race, and ethnicity.* Journal of Clinical Endocrinology & Metabolism 87(7): 3047-3050.

59. Hasserius R, Karlsson MK et al (2003) *Prevalent vertebral deformities predict increased mortality and increased fracture rate in both men and women: A 10-year population-based study of 598 individuals from the Swedish cohort in the European Vertebral Osteoporosis Study.* Osteoporosis International 14: 61-68.

60. Cooper C, Fall C, et al (1997) *Growth in infancy and bone mass in later life.* Annals of the Rheumatic Diseases 56(1): 17-21.

61. Marcus R, Majumber S (2001) *The nature of osteoporosis*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J, San Diego: Academic Press, pages3-17.

62. Doran PM, Khosla S (2000) *Senile Osteoporosis,* in *The Osteoporosis Primer.* Edited by Henderson JE and Goltzman D. Cambridge: Cambridge University Press, pages 225-235.

63. Prestwood KM, Raisz LG (2000) *Consequences of alterations in bone remodeling*, in *The Osteoporosis Primer.* Edited by Henderson JE and Goltzman D. Cambridge: Cambridge University Press, pages 199-210.

64. Mazess RB (1982*) On aging bone loss.* Clinical Orthopaedics and Related Research 165: 239-252.

65. Hirotoshi M, Genant HK (1998) *Statement on the diagnosis and management of osteoporosis from the Consensus Development Conference at the second International Conference on Osteoporosis, Osaka 1997.* Journal of Bone and Mineral Research 16: 206-214.

66. Burr D, Forwood M, et al (1997) *Bone microdamage and skeletal fragility in osteoporotic and stress fractures*, Journal of Bone and Mineral Research 12(1): 6-15.

67. Raisz LG (1999) *Osteoporosis: Current approaches and future prospects in diagnosis, pathogenesis, and management.* Journal of Bone and Mineral Metabolism 17: 79-89.

68. Parkkari J, Kannus P, Palvanen M, et al (1999) *Majority of hip fractures occur as a result of a fall and impact on the trochanter of the femur: A prospective controlled hip fracture study with 206 consecutive patients.* Calcified Tissue International 65(3): 183-187.

69. Uitterlinden AG, van Leeuwen J, Pols H (2001) *Gentics and genomics of osteoporosis*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J, San Diego: Academic Press, pages 639-667.

70. World Health Organization (1994) *Assessment of fracture risk and its application to screening for postmenopausal osteoporosis.* WHO technical report series 843. Geneva: WHO.

71. Marshall D, Johnell O, Wedel H (1996) *Meta-analysis of how well measures of bone density predict occurrence of osteoporotic fractures.* British Medical Journal 312: 1254-1249.

72. Cummings SR, Black DM, et al, for the Study of Osteoporotic Fractures Research Group (1993) *Bone density at various sites for prediction of hip fracture.* Lancet 341: 72-75.

73. Hayes WC, Piazza SJ, Zysset PK (1991) *Biomechanics of fracture risk prediction of the hip and spine by quantitative computerized tomography.* Radiological Clinics of North America 29: 1-18.

74. Slemenda CW, Johnson CC, Hui SL (2001) *Assessment of fracture risk*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J, San Diego: Academic Press, pages 809-817.

75. Melton LJ III, Atkinson EJ, et al (1999) *Vertebral fractures predict subsequent fractures.* Osteoporosis International 10: 214-221.

76. Albagha OME, Ralston SH (2003) *Genetic determinants of susceptibility to osteoporosis.* Endocrinology and Metabolism Clinics of North America 32(1): 65-81.

77. Patel MS, Rubin LA, Cole DEC (2000) *Genetic determinants of osteoporosis*, in *The Osteoporosis Primer.* Edited by Henderson JE and Goltzman D. Cambridge: Cambridge University Press, pages 131-146.

78. Larkin M (2002) *Boning up on osteoporosis.* The Lancet 359, page 271.

79. Ross PD (1996) *Osteoporosis: Frequency, consequences, and risk factors.* Archives of Internal Medicine 156: 1399-1411.

80. Wu F, Mason B, Horne, et al (2002) *Fractures between the ages of 20 and 50 years increase women's risk of subsequent fractures.* Archives of Internal Medicine 162(1): 33-36.

81. Hosmer WD, Genant HK, Browner WS (2002) *Fractures before menopause: A red flag for physicians.* Osteoporosis International 13(3): 337-341.

82. Jaglal SB, Carroll J, et al (2003) *How are family physicians managing osteoporosis?* Canadian Family Physician 49: 462-468.

83. McKercher HG, Crilly RG, Kloseck M (2000) *Osteoporosis management in long-term care.* Canadian Family Physician 46: 2228-2235.

84. Hajcsar EE, Hawker G, Bogoch ER (2000) *Investigation and treatment of osteoporosis in patients with fragility fractures.* Canadian Medical Association Journal 163: 819-822.

85. Castel H, Bonneh DY, Sherf M, Liel Y (2001) *Awareness of osteoporosis and compliance with management guidelines in patients with newly diagnosed low-impact fractures.* Osteoporosis International 12(7): 559- 564.

86. Pal B (1999) *Questionnaire survey of advice given to patients with fractures.* British Journal of Medicine 318: 500-501.

87. Torgerson DJ, Dolan P (1998) *Prescribing by general practioners after an osteoporotic fracture.* Annals of the Rheumatic Diseases 57: 378-379.

88. Bellantonio S, Fortinsky R, Prestwood K (2001) *How well are community-living women treated for osteoporosis after hip fracture?* Journal of the American Geriatric Society 49(9): 1197-1204.

89. Eddy DM, Johnston CC, et al (1998) *Osteoporosis: Review of the evidence for prevention, diagnosis, and treatment and cost-effectiveness analysis.* Osteoporosis International 8 (Suppl 4): 1005-1010.

90. Kanis JA, Delmas P, Burckhardt P, Cooper C, Torgerson D (1997) *Guidelines for diagnosis and management of osteoporosis. The European Foundation for Osteoporosis and Bone Disease.* Osteoporosis International 7: 390–406.

91. Häuselmann HJ, Rizzoli R (2003) *A comprehensive review of treatments for postmenopausal osteoporosis.* Osteoporosis International 14(1): 2-12.

92. Writing Group for the Women's Health Initiative (2002) *Risks and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results from the Women's Health Initiative Randomized Controlled Trial.* Journal of the American Medical Association 288(3): 321-333.

93. Black DM, Thompson DE, et al (2000) *Fracture risk reduction with alendronate in women with osteoporosis: The Fracture Intervention Trial (FIT).* Journal of Clinical Endocrinology and Metabolism 85: 4118-4124.

94. Harris ST, Watts NB, et al (1999) *Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: A randomized controlled trial.* Journal of the American Medical Association 282: 1344-1352.

95. International Osteoporosis Foundation and European Society of Musculoskeletal Radiology (2003) *Vertebral fracture initiative resource document*, www.osteofound.org/ health_professionals/education_radiologists/ pdf/resource_document.pdf, 39 pages.

96. Jackson SA, Tenenhouse A, Robertson L, and the CaMos Study Group (2000) *Vertebral fracture definition from population-based data: Preliminary results from the CaMos.* Osteoporosis International 11: 680-687.

97. Eastell R, Cedel SL, et al (1991) *Classification of vertebral fractures.* Journal of Bone and Mineral Research 6: 207-215.

98. McCloskey EV, Spector TD, et al (1993) *The assessment of vertebral deformity: A method for use in population studies and clinical trials.* Osteoporosis International 3(3): 138-147.

99. Heggeness MH, Mathis KB (2001) *An orthopaedic perspective of osteoporosis*, in *Osteoporosis, Second Edition*, Volume 2. Edited by Marcus R, Feldman D, Kelsey J. San Diego: Academic Press, pages 485-499.

100. Genant HK, Jergas M (2003) *Assessment of prevalent and incident vertebral fractures in osteoporosis research.* Osteoporosis International 14(Suppl 3): S43-S55.

101. Faciszewski T, McKiernan F (2002) *Calling all vertebral fractures. Classification of vertebral compression fractures: A consensus for comparison of treatment and outcome.* Journal of Bone and Mineral Research 17(2): 185-191.

102. Gehlbach SH, Bigelow C, et al (2000) *Recognition of vertebral fracture in a clinical setting.* Osteoporosis International 11(7): 577-582.

103. Roux C (2001) *Do osteoporotic vertebral deformities deserve medical attention.* Joint Bone Spine 68: 194-197.

104. Ensrud KE, Nevitt MC, et al (1999) *What proportion of incident morphometric vertebral fractures are clinically diagnosed and vice versa?* Journal of Bone and Mineral Research 14(S1): S138.

105. National Osteoporosis Foundation Working Group on Vertebral Fractures (1995) *Report: Assessing vertebral fractures*. Journal of Bone and Mineral Research 10(4): 518-522.

106. Melton LJ III, Cooper C (2001) *Magnitude and impact of osteoporosis and fractures*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J. San Diego: Academic Press, pages 557-567.

107. O'Neill TW, Felsenberg D, et al (1996) *The prevalence of vertebral deformity in European men and women: The European Vertebral Osteoporosis Study (EVOS)*. Journal of Bone and Mineral Research 11: 1010-1017.

108. Ismail AA, Cockerill A, et al (2001) *Prevalent vertebral deformity predicts incident hip though not distal forearm fracture: Results from the EPOS*. Osteoporosis International 12: 85-90.

109. Ettinger B, Black DM, Mitlak BH, et al (1999) *Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: Results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators*. Journal of the American Association 28: 637-645.

110. Reginster J, Minne HW, et al (2000) *Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis*. Vertebral Efficacy with Risedronate Therapy (VERT). Osteoporosis International 11: 83-91.

111. Ismail AA, O'Neill TW, et al (1998) *Mortality associated with vertebral deformity with men and women: Results from the European Prospective Osteoporosis Study (EPOS)*. Osteoporosis International 8(3): 291-297.

112. Bouxsein ML (2001) *Biomechanics of age-related fractures*. In: *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J. San Diego: Academic Press, pages 509-531.

113. Myers E, Wilson S (1997) *Biomechanics of osteoporosis and vertebral fractures*. Spine 22, 25S-31S.

114. Papaioannou A, Watts NB, et al (2002) *Diagnosis and management of vertebral fractures in elderly adults*. American Journal of Medicine 113: 220-228.

115. Ross PD (1997) *Clinical consequences of vertebral fractures*. American Journal of Medicine 102(Suppl 6A): S30-S43.

116. Huang C, Ross PD, Wasnich RD (1996) *Vertebral fractures and other predictors of physical impairment and health care utilization*. Archives of Internal Medicine 156: 2469-2475.

117. Cockerill W, Ismail AA, et al (2000) *Does location of vertebral deformity within the spine influence back pain and disability?* European Vertebral Osteoporosis Study EVOS Group. Annals of Rheumatic Diseases 59: 368- 371.

118. Ross PD, Davis JW, Epstein RS, Wasnich RD (1994) *Pain and disability associations with new vertebral fractures and other spinal conditions.* Journal of Clinical Epidemiology 47: 231-239.

119. Johnell O, Gullberg B, Kanis JA (1997) *The hospital burden of vertebral fracture in Europe: A study of National Register Resources.* Osteoporosis International 7: 138-144.

120. Papaioannou A, Adachi JD, et al (2001) *Lengthy hospitalization associated with vertebral fractures despite control for comorbid conditions.* Osteoporosis International 12(10): 870-874.

121. Nevitt MC, Thompson DE, et al, and The Fracture Intervention Trial Research Group (2000) *Effect of alendronate on limited-activity days and bed-disability days caused by back pain in postmenopausal women with existing vertebral fractures.* Archives of Internal Medicine 160(1): 77-85.

122. Greendale GA, Barrett-Connor E, Ingles S, Haile R (1995) *Late physical and functional effects of osteoporotic fractures in women: The Rancho Bernardo Study.* Journal of the American Geriatric Society 43: 955-961.

123. Kado DM, Browner WS (1999) *Vertebral fractures and mortality in older women.* Archives of Internal Medicine 159:1215-1220.

124. Schlaich C, Minne HW, et al (1998) *Reduced pulmonary function in patients with spinal osteoporotic fractures.* Osteoporosis International 8(3): 261-267.

125. Nevitt MC, Ettinger B, et al (1998) *The association of radiographically detected vertebral fractures with back pain and function: A prospective study.* Annals of Internal Medicine 128(10): 793-800.

126. Burger H, Van Daele PLA, et al (1997) *Vertebral deformities and functional impairment in men and women.* Journal of Bone and Mineral Research 12: 152-157.

127. Pluijm SMF, Tromp AM, et al, (2000) *Consequences of vertebral deformities in older men and women.* Journal of Bone and Mineral Research 15: 1564-1572.

128. Leidig G, Minne HW, et al (1990) *A study of complaints and their relation to vertebral destruction in patients with osteoporosis.* Bone and Mineral 8: 217-229.

129. Adachi JD, Ioannidis G, et al, and the CaMOs Research Group (2001) *The influence of osteoporotic fractures on health-related quality of life in community-dwelling men and women across Canada.* Osteoporosis International 12: 903-908.

130. Tosteson A, Gabriel S, et al (2001) *Impact of hip and vertebral fractures on quality-adjusted life years*. Osteoporosis International 12: 1042-1049.

131. Kado DM, DuongT, et al (2003) *Incident vertebral fractures and mortality in older women: A prospective study*. Osteoporosis International 14(7): 589-594.

132. Guermazi A, Mohr A, et al (2002) *Identification of vertebral fractures in osteoporosis*. Seminars in Musculoskeletal Radiology 6(2): 241-252.

133. De Laet C, Reeve J (2001) *Epidemiology of osteoporotic fractures in Europe*, in *Osteoporosis, Second Edition*, Volume 1. Edited by Marcus R, Feldman D, Kelsey J. San Diego: Academic Press, pages 585-597.

134. Ross PD, Davis JW, Epstein RS, Wasnich (1991) *Pre-existing fractures and bone mass predict vertebral fracture incidence in women*. Annals of Internal Medicine 114: 919-923

135. Ettinger B, Cooper C (1995) *Clinical assessment of osteoporotic vertebral fractures*. In: *Vertebral fracture in osteoporosis*, edited by Genant HK, Jergas M, van Kuijk C. San Francisco: Radiology Research and Education Foundation.

136. Jonsson B, Gardsell P, et al (1994) *Remembering fractures: Fracture registration and proband recall in southern Sweden*. Journal of Epidemiology and Community Health 48: 489- 490.

137. Nevitt MC, Cummings SR, et al (1992) *The accuracy of self-report of fractures in elderly women: Evidence from a prospective study*. American journal of Epidemiology 135: 490- 499.

138. Ferrar L, Jiang G, Eastell R, and NFA PEEL (2003) *Visual identification of vertebral fractures in osteoporosis using morphometric X-ray absorptiometry*. Journal of Bone and Mineral Research 18(5): 933-938.

139. Delmas PD, Watts N, et al (2001) *Underdiagnosis of vertebral fractures is a worldwide problem: The IMPACT study*. Journal of Bone and Mineral Research 16(Suppl 1): S139.

140. Genant HK, Wu CY, Van Kuij KS, Nevitt MC (1993) *Vertebral fracture assessment using a semi quantitative technique*. Journal of Bone and Mineral Research 8: 1137-1348.

141. Smith-Bindman R, Cummings SR, et al (1991) *A comparison of normal morphometric definitions of vertebral fracture*. Journal of Bone and Mineral Research 6: 25-34.

142. Melton LJ III (1987) *Epidemiology of vertebral fractures*. In: *Osteoporosis*, edited by Christensen C, Johansen JS, Riis B. Copenhagen.

143. Minne HW, Leidig G, et al (1988) *A newly developed spinal deformity index (SDI) to quantitate vertebral crush fractures in patients with osteoporosis*. Bone and Mineral 3(4): 335-349.

144. Black DM, Cummings SR, et al (1991) *A new approach to defining vertebral dimensions*. Journal of Bone and Mineral Research 6: 883-891.

145. Wu CY, Li J, Jergas M, Genant HK (1995) *Comparison of semiquantitative and quantitative techniques for the assessment of prevalent and incident vertebral fractures*. Osteoporosis International 5(5): 354-370.

146. Black DM, Palermo L, et al (1999) *Defining incident vertebral deformity: A prospective comparison of several approaches*. Journal of Bone and Mineral Research 14(1): 90-101.

147. Leidig-Bruckner G, Genant HK (1994) *Comparison of a semiquantitative and quantitative method for assessing vertebral fractures in osteoporosis*. Osteoporosis International 4(3):154-161.

148. Davies KM, Stegman MR (1996) *Prevalence and severity of vertebral fracture: The Saunders County Bone Quality Study*. Osteoporosis International 6: 160-165.

149. Melton LJ, Kan SH (1989) *Epidemiology of vertebral fractures in women*. American Journal of Epidemiology 129:1000–1011.

150. Kitazawa A, Kazuhiro K, Yamazaki K, Inoue T (2001) *Prevalence of vertebral fractures in a population-based sample from Japan*. Journal of Bone and Mineral Metabolism 19: 115-118.

151. The European Prospective Osteoporosis Study (EPOS) Group (2002) *Incidence of vertebral fracture in Europe: Results from the EPOS*. Journal of Bone and Mineral Research 17(4): 716-724.

152. Van der Klift M, De Laet C, et al (2002) *The incidence of vertebral fractures in men and women: The Rotterdam Study*. Journal of Bone and Mineral Research 17(2): 1051-1056.

153. Cooper C, Atkinson EJ, O'Fallon WM, Melton LJ III (1992) *Incidence of clinically diagnosed vertebral fractures: A population-based study in Rochester, Minnesota, 1985-1989*. Journal of Bone and Mineral Research 7(2): 221-227.

154. Jones G, Nguyen T, et al (1994) *Symptomatic fracture incidence in elderly men and women: The Dubbo Osteoporosis Epidemiology Study (DOES)*. Osteoporosis International 4: 277-282.

155. Sanders KM, Seeman E, et al (1999) *Age and gender specific rate of fractures in Australia: A population based study*. Osteoporosis International 10: 240-247.

156. Touloumi G, Babiker AG, Pocock SJ, Darbyshire JH (2001) *Impact of missing data due to drop-outs on estimators for rates of change in longitudinal studies: A simulation study*. Statistics in Medicine 20: 3715-3728.

157. Greenland S, Finkle W (1995) *A critical look at methods for handling missing covariates in epidemiologic regression analyses*, American Journal of Epidemiology, 142(12): 1255-1264.

158. Rubin DB (1987) *Multiple imputation for nonresponse surveys*. Wiley Series in Probability and Mathematical Statistics: John Wiley & Sons, Inc., 258 pages.

159. Rubin DB (1976) *Inference and missing data*. Biometrika 63: 581-592.

160. Allison PD (2001) *Missing data*. Sage University Papers Series: Quantitative Applications on the Social Sciences, 07-136, Sage Publications, 93 pages.

161. Schafer JL (1997) *Analysis of incomplete multivariate data*. Monographs on Statistics and Applied Probability 72. London: Chapman & Hall, 430 pages.

162. Little JA, Rubin DB (2002) *Statistical analysis with missing data, Second edition*. Hoboken, New Jersey: Wiley & Sons Inc., 381 pages.

163. Schafer JL (1999) *Multiple imputation: A primer*. Statistical Methods in Medical Research 8(1): 3-15.

164. Arnold AM, Kronmal RA (2003) *Multiple imputation of baseline data in the cardiovascular health study*. American Journal of Epidemiology 157(1): 74-84.

165. Little RJA (1992) *Regression with missing X's: A review*. Journal of the American Statistical Association 87(420): 1227-1236.

166. Rubin DB, Schenker N (1991) *Multiple imputation in health-care databases: An overview and some applications*. Statistics in Medicine 10: 585-598.

167. Rubin DB (1978) *Multiple imputations in sample surveys – A phenomenological Bayesian approach to nonresponse*. In: Proceedings of the Survey Research Methods Section, American Statistical Association, pages 20-34.

168. Rubin DB (1996) *Multiple imputation after 18+ years*. Journal of the American Statistical Association, 91(434): 473-89.

169. Twisk J, de Vente W (2002) *Attrition in longitudinal studies: How to deal with missing data*. Journal of Clinical Epidemiology 55: 329-337.

170. Barnard J, Meng XL (1999) *Applications of multiple imputation in medical studies: From AIDS to NHANES*. Statistical Methods in Medical Research 8(1): 17-36.

171. Zhou XH, Eckert GJ, Tierney WM (2001) *Multiple imputation in public health research*. Statistics in Medicine 20: 1541-1549.

172. Faris PD, Ghali WA, et al (2002) *Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses*. Journal of Clinical Epidemiology 55: 184-191.

173. Ware JE, Sherbourne CD (1992) *The MOS 36-itme short form health survey (SF-36): 1) Conceptual framework and item selection*. Medical Care 30: 473-483.

174. Feeny D, Torrance G, Goldsmith C, Furlong W, Boyle M (1994) *A multi-attribute approach to population health status*. Working paper #94-5, Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario.

175. Kreiger N, Tenenhouse A, Joseph L, et al (1999) *The Canadian Multicentre Osteoporosis Study (CaMos): Background, rationale, methods*. Canadian Journal on Aging 18(3):376-387.

176. Tenenhouse A, Kreiger N, Hanley D (2000) *Canadian Multicentre Osteoporosis Study (CaMos)*. Drug Development Research 49: 201-205.

177. Brooks SP, Gelman A (1998) *Alternative methods for monitoring convergence of iterative simulations*. Journal of Computational and Graphical Statistics 7: 434-455.

178. Raftery, AE (1995). Bayesian model selection in social research (with Discussion). In *Sociological Methodology 1995,* (Peter V. Marsden, ed.), Cambridge, Massachusetts. Blackwells: 111-196.

179. Kass R, Raftery A (1995) Bayes factors. Journal of the American Statistical Association 90: 773-795.

180. Raftery, AE, Volinsky C. Bayesian Model Averaging Home Page. ***http://www.research.att.com/~volinsky/bma.html*** . Last accessed November, 2004.

APPENDIX 1

WinBUGS program used to estimate crude incidence rates in women, excluding consideration of pre-existing vertebral deformities. A similar program was used to estimate male rates.

```
# x[] is the binary variable for incidence, and PY[] the number of person-years for each individual

model
{
for (i in 1:634)          # female Respondent data, aged 55-64
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.55to64/1000;
}
rate.55to64 ~ dgamma(0.001,0.001);

for (i in 635:1432)  # female Respondent data, aged 65-74, stacked under the 55-64 group
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.65to74/1000;
}
rate.65to74 ~ dgamma(0.001,0.001);

for (i in 1433:1982)   # data for those aged 75+, stacked under the 65-74 group
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.75/1000;
}
rate.75 ~ dgamma(0.001,0.001);

rate55.weighted.to.Census <- weight1 * rate.55to64 + weight2 * rate.65to74 + weight3 *
rate.75
}

# Initial values for 3 chains
list(rate.55to64 = 0, rate.65to74=0, rate.75=0)
list(rate.55to64 = 1, rate.65to74=1, rate.75=1)
list(rate.55to64 = 100, rate.65to74=100, rate.75=100)

#  Data: 1996 Canadian census weights
list(weight1 = 0.3883467, weight2 = 0.3375752, weight3 =0.2740781)
```

APPENDIX 2

WinBUGS program used to estimate crude incidence rates in women, stratified by the absence and presence of vertebral deformities at baseline. A similar program was used to estimate the male rates.

```
model
{
for (i in 1:634)              # data for women aged 55-64
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.55to64[i]/1000;
log(rate.55to64[i]) <- alpha1 + beta1*prev[i] ;
}
alpha1 ~ dnorm(0, 0.0001)
beta1 ~ dnorm(0, 0.0001)

for (i in 635:1432)           # data for women aged 65-75
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.65to74[i]/1000;
log(rate.65to74[i]) <- alpha2 + beta2*prev[i] ;
}
alpha2 ~ dnorm(0, 0.0001)
beta2 ~ dnorm(0, 0.0001)

for (i in 1433:1982)          # data for women aged 75+
{
x[i] ~ dpois(lambda[i]);
lambda[i] <- PY[i]*rate.75[i]/1000;
log(rate.75[i]) <- alpha3 + beta3*prev[i] ;
}
alpha3 ~ dnorm(0, 0.0001)
beta3~ dnorm(0, 0.0001)

# rate for 55+, weighted to the 1996 Canadian Census - result should be similar that
obtained with simple Poisson:

rate55.weighted.to.Census <-  exp(alpha1)*w1i + exp(alpha1 + beta1)*w1ii
                              + exp(alpha2)*w2i + exp(alpha2 + beta2)*w2ii
                              + exp(alpha3)*w3i + exp(alpha3 + beta3)*w3ii
```

```
# rate55+ for those without prev deformities:

rate.55.noprev <-  exp(alpha1)*w1iadj + exp(alpha2)*w2iadj + exp(alpha3)*w3iadj


# rate for 55+ for those with >=1 prev deformities:

rate.55.wprev <-  exp(alpha1 + beta1)*w1iiadj + exp(alpha2 + beta2)*w2iiadj +
                  exp(alpha3 + beta3)*w3iiadj
}

# Initial values for 2 chains

list(alpha1 = 0, alpha2 = 0, alpha3 = 0, beta1=0, beta2=0, beta3=0)
list(alpha1 = 5, alpha2 = 5, alpha3 = 5, beta1=1, beta2=1, beta3=1)


#  Data: weights derived from the 1996 Canadian census

list(w1i = 0.3730333,  w2i = 0.3011949,  w3i = 0.2187642,
         w1ii = 0.0153134, w2ii = 0.0363803, w3ii = 0.0553139,
        w1iadj = 0.417734, w2iadj = 0.3372872, w3iadj = 0.2449788,
        w1iiadj = 0.1431057, w2iiadj = 0.3399787, w3iiadj  =0.5169156)
```

APPENDIX 3

S-Plus program used to impute sex for the TR2 group who had data for the number of individuals ≥25 within the household (HH25). A similar program was used for TR2 nonparticipants with missing values for HH25.

```
female.sex.wHH25 <- function (xsex, alpha, beta.tel, beta.HH25, city.coef)
{
   inv.prob <- rep(NA, length(xsex[,1]))
   prob <- rep(NA, length(xsex[,1]))
   for (i in 1:length(xsex[,1]) )
        {
         inv.prob[i] <- alpha + beta.tel*xsex[i,1] + beta.HH25*xsex[i,2]
         for (j in 1:8)
            {
              if(xsex$city[i]= =j) inv.prob[i] <- inv.prob[i] + city.coef[j]
            }
          prob[i] <- exp(inv.prob[i]) / ( 1 + exp(inv.prob[i]) )
        }
   female <- rbinom(length(xsex[,1]), 1, prob)
   return(female)
}
```

APPENDIX 4

S-Plus program used to impute age for the TR2 group who had data for the number of individuals ≥25 within the household (HH25). A similar program was used for TR2 nonparticipants with missing values for HH25.

```
# TRage.wHH25 is the dataset with subjects without age values, and with HH25 data
# available
# TRage is the dataset of subjects with age available (and HH25), from which the
# coefficients are estimated
# res.SE is the residual SE computed from the model run with the predicted age as the
# response, on the TRage dataset


options(contrasts = c(factor = "contr.treatment", ordered = "contr.poly"))

female.age.wHH25 <- function (TRage.wHH25, res.SE)
 {
  age.lm <- lm(age ~ tel + HH25 + city, data = TRage, na.action = na.omit)
  age.pred <- predict.lm(age.lm,TRage.wHH25, se.fit = T )
  mean.age.pred <- age.pred$fit
  sd.mean.pred <- age.pred$se.fit
  sd.ind.pred <- res.SE*(sqrt( (sd.mean.pred/res.SE)^2 + 1) )
  pred.final <- rep(NA, length(TRage.wHH25[,1]) )
  for (i in 1:length(TRage.wHH25[,1]) )
     {
       pred.final[i] <- rnorm(1, mean.age.pred[i], sd.ind.pred[i])
      }
  return(pred.final)
}
```

APPENDIX 5

WinBUGS program used for the multiple imputation of incidence in the different participation groups, and the estimation of adjusted incidence rates in women. A similar program with different imputation models was used for the male estimates.

```
# Use this program for:
#       (1) Multiple imputation of incidence in the Censored and Nonresponders
#           (RQ,TR1,TR2)
#       (2) Estimation of MI-adjusted rates for attrition bias alone (LFU+Resp)
#       (3) Estimation of the MI-adjusted rate for Nonresponse & attrition biases (Resp +
#           LFU + RQ + TR1 + TR2)

# Multiple Imputation models for each Participation Group are unique
# Initial values for each coefficient are derived from taking a range of values about MLE,
# generated with S-Plus

# The following rates are available for monitoring:
#  (Resp = respondents, LFU = lost to follow-up, Ign = ignorable model,
# NI = nonignorable models)

        # Estimates for the 55+ groups

#               rate.55.LFU             overall.rate.Ign.55
#               rate.55.RQ              overall.rate.NI.delta.55
#               rate.55.TR1             overall.rate.NI.0.5delta.55
#               rate.55.TR2             overall.rate.NI.minus.55

        # Estimates for the 55to64 groups
                                        overall.rate.Ign.55to64
#               rate.LFU.55to64         overall.rate.NI.delta.55to64
#               rate.RQ.55to64          overall.rate.NI.0.5delta.55to64
#               rate.TR1.55to64         overall.rate.NI.minus.55to64
#               rate.TR2.55to64

        # Estimates for the 65to74 groups
                                        overall.rate.Ign.65to74
#               rate.LFU.65to74         overall.rate.NI.delta.65to74
#               rate.RQ.65to74          overall.rate.NI.0.5delta.65to74
#               rate.TR1.65to74         overall.rate.NI.minus.65to74
#               rate.TR2.65to74

        # Estimates for the 75+ groups
#                               overall.rate.Ign.75
```

```
#              rate.LFU.75          overall.rate.NI.delta.75
#              rate.RQ.75                overall.rate.NI.0.5delta.75
#              rate.TR1.75          overall.rate.NI.minus.75
#              rate.TR2.75

# Monitor the following rates for attrition bias adjustment analyses:

#      LFU.rate.Ign.55to64              LFU.rate.Ign.65to74
#      LFU.rate.Ign.75                  LFU.rate.Ign.55
#      LFU.rate.NI.delta.55to64    LFU.rate.NI.delta.65to74    LFU.rate.NI.delta.75
#      LFU.rate.NI.delta.55
#      LFU.rate.NI.0.5delta.55to64 LFU.rate.NI.0.5delta.65to74
#      LFU.rate.NI.0.5delta.75     LFU.rate.NI.0.5delta.55
#      LFU.rate.NI.minus.55to64    LFU.rate.NI.minus.65to74    LFU.rate.NI.minus.75
#      LFU.rate.NI.minus.55



model
{
                                    # 55-64 age group
                                # Censored
for (i in 1:1283)       # number aged 55-64; first 657 are Censored, followed by
                    # Respondents, so that observed data are used to estimate
                    #  missing incidence data by MI
{
 # imputation model for Censored women aged 55-64
  logit(p.1[i]) <- alpha1  + beta1.prev*prev[i] + beta1.age*age[i]
                    + beta1.HxFr*HxFr[i] + beta1.FamHx*FamHx[i]
                     + beta1.BMD*BMD[i]
  Inc[i] ~ dbern(p.1[i])  # the imputed binary incidence for the censored
 }
alpha1 ~  dnorm(0, 0.0001)            # prior distribution specifications
beta1.prev ~  dnorm(0, 0.0001)
beta1.age ~  dnorm(0, 0.0001)
beta1.HxFr ~  dnorm(0, 0.0001)
beta1.FamHx ~ dnorm(0, 0.0001)
beta1.BMD ~ dnorm(0, 0.0001)

lambda.LFU.55to64 <- sum(Inc[1:657])              # number of cases in Censored
imputed.inc.LFU.55to64 ~  dpois(lambda.LFU.55to64)  # likelihood function for rate

# approximating the PY for Censored
Inc0.LFU.55to64 <-  657  - lambda.LFU.55to64  # number of non-cases in
                                                #   the Censored
PY.LFU.55to64 <-  lambda.LFU.55to64*2.5  + Inc0.LFU.55to64*5
```

```
rate.LFU.55to64 <- imputed.inc.LFU.55to64*1000 / PY.LFU.55to64  # crude rate in
                                                                # Censored, aged 55-64


                        # RQ group
for (i in 1284:2559)    # number aged 55-64; first data stack are RQ, followed by the
                                # Respondents
{

# imputation model for female RQ aged 55-64
 logit(p.2[i]) <-  alpha2 + beta2.age*age[i] + beta2.HxFr*HxFr[i]
                + beta2.FamHx*FamHx[i]
 Inc[i] ~ dbern(p.2[i])
}
alpha2 ~ dnorm(0, 0.0001)
beta2.age ~ dnorm(0, 0.0001)
beta2.HxFr ~ dnorm(0, 0.0001)
beta2.FamHx ~ dnorm(0, 0.0001)

lambda.RQ.55to64 <- sum(Inc[1284:1930])
imputed.inc.RQ.55to64 ~ dpois(lambda.RQ.55to64)
Inc0.RQ.55to64 <-  647  -  lambda.RQ.55to64
PY.RQ.55to64 <-  lambda.RQ.55to64*2.5 + Inc0.RQ.55to64*5
rate.RQ.55to64 <- imputed.inc.RQ.55to64*1000 / PY.RQ.55to64

                        # TR1 group
for (i in 2560:3394)    # number aged 55-64; first data stack are TR1, followed by the
                                # Respondents

{
 logit(p.3[i]) <-  alpha3+ beta3.age*age[i]  + beta3.HH25*HH25[i]
 Inc[i] ~ dbern(p.3[i])
}

alpha3 ~ dnorm(0, 0.0001)
beta3.age ~ dnorm(0, 0.0001)
beta3.HH25 ~ dnorm(0, 0.0001)

lambda.TR1.55to64 <- sum(Inc[2560:2762])
imputed.inc.TR1.55to64 ~  dpois(lambda.TR1.55to64)
Inc0.TR1.55to64 <-  203  -  lambda.TR1.55to64
PY.TR1.55to64 <-  lambda.TR1.55to64*2.5 + Inc0.TR1.55to64*5
rate.TR1.55to64 <-  imputed.inc.TR1.55to64*1000 / PY.TR1.55to64
```

```
                    # TR2 group
for (i in 3395:4544)  # number aged 55-64; first data stack are TR2, followed by the
                    # Respondents

{
 logit(p.4[i]) <-  alpha4 + beta4.age*age[i] + beta4.tel*tel[i]
 Inc[i] ~ dbern(p.4[i])
}
alpha4 ~ dnorm(0, 0.0001)
beta4.age ~ dnorm(0, 0.0001)
beta4.tel ~ dnorm(0, 0.0001)

lambda.TR2.55to64 <- sum(Inc[3395:3911])
imputed.inc.TR2.55to64 ~  dpois(lambda.TR2.55to64)
Inc0.TR2.55to64 <- 517 - lambda.TR2.55to64
PY.TR2.55to64 <-  lambda.TR2.55to64*2.5 + Inc0.TR2.55to64*5
rate.TR2.55to64 <- imputed.inc.TR2.55to64*1000 / PY.TR2.55to64

# estimating the Respondent rate for those aged 55-64, as it will be needed for the 3
# non-ignorable models:
lambda.Resp.55to64 <- sum(Inc[3912:4544])
Inc.Resp.55to64 ~ dpois(lambda.Resp.55to64)
Inc0.Resp.55to64 <-  633 - lambda.Resp.55to64
PY.Resp.55to64 <-  lambda.Resp.55to64*2.5 + Inc0.Resp.55to64*5
rate.Resp.55to64 <- Inc.Resp.55to64*1000 / PY.Resp.55to64

# ignorable model, adjusted rate for those aged 55-64 for NR & attrition biases:
overall.rate.Ign.55to64 <-
weight1.55to64*rate.LFU.55to64 + weight2.55to64*rate.RQ.55to64 +
weight3.55to64*rate.TR1.55to64 + weight4.55to64*rate.TR2.55to64 +
weight5.55to64*rate.Resp.55to64

# nonignorable model + delta, adjusted rate for those aged 55-64:
overall.rate.NI.delta.55to64 <-
  weight1.55to64*( rate.LFU.55to64 +(abs(rate.LFU.55to64 - rate.Resp.55to64)))
+ weight2.55to64*( rate.RQ.55to64  - (abs(rate.RQ.55to64 - rate.Resp.55to64))) +
weight3.55to64*( rate.TR1.55to64 + (abs(rate.TR1.55to64 - rate.Resp.55to64))) +
weight4.55to64*( rate.TR2.55to64 + (abs(rate.TR2.55to64 - rate.Resp.55to64))) +
weight5.55to64* rate.Resp.55to64

# nonignorable model + ½ delta, adjusted rate for those aged 55-64
overall.rate.NI.0.5delta.55to64 <-
weight1.55to64*(rate.LFU.55to64 + 0.5*(abs(rate.LFU.55to64 -  rate.Resp.55to64)))
+ weight2.55to64*(rate.RQ.55to64 - 0.5*(abs(rate.RQ.55to64 -  rate.Resp.55to64)))
+ weight3.55to64*(rate.TR1.55to64 + 0.5*(abs(rate.TR1.55to64 -  rate.Resp.55to64)))
+ weight4.55to64*(rate.TR2.55to64 + 0.5*(abs(rate.TR2.55to64 -  rate.Resp.55to64)))
 + weight5.55to64*rate.Resp.55to64
```

```
# nonignorable model - ½ delta, rate for those aged 55-64
overall.rate.NI.minus.55to64 <-
weight1.55to64*(rate.LFU.55to64 - 0.5*(abs(rate.LFU.55to64 - rate.Resp.55to64)))
 + weight2.55to64*(rate.RQ.55to64 + 0.5*(abs( rate.RQ.55to64 -  rate.Resp.55to64)))
+ weight3.55to64*(rate.TR1.55to64 - 0.5*(abs(rate.TR1.55to64 - rate.Resp.55to64)))
+ weight4.55to64*(rate.TR2.55to64 - 0.5*(abs(rate.TR2.55to64 - rate.Resp.55to64)))
+ weight5.55to64*rate.Resp.55to64


                                    # 65-74 age group
                        # Censored
for (i in 4545:6290)
{
logit(p.5[i]) <- alpha5  + beta5.prev*prev[i] + beta5.age*age[i] +
beta5.HxFr*HxFr[i] + beta5.FamHx*FamHx[i]
Inc[i] ~ dbern(p.5[i])
}
alpha5 ~  dnorm(0, 0.0001)
beta5.prev ~  dnorm(0, 0.0001)
beta5.age ~  dnorm(0, 0.0001)
beta5.HxFr ~  dnorm(0, 0.0001)
beta5.FamHx ~ dnorm(0, 0.0001)

lambda.LFU.65to74 <- sum(Inc[4545:5506])
imputed.inc.LFU.65to74 ~    dpois(lambda.LFU.65to74)
Inc0.LFU.65to74 <-  962  - lambda.LFU.65to74
PY.LFU.65to74<- lambda.LFU.65to74*2.5 + Inc0.LFU.65to74*5
rate.LFU.65to74 <-   imputed.inc.LFU.65to74*1000 / PY.LFU.65to74


                        # RQ group
for (i in 6291:8189)
{
logit(p.6[i]) <-   alpha6 + beta6.age*age[i]+ beta6.HxFr*HxFr[i]   +
                beta6.FamHx*FamHx[i]
Inc[i] ~ dbern(p.6[i])
}
alpha6 ~  dnorm(0, 0.0001)
beta6.age ~  dnorm(0, 0.0001)
beta6.HxFr ~  dnorm(0, 0.0001)
beta6.FamHx ~ dnorm(0, 0.0001)

lambda.RQ.65to74 <- sum(Inc[6291:7405])
imputed.inc.RQ.65to74 ~      dpois(lambda.RQ.65to74)
Inc0.RQ.65to74<-  1115  - lambda.RQ.65to74
```

```
PY.RQ.65to74 <- lambda.RQ.65to74*2.5 + Inc0.RQ.65to74*5
rate.RQ.65to74 <- imputed.inc.RQ.65to74*1000 / PY.RQ.65to74

                    # TR1 group
for (i in 8190:9343)
{
logit(p.7[i]) <- alpha7 + beta7.age*age[i] + beta7.HH25*HH25[i]
Inc[i] ~ dbern(p.7[i])
}
alpha7 ~ dnorm(0, 0.0001)
beta7.age ~ dnorm(0, 0.0001)
beta7.HH25 ~ dnorm(0, 0.0001)

lambda.TR1.65to74 <- sum(Inc[8190:8546])
imputed.inc.TR1.65to74 ~ dpois(lambda.TR1.65to74)
Inc0.TR1.65to74 <- 357 - lambda.TR1.65to74
PY.TR1.65to74 <- lambda.TR1.65to74*2.5 + Inc0.TR1.65to74*5
rate.TR1.65to74 <- imputed.inc.TR1.65to74*1000 / PY.TR1.65to74


                    # TR2 group
for (i in 9344:10672)
{
logit(p.8[i]) <- alpha8 + beta8.age*age[i] + beta8.tel*tel[i]
Inc[i] ~ dbern(p.8[i])
}
alpha8 ~ dnorm(0, 0.0001)
beta8.age ~ dnorm(0, 0.0001)
beta8.tel ~ dnorm(0, 0.0001)

lambda.TR2.65to74 <- sum(Inc[9344:9875])
imputed.inc.TR2.65to74 ~ dpois(lambda.TR2.65to74)
Inc0.TR2.65to74 <- 532 - lambda.TR2.65to74
PY.TR2.65to74 <- lambda.TR2.65to74*2.5 + Inc0.TR2.65to74*5
rate.TR2.65to74 <-imputed.inc.TR2.65to74*1000 / PY.TR2.65to74

          # Respondents aged 65-74
lambda.Resp.65to74 <- sum(Inc[9876:10672])
Inc.Resp.65to74 ~ dpois(lambda.Resp.65to74)
Inc0.Resp.65to74 <- 797 - lambda.Resp.65to74
PY.Resp.65to74 <- lambda.Resp.65to74*2.5 + Inc0.Resp.65to74*5
rate.Resp.65to74 <- Inc.Resp.65to74*1000 / PY.Resp.65to74

overall.rate.Ign.65to74 <- weight1.65to74*rate.LFU.65to74 +
weight2.65to74* rate.RQ.65to74 + weight3.65to74*rate.TR1.65to74 +
weight4.65to74*rate.TR2.65to74 + weight5.65to74*rate.Resp.65to74
```

```
overall.rate.NI.delta.65to74  <-
  weight1.65to74*( rate.LFU.65to74 + (abs(rate.LFU.65to74 - rate.Resp.65to74)))
+ weight2.65to74*( rate.RQ.65to74 - (abs(rate.RQ.65to74 - rate.Resp.65to74)))
+ weight3.65to74*( rate.TR1.65to74 - (abs(rate.TR1.65to74 - rate.Resp.65to74)))
+ weight4.65to74*( rate.TR2.65to74 + (abs(rate.TR2.65to74 - rate.Resp.65to74)))
+ weight5.65to74* rate.Resp.65to74

overall.rate.NI.0.5delta.65to74  <-
weight1.65to74*(rate.LFU.65to74 + 0.5*( abs(rate.LFU.65to74 -  rate.Resp.65to74)))
+ weight2.65to74*( rate.RQ.65to74 -  0.5*( abs(rate.RQ.65to74 -  rate.Resp.65to74)))
 + weight3.65to74*(rate.TR1.65to74 - 0.5*( abs(rate.TR1.65to74 -  rate.Resp.65to74)))
 + weight4.65to74*(rate.TR2.65to74 + 0.5*( abs(rate.TR2.65to74 -  rate.Resp.65to74)))
 + weight5.65to74*rate.Resp.65to74

overall.rate.NI.minus.65to74  <-
weight1.65to74*(rate.LFU.65to74 - 0.5*( abs(rate.LFU.65to74 - rate.Resp.65to74)))
 +  weight2.65to74*(  rate.RQ.65to74 +  0.5*( abs(rate.RQ.65to74 -  rate.Resp.65to74)))
+  weight3.65to74*( rate.TR1.65to74 + 0.5*( abs(rate.TR1.65to74 - rate.Resp.65to74))
 +  weight4.65to74*( rate.TR2.65to74 - 0.5*( abs(rate.TR2.65to74 -
rate.Resp.65to74)))
 +  weight5.65to74* rate.Resp.65to74


                                # 75+ age group
                          # Censored
for (i in 10673:11713)
{
logit(p.9[i]) <- alpha9  + beta9.prev*prev[i] + beta9.age*age[i]
                + beta9.HxFr*HxFr[i]    + beta9.Rx*Rx[i]
                  + beta9.BMD*BMD[i]
Inc[i] ~ dbern(p.9[i])
}
alpha9 ~  dnorm(0, 0.0001)
beta9.prev ~  dnorm(0, 0.0001)
beta9.age ~  dnorm(0, 0.0001)
beta9.HxFr ~  dnorm(0, 0.0001)
beta9.Rx ~ dnorm(0, 0.0001)
beta9.BMD ~ dnorm(0, 0.0001)

lambda.LFU.75 <- sum(Inc[10673:11166])
Inc.LFU.75 ~    dpois(lambda.LFU.75)
Inc0.LFU.75 <-  494  - lambda.LFU.75
PY.LFU.75 <-  lambda.LFU.75*2.5+ Inc0.LFU.75*5
rate.LFU.75 <-     Inc.LFU.75*1000 / PY.LFU.75

                          # RQ Group
for (i in 11714:13237)
```

```
{
logit(p.10[i]) <-  alpha10+ beta10.age*age[i] + beta10.HxFr*HxFr[i]  +
beta10.FamHx*FamHx[i]
Inc[i] ~ dbern(p.10[i])
}

alpha10 ~  dnorm(0, 0.0001)
beta10.age ~  dnorm(0, 0.0001)
beta10.HxFr ~  dnorm(0, 0.0001)
beta10.FamHx ~ dnorm(0, 0.0001)

lambda.RQ.75 <- sum(Inc[11714:12695])
imputed.inc.RQ.75 ~  dpois(lambda.RQ.75)
Inc0.RQ.75 <-  982  -  lambda.RQ.75
PY.RQ.75 <- lambda.RQ.75*2.5 + Inc0.RQ.75*5
rate.RQ.75 <-  imputed.inc.RQ.75*1000 / PY.RQ.75

                        # TR1 group
for (i in 13238:14178)
{
logit(p.11[i]) <-  alpha11 +  beta11.age*age[i]  +  beta11.HH25*HH25[i]
Inc[i] ~ dbern(p.11[i])
}

alpha11 ~  dnorm(0, 0.0001)
beta11.age ~  dnorm(0, 0.0001)
beta11.HH25 ~  dnorm(0, 0.0001)

lambda.TR1.75 <- sum(Inc[13238:13629])
imputed.inc.TR1.75 ~    dpois(lambda.TR1.75)
Inc0.TR1.75 <-  392  -  lambda.TR1.75
PY.TR1.75 <-  lambda.TR1.75*2.5 + Inc0.TR1.75*5
rate.TR1.75 <-     imputed.inc.TR1.75*1000 / PY.TR1.75

                        # TR2 group
for (i in 14179:15299)
{
logit(p.12[i]) <-  alpha12+  beta12.age*age[i]  +  beta12.tel*tel[i]
Inc[i] ~ dbern(p.12[i])
}

alpha12 ~  dnorm(0, 0.0001)
beta12.age ~  dnorm(0, 0.0001)
beta12.tel ~  dnorm(0, 0.0001)

lambda.TR2.75 <- sum(Inc[14179:14749])
imputed.inc.TR2.75 ~    dpois(lambda.TR2.75)
```

```
Inc0.TR2.75 <-  571 - lambda.TR2.75
PY.TR2.75 <- lambda.TR2.75*2.5 + Inc0.TR2.75*5
rate.TR2.75 <-     imputed.inc.TR2.75*1000 / PY.TR2.75

        # Respondents aged 75+
lambda.Resp.75 <-  sum(Inc[14750:15299])
Inc.Resp.75 ~    dpois(lambda.Resp.75)
Inc0.Resp.75 <-   550 -  lambda.Resp.75
PY.Resp.75 <-     lambda.Resp.75*2.5 + Inc0.Resp.75*5
rate.Resp.75 <-    Inc.Resp.75*1000 / PY.Resp.75


overall.rate.Ign.75  <-  weight1.75*rate.LFU.75  + weight2.75*  rate.RQ.75
+ weight3.75* rate.TR1.75  + weight4.75* rate.TR2.75  +
  weight5.75*rate.Resp.75

overall.rate.NI.delta.75  <-
weight1.75*( rate.LFU.75 +   (abs(rate.LFU.75 - rate.Resp.75)))
+ weight2.75*(rate.RQ.75  - (abs(rate.RQ.75 - rate.Resp.75)))
+ weight3.75*(rate.TR1.75 + (abs(rate.TR1.75 - rate.Resp.75)))
+ weight4.75*( rate.TR2.75 + (abs(rate.TR2.75 - rate.Resp.75)))
+ weight5.75* rate.Resp.75

overall.rate.NI.0.5delta.75  <-
weight1.75*(rate.LFU.75 + 0.5*( abs(rate.LFU.75-  rate.Resp.75)))
+ weight2.75*(rate.RQ.75 - 0.5*( abs(rate.RQ.75 -  rate.Resp.75)))
+ weight3.75*(rate.TR1.75 + 0.5*( abs(rate.TR1.75 - rate.Resp.75)))
+ weight4.75*(rate.TR2.75 + 0.5*( abs(rate.TR2.75 -  rate.Resp.75)))
+ weight5.75*rate.Resp.75

overall.rate.NI.minus.75  <-
weight1.75*(rate.LFU.75 - 0.5*( abs(rate.LFU.75 - rate.Resp.75)))
+  weight2.75*( rate.RQ.75 + 0.5*(abs(rate.RQ.75 - rate.Resp.75)))
+  weight3.75*(rate.TR1.75 - 0.5*( abs(rate.TR1.75 -  rate.Resp.75)))
+  weight4.75*(rate.TR2.75 - 0.5*( abs(rate.TR2.75 -  rate.Resp.75)))
+  weight5.75*rate.Resp.75



# Rates for the 55+ age group for the different participation groups:
rate.55.Resp <- w.55to64*rate.Resp.55to64 +  w.65to74*rate.Resp.65to74 +
                  w.75 *rate.Resp.75

rate.55.LFU <- w.55to64*rate.LFU.55to64 +  w.65to74*rate.LFU.65to74 +
                  w.75 *rate.LFU.75

rate.55.RQ <-  w.55to64*rate.RQ.55to64 +  w.65to74*rate.RQ.65to74 +
```

```
                w.75*rate.RQ.75

rate.55.TR1 <- w.55to64*rate.TR1.55to64 +  w.65to74*rate.TR1.65to74 +
                w.75*rate.TR1.75

rate.55.TR2 <- w.55to64*rate.TR2.55to64 + w.65to74*rate.TR2.65to74 +
               w.75 *rate.TR2.75

# Rates for the 55+ age group in the 4 missingness models
overall.rate.Ign.55 <-  w.55to64*overall.rate.Ign.55to64 +
w.65to74*overall.rate.Ign.65to74 +  w.75 *overall.rate.Ign.75

overall.rate.NI.delta.55 <-  w.55to64*overall.rate.NI.delta.55to64 +
w.65to74*overall.rate.NI.delta.65to74 +  w.75*overall.rate.NI.delta.75

overall.rate.NI.0.5delta.55 <-  w.55to64*  overall.rate.NI.0.5delta.55to64 +
w.65to74*  overall.rate.NI.0.5delta.65to74 + w.75*  overall.rate.NI.0.5delta.75

overall.rate.NI.minus.55 <-  w.55to64* overall.rate.NI.minus.55to64 + w.65to74*
overall.rate.NI.minus.65to74 + w.75 *  overall.rate.NI.minus.75

# adjusted rates for attrition bias for each age group:
LFU.rate.Ign.55to64  <- 0.5120811*rate.LFU.55to64 + 0.4879189*rate.Resp.55to64

LFU.rate.Ign.65to74  <- 0.5509737*rate.LFU.65to74 + 0.4490264*rate.Resp.65to74

LFU.rate.Ign.75 <- 0.4745437*rate.LFU.75 + 0.5254563*rate.Resp.75

LFU.rate.Ign.55 <-     w.55to64* LFU.rate.Ign.55to64 + w.65to74* LFU.rate.Ign.65to74
+  w.75* LFU.rate.Ign.75

LFU.rate.NI.delta.55to64  <-  0.5120811*(rate.LFU.55to64 + (rate.LFU.55to64 -
rate.Resp.55to64))+ 0.4879189*rate.Resp.55to64

LFU.rate.NI.delta.65to74  <-  0.5509737*(rate.LFU.65to74 + (rate.LFU.65to74 -
rate.Resp.65to74)) + 0.4490264*rate.Resp.65to74

LFU.rate.NI.delta.75  <-  0.4745437*(rate.LFU.75 + (rate.LFU.75 - rate.Resp.75))
+ 0.5254563*rate.Resp.75

LFU.rate.NI.delta.55 <- w.55to64* LFU.rate.NI.delta.55to64 + w.65to74*
LFU.rate.NI.delta.65to74 +  w.75* LFU.rate.NI.delta.75

LFU.rate.NI.0.5delta.55to64  <-  0.5120811*(rate.LFU.55to64 + 0.5*(rate.LFU.55to64 -
rate.Resp.55to64)) + 0.4879189*rate.Resp.55to64
```

```
LFU.rate.NI.0.5delta.65to74  <- 0.5509737*(rate.LFU.65to74 + 0.5*(rate.LFU.65to74 -
rate.Resp.65to74)) + 0.4490264*rate.Resp.65to74

LFU.rate.NI.0.5delta.75  <- 0.4745437*(rate.LFU.75 + 0.5*(rate.LFU.75 -
rate.Resp.75)) + 0.5254563*rate.Resp.75

LFU.rate.NI.0.5delta.55 <-    w.55to64* LFU.rate.NI.0.5delta.55to64 + w.65to74*
LFU.rate.NI.0.5delta.65to74 +  w.75* LFU.rate.NI.0.5delta.75

LFU.rate.NI.minus.55to64  <- 0.5120811*(rate.LFU.55to64 - 0.5*(rate.LFU.55to64 -
rate.Resp.55to64)) + 0.4879189*rate.Resp.55to64

LFU.rate.NI.minus.65to74  <- 0.5509737*(rate.LFU.65to74
- 0.5*(rate.LFU.65to74 - rate.Resp.65to74)) +  0.4490264*rate.Resp.65to74

LFU.rate.NI.minus.75  <- 0.4745437*(rate.LFU.75  -  0.5*(rate.LFU.75 - rate.Resp.75))
+ 0.5254563*rate.Resp.75

LFU.rate.NI.minus.55 <-      w.55to64* LFU.rate.NI.minus.55to64 +
w.65to74* LFU.rate.NI.minus.65to74  +   w.75* LFU.rate.NI.minus.75

}


# Initial values for 2 chains, which were derived by taking a range of the MLE
#Chain 1
list(alpha1 = -5.6,  beta1.prev = 0.1 , beta1.age = -0.001,  beta1.HxFr = 0.1,  beta1.FamHx
= 0.1, beta1.BMD = 0.05,
alpha2 = -8, beta2.age = -0.005,  beta2.HxFr = 0.1, beta2.FamHx = 0.3,
alpha3 = -10, beta3.age = 0.005, beta3.HH25 = -0.5,
alpha4 = -10, beta4.age = 0.01 , beta4.tel = -0.05,
alpha5 = -20, beta1.prev = 0.5, beta5.age = 0.005, beta5.HxFr = 0.1, beta5.FamHx = 0.1,
alpha6 = -20, beta6.age = 0.01,  beta6.HxFr = 0.1,  beta6.FamHx = 0.1,
alpha7 = -20,  beta7.age = 0.01, beta7.HH25 = -1,
alpha8 = -20, beta8.age = 0.05, beta8.tel = -0.005,
alpha9 = -0.5,  beta9.prev = 0.3 , beta9.age = 0,  beta9.HxFr = 0.23,
beta9.BMD = -3, beta9.Rx = 0.4,
alpha10 = -5, beta10.age = 0,  beta10.HxFr = 0.25,  beta10.FamHx = 0.2,
alpha11 = -5.2, beta11.age = 0, beta11.HH25 = -0.5,
alpha12 = -6, beta12.age = 0, beta12.tel = -0.05)

# Initial values
# Chain2
list(alpha1 = -1,  beta1.prev = 1 , beta1.age = 0.01,  beta1.HxFr = 1,  beta1.FamHx = 1,
beta1.BMD = 0.5,
alpha2 = -2, beta2.age = 0.05,  beta2.HxFr = 1,  beta2.FamHx = 2,
alpha3 = -1,  beta3.age = 0.5, beta3.HH25 = -0.05,
```

alpha4 = -1, beta4.age = 0.08 , beta4.tel = -0.2,
alpha5 = -5, beta5.prev = 2 , beta5.age = 0.2, beta5.HxFr = 0.8, beta5.FamHx = 0.8,
alpha6 = -5, beta6.age = 0.5, beta6.HxFr = 1, beta6.FamHx = 1,
alpha7 = -5, beta7.age = 0.5, beta7.HH25 = -0.1,
alpha8 = -5, beta8.age = 0.5 , beta8.tel = -0.2 ,
alpha9 = -0.05, beta9.prev = 1.5 , beta9.age = -0.03, beta9.HxFr = 1,
beta9.BMD = -0.5, beta9.Rx = 2,
alpha10 = -1, beta10.age = 0.01, beta10.HxFr = 1, beta10.FamHx = 0.9,
alpha11 = -0.5, beta11.age = 0.1 , beta11.HH25 = -0.05,
alpha12 = -0.5, beta12.age = 0.05 , beta12.tel = -0.002)

 # Data
list(weight1.55to64 = 0.2475509 ,weight2.55to64 = 0.243793 , weight3.55to64 = 0.0764883, weight4.55to64 = 0.1948003, weight5.55to64 = 0.2373775, weight1.65to74 = 0.2555792 , weight2.65to74 = 0.2962274, weight3.65to74 = 0.0948459, weight4.65to74 = 0.141339, weight5.65to74 = 0.2120085, weight1.75 = 0.1652727 , weight2.75 = 0.328538 , weight3.75 = 0.1311475, weight4.75 = 0.1910338 , weight5.75 = 0.184008, w.55to64 = 0.3883467 , w.65to74 = 0.3375752 , w.75 = 0.2740781)

#The last 3 weights are derived from the 1996 Canadian census