8-11-2015

The Role of Health Literacy and Numeracy on Exercise Self-efficacy and Exercise Behavior in the PAADRN Bone Health Intervention

Elizabeth A. Fallon

ABSTRACT

THE ROLE OF HEALTH LITERACY AND NUMERACY ON EXERCISE SELF-EFFICACY AND EXERCISE BEHAVIOR IN THE PAADRN BONE HEALTH INTERVENTION

By

ELIZABETH ANNE FALLON

JULY 21, 2015

INTRODUCTION: Osteoporotic bone fracture is a major cause of hospitalization, disability, loss of independent living capacity, and mortality among aging adults. Although physical exercise may sustain bone mineral density and prevent falls and fractures among individuals at risk for low bone mineral density, adherence to exercise recommendations is low. Increasing efficacy and effectiveness of treatment for osteoporosis would benefit from examination of heterogeneity of treatment effects. Previous research indicates that poor health literacy (HL) and health numeracy (HN) may be associated with less exercise behavior and heterogeneity of treatment effects may be evident across high and low level of health literacy and/or health numeracy.

AIM: Examine heterogeneity of treatment effects due to HL/HN on post-intervention exercise self-efficacy and exercise behavior among older adults enrolled in a large, multi-site randomized controlled trial designed to increase exercise as part of osteoporosis guideline concordant care.

METHODS: A secondary data analysis was conducted, utilizing a modified intent-to-treat approach. Linear mixed models with and without covariates were conducted to examine heterogeneity of treatment effects by incorporating the treatment by high/low HL/HN interaction. Analyses were conducted using pooled, as well as site-specific samples.

EMBARGO

THE ROLE OF HEALTH LITERACY AND NUMERACY ON EXERCISE SELF-EFFICACY AND EXERCISE BEHAVIOR IN THE PAADRN BONE HEALTH INTERVENTION

by

ELIZABETH A. FALLON

B.S., UNIVERSITY OF FLORIDA
M.S.E.S.S., UNIVERSITY OF FLORIDA
Ph.D., UNIVERSITY OF FLORIDA

A Thesis Submitted to the Graduate Faculty
of Georgia State University in Partial Fulfillment
of the
Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA
30303

APPROVAL PAGE

THE ROLE OF HEALTH LITERACY AND NUMERACY ON EXERCISE SELF-EFFICACY AND EXERCISE BEHAVIOR IN THE PAADRN BONE HEALTH INTERVENTION

by

ELIZABETH A. FALLON

Approved:

____
Douglas Roblin, Ph.D., Georgia State University
Committee Chair

____
Fredric Wolinsky, Ph.D., University of Iowa
Committee Member

Date: July 21, 2015

DEDICATION

To Brett -

For loving me every day,
and being my true north in every new adventure.

ACKNOWLDEGEMENTS

First and foremost, thank you to my committee chair, Dr. Douglas Roblin (Georgia State University), who embodies the noblest ideals of science and academia. His focus on good science through interdisciplinary collaborations has challenged me to broaden my knowledge base, while encouraging me to contribute from my existing areas of expertise. I could not have imagined a better mentor during this time of professional growth and transition.

Thank you to Dr. Fredric Wolinsky (Thesis Committee Member, University of Iowa), for your brilliant ideas for solving data access problems, and your valuable time in reading this document, and engaging in my thesis defense.

To the entire PAADRN Steering Committee: Dr. Wolinsky, Dr. Roblin, Dr. Peter Cram, Dr. Ken Saag - thank you for your openness to a new collaborator at such a demanding time in the PAADRN trial.

To the members of the PAADRN team (Stephanie Edmonds, Dr. Yiyue Lou, and Thuy Nguyen) who acclimate me to the PAADRN data - Thank you for sharing data dictionaries, creating simulated data sets, and ultimately, running my SAS code to send the output.

To the GSU biostatistics faculty (Dr. Lai, Dr. Luo, and Dr. Hayat), who put so much time and effort into an evolving and excellent curriculum. I'm already putting my new skills into practice, everyday!

To Tracy Ayers – thank you for being the best SAS teacher, ever.

Thank you to Jessica Pratt, Gina Sample, and Dr. Maggie Tolan who shepherded me (and all my paperwork and petitions) through the MPH process, from admission to graduation. Your passion, work ethic, and thoughtful advice for professional development have been invaluable.

Finally, to my wonderful family, friends and colleagues: George Fallon, Keith Fallon, Sara Wilcox, Bryan Blissmer, Debra Riebe, Colleen Redding, Ninoska Peterson, Cheryl Der Ananian, Melissa Bopp, Brandonn Harris, Emily Murphy, Kim Fournier, Paige Johnson, and Joey Lightner – Thank you for your patience, support, laughter, guidance, mentorship, and countless letters of recommendation.

Author's Statement Page

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

Author: Elizabeth A. Fallon

____

Signature of Author

TABLE OF CONTENTS

List of Tables

List of Figures

Introduction

Osteoporosis is characterized by having bone mineral density (BMD) more than 2.5 standard deviations below normal for a healthy 30-year old adult [1,2]. Osteopenia is having a BMD value more than one standard deviation below normal for a healthy 30-year old adult, but less than the 2.5 standard deviations required for osteoporosis diagnosis. The major deleterious health outcome of osteoporosis is bone fracture, typically of the hip, vertebrae, and wrist, and is a major cause of hospitalization, loss of productivity, loss of independent living capacity, and disability among aging adults [3]. Furthermore, among those sustaining hip fractures, mortality rate within one year of fracture is 20%[4].

In the United States, 10.2 million adults have osteoporosis, with osteopenia affecting an additional 43.4 million[5]. Currently, annual United States osteoporotic-related fracture costs are an estimated $US 16.9 billion, with annual costs projected to rise to $US 25 billion by 2025. Among women 55 years and older, costs associated with osteoporotic-related bone fracture are more costly than myocardial infarction, stroke, or breast cancer[6].

Internationally, an estimated 49 million individuals over 50 years have osteoporosis, with higher prevalence among women (9% - 38%), compared to men (1% - 8%)[7]. Similar to the United States, costs associated with osteoporotic-related fracture are comparable to that of stroke and myocardial infarction[8], and are projected to increase six-fold by 2050 [9].

Thus, due to the significant morbidity, mortality, and economic costs of osteoporotic-fracture and the projected population growth among individuals over 65 years, effective public health interventions designed to improve bone health and ultimately reduce risk of falls and bone fracture are needed.

LITERATURE REVIEW

National Osteoporosis Foundation Recommendations for Prevention and Treatment of Osteoporosis. The National Osteoporosis Foundation provides detailed recommendations for the prevention, risk assessment, diagnosis, and treatment of osteoporosis in clinical settings[10]. Briefly, prevention and risk assessment recommendations state that all men age ≥50 years and postmenopausal women should be evaluated for osteoporosis risk and universally counseled on prevention actions (e.g., disease risk, fracture risk, falls prevention, calcium and vitamin D dietary requirements, necessary weight-bearing, muscle strengthening, flexibility and balance exercise, smoking cessation, and limiting alcohol consumption). Additionally, height should be measured annually, using a stadiometer. Approved pharmacologic treatment (e.g., bisphosphonates, calcitonin, estrogen agonist/antagonist, estrogen or hormone therapy) should be implemented for:

(a) individuals with BMD ≤ -2.5 standard deviations below normal for a healthy 30-year old adult,

(b) men aged ≥50 years and postmenopausal women with low bone mineral density (-1.0 to -2.5 standard deviations below normal for a healthy 30-year old adult) and a 10-year fracture risk ≥3%, or

(c) a 10-year fracture risk score ≥20%.

Finally, according to National Osteoporosis Foundation recommendations, regular monitoring (minimum every 1-2 years [10]) using BMD testing and biochemical markers should be conducted to determine the efficacy of individual treatment programs, and use as basis of treatment initiation and/or modification when indicated [10].

Role of Physical Exercise for Bone Health. The National Osteoporosis Foundation recommends provider counseling, and subsequent patient self-management of, and adherence to, physical exercise. Exercise is an important component of osteoporosis treatment as scientific evidence suggests that exercise interventions reduce falls [11], sustain and may even modestly increase BMD[10] among aging adults (See Appendix A[12-27]).

For the purposes of this thesis, systematic literature reviews and meta-analytic reviews exploring the effect of various exercise modalities on bone mineral density are summarized in Appendix A[12-27]. Because osteoporosis prevalence is higher among women, compared to men, 70.6% (12/17) of reviews have focused on the effect of exercise on BMD among adult women, with the majority (9/12, 75%) focused exclusively on postmenopausal women. Across the 12 reviews focusing on women, there is an overall positive effect of exercise on BMD for the:

(a) lumbar spine (total reporting = 9 reviews; positive effect = 8, no effect = 1, negative effect = 0) and

(b) femur/femoral neck (total reporting = 12 reviews w/13 effect sizes; positive effect = 7, no effect = 6, negative effect = 0).

The effect of exercise on other skeletal sites (e.g., total hip, vertebral, radius, whole body) has also been systematically reviewed within the literature, but to a lesser extent. Despite fewer studies assessing these other skeletal sites, authors generally note that the findings are similar to that of the lumbar spine and femoral neck, suggesting that physical exercise likely produces a small, but clinically relevant increase in BMD among women.

In contrast to the large number of reviews focusing on women, only three literature reviews have focused exclusively on adult men, all reporting a positive effect of exercise on the femur/femoral neck. Only one review concluded that exercise significantly increases BMD for the lumbar spine (no effect = 2).

Regardless of the gender focus of the research, several reviews have noted the lack of methodological quality among controlled trials examining the effect of exercise on bone density. Many trials do not have sufficient power to detect change, are heterogeneous in their exercise protocols, fail to report randomization methods or use blinding, fail to report compliance/drop-out, and fail to conduct intent-to-treat analysis and assess/incorporate important confounders (e.g., use of contraception, hormone replacement therapy). Furthermore, few studies have provided long-term follow-up data.

In light of these methodological limitations within the literature, perhaps the best approach is to rely on the following conclusions from the most comprehensive and rigorous systematic literature review of randomized controlled trials examining exercise on BMD[28]. These conclusions are that:

a) low intensity resistance training showed no effect on BMD at any skeletal site,
b) high intensity resistance training showed a significant effect for the spine and femoral neck, but not total hip,
c) low intensity weight bearing aerobic exercise (e.g., walking, tai chi) had a significant effect on the spine, while
d) high intensity weight bearing aerobic exercise (e.g., jogging, jumping) had a significant effect on BMD for total hip and trochanter, but no significant effect for femoral neck spin, med femur, or tibia [28].

An unfortunate consequence of poor methodological quality within this area of research is the wide variation in physical activity recommendations adopted across professional organizations interested in aging adults, exercise, and bone health (See Appendix B) [10,29-35]. While there is agreement that weight-bearing aerobic and anaerobic exercise are beneficial for bone health, future interdisciplinary collaboration is greatly needed to assess recent advances in exercise and bone health research and based on these assessments, create and promote consistent set of recommendations for physical activity and optimal bone health. Specifically, patients would benefit from health promotion messages that provide:

(a) minimum exercise duration needed for treatment efficacy and

(b) specific guidance regarding frequency, intensity, time per session, number of sets/repetitions, mode, or specific exercises most beneficial to improve BMD in general or at specific body sites.

In conclusion, individuals utilizing physical exercise to improve bone health are best advised to engage in a combination of moderate to high intensity weight-bearing aerobic activities (e.g., jogging, jumping, stair climbing) and a whole-body resistance-training regimen. Regular reassessments of exercise regimen should be conducted to ensure variety, and continued fitness progression. Furthermore, future clinical trials examining the effect of exercise on BMD should better adhere to recommendations for progressive resistance training, cardiovascular health, and older adults [32,36,37] as the foundational starting point for developing experimental exercise protocols for bone health. Finally, public health interventions promoting exercise among older adults with osteoporosis or osteopenia should:

(a) separately promote and measure aerobic weight bearing exercise and resistance/strength training exercise,

(b) stratify aerobic, strength training, and combined aerobic/strength training exercise behavior by two (not meeting/meeting recommendations) or three groups (sedentary/insufficient activity/meets recommendations), and

(c) use both self-report and objective measures of physical activity, where feasible.

Public Health Interventions for Exercise among Individuals at risk of Osteoporosis. Because physical exercise can sustain and increase BMD, establishing efficacious public health interventions to initiate and maintain sufficient levels of physical activity among aging adults at risk of osteoporosis and osteoporotic fracture becomes paramount.

The earliest reviews of osteoporosis disease management identified only 2 and 3 randomized control trials using exercise for disease management [38] and prevention of osteoporotic fractures [39], respectively. Lock et al. [39] concluded from their review that these interventions resulted in a lower risk of spinal fractures, but this finding was not statistically significant [total 322 participants; RR = 0.52, 95% CI: 0.17, 1.60]. In 2010, Lai et al. [40] reviewed a total of 24 randomized control trials of healthcare interventions for community-dwelling postmenopausal women with osteoporosis and found that 80% (4/5 studies) showed improvement in calcium intake, while only 25% (1/4 studies) showed increases in exercise. Most recently, a systematic review of multifaceted osteoporosis group education interventions summarized a total of seven studies, and only four reporting outcome measures for physical activity behavior and/or physical fitness. Of these 75% (3/4) reported significant, positive changes in exercise behavior and/or fitness.

Fortunately, in the absence of literature specifically focused on lifestyle interventions to improve bone health among older adults, there is much to be learned from the plethora of recent systematic and meta-analytic reviews summarize public health physical activity interventions aimed at aging adults, in general [41-45], and frail older adults at risk for falls [46,47]. Overall, these reviews provide substantial evidence in support of:

(a) theory-based psychosocial and behavioral interventions [41,44],

(b) the importance of self-efficacy in the behavior change process [41],

(c) the use of “mediated” (non-face-to-face via mailed print materials), “technology mediated” (telephone, internet) interventions [44] and remote feedback interventions [42],

(d) the benefit of health care provider and health clinic-based interventions [45,48], and

(e) successful long-term maintenance of physical activity at 12 months follow-up, but not 24 months follow-up [43].

Thus, behavioral and public health scientists have established a robust foundation from which scientists and practitioners can learn about efficacious interventions to increase physical activity and improve health outcomes of older adults.

Public health interventions have some limitations. There is limited information on factors that might effect external validity [47]. Little is known about the successful initiation and maintenance of resistance training in older adults [49]; an exceptional oversight given the large number of older adults at risk for low BMD and osteoporotic fractures, who would benefit greatly from moderate- to high-intensity resistance training[28]. Finally, more research is needed to better understand mediators of treatment effects and factors leading to heterogeneity of treatment effects, which may enable increased efficacy of health communication approaches, individualized tailoring, and potentially improve cost-effectiveness.

Mediators and heterogeneity of physical activity intervention effects. Due to the extensive number of physical activity interventions aimed at aging adults, systematic and meta-analytic reviews have examined heterogeneity of treatment effects and reported that intervention-induced changes in physical activity behavior were unrelated to gender [50,51], age [50], race/ethnicity [50], socioeconomic distribution [50], and delivery method (e.g., home vs. center-based, telephone, mail) [50].

Compared to the examination of factors leading to heterogeneity of treatment effects, mediation analyses have only recently been widely implemented for physical activity interventions. Thus, despite heterogeneity of behavioral interventions for physical activity, only a small number of mediators have been examined. A review of 23 reported studies revealed that insufficient data was available to advocate for one behavioral theory over another, but did note that self-regulation constructs had the best empirical evidence for successful mediation [52].

Other than socio-economic and demographic characteristics (that are not mutable), health literacy and health numeracy are two constructs that potentially influence the efficacy and effectiveness of healthcare and lifestyle interventions for chronic disease. Self-management regimens required for chronic diseases (e.g., type II diabetes, heart disease, cancer and osteoporosis) are often complex, and require patient comprehension of treatment options, health insurance benefits, provider instructions regarding medication side effects and adherence, and lifestyle changes (e.g., diet and exercise). As such, lower health literacy has been consistently associated with greater emergency room admissions and hospitalization, lower use of evidence-based preventive services (e.g., mammography screenings and flu vaccines), lower medication adherence, poorer ability to interpret labels and health messages, and poorer overall health status, and higher mortality rates among aging adults [53].

Health literacy/numeracy. The most widely accepted definition of health literacy is "the degree to which individuals have the capacity to obtain, process, and understand basic information and services needed to make appropriate decisions regarding their health" [54,55]. The Institute of Medicine further expands health literacy into four domains [56] considered to be aspects of "health knowledge":

(1) cultural and conceptual knowledge,

(2) oral literacy, including speaking and listening skills,

(3) print literacy, including writing and reading skills, and

(4) numeracy.

Over time, some researchers have also proposed to expand this rather "patient-centric" perspective to include the dynamics of patient-provider interactions, and patient-healthcare system interactions. If the healthcare system is more sensitive to low health literacy, then health outcomes could be improved by the wide scale adoption of effective communication strategies to mitigate its deleterious influence [57].

Health numeracy has gained increased attention as an independent, complementary concept to health literacy because some research has suggested it influences comprehension of disease risk, food labels, health monitoring tools and tests (e.g., blood glucose levels), and medication adherence [53]. Merging definitions from different sources, health numeracy is defined as "the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions." [58] Golbeck and colleagues [58] go on to propose four, overlapping categories:

(1) basic – skills to identify numbers, and make sense of numerical information that does not require any manipulation

(2) computational – ability to count, quantify, compute, use simple numerical manipulations

(3) analytical – the ability to make sense of through inference, estimation, proportions, percentages, frequencies

(4) statistical – the ability to understand probability statements, critically analyze quantitative health information, and understand concepts such as "randomization"

According to the National Assessment of Health Literacy [59], 36% of United States adults have limited health literacy, 22% have only basic health literacy, and 14% have below basic health literacy. Across 85 empirical studies examining health literacy, prevalence of low health literacy was 26% [95% CI: 22%, 29%], and prevalence of marginal health literacy was 20% [95% CI: 16%, 23%] [60]. Salient to osteoporosis research is the finding that age is strongly, negatively associated with health literacy and numeracy skills [61].

Health literacy/numeracy and health outcomes. A review of health literacy/numeracy on health outcomes included 98 articles on health literacy, 22 articles examining health numeracy, and 9 studies measuring both literacy and numeracy [53]. Results revealed that low health literacy was consistently associated with greater emergency room admissions and hospitalization, lower use of evidence-based preventive services (e.g., mammography screenings and flu vaccines), lower medication adherence, poorer ability to interpret labels and health messages, and poorer overall health status and higher mortality rates among aging adults. Salient to this study, no attempt was made to isolate the effect of health literacy on exercise behavior, specifically. Furthermore, the strength of the evidence for a relationship between health literacy and health behaviors in general, was deemed low and/or insufficient because there were too few studies to confidently calculate an effect [53].

Regarding the relationship between health numeracy and health outcomes, there were often too few studies to conduct a meta-analytic review, and when analyses were conducted (e.g., asthma management, healthcare utilization), the results were generally inconclusive. Furthermore, the authors noted the low or poor quality of many studies, specifically citing that most studies used cross-sectional designs, often had small sample sizes, used convenience samples, and many relied unadjusted analyses.

While research examining the relationship between health literacy/numeracy and health outcomes continues to be important, there is also need to better understand the mechanism(s) by which health literacy/numeracy effects health outcomes[62-64]. Paasche-Orlow & Wolf[63] proposed a model by which intrinsic self-care factors of patients (i.e., knowledge, skills, self-efficacy, problem-solving, motivation) mediate the relationship between health literacy and health outcomes. von Wagner and colleagues[64] use aspects of various theories within health psychology (e.g, health belief model, theory of planned behavior, transtheoretical model) expand to further expand the original model by separating Paasche-Orolow's intrinsic self-care factors into two distinct phases: motivational and volitional. Motivational phase includes the ability of health literacy to influence knowledge, understanding, beliefs, and attitudes. The volitional phase (also called action control) focuses on the ability of health literacy to influence implementation skills such as planning, organizing, and carrying out an action or behavior. Preliminary support for the von Wagner et al. model has been demonstrated through a recent literature review focusing on the association between health literacy and diabetes self-management behaviors[62]. Although no similar review was identified specifically for osteoporosis self-management or physical activity, Fransen et al.[62] provide insight into the value of von Wagner's model, and provides empirical support for the importance of examining the relationship between health literacy/numeracy and psychosocial constructs known to influence health behaviors.

Health literacy/numeracy and musculoskeletal health. Loke et al. [65] reviewed 8 studies examining health literacy and health outcomes among patients with musculoskeletal disease. Four studies (50%) focused on patients with arthritis, and four (50%) failed to distinguish patient diagnosis. None explicitly examined patients at risk for osteopenia/osteoporosis. The authors concluded there was no consistent relationship between health literacy and disease-specific outcome measures in patients with chronic musculoskeletal conditions. Similar to the Berkman et al. [53] review, however, the authors commented on the poor methodological quality of the studies included in the review. Specifically, the majority of studies were cross-sectional in design, contained low sample sizes, and several combined patients with various musculoskeletal diagnoses. Thus, relevant to this study, there is little evidence to support a relationship between health literacy/numeracy and health outcomes for osteopenia/osteoporosis.

Health Literacy/Numeracy and Exercise Behavior. Few studies have examined the relationship between health literacy/numeracy and exercise behavior. In their review, Berkman et al. [53] identified only five studies examining health literacy and health behaviors (e.g., grouping together healthy eating habits, exercise, and seat belt use), and only one study examining health numeracy and health behaviors.

A systematic literature review conducted for the purposes of this thesis (See Appendix C) [66-82] identified 17 studies that (a) directly measured health literacy and/or health numeracy, (b) assessed physical activity/exercise behavior, and (c) presented statistical analysis examining the relationship between these variables. Of these, 11 were cross-sectional, 2 were longitudinal (without randomization), and 4 were secondary data analyses from randomized controlled trials intended to change physical activity behavior. Furthermore, 82.3% (14/17) measured only health literacy, 1/17 measured only health numeracy, 1/17 measured both health literacy and health numeracy, separately, and 1/17 used a combined measure of health literacy and numeracy. Thus, while the research examining the relationship between health literacy and exercise behavior is growing, there is extremely little empirical research for:

(a) the relationship between health numeracy and exercise behavior,
(b) the unique influence of health literacy and health numeracy on exercise behavior,
(c) the interactive influence of these constructs on exercise behavior.

Among the cross-sectional studies examining health literacy, 55.6% (5/9) found no relationship between health literacy and physical activity behavior, 33.3% (3/9) found a statistically significant positive relationship between health literacy and physical activity, and 11% (1/9) revealed a negative relationship between health literacy and physical activity. The single cross-sectional study examining health numeracy, showed no relationship with physical activity behavior.

In three cross-sectional studies [72,73,76], authors examined theoretical frameworks proposing that knowledge mediates the relationship between health literacy and self-management behaviors [60,62,83]. Of these, two reported that the relationship between health literacy and physical activity was *not* mediated by knowledge [73,76].

Both studies utilizing longitudinal (non-randomized designs), found a small but statistically significant relationship between health literacy and physical activity behavior [77,84]. Only one of these studies employed multivariate mediation methods, and found that the health literacy-physical activity relationship was mediated by self-efficacy [77].

Finally, four randomized controlled trials examined the heterogeneity of treatment effects due to health literacy on intervention-induced changes to physical activity behavior. While none of these studies revealed heterogeneity of treatment effects due to health literacy [79-82], it is important to note that several of these studies varied in quality, limited by small sample sizes and poor reporting of statistical methodology.

In conclusion, while low health literacy has consistently been associated with poorer general health outcomes (e.g., greater emergency room admissions and hospitalization, lower use of evidence-based preventive services, lower medication adherence, poorer ability to interpret labels and health messages, and poorer overall health status and higher mortality rates among elderly adults), there is insufficient evidence to support a relationship between health literacy and bone health outcomes or physical activity behavior, specifically. Furthermore, even less empirical evidence is available to understand the role of low health numeracy for general health outcomes, musculoskeletal disease, osteoporosis, or exercise behavior.

PURPOSE AND HYPOTHESES

Due to the complexity of osteoporosis self-management regimens, the importance of physical activity behavior in sustaining and increasing BMD, and the lack of research examining the role of health literacy/numeracy in osteoporosis self-management and specifically for physical activity behavior, the purpose of this study is to examine the heterogeneity of treatment effects due to of health literacy/numeracy on post-intervention exercise self-efficacy and exercise behavior among older adults enrolled in a large, multi-site randomized controlled trial designed to increase osteoporosis guideline concordant care. Specific research objectives, with corresponding hypotheses are:

<u>Research Objective 1:</u> To determine whether there is variation in exercise behavior across high and low levels of health literacy/numeracy. For each of these hypotheses,

the dependent variable (exercise behavior) will be assessed using both a continuous and a categorical measure of exercise behavior.

Hypothesis H1.A.1. Compared to the high health literacy group, the low health literacy group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.

Hypothesis H1.A.2. Compared to the high health literacy group, the low health literacy group will have a lower proportion of individuals with exercise behaviors that are consistent with National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.

Hypothesis H1.B.1. Compared to the high numeracy ability group, the low health numeracy ability group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.

Hypothesis H1.B.2. Compared to the high numeracy ability group, the low health numeracy ability group will have a lower proportion of individuals with exercise behaviors that are consistent with National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.

Hypothesis H1.C.1. Compared to the high preference for numerical display group, the low preference for numerical display group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.

Hypothesis H1.C.2. Compared to the high preference for numerical display group, the low preference for numerical display group will have a lower proportion of individuals with exercise behaviors that are consistent with National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.

Research Objective 2: To determine whether there is variation in exercise self-efficacy across high and low levels of health literacy/numeracy. For each of these hypotheses, the dependent variable (exercise self-efficacy) will be assessed using a continuous measure.

Hypothesis H2.A. Compared to the high health literacy group, the low health literacy group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.

Hypothesis H2.B. Compared to the high numeracy ability group, the low numeracy ability group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.

Hypothesis H2.C. Compared to the high preference for numeric display group, the low preference for numeric display group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.

Research Objective 3: Pending the identification of the hypothesized difference for exercise behavior across low and high levels of health literacy/numeracy (Research Objective 1), research objective three is to test whether level of health literacy/numeracy results in heterogeneity of treatment effects for exercise behavior outcomes. Each analysis will be conducted for the pooled sample, as well as for each treatment site (i.e., University of Iowa, University of Alabama Birmingham, and Kaiser Permanente, Georgia).

Hypothesis H3.A.1. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for 12-week exercise behavior. Specifically, individuals with low health literacy in the control group will have the lowest levels of exercise behavior, compared to individuals with low health literacy in the treatment group, and compared to individuals with high health literacy (regardless of treatment condition).

Hypothesis H3.A.2. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for meeting exercise recommendations at 12-week follow-up. Specifically, individuals with low health literacy in the control group will have the lowest proportion of individuals with exercise behaviors that are consistent with National Osteoporosis Foundation guidelines, compared to individuals with low health literacy in the treatment group, and compared to individuals with high health literacy (regardless of treatment condition).

Hypothesis H3.B.1. Heterogeneity of treatment effects will be evident across high and low levels of numeracy ability for 12-week exercise behavior. Specifically, individuals with low numeracy ability in the control group will have the lowest levels of exercise behavior, compared to individuals with low numeracy ability in the treatment group, and compared to individuals with high numeracy ability (regardless of treatment condition).

Hypothesis H3.B.2. Heterogeneity of treatment effects will be evident across high and low levels of numeracy ability for exercise behaviors that are consistent with exercise recommendations at 12-week follow-up. Specifically, individuals with low numeracy ability in the control group will have the lowest proportion of individuals with exercise behaviors that are consistent with National Osteoporosis Foundation guidelines for exercise behavior, compared to individuals with low numeracy ability in the treatment group, and compared to individuals with high numeracy ability (regardless of treatment condition).

Hypothesis H3.C.1. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for 12-week exercise behavior. Specifically, individuals with low preference for numerical display in the control group will have the lowest levels of exercise behavior, compared to individuals with low preference for numerical display in the treatment group, and compared to individuals with high preference for numerical display (regardless of treatment condition).

Hypothesis H3.C.2. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for exercise behaviors that are consistent with exercise recommendations at 12-week follow-up. Specifically, individuals with low preference for numerical display in the control group will have the lowest proportion of individuals reporting exercise behaviors that are consistent with National Osteoporosis Foundation guidelines for exercise behavior, compared to individuals with low preference for numerical display in the treatment group, and compared to individuals with high preference for numerical display (regardless of treatment condition).

Research Objective 4: Pending the identification of the hypothesized difference for exercise self-efficacy across low and high levels of health literacy/numeracy (Research Objective 2), research objective four is to test whether level of health literacy/numeracy results in heterogeneity of treatment effects for 12-week self-efficacy outcomes. Each analysis will be conducted for the pooled sample, as well as for each treatment site (i.e., University of Iowa, University of Alabama Birmingham, and Kaiser Permanente, Georgia).

Hypothesis H4.A. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for 12-week exercise self-efficacy. Specifically, individuals in the control group with low health literacy will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high health literacy and compared to individuals in the treatment group (regardless of health literacy level).

Hypothesis H4.B. Heterogeneity of treatment effects will be evident across high and low levels of numeracy ability for 12-week exercise self-efficacy. Specifically, individuals in the control group with low numeracy ability will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high numeracy ability and compared to individuals in the treatment group (regardless of numeracy ability level).

Hypothesis H4.C. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for 12-week exercise self-efficacy. Specifically, individuals in the control group with low preference for numerical display will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high preference for numerical display and compared to individuals in the treatment group (regardless of preference of numerical display).

METHODS AND PROCEDURES

Description of the Data. Data for this study was obtained from the Patient Activation After DXA Notification (PAADRN; ClinicalTrials.gov identifier: NCT01507662) study [85]. PAADRN was designed to be a pragmatic, scalable intervention targeting adults undergoing DXA screening for osteoporosis within three sites located in Iowa, Georgia, and Alabama. The primary intervention modality was a printed mailing [86], designed to better communicate patient-specific:

a) risk status for osteopenia/osteoporosis using DXA,
b) risk status for fracture risk status using FRAX score, and
c) promote actions that reduce risk of fracture and sustain bone health (e.g., medication options, physical activity, and dietary change).

Consistent with the goals of pragmatic clinical trials, exclusion criteria for the original trial were minimal, such that only those with the inability to ready or speak English, prisoners, and individuals with substantial mental, visual, or auditory impairments were excluded from participation. Eligible participants were patients at participating clinical sites, scheduled for bone density scans [85,86].

Protection of human subjects. The primary PAADRN intervention originally received approval by Institutional Review Boards at University of Iowa (organizing site), University of Alabama, Birmingham, and Kaiser Permanente, Southeast. The

Institutional Review Board at Georgia State University approved this study, a secondary analysis of the PAADRN data. Informed consent was attained within the original study protocols, and all data were managed and secured by the PAADRN research staff at University of Iowa, per the human subjects protocols. Consistent with the PAADRN steering committee's processes and procedures, a research proposal containing detailed research questions, specific variables being requested, analysis protocols, and hypotheses were submitted for approval before analysis began.

Measures

Covariates. Baseline interviews at the time of the DXA screening were used to assess the demographic variables gender (male or female), race (white or minority), age (< 65 years, 65-75 years, and 75 years or older) and education level (≤ high school/GED, some college/4-year degree, and some graduate school). Health behaviors that were assessed included drinking any alcohol (yes or no), and being a former smoker (yes or no) or current smoker (yes or no). The following bone health risk factors were also assessed: Frax risk score (low risk, moderate risk, and high risk), past diagnosis of osteopenia (yes or no) or osteoporosis (yes or no), having a prior bone density scan (yes or no), having a prior hip fracture (yes or no), and currently or formerly on medication for bone health (yes or no). Finally, having comorbid chronic health conditions such as chronic obstructive pulmonary disease (yes or no), depression (yes or no), breast (yes or now), and/or prostate cancer (yes or no) were documented.

Health Literacy. The Single Item Literacy Screener [87] was used to assess perceived health literacy. Using a 5-point Likert-type scale (1-Always, 2-Often, 3-Sometimes, 4-Rarely, and 5-Never), each participant answered the question "How often do you need to have someone help you when you read instructions, pamphlets, or other written material from your doctor or pharmacy?" Lower scores are indicative of lower perceived health literacy. This 1-item tool has been validated [87] using the Short Test of Functional Health Literacy in Adults [88,89]. The recommended cut-off for defining low health literacy is ≤ 3. For this study, two categories (high and low health literacy) were created using this method.

Health Numeracy. The Subjective Numeracy Scale [90,91] was used to assess two facets of health numeracy; numeric ability and preference for display of numeric information. This instrument has been validated using objective numeracy tests and has been shown to predict disease risk comprehension [90,91]. Four items assessed numeracy ability on a 6-point Likert-type scale with endpoints "Not at all good" (1) and "Extremely good" (6). Internal consistency (Cronbach's alpha) for the numeric ability subscale in this study was good ($\alpha = 0.89$). Scores for the four items are averaged with lower scores indicative of lower numeric ability [92]. The mean numeric ability score was used to create a dichotomous variable (high and low numeracy ability) for the purposes of regression.

Four items assessed preference for display of numeric information on a 6-point Likert-type scale. After reverse coding the necessary item, the mean preference for numeric display score was computed, with lower scores indicative of lower preference for numerical display [92]. Internal consistency (Cronbach's alpha) for preference for numerical display in this study was weak (α = 0.62). Consistent with methods for numeracy ability, the mean score was used to create a dichotomous variable (high and low preference for numeric display) for the purposes of regression analysis.

Exercise Behavior. Based on the Behavioral Risk Factor Surveillance System [93] weight-bearing aerobic and resistance/strength exercise behaviors were assessed. Specifically for weight-bearing aerobic exercise, participants were asked "In the past 30 days, how many times per week were you engaged in aerobic activity? (This includes walking, hiking, jogging, aerobic classes or video, stair climbing, elliptical machine, dancing, and biking (not on a stationary bike). Please DO NOT include swimming." Possible response options included: none, 1-2 times per week, 3-4 times per week, and 5 or more times per week. Using the same response options strength training exercise was assessed by asking "In the past 30 days, how many times per week were you engaged in strength training? (This includes lifting weights, using elastic or resistance bands, lifting your own weight (push-up or crunches), using weight machines, Pilates and yoga.)"

Responses were then used to create two exercise variables. First, weighted scores were assigned to each response option (See Table 1), allowing the creation of a single continuous exercise variable, which combines weight-bearing and resistance exercise on a scale from 0 (completely sedentary) to 10 (performs aerobic and resistance training 5 or more days per week). Second, using this combined score, a dichotomous variable was created (<5: consistent with National Osteoporosis Foundation recommendations for exercise; ≥5: inconsistent with National Osteoporosis Foundation recommendations for exercise).

Exercise Self-Efficacy. The Osteoporosis Self-Efficacy Scale [94] contains 10 items assessing confidence to exercise. Responses were assessed on a 10-point Likert-type scale, anchored by "not at all confident" (1) to "very confident" (10). Mean scores are calculated, with higher scores indicative of greater self-efficacy for exercise. Internal consistency (Cronbach's alpha) of the scale within this study at baseline ($\alpha = 0.97$) and 12-week ($\alpha = 0.97$) follow-up was excellent.

Statistical Analysis

A subset of the PAADRN cohort, containing only individuals with complete data for treatment, dependent, and moderating variables was created (N = 6591; 85.1% of original PAADRN sample size). The analytic approach was consistent with the PAADRN primary outcomes papers[85,86]. Statistical analysis began by examining whether there were differences between treatment and control groups for baseline demographics and relevant covariates. Normality of continuous variables was assessed using the Kolmogorov-Smirnov statistic. For normally distributed continuous variables, independent sample t-tests were conducted, using pooled or Satterthwaite statistics, as indicated by the equality of variances test. For non-normally distributed continuous variables, Wilcoxon Rank Sum analyses were conducted. For dichotomous categorical variables, chi-square analyses were conducted. For ordinal categorical variables with more than two categories, the Mantel-Haenszel chi-square was computed.

Second, the following three sets of unadjusted pairwise comparisons were conducted to better understand the role of treatment group, sociodemographic variables, and proposed moderators on exercise self-efficacy and behavior:

(1) differences in baseline and 12-week exercise behavior and exercise self-efficacy by each study covariate,

(2) differences in baseline and 12-week exercise behavior and exercise self-efficacy were examined *within* each covariate sub-category, and

(3) differences in exercise self-efficacy and behavior *change* by each study covariate.

For continuous outcome variables, this analysis used independent samples and paired t-tests, as appropriate. For the dichotomous exercise outcome variable (meets/does not meet recommendations), chi-square analyses were conducted, and generalized linear models with Tukey's post hoc tests were used when covariates had more than two categories.

Third, Pearson correlations were conducted among the study's independent, dependent and proposed moderating variables. Finally, generalized linear mixed models were used to conduct unadjusted and adjusted examinations of the moderating influence of health literacy (2 categories: low vs. high) and health numeracy (2 categories: low vs. high numeracy ability; 2 categories: low vs. high preference for numerical display) on post-intervention exercise self-efficacy and exercise behavior.

Conceptually, the moderation analysis was guided by the Baron and Kenny[95] model (See Figure 1). When the effect of the intervention arm on exercise behavior (path a) is statistically significant, and the effect of health literacy (or health numeracy) on exercise behavior (path b) is statistically significant, moderation exists when the interaction of intervention arm and health literacy (or numeracy; path c) are entered into the model and found to be statistically significant. Thus, there is a moderating effect when the interaction of the treatment arm and health literacy (or numeracy) significantly changes the strength or direction of the intervention effect on exercise behavior/self-efficacy (path a').

Analytically, two linear mixed models using PROC GLIMMIX in SAS (version 9.4) were conducted to test for moderation, according to the Baron and Kenny method. The reduced model included the 12-week exercise behavior variable served as the dependent variable, and treatment group (2 categories: treatment vs. control), literacy (or numeracy) variable (2 categories: high vs. low), and the treatment by moderating variable interaction (See Figure 1) were included in the model, while controlling for baseline exercise behavior/self-efficacy. The full model included the independent and dependent variables as described above for the reduced model but also included the following covariates: age (3 groups), gender, race, education, COPD, depression, prostate cancer, breast cancer, current and past smoking behavior, current alcohol use, self-reported health status (5 categories), receiving prior DXA, history of osteoporosis, history of osteoporosis treatment, glucocorticoids use, bone density risk (3 categories: normal, low, osteoporosis), study site, lowest t-score, and 10-year FRAX score. This analytical method was also conducted using 12-week exercise self-efficacy as the dependent variable, controlling for baseline exercise self-efficacy. Furthermore, consistent with the PAADRN primary outcomes papers, full and reduced regression models were used to assess the pooled effect (including all three study sites), as well as the site-specific effect.

Results

Due to missing data for key covariates, participant numbers included in the analysis for reduced and full models will vary. When examining differences for covariates by treatment group (See Table 2), results showed that the control group had significantly greater numbers of individuals reporting breast cancer diagnosis (χ^2 = 45.60, df = 2, p < 0.001), more individuals reporting past diagnosis of osteoporosis (χ^2 = 10.99, df = 2, p = 0.004), and treatment for osteoporosis (χ^2 = 4.43, df = 1, p = 0.04). Based on DXA results, there were fewer individuals in the control group diagnosed as 'normal' bone health (χ^2 = 17.06, df = 1, p < 0.001). Individuals in the control group had significantly lower t-scores for bone mineral density (z = -4.27, p < 0.001), and significantly higher 10-year fracture risk (FRAX score; z = 2.44, p = 0.01).

Research Objective 1. Results of all pairwise comparisons, conducted for the first two research objectives are displayed in Tables 3 and 4. There is sufficient evidence to reject the null hypothesis that high and low health literacy groups have equal levels of exercise at baseline and 12-week time-points (Hypothesis H1.A.1, See Table 3). Therefore, compared to individuals in the high health literacy group, individuals in the low health literacy group have lower levels of exercise at baseline (t = -4.01, df = 6589, p < 0.001) and 12-week follow-up (t = -4.22, df = 6589, p < 0.001). There was no difference between high and low health literacy groups for change in exercise behavior (t = -0.30, p = 0.77; Cohen's d = 0.011).

Similarly, there is sufficient evidence to reject the null hypothesis that high and low health literacy groups have equal proportions of individuals exercising consistent with National Osteoporosis recommendation (Hypothesis H1.A.2, See Table 3). Compared to those with high health literacy, there are a lower proportion of individuals with low health literacy exercising at a level consistent with National Osteoporosis Foundation recommendations at baseline (χ^2 = 5.91, df = 1, p = 0.02) and at 12-week follow-up (χ^2 = 5.00, df = 1, p = 0.03).

For numeracy ability, there is sufficient evidence to reject the null hypothesis that compared to those with high numeric ability, individuals with low numeric ability have lower levels of exercise at baseline and 12-week follow-up (Hypothesis H1.B.1, See Table 3). Compared to individuals with high numeric ability, individuals with low numeric ability reported lower levels of exercise at baseline (t = -8.93, df = 2025.3, p < 0.001) and 12-week follow-up (t = -8.79, df = 6589, $p < 0.001$). There was no difference between high and low numeric ability groups for change in exercise behavior (t = -0.29, df = 6589, p = 0.77; Cohen's d = -0.048).

Similarly, there is sufficient statistical evidence to reject the null hypothesis that high and low numeric ability groups have equal proportions of individuals exercising consistent with National Osteoporosis recommendation (Hypothesis H1.B.2, See Table 3). Compared to the high numeric ability group, the low numeric ability group had a lower proportion of individuals exercising at a level consistent with National Osteoporosis Foundation recommendations at baseline (χ^2 = 32.43, df = 1, $p < 0.001$) and at 12-week follow-up (χ^2 = 42.24, df = 1, $p < 0.001$).

Finally, for preference for numerical display, there is sufficient evidence to reject the null hypothesis that compared to those with high preference for numeric display, individuals with low preference for numeric display have lower levels of exercise at baseline and 12-week follow-up (Hypothesis H1.C.1, See Table 3). Results revealed that individuals with low preference for numerical display have lower levels of exercise behaviors at baseline (t = -7.28, df = 1699.6, $p < 0.001$) and 12-week follow-up (t = -8.79, df = 1676.5, $p < 0.001$). There was no difference between high and low preference for numeric display groups for change in exercise behavior (t = -1.43, p = 0.15; Cohen's d = -0.007).

Additionally, there is sufficient statistical evidence to reject the null hypothesis that high and low preference for numeric display groups have equal proportions of individuals exercising consistent with National Osteoporosis recommendation (Hypothesis H1.C.2, See Table 3). Compared to those with high preference for numerical display, a lower proportion of individuals with low preference for numerical display are exercising at a level consistent with National Osteoporosis Foundation recommendations at baseline ($\chi^2 = 29.27$, df = 1, $p < 0.001$) and at 12-week follow-up ($\chi^2 = 43.27$, df = 1, $p < 0.001$).

Research Objective 2. Regarding self-efficacy for exercise, there is sufficient statistical evidence to reject the null hypothesis that compared to those with high health literacy, individuals with low health literacy have lower levels of exercise self-efficacy at baseline and 12-week follow-up (Hypothesis H2.A, See Table 4). Compared to individuals with high health literacy, individuals with low health literacy had lower levels of exercise self-efficacy at baseline ($t = -7.71$, df = 1380.6, $p < 0.001$) and 12-week follow-up ($t = -7.50$, df = 1406.2, $p < 0.001$). There was no difference between high and low health literacy groups for change in exercise behavior ($t = 0.42$, df = 1364.3, $p = 0.67$; Cohen's d = 0.010).

There is sufficient statistical evidence to reject the null hypothesis that compared to those with high numeric ability, individuals with low numeric ability have lower levels of exercise self-efficacy at baseline and 12-week follow-up (Hypothesis H2.B, See Table 4). Those with low numeric ability reported lower levels of exercise self-efficacy at baseline (t = -12.85, df = 1738.4, p < 0.001) and 12-week follow-up (t = -10.08, df = 1815.1, p < 0.001). A statistically significant difference between high and low numeric ability groups for change in exercise self-efficacy was revealed (t = 3.45, df = 1680.1, p < 0.001; Cohen's d = 0.125). Individuals with high numeric ability decreased in self-reported exercise self-efficacy (mean change = -0.10, *sd* = 1.74; p < 0.001) while individuals with low numeric ability increased in self-reported exercise self-efficacy (mean change = 0.13, *sd* = 2.24; p = 0.04).

There is sufficient statistical evidence to reject the null hypothesis that compared to those with high preference for numeric display, individuals with low preference for numeric display have lower levels of exercise self-efficacy at baseline and 12-week follow-up (Hypothesis H2.C, See Table 4). Those with low preference for numerical display had lower levels of exercise self-efficacy at baseline (t = -11.71, df = 1456.2, p < 0.001) and 12-week follow-up (t = -8.84, df = 1510.5, 1406.2, p < 0.001). There was, however, a statistically significant difference between high and low preference for numerical display groups for change in self-reported self-efficacy (t = 3.63, df = 1443.8, p < 0.001; Cohen's d = 0.125). Individuals with high preference for numeric display decreased in self-reported self-efficacy (mean change = -0.10, *sd* = 1.78; p < 0.001) while individuals with low preference for numeric display increased in self-efficacy (mean change = 0.15, *sd* = 2.19; p = 0.02).

Research Objective 3. Reduced linear mixed models (without covariates) indicated that health literacy was significantly, positively associated with 12-week exercise behavior for the entire sample (See Table 7; $n = 6591$, $t = 2.23$, $p = 0.03$), as well as University of Iowa ($n = 1532$, $t = 2.07$, $p = 0.04$), and the University of Alabama, Birmingham ($n = 2714$, $t = 2.56$, $p = 0.01$) samples. This effect did not remain statistically significant, however, in the full models containing covariates (pooled: $n = 6542$, $t = 0.39$, $p = 0.70$; University of Iowa: $n = 1516$, $t = 0.76$, $p = 0.45$; University of Alabama, Birmingham: $n = 2698$, $t = 1.55$, $p = 0.12$). Furthermore, there is insufficient evidence to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of health literacy for 12-week exercise behavior (Hypothesis H3.A.1, See Table 7). Full models (including covariates) revealed no heterogeneity of treatment effects of health literacy on 12-week exercise behavior (full model, pooled effect: $n = 6591$, $t = -0.73$, $p = 0.47$; full model, University of Alabama, Birmingham only: $n = 2698$, $t = -0.35$, $p = 0.72$; full model, Kaiser Permanente, Georgia only: $n = 2328$, $t = -0.03$, $p = 0.98$; full model, University of Iowa only: $n = 1516$, $t = -0.98$, $p = 0.33$). Covariates that were significantly associated with 12-week follow-up exercise behavior in pooled analyses ($n = 6542$) included: self-reported health status ($t = -8.13$, $p < 0.001$), depression ($t = -3.64$, $p < 0.001$), being a current smoker ($t = -2.46$, $p = 0.01$), using alcohol ($t = 2.71$, $p = 0.006$), having some graduate school ($t = 4.73$, $p < 0.001$), and being male ($t = 4.16$, $p < 0.001$).

Furthermore, there is insufficient statistical evidence to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of health literacy for exercising at a level consistent with National Osteoporosis Foundation recommendations at 12-weeks (Hypothesis H3.A.2, See Table 7). There was no association between health literacy and exercising at levels consistent with National Osteoporosis Foundation recommendations for exercise behavior in the reduced model (pooled: n = 6591, t = 1.28, p = 0.20; University of Iowa: n = 1532, t = 1.54, p = 0.12; University of Alabama, Birmingham: n = 2714, t = 1.08, p = 0.28; Kaiser Permanente, Georgia: n = 2345, t = -0.65, p = 0.51), or the full model with covariates (pooled: n = 6542, t = -0.57, p = 0.57; University of Iowa: n = 1516, t = 0.37, p = 0.71; University of Alabama, Birmingham: n = 2698, t = 0.17, p = 0.87; Kaiser Permanente, Georgia: n = 2328, t = -1.45, p = 0.14). Covariates that were significantly associated with exercising at levels consistent with National Osteoporosis Foundation recommendations at 12-week follow-up in pooled analyses (n = 6542) included: self-reported health status (t = -9.20, p < 0.001), depression (t = -2.78, p =0.01), being a current smoker (t = -2.21, p = 0.03), using alcohol (t = 3.60, p < 0.001), having some graduate school (t = 4.06, p < 0.001), and being male (t = 3.82, p < 0.001), being on osteoporosis medications in the past (t = 2.34, p = 0.02), and being above 75 years (t = -3.35, p < 0.001).

While reduced mixed models (without covariates) revealed a significant positive relationship between preference for numerical display and 12-week continuous exercise behavior for the total sample (See Table 8a; reduced model, pooled effect: n = 6591, t = 3.46, p < 0.001), as well as site-specific effects for University of Alabama, Birmingham (n = 2714, t = 2.77, p = 0.006) and Kaiser Permanente, Georgia (n = 2345, t = 1.72, p = 0.08). These associations were no longer significant, in the full models with covariates (See Table 8a; full model, pooled effect: n = 6542, t = 1.13, p = 0.26; full model, University of Alabama, Birmingham only: $n = 2698$, $t = 1.25$, $p = 0.21$; full model, Kaiser Permanente, Georgia only: $n = 2328$, $t = 0.58$, $p = 0.56$). Additionally, there is insufficient statistical evidence to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of preference for numerical display for 12-week exercise behavior (Hypothesis H3.C.1, See Table 8a). Full models (including covariates) revealed no heterogeneity of treatment effects of preference for numerical display on 12-week exercise behavior (full model, pooled effect: $n = 6542$, $t = 0.61$, $p = 0.55$; full model, University of Alabama, Birmingham only: $n = 2698$, $t = -0.14$, $p = 0.89$; full model, Kaiser Permanente, Georgia only: $n = 2328$, $t = 0.46$, $p = 0.65$). Covariates that were significantly associated with exercise behavior at 12-week follow-up in pooled analyses ($n = 6542$) included: self-reported health status ($t = -8.01$, $p < 0.001$), depression ($t = -3.59$, $p < 0.001$), being a current smoker ($t = -2.49$, $p = 0.01$), using alcohol ($t = 2.58$, $p = 0.01$), having some graduate school ($t = 4.53$, $p < 0.001$), and being male ($t = 4.03$, $p < 0.001$), and being above 75 years ($t = -3.15$, $p = 0.002$).

Similar to the results for 12-week exercise behavior, the reduced model (without covariates) revealed a significant, positive association between preference for numerical display and exercising at levels consistent with the National Osteoporosis Foundation recommendations for exercise at 12-week follow-up (See Table 8a; pooled: $n = 6591$, $t = 2.94$, $p = 0.003$, University of Iowa: $n = 1532$, $t = 1.38$, $p = 0.17$; University of Alabama, Birmingham: $n = 2714$, $t = 1.40$, $p = 0.16$, Kaiser Permanente, Georgia: $n = 2345$, $t = 2.12$, $p = 0.03$). Again, these results were no longer statistically significant after including covariates in the full model (full model, pooled effect: *n* = 6542, *t* = 0.66, *p* = 0.51; full model, University of Iowa only: *n* = 1516, *t* = 0.21, *p* = 0.83; full model, University of Alabama, Birmingham only: *n* = 2698, *t* = 0.12, *p* = 0.91; full model, Kaiser Permanente, Georgia only: *n* = 2328, *t* = 0.82, *p* = 0.41). Additionally, there is insufficient statistical evidence to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of preference for numerical display for exercising at levels consistent with the National Osteoporosis Foundation recommendations for exercise at 12-week follow-up (Hypothesis H3.C.2, See Table 8a). Full regression models (with covariates) showed insufficient statistical evidence for heterogeneity of treatment effects due to preference for numerical display for meeting exercise recommendations at 12-weeks (full model, pooled effect: *n* = 6542, *t* = 0.65, *p* = 0.52; full model, University of Iowa only: *n* = 1516, t = 0.69, *p* = 0.49; full model, University of Alabama, Birmingham only: *n* = 2698, *t* = 0.74, *p* = 0.46; full model, Kaiser Permanente, Georgia only: *n* = 2328, *t* = -0.30, *p* = 0.76). Covariates that were significantly associated with exercising consistent with National Osteoporosis Foundation recommendations at 12-week follow-up in pooled analyses (*n* = 6542) included: self-reported health status (*t* = -9.03, $p < 0.001$), depression (*t* = -2.67, *p* = 0.008), being a current smoker (*t* = -2.25, *p* = 0.02), using alcohol (*t* = 3.46, $p < 0.001$), having some graduate school (*t* = 3.84, $p < 0.001$), and being male (*t* = 3.80, $p < 0.001$), being non-white (*t* = 2.10, *p* = 0.04), being on osteoporosis medications in the past (*t* = 2.33, *p* = 0.02), and being above 75 years (*t* = -3.18, *p* = 0.002).

When examining the association of numeric ability and 12-week exercise behavior (See Table 8b), reduced models (without covariates) revealed a significant positive association for the pooled sample (t = 3.58, p < 0.001), as well as the University of Iowa (t = 2.50, p = 0.01), and Kaiser Permanente, Georgia (t = 2.50, p = 0.01) sites. Similar to the results for health literacy and preference for numerical display, these associations were no longer statistically significant after including covariates in the full model (full model, pooled effect: n = 6542, t = 0.93, p = 0.35; full model, University of Iowa only: n = 1516, t = 0.75, p = 0.46; full model, Kaiser Permanente, Georgia only: n = 2328, t = 1.20, p = 0.23). Furthermore, after including covariates in the full models, there was insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of numeric ability for exercise behavior at 12-week follow-up (Hypothesis H3.B.1, See Table 8b). Full regression models (with covariates) showed insufficient statistical evidence for heterogeneity of treatment effects due to numeric ability for 12-week continuous exercise behavior (See Table 8b; full model, pooled effect: n = 6542, t = -0.15, p = 0.88; full model, University of Iowa only: n = 1516, t = -0.38, p = 0.70; full model, University of Alabama, Birmingham only: n = 2698, t = 0.13, p = 0.89; full model, Kaiser Permanente, Georgia only: n = 2328, t = 0.04, p = 0.97). Covariates that were significantly associated with exercising consistent with exercise behavior in pooled analyses (n = 6542) included: self-reported health status (t = -8.03, p < 0.001), depression (t = -3.56, p < 0.001), being a current smoker (t = -2.46, p = 0.01), using alcohol (t = 2.65, p = 0.008), having some graduate school (t = 4.64, p < 0.001), and being male (t = 4.08, p < 0.001), and being above 75 years (t = -3.24, p < 0.001).

Overall, analyses examining the association between numeric ability and meeting National Osteoporosis Foundation recommendations for exercise were similar to the results for continuous exercise, above. Reduced models (without covariates) showed a significant positive association for the pooled sample (n = 6591, t = 3.24, p = 0.001), and University of Iowa (n = 1532, t = 2.07, p = 0.04). After including covariates in the full models, however, these results were no longer statistically significant (full model, pooled effect: n = 6542, t = 0.44, p = 0.66; full model, University of Iowa only: n = 1516, t = 0.46, p = 0.65). Furthermore, after adjusting for covariates, there is insufficient statistical evidence to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of numeric ability for exercising at levels consistent with the National Osteoporosis Foundation recommendations for exercise at 12-week follow-up (Hypothesis H3.B.2; See Table 8b). Full regression models (with covariates) showed insufficient statistical evidence for heterogeneity of treatment effects due to numeric ability for meeting exercise recommendations at 12-weeks (full model, pooled effect: n = 6542, t = 0.11, p = 0.91; full model, University of Iowa only: n = 1516, t = -0.33, p = 0.74; full model, University of Alabama, Birmingham only: n = 2698, t = -0.33, p = 0.74; full model, Kaiser Permanente, Georgia only: n = 2328, t = 0.85, p = 0.40). Covariates that were significantly associated with exercising consistent with National Osteoporosis Foundation recommendations at 12-week follow-up in pooled analyses (n = 6542) included: self-reported health status (t = -9.04, $p < 0.001$), depression (t = -2.66, p = 0.008), being a current smoker (t = -2.23, p = 0.03), using alcohol (t = 3.52, $p < 0.001$), having some graduate school (t = 3.93, $p < 0.001$), being male ($t = 3.84$, $p < 0.001$), being non-white ($t = 2.05$, $p = 0.04$), being on osteoporosis medications in the past ($t = 2.34$, $p = 0.02$), and being above 75 years ($t = -3.24$, $p = 0.001$).

Research Objective 4. Reduced (without covariates) linear mixed models revealed that health literacy was significantly, positively associated with 12-week exercise self-efficacy for the entire sample (See Table 9; $n = 6591$, $t = 2.84$, $p = 0.004$), for the University of Alabama, Birmingham site ($n = 2714$, $t = 2.04$, $p = 0.04$), and the Kaiser Permanente site ($n = 2345$, $t = 2.52$, $p = 0.01$). This effect did not remain statistically significant, however, after including covariates (full model, pooled effect: $n = 6542$, $t = 1.51$, $p = 0.13$; full model, University of Alabama, Birmingham only: $n = 2698$, $t = 1.58$, $p = 0.11$; full model, Kaiser Permanente, Georgia only: $n = 2328$, $t = 1.38$, $p = 0.17$). Furthermore, there is insufficient statistical evidence to reject the null hypothesis (Hypothesis H4.A) that there is no evidence for heterogeneity of treatment effects across high and low levels of health literacy for 12-week exercise self-efficacy for the pooled effect or the site-specific effect at University of Alabama, Birmingham and University of Iowa (full model, pooled effect: $n = 6542$, $t = -0.91$, $p = 0.37$; full model, University of Alabama, Birmingham only: $n = 2698$, $t = 0.23$, $p = 0.81$; full model, University of Iowa only: $n = 1516$, $t = 0.56$, $p = 0.57$). There is sufficient statistical evidence to reject the null hypothesis for the site-specific effect at the Kaiser Permanente, Georgia site, indicating the presence of heterogeneous treatment effect for health literacy on self-efficacy outcomes (reduced model: $n = 2328$, $t = -2.44$, $p = 0.01$; full model: $n = 2328$, $t = -2.21$, $p = 0.03$). Covariates that were significantly associated with exercise self-efficacy at 12-week follow-up in pooled analyses (n = 6542) included: self-reported health status (t = -9.57, p < 0.001), depression (t = -5.29, p < 0.001), being a current smoker (t = -2.22, p = 0.03), using alcohol (t = 3.16, p = 0.002), and being above 75 years (t = -2.09, p = 0.04).

Reduced (without covariates) linear mixed models showed that numeric ability was significantly, positively associated with 12-week exercise self-efficacy for the entire sample (See Table 10; reduced model, pooled: n = 6591, t = 2.88, p = 0.004), and for the Kaiser Permanente, Georgia site (n = 2345, t = 2.67, p = 0.008). This effect did not remain statistically significant, however, after including covariates (full model, pooled effect: n = 6542, t = 0.85, p = 0.39; full model, Kaiser Permanente, Georgia only: n = 2328, t = 1.74, p = 0.08). There was insufficient evidence to reject the null hypothesis (Hypothesis H4.B) that there is no heterogeneity of treatment effects across high and low levels of numeric ability for 12-week exercise self-efficacy (reduced model, pooled effect: n = 6591, t = -1.75, p = 0.08; full model, pooled effect: n = 6542, t = -1.42, p = 0.16; reduced model, University of Iowa: n = 1532, t = -0.05, p = 0.96; full model, University of Iowa: n = 1516, t = 0.42, p = 0.67; reduced model, University of Alabama Birmingham: n = 2714, t = -0.74, p = 0.45; full model, University of Alabama, Birmingham: n = 2698, t = -0.78, p = 0.43). For the Kaiser Permanente, Georgia site, however, there was sufficient statistical evidence to reject the null hypothesis (Hypothesis H4.B) and conclude there is heterogeneity of treatment effects of numeric ability on self-efficacy (reduced model: n = 2345, t = -2.22, p = 0.03; full model: n = 2328, t = -1.98, p = 0.05). Other significant covariates in the model included: self-reported health status ($t = -5.52$, $p < 0.001$), depression ($t = -3.41$, $p < 0.001$), and age above 75 years ($t = 2.61$, $p = 0.009$).

Neither reduced nor full models revealed a significant relationship between preference for numerical display and 12-week exercise self-efficacy for the pooled sample (See Table 10; reduced model: $n = 6591$, $t = 0.29$, $p = 0.77$; full model: $n = 6542$, $t = -0.86$, $p = 0.39$), the University of Iowa site (reduced model: $n = 1532$, $t = 1.28$, $p = 0.20$; full model: $n = 1516$, $t = 0.29$, $p = 0.77$), or the Kaiser Permanente, Georgia site (reduced model: $n = 2345$, $t = 0.62$, $p = 0.54$; full model: $n = 2328$, $t = -0.06$, $p = 0.95$). There was also insufficient evidence to reject the null hypothesis (Hypothesis H4.C) that preference for numerical display does not result in heterogeneity of treatment effects for 12-week exercise self-efficacy for the pooled sample (reduced model: $n = 6591$, $t = 1.29$, $p = 0.20$; full model: $n = 6542$, $t = 1.32$, $p = 0.19$), the University of Iowa site (reduced model: $n = 1532$, $t = -0.47$, $p = 0.64$; full model: $n = 1516$, $t = -0.39$, $p = 0.69$), or the Kaiser Permanente, Georgia site (reduced model: $n = 2345$, $t = -0.16$, $p = 0.87$; full model: $n = 2328$, $t = -0.27$, $p = 0.79$). For the University of Alabama, Birmingham site, however, there is sufficient statistical evidence to reject the null hypothesis and conclude heterogeneity of treatment effects across high and low levels of preference for numerical display on 12-week exercise self-efficacy (reduced model: $n = 2714$, $t = 2.37$, $p = 0.02$; full model: $n = 2698$, $t = 2.33$, $p = 0.02$). Other significant covariates in the model included: self-reported health status ($t = -5.86$, $p < 0.001$), depression ($t = -2.95$, $p = 0.003$), having COPD ($t = -2.73$, $p = 0.006$), being male ($t = 2.48$, $p = 0.01$), and alcohol consumption ($t = 2.12$, $p = 0.03$).

Discussion

Although health literacy and health numeracy have consistently been related to health outcomes such as emergency room admissions and hospitalization, use of evidence-based preventive services (e.g., mammography screenings and flu vaccines), medication adherence, and overall health status, relatively little research has examined their role in (a) influencing exercise self-efficacy/exercise behavior, or (b) whether health literacy/health numeracy result in heterogeneity of treatment effects. Therefore, through the examination of a number of intermediary and culminating hypotheses (See Table 11 for summary) this study adds to the literature in the several ways.

First, after adjusting for covariates in the full models, the relationship between health literacy and health numeracy and 12-week exercise self-efficacy and 12-week exercise behavior is not statistically significant. This finding supports previous literature where, after controlling for covariates, the association between health literacy and exercise behavior is not statistically significant[66-69,71]. Second, in pooled analyses, there was no indication of heterogeneity of treatment effects for health literacy, numeric ability, or preference for numeric display on exercise behavior. This corresponds with findings from other randomized controlled trials that have examined heterogeneity of treatment effects across levels of health literacy on exercise behavior[79-82]. Taking these first two points together, this study adds to the notion that a direct association between health literacy/numeracy and exercise behavior, or moderating influence of health literacy/numeracy on exercise behavior is questionable.

Third, no known previous research has examined heterogeneity of treatment effects across levels of health literacy/numeracy on self-efficacy for exercise behavior. While not evident in the pooled sample, there was statistically significant heterogeneity of treatment effects across high and low levels of health literacy for self-efficacy within the Kaiser Permanente, Georgia sample. Additionally, there was a statistically significant heterogeneity of treatment effects across high and low levels of preference for numeric display for self-efficacy. Thus, the treatment's effect across the three study sites was different for level of health literacy/preference for numeric display. Future analysis will incorporate data visualization methods to more completely and precisely interpret the nature of these heterogeneous effects.

One plausible explanation for this finding is that compared to individuals with high numeracy/literacy, those with low numeracy/literacy lack the behavioral skills, and/or environmental supports needed to feel confident they can initiate and maintain their exercise behavior. Another explanation is related to readiness to change exercise, such that individuals with higher health literacy/numeracy may already be in a higher stage of change for exercise, and therefore self-efficacy increases after the letter intervention. In contrast, individuals with low literacy/numeracy who may be in lower stages of change, where after receiving their intervention letter, self-efficacy can remain relatively stable or decrease, across early stages of readiness to change[96-98]. Lastly, it is important to consider whether the nature of the relationship between health literacy/numeracy and self-efficacy is best characterized as a moderating relationship, as proposed in this study, or if there is a mediated relationship, as proposed in previous literature[62,63,68]. It is possible that the relationship between health literacy/numeracy and exercise behavior is mediated by self-efficacy.

There may be several reasons for the lack of a statistically significant heterogeneity of treatment effects for exercise behavior and exercise self-efficacy in the pooled sample. It is important to note the limitations of several measurement tools used in this study. For example, there was low inter-item reliability of the preference for numeric display subscale of the subjective numeracy scale. Future analysis of this data may consider analyzing single items of this scale, conducting secondary analyses to better understand why there is low inter-item reliability, and/or to limit use of this data due to poor psychometric qualities. Additionally, continued research to explore the theoretical definition of numeracy, and the operational measurement of these constructs is needed.

In addition to the measurement concerns raised for preference of numeric display, future pragmatic trials should consider incorporating both self-report and objective measures of exercise [99-101] (e.g., pedometers, accelerometers), incorporating physical activity measures that are sensitive to change for sedentary behavior, exercise behavior, and to the degree it is relevant to the research context, measurement tools that are sensitive to changes in contextual physical activity (e.g., home-based, occupational, transportation).

Another potential reason for the null finding for the pooled analyses in this study is the rather small effect of the intervention at 12-weeks - a relatively short time for behavior change to occur. Built upon principles of health communication and pragmatic randomized controlled trials, the intervention was low intensity and ultimately, insufficient to create larger changes in exercise behavior, compared to the control group. Exercise can be a difficult behavior to change, even when subjected to substantially more intense interventions [102]. Future studies using low intensity interventions, such as the PAADRN intervention, should focus on more proximal outcomes in the behavior change process (e.g., knowledge, readiness to change) and/or combine newly designed health risk communication materials with referrals to internet-based or community-based resources that support behavior change.

Finally, research is needed to examine whether, (a) health numeracy can explain unique variance in health outcomes in addition to health literacy, and (b) whether the interaction of health literacy and health numeracy is influential in explaining, or moderating intervention effects on exercise self-efficacy/exercise behavior. The largest gains in exercise behavior change are likely to be made by individuals who have both low health literacy and low health numeracy. To date, however, no research has explicitly examined the interactive effect of these constructs on exercise behavior change.

Despite the overall null finding of this study, it is a valuable contribution to the literature for several reasons. First, compared to other health outcomes, few studies have examined the association of health literacy with exercise behavior. Furthermore, no known studies have examined (a) the association between health numeracy and exercise self-efficacy or exercise behavior, or (b) the moderating effect of health literacy or health numeracy on exercise self-efficacy (See Appendix C). Understanding these variable relationships are important, as self-efficacy is one of the most powerful psychosocial constructs known to influence exercise initiation and maintenance [103]. Second, the available research is biased by relatively small samples, reliance on cross-sectional study designs, and failure to adjust analyses for covariates. Thus, this study is a benefit to the literature due to is methodological improvements and lower risk of biased results. Finally, no known research has examined the influence of health literacy/numeracy on exercise self-efficacy and exercise behavior within the context of bone health, specifically.

Future studies will benefit by employing both self-report and objective measures of exercise [104]. Furthermore, because (a) there is low prevalence of sufficient resistance training behavior among older adults [105,106], (b) increases in bone health require a combination of weight-bearing aerobic and resistance/strength training [28] and (c) the general lack of intervention research explicitly promoting resistance training [49,106], there is great need to develop, test and disseminate efficacious interventions designed to explicitly initiate and maintain sufficient levels of *both* weight-bearing and resistance training exercise for bone health. Additionally, pragmatic interventions with low-intensity interventions could benefit by incorporating the stages of change construct from the Transtheoretical Model, either in measurement, as part of the intervention, or both. Measuring stage of change will allow the detection of readiness to change exercise, even though behavior itself has not yet changed. Considering that the majority of aging adults are sedentary or insufficiently active, interventions that create small changes in cognitive processes of change and/or stages of change are quite relevant to public health.

In conclusion, this secondary analysis of the PAADRN bone health intervention did not support the potential for heterogeneity of treatment effects for health literacy or health numeracy on exercise self-efficacy or exercise behavior. This study, in combination with previous research also reporting no heterogeneity of treatment effects of health literacy on exercise behavior[79-82] suggests that focusing on mediated pathways among these variables may prove more beneficial than continued exploration of heterogeneity of treatment effects. Furthermore, research is needed to better understand whether health literacy and health numeracy explain significant levels of unique variance in exercise self-efficacy and exercise behavior, and whether there is an interactive effect of these concepts on exercise self-efficacy and/or exercise behavior.

References

1. Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville (MD); 2004.
2. United States Department of Health and Human Services (USDHHS). The Surgeon General's Report on Bone Health and Osteoporosis: What It Means To You. In: U.S. Department of Health and Human Services, Offices of the Surgeon General, editor; 2012.
3. Harvey N, Dennison E, Cooper C. Osteoporosis: impact on health and economics. *Nature reviews Rheumatology* 2010; 6(2): 99-105.
4. Leibson CL, Tosteson AN, Gabriel SE, Ransom JE, Melton LJ. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. *Journal of the American Geriatrics Society* 2002; 50(10): 1644-50.
5. Wright NC, Looker AC, Saag KG, et al. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 2014; 29(11): 2520-6.
6. Singer A, Exuzides A, Spangler L, et al. Burden of illness for osteoporotic fractures compared with other serious diseases among postmenopausal women in the United States. *Mayo Clinic proceedings* 2015; 90(1): 53-62.
7. Wade SW, Strader C, Fitzpatrick LA, Anthony MS, O'Malley CD. Estimating prevalence of osteoporosis: examples from industrialized countries. *Archives of osteoporosis* 2014; 9(1): 182.

8. Piscitelli P, Iolascon G, Argentiero A, et al. Incidence and costs of hip fractures vs strokes and acute myocardial infarction in Italy: comparative analysis based on national hospitalization records. *Clinical interventions in aging* 2012; 7: 575-83.

9. Bleibler F, Konnopka A, Benzinger P, Rapp K, Konig HH. The health burden and costs of incident fractures attributable to osteoporosis from 2010 to 2050 in Germany--a demographic simulation model. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2013; 24(3): 835-47.

10. Cosman F, de Beur SJ, LeBoff MS, et al. Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2014; 25(10): 2359-81.

11. Michael YL, Whitlock EP, Lin JS, Fu R, O'Connor EA, Gold R. Primary care-relevant interventions to prevent falling in older adults: a systematic evidence review for the U.S. Preventive Services Task Force. *Annals of internal medicine* 2010; 153(12): 815-25.

12. Gomez-Cabello A, Ara I, Gonzalez-Aguero A, Casajus JA, Vicente-Rodriguez G. Effects of training on bone mass in older adults: a systematic review. *Sports medicine (Auckland, NZ)* 2012; 42(4): 301-25.

13. Kelley GA, Kelley KS, Kohrt WM. Effects of ground and joint reaction force exercise on lumbar spine and femoral neck bone mineral density in

postmenopausal women: a meta-analysis of randomized controlled trials. *BMC musculoskeletal disorders* 2012; 13: 177.

14. Kelley GA, Kelley KS, Kohrt WM. Exercise and bone mineral density in men: a meta-analysis of randomized controlled trials. *Bone* 2013; 53(1): 103-11.
15. Ma D, Wu L, He Z. Effects of walking on the preservation of bone mineral density in perimenopausal and postmenopausal women: a systematic review and meta-analysis. *Menopause (New York, NY)* 2013; 20(11): 1216-26.
16. Martyn-St James M, Carroll S. Meta-analysis of walking for preservation of bone mineral density in postmenopausal women. *Bone* 2008; 43(3): 521-31.
17. Martyn-St James M, Carroll S. A meta-analysis of impact exercise on postmenopausal bone loss: the case for mixed loading exercise programmes. *British journal of sports medicine* 2009; 43(12): 898-908.
18. Martyn-St James M, Carroll S. Effects of different impact exercise modalities on bone mineral density in premenopausal women: a meta-analysis. *Journal of bone and mineral metabolism* 2010; 28(3): 251-67.
19. Kelley GA, Kelley KS. Efficacy of resistance exercise on lumbar spine and femoral neck bone mineral density in premenopausal women: a meta-analysis of individual patient data. *Journal of women's health* 2004; 13(3): 293-300.
20. Kelley GA, Kelley KS. Exercise and bone mineral density at the femoral neck in postmenopausal women: a meta-analysis of controlled clinical trials with individual patient data. *American journal of obstetrics and gynecology* 2006; 194(3): 760-7.

21. Kelley GA, Kelley KS, Tran ZV. Resistance training and bone mineral density in women: a meta-analysis of controlled trials. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists* 2001; 80(1): 65-77.

22. Kelley GA, Kelley KS, Tran ZV. Exercise and lumbar spine bone mineral density in postmenopausal women: a meta-analysis of individual patient data. *The journals of gerontology Series A, Biological sciences and medical sciences* 2002; 57(9): M599-604.

23. Martyn-St James M, Carroll S. High-intensity resistance training and postmenopausal bone loss: a meta-analysis. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2006; 17(8): 1225-40.

24. Palombaro KM. Effects of walking-only interventions on bone mineral density at various skeletal sites: a meta-analysis. *Journal of geriatric physical therapy (2001)* 2005; 28(3): 102-7.

25. Kelley GA, Kelley KS, Tran ZV. Exercise and bone mineral density in men: a meta-analysis. *Journal of applied physiology* 2000; 88(5): 1730-6.

26. Zehnacker CH, Bemis-Dougherty A. Effect of weighted exercises on bone mineral density in post menopausal women. A systematic review. *Journal of geriatric physical therapy* 2007; 30(2): 79-88.

27. Bolam KA, van Uffelen JG, Taaffe DR. The effect of physical exercise on bone density in middle-aged and older men: a systematic review. *Osteoporosis international : a journal established as result of cooperation between the European*

Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2013; 24(11): 2749-62.

28. Howe TE, Shea B, Dawson LJ, et al. Exercise for preventing and treating osteoporosis in postmenopausal women. *The Cochrane database of systematic reviews* 2011; (7): Cd000333.
29. Kohrt WM, Bloomfield SA, Little KD, Nelson ME, Yingling VR, American College of Sports M. American College of Sports Medicine Position Stand: physical activity and bone health. *Medicine and science in sports and exercise* 2004; 36(11): 1985-96.
30. National Osteoporosis Foundation. Exercise for strong bones. 2015.
31. Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 2007; 116(9): 1081-93.
32. American College of Sports M, Chodzko-Zajko WJ, Proctor DN, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Medicine and science in sports and exercise* 2009; 41(7): 1510-30.
33. Nelson ME, Rejeski WJ, Blair SN, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Circulation* 2007; 116(9): 1094-105.
34. United States Department of Health & Human Services (USDHHS). 2008 Physical Activity Guidelines for Americans, Committee Report. 2008. http://www.health.gov/paguidelines/Report/pdf/CommitteeReport.pdf.

35. United States Department of Health & Human Services (USDHHS). 2008 Physical Activity Guidelines for Americans. 2008.

36. Garber CE, Blissmer B, Deschenes MR, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and science in sports and exercise* 2011; 43(7): 1334-59.

37. American College of Sports M. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Medicine and science in sports and exercise* 2009; 41(3): 687-708.

38. Kastner M, Straus SE. Clinical decision support tools for osteoporosis disease management: a systematic review of randomized controlled trials. *Journal of general internal medicine* 2008; 23(12): 2095-105.

39. Lock CA, Lecouturier J, Mason JM, Dickinson HO. Lifestyle interventions to prevent osteoporotic fractures: a systematic review. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2006; 17(1): 20-8.

40. Lai P, Chua SS, Chan SP. A systematic review of interventions by healthcare professionals on community-dwelling postmenopausal women with osteoporosis. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2010; 21(10): 1637-56.

41. Chase JA. Interventions to Increase Physical Activity Among Older Adults: A Meta-Analysis. *The Gerontologist* 2014.

42. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. *Patient education and counseling* 2013; 91(1): 14-24.

43. Hobbs N, Godfrey A, Lara J, et al. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. *BMC medicine* 2013; 11: 75.

44. Muller AM, Khoo S. Non-face-to-face physical activity interventions in older adults: a systematic review. *The international journal of behavioral nutrition and physical activity* 2014; 11(1): 35.

45. Stevens Z, Barlow C, Kendrick D, et al. Effectiveness of general practice-based physical activity promotion for older adults: systematic review. *Primary health care research & development* 2014; 15(2): 190-201.

46. Gine-Garriga M, Roque-Figuls M, Coll-Planas L, Sitja-Rabert M, Salva A. Physical exercise interventions for improving performance-based measures of physical function in community-dwelling, frail older adults: a systematic review and meta-analysis. *Archives of physical medicine and rehabilitation* 2014; 95(4): 753-69 e3.

47. McMahon S, Fleury J. External validity of physical activity interventions for community-dwelling older adults with fall risk: a quantitative systematic literature review. *Journal of advanced nursing* 2012; 68(10): 2140-54.

48. Neidrick TJ, Fick DM, Loeb SJ. Physical activity promotion in primary care targeting the older adult. *Journal of the American Academy of Nurse Practitioners* 2012; 24(7): 405-16.

49. Winett RA, Williams DM, Davy BM. Initiating and maintaining resistance training in older adults: a social cognitive theory-based approach. *British journal of sports medicine* 2009; 43(2): 114-9.

50. Conn VS, Hafdahl AR, Brown SA, Brown LM. Meta-analysis of patient education interventions to increase physical activity among chronically ill adults. *Patient education and counseling* 2008; 70(2): 157-72.

51. Conn VS, Hafdahl AR, Moore SM, Nielsen PJ, Brown LM. Meta-analysis of interventions to increase physical activity among cardiac subjects. *International journal of cardiology* 2009; 133(3): 307-20.

52. Rhodes RE, Pfaeffli LA. Mediators of physical activity behaviour change among adult non-clinical populations: a review update. *The international journal of behavioral nutrition and physical activity* 2010; 7: 37.

53. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Annals of internal medicine* 2011; 155(2): 97-107.

54. Healthy People 2010. Washington, D.C.: U.S. Dept. of Health and Human Services, 2000.

55. Nielsen-Bohlman L, Institute of Medicine (U.S.). Committee on Health Literacy. Health literacy : a prescription to end confusion. Washington, D.C.: National Academies Press; 2004.

56. Baker DW. The meaning and the measure of health literacy. *Journal of general internal medicine* 2006; 21(8): 878-83.

57. Nutbeam D. The evolving concept of health literacy. *Social science & medicine* 2008; 67(12): 2072-8.

58. Golbeck AL, Ahlers-Schmidt CR, Paschal AM, Dismuke SE. A definition and operational framework for health numeracy. *American journal of preventive medicine* 2005; 29(4): 375-6.

59. Kutner MA, United States. Department of Education., National Center for Education Statistics. The health literacy of America's adults : results from the 2003 National Assessment of Adult Literacy. Washington, DC.: United States Department of Education ; National Center for Education Statistics; 2006.

60. Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR. The prevalence of limited health literacy. *Journal of general internal medicine* 2005; 20(2): 175-84.

61. Kobayashi LC, Wardle J, Wolf MS, von Wagner C. Aging and Functional Health Literacy: A Systematic Review and Meta-Analysis. *The journals of gerontology Series B, Psychological sciences and social sciences* 2014.

62. Fransen MP, von Wagner C, Essink-Bot ML. Diabetes self-management in patients with low health literacy: Ordering findings from literature in a health literacy framework. *Patient Education & Counseling* 2012; 88(1): 44-53.

63. Paasche-Orlow MK, Wolf MS. Promoting health literacy research to reduce health disparities. *Journal of health communication* 2010; 15 Suppl 2: 34-41.

64. von Wagner C, Steptoe A, Wolf MS, Wardle J. Health literacy and health actions: a review and a framework from health psychology. *Health education & behavior : the official publication of the Society for Public Health Education* 2009; 36(5): 860-77.

65. Loke YK, Hinz I, Wang X, Rowlands G, Scott D, Salter C. Impact of health literacy in patients with chronic musculoskeletal disease--systematic review. *PloS one* 2012; 7(7): e40210.

66. Bains SS, Egede LE. Associations between health literacy, diabetes knowledge, self-care behaviors, and glycemic control in a low income population with type 2 diabetes. *Diabetes technology & therapeutics* 2011; 13(3): 335-41.

67. Cavanaugh K, Huizinga MM, Wallston KA, et al. Association of numeracy and diabetes control. *Annals of internal medicine* 2008; 148(10): 737-46.

68. von Wagner C, Knight K, Steptoe A, Wardle J. Functional health literacy and health-promoting behaviour in a national sample of British adults. *Journal of epidemiology and community health* 2007; 61(12): 1086-90.

69. Wolf MS, Gazmararian JA, Baker DW. Health literacy and health risk behaviors among older adults. *American journal of preventive medicine* 2007; 32(1): 19-24.

70. Adams RJ, Piantadosi C, Ettridge K, et al. Functional health literacy mediates the relationship between socio-economic status, perceptions and lifestyle behaviors related to cancer risk in an Australian population. *Patient education and counseling* 2013; 91(2): 206-12.

71. McCleary-Jones V. Health literacy and its association with diabetes knowledge, self-efficacy and disease self-management among African Americans with

diabetes mellitus. *The ABNF journal : official journal of the Association of Black Nursing Faculty in Higher Education, Inc* 2011; 22(2): 25-32.

72. Osborn CY, Paasche-Orlow MK, Bailey SC, Wolf MS. The mechanisms linking health literacy to behavior and health status. *American journal of health behavior* 2011; 35(1): 118-28.

73. Brega AG, Ang A, Vega W, et al. Mechanisms underlying the relationship between health literacy and glycemic control in American Indians and Alaska Natives. *Patient education and counseling* 2012; 88(1): 61-8.

74. Lai AY, Ishikawa H, Kiuchi T, Mooppil N, Griva K. Communicative and critical health literacy, and self-management behaviors in end-stage renal disease patients with diabetes on hemodialysis. *Patient education and counseling* 2013; 91(2): 221-7.

75. Husson O, Mols F, Fransen MP, van de Poll-Franse LV, Ezendam NP. Low subjective health literacy is associated with adverse health behaviors and worse health-related quality of life among colorectal cancer survivors: results from the profiles registry. *Psycho-oncology* 2014.

76. van der Heide I, Uiters E, Rademakers J, Struijs JN, Schuit AJ, Baan CA. Associations among health literacy, diabetes knowledge, and self-management behavior in adults with diabetes: results of a dutch cross-sectional study. *Journal of health communication* 2014; 19 Suppl 2: 115-31.

77. Geboers B, de Winter AF, Luten KA, Jansen CJ, Reijneveld SA. The association of health literacy with physical activity and nutritional behavior in older adults, and its

social cognitive mediators. *Journal of health communication* 2014; 19 Suppl 2: 61-76.

78. Kim S, Love F, Quistberg DA, Shea JA. Association of health literacy with self-management behavior in patients with diabetes. *Diabetes care* 2004; 27(12): 2980-2.

79. Cavanaugh K, Wallston KA, Gebretsadik T, et al. Addressing literacy and numeracy to improve diabetes care: two randomized controlled trials. *Diabetes care* 2009; 32(12): 2149-55.

80. Eckman MH, Wise R, Leonard AC, et al. Impact of health literacy on outcomes and effectiveness of an educational intervention in patients with chronic diseases. *Patient education and counseling* 2012; 87(2): 143-51.

81. Dominick GM, Dunsiger SI, Pekmezi DW, Marcus BH. Health literacy predicts change in physical activity self-efficacy among sedentary Latinas. *Journal of immigrant and minority health / Center for Minority Public Health* 2013; 15(3): 533-9.

82. Glasgow RE, Kurz D, King D, et al. Twelve-month outcomes of an Internet-based diabetes self-management support program. *Patient education and counseling* 2012; 87(1): 81-92.

83. Bailey SC, McCormack LA, Rush SR, Paasche-Orlow MK. Highlighting progress in the field of health literacy research. *Journal of health communication* 2014; 19 Suppl 2: 5-9.

84. Kim SH. Health literacy and functional health status in Korean older adults. *Journal of clinical nursing* 2009; 18(16): 2337-43.

85. Edmonds SW, Wolinsky FD, Christensen AJ, et al. The PAADRN study: a design for a randomized controlled practical clinical trial to improve bone health. *Contemporary clinical trials* 2013; 34(1): 90-100.

86. Edmonds SW, Solimeo SL, Lu X, Roblin DW, Saag KG, Cram P. Developing a bone mineral density test result letter to send to patients: a mixed-methods study. *Patient Prefer Adherence* 2014; 8: 827-41.

87. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. *BMC family practice* 2006; 7: 21.

88. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient education and counseling* 1999; 38(1): 33-42.

89. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *Journal of general internal medicine* 1995; 10(10): 537-41.

90. Witteman HO, Fagerlin A, Zikmund-Fisher BJ, et al. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Journal of medical Internet research* 2007; 27(5): 672-80.

91. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical decision making : an international journal of the Society for Medical Decision Making* 2007; 27(5): 663-71.

92. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Medical decision making : an international journal of the Society for Medical Decision Making* 2007; 27(5): 672-80.

93. Behavioral Risk Factor Surveillance System Questionnaire. In: (USDHHS) United States Department of Health and Human Services, editor; 2011.

94. Horan ML, Kim KK, Gendler P, Froman RD, Patel MD. Development and evaluation of the Osteoporosis Self-Efficacy Scale. *Research in nursing & health* 1998; 21(5): 395-403.

95. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 1986; 51(6): 1173-82.

96. Keefe FJ, Lefebvre JC, Kerns RD, et al. Understanding the adoption of arthritis self-management: stages of change profiles among arthritis patients. *Pain* 2000; 87(3): 303-13.

97. Sarkin JA, Johnson SS, Prochaska JO, Prochaska JM. Applying the transtheoretical model to regular moderate exercise in an overweight population: validation of a stages of change measure. *Preventive medicine* 2001; 33(5): 462-9.

98. Herrick AB, Stone WJ, Mettler MM. Stages of change, decisional balance, and self-efficacy across four health behaviors in a worksite environment. *American journal of health promotion : AJHP* 1997; 12(1): 49-56.

99. Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults:

a systematic review and data analysis. *European review of aging and physical activity : official journal of the European Group for Research into Elderly and Physical Activity* 2014; 11: 35-49.

100. Lee IM, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. *British journal of sports medicine* 2014; 48(3): 197-201.

101. Trost SG, O'Neil M. Clinical use of objective measures of physical activity. *British journal of sports medicine* 2014; 48(3): 178-81.

102. King AC, Castro CM, Buman MP, Hekler EB, Urizar GG, Jr., Ahn DK. Behavioral impacts of sequentially versus simultaneously delivered dietary plus physical activity interventions: the CALM trial. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine* 2013; 46(2): 157-68.

103. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW. Correlates of physical activity: why are some people physically active and others not? *Lancet* 2012; 380(9838): 258-71.

104. Strath SJ, Kaminsky LA, Ainsworth BE, et al. Guide to the assessment of physical activity: Clinical and research applications: a scientific statement from the American Heart Association. *Circulation* 2013; 128(20): 2259-79.

105. Adult participation in aerobic and muscle-strengthening physical activities--United States, 2011. *MMWR Morbidity and mortality weekly report* 2013; 62(17): 326-30.

106. Vezina JW, Der Ananian CA, Greenberg E, Kurka J. Sociodemographic correlates of meeting US Department of Health and Human Services muscle strengthening

recommendations in middle-aged and older adults. *Preventing chronic disease* 2014; 11: E162.

Table 1. Scoring of the exercise measure.

Item Response	Weight-bearing aerobic exercise	Resistance/Strength training exercise
none	0	0
1-2 days per week	1.5	1.5
3-4 days per week	3.5	3.5
5 or more days per week	5	5

Table 2: Baseline characteristics among PAADRN participants included in this study (N = 6591).

	missing n	Intervention (N = 3337)	Control (N = 3254)	p-value
Sociodemographics				
Age, mean (SD)	0	66.31 (8.23)	66.53 (8.12)	0.26
Women, n (%)	0	2815 (84.36)	2758 (84.76)	0.65
Race/Ethnicity				
White, n (%)	0	2635 (78.96)	2575 (79.13)	0.87
Education				
HS completed or less, n (%)		808 (24.21)	805 (24.21)	
Completed or Some college, n (%)	1	1794 (53.76)	1794 (53.76)	0.67
Graduate school, n (%)		734 (22.00)	734 (22.00)	
Comorbid Conditions				
COPD, n (%)	9	220 (6.59)	217 (6.67)	0.93
Depression, n (%)	5	785 (23.52)	749 (23.02)	0.35
Prostate cancer, n (%)	5575§	96 (2.88)	70 (2.15)	0.31
Breast cancer, n (%)	1	365 (10.94)	540 (16.59)	< 0.001
Health Habits				
Current smoker, n (%)	1	251 (7.52)	235 (7.22)	0.53
Past smoker, n (%)	8	1268 (38.00)	1172 (36.02)	0.09
Current alcohol user, n (%)	1	1538 (46.09)	1547 (47.54)	0.29
Self-reported Health Status				
Excellent, (%)		394 (11.81)	435 (13.37)	
Very Good, n (%)		1119 (33.53)	1093 (33.59)	
Good, n (%)	4	1236 (37.04)	1132 (34.79)	0.22
Fair, n (%)		460 (13.78)	465 (14.29)	
Poor, n (%)		125 (3.75)	128 (3.93)	
Bone Health				
Prior DXA, n (%)	1	2276 (68.20)	2266 (69.64)	0.29
History of OP, n (%)	59	674 (20.20)	767 (23.57)	0.004
History of OP treatment, n (%)	0	1254 (37.58)	1305 (40.10)	0.04
Glucocorticoids Use, n (%)	0	510 (15.28)	503 (15.46)	0.84
Study DXA Results				
Normal, n (%)		972 (29.13)	819 (25.17)	
Low BMD, n (%)	0	1763 (52.83)	1752 (53.84)	< 0.001
Osteoporosis, n (%)		602 (18.04)	602 (18.04)	
Lowest T-Score	0	-1.54 (1.08)	-1.64 (1.07)	< 0.001
10-year Fracture Risk (FRAX), mean % (SD)	0	11.90 (9.07)	12.35 (8.98)	0.01

Note: §Women were coded as 'missing.'

Table 3. Baseline and 12-week follow-up exercise behavior by treatment assignment, covariates, health literacy, and health numeracy (N = 6591).

	Continuous Exercise Measure				Dichotomous Exercise Measure		
	Baseline Mean (sd)	Follow-up Mean (sd)	p-value (Row)	Change Mean (sd)	Baseline Percent	Follow-up Percent	p-value (Row)
Treatment Assignment							
Letter	3.34 (2.90)	3.84 (2.94)	< 0.001	0.50 (2.68)	38.81	46.93	< 0.001
Control	3.38 (2.91)	3.79 (2.97)	< 0.001	0.41 (2.78)	39.98	45.79	< 0.001
p-value (Column)	0.62	0.48	--	0.20	0.33	0.35	--
Sociodemographics							
Age Group							
50 <65 years	3.45 (2.91)	3.92 (2.91)	< 0.001	0.46 (2.73)	40.99	46.84	< 0.001
65 to < 75 years	3.34 (2.89)	3.84 (2.98)	< 0.001	0.50 (2.71)	38.49	47.40	< 0.001
≥75 years	3.21 (2.91)	3.53 (3.00)	< 0.001	0.32 (2.79)	37.94	42.64	< 0.001
p-value (Column)	0.05	< 0.001	--	0.17	0.04	0.06	--
Gender							
Female	3.31 (2.88)	3.75 (2.94)	< 0.001	0.45 (2.71)	38.58	45.38	< 0.001
Male	3.66 (3.01)	4.17 (3.04)	< 0.001	0.52 (2.87)	43.65	51.77	< 0.001
p-value (Column)	< 0.001	< 0.001	--	0.46	0.003	< 0.001	--
Race/Ethnicity							
White	3.45 (2.93)	3.90 (2.99)	< 0.001	0.45 (2.73)	40.94	47.56	< 0.001
Non-White	3.02 (2.79)	3.50 (2.82)	< 0.001	0.48 (2.75)	33.53	41.85	< 0.001
p-value (Column)	< 0.001	< 0.001	--	0.68	< 0.001	< 0.001	--

Table 3 (continued). Baseline and 12-week follow-up exercise behavior by treatment assignment, covariates, health literacy, and health numeracy (N = 6591).

	Continuous Exercise Measure				Dichotomous Exercise Measure		
	Baseline Mean (sd)	Follow-up Mean (sd)	p-value (Row)	Change Mean (sd)	Baseline Percent	Follow-up Percent	p-value (Row)
Health Literacy							
Low: ≤ 3 on 5-point scale	3.03 (2.90)	3.46 (2.92)	< 0.001	0.43 (2.87)	36.02	43.20	< 0.001
High: > 3 on 5-point scale	3.42 (2.90)	3.89 (2.96)	< 0.001	0.46 (2.71)	40.02	46.96	< 0.001
p-value (Column)	< 0.001	< 0.001	--	0.77	0.02	0.03	--
Health Numeracy: Preference for Display of Numerical Information							
Low: Bottom 2 quartiles	2.79 (2.72)	3.14 (2.82)	< 0.001	0.35 (2.71)	32.21	37.46	< 0.001
High: Top 3 quartiles	3.48 (2.93)	3.96 (2.97)	< 0.001	0.48 (2.74)	40.86	48.20	< 0.001
p-value (Column)	< 0.001	< 0.001	--	0.15	< 0.001	< 0.001	--
Health Numeracy: Cognitive Ability							
Low: Bottom 2 quartiles	2.74 (2.78)	3.17 (2.88)	< 0.001	0.44 (2.66)	32.43	38.26	< 0.001
High: Top 3 quartiles	3.51 (2.91)	3.98 (2.96)	< 0.001	0.46 (2.75)	41.07	48.33	< 0.001
p-value (Column)	< 0.001	< 0.001	--	0.77	< 0.001	< 0.001	--

Table 4. Baseline and 12-week follow-up exercise self-efficacy by treatment assignment, covariates, health literacy, and health numeracy (N = 6591).

	Continuous Exercise Self-Efficacy Measure			
	Baseline Mean (sd)	Follow-up Mean (sd)	p-value (Row)	Change Mean (sd)
Treatment Assignment				
Letter	7.78 (2.15)	7.74 (2.15)	0.03	-0.07 (1.87)
Control	7.81 (2.16)	7.74 (2.16)	0.131	-0.05 (1.84)
p-value (Column)	0.59	0.91	--	0.62
Sociodemographics				
Age Group				
50 <65 years	7.83 (2.10)	7.72 (2.13)	0.002	-0.12 (1.89)
65 to < 75 years	7.90 (2.10)	7.86 (2.11)	0.26	-0.04 (1.81)
≥75 years	7.47 (2.35)	7.48 (2.29)	0.77	0.02 (1.88)
p-value (Column)	< 0.001	< 0.001	--	0.10
Gender				
Female	7.77 (2.18)	7.72 (2.18)	0.04	-0.05 (1.88)
Male	7.97 (1.99)	7.87 (2.03)	0.06	-0.10 (1.69)
p-value (Column)	0.004	0.03	--	0.41
Race/Ethnicity				
White	7.79 (2.15)	7.74 (2.15)	0.05	-0.05 (1.80)
Non-White	7.84 (2.14)	7.75 (2.15)	0.08	-0.09 (2.03)
p-value (Column)	0.41	0.88	--	0.46
Health Literacy				
Low: ≤ 3 on 5-point scale	7.30 (2.33)	7.26 (2.27)	0.58	-0.04 (2.06)
High: > 3 on 5-point scale	7.89 (2.10)	7.83 (2.12)	0.009	-0.06 (1.81)
p-value (Column)	< 0.001	< 0.001	--	0.67
Health Numeracy: Preference for Display of Numerical Information				
Low: Bottom 2 quartiles	7.04 (2.47)	7.18 (2.35)	0.02	0.15 (2.19)
High: Top 3 quartiles	7.96 (2.05)	7.85 (2.09)	< 0.001	-0.10 (1.78)
p-value (Column)	< 0.001	< 0.001	--	< 0.001
Health Numeracy: Cognitive Ability				
Low: Bottom 2 quartiles	7.04 (2.45)	7.16 (2.33)	0.04	0.13 (2.24)
High: Top 3 quartiles	7.98 (2.03)	7.88 (2.09)	< 0.001	-0.10 (1.74)
p-value (Column)	< 0.001	< 0.001	--	< 0.001

Table 5a. Correlations of 12-week follow-up exercise behavior (continuous, dependent variable) with baseline exercise behavior (continuous), treatment assignment (independent variable), health literacy and health numeracy (moderator variables; N = 6591).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
12-week Follow-up Exercise: Continuous (1)	3.82 (2.96)							
Baseline Exercise: Continuous (2)	0.565 (< 0.001)	3.36 (2.90)						
Letter Intervention vs. Control (3)	0.009 (0.480)	-0.006 (0.624)	--					
Literacy: Continuous (4)	0.049 (< 0.001)	0.039 (0.002)	-0.005 (0.668)	4.33 (0.85)				
Literacy: Dichotomous Low vs. Other (5)	0.052 (< 0.001)	0.049 (< 0.001)	0.001 (0.916)	0.809 (< 0.001)	--			
Numeracy - Preference: Continuous (6)	0.139 (< 0.001)	0.121 (< 0.001)	0.018 (0.137)	0.214 (< 0.001)	0.210 (< 0.001)	4.49 (1.15)		
Numeracy - Preference: Dichotomous Low vs. Other (7)	0.104 (< 0.001)	0.089 (< 0.001)	0.028 (0.022)	0.172 (< 0.001)	0.188 (< 0.001)	0.748 (< 0.001)	--	
Numeracy - Ability: Continuous (8)	0.130 (< 0.001)	0.127 (< 0.001)	-0.003 (0.795)	0.279 (< 0.001)	0.268 (< 0.001)	0.516 (< 0.001)	0.403 (< 0.001)	4.76 (1.19)
Numeracy - Ability: Dichotomous Low vs. Other (9)	0.108 (< 0.001)	0.106 (<0.001)	-0.014 (0.266)	0.216 (< 0.001)	0.210 (< 0.001)	0.409 (< 0.001)	0.346 (< 0.001)	0.816 (< 0.001)

Note: Mean and standard deviations (in parentheses) are along the diagonal. Pearson correlations and p-values are in cells below the diagonal.

Table 5b. Correlations of 12-week follow-up exercise behavior (dichotomous, dependent variable) with baseline exercise behavior (dichotomous), treatment assignment (independent variable), health literacy and health numeracy (moderator variables; N = 6591).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
12-week Follow-up Exercise: Dichotomous (1)	--							
Baseline Exercise: Dichotomous (2)	0.445 (< 0.001)	--						
Letter Intervention vs. Control (3)	0.011 (0.354)	-0.012 (0.329)	--					
Literacy: Continuous (4)	0.031 (0.013)	0.026 (0.036)	-0.005 (0.668)	4.33 (0.85)				
Literacy: Dichotomous Low vs. Other (5)	0.028 (0.025)	0.030 (0.015)	0.001 (0.926)	0.809 (< 0.001)	--			
Numeracy - Preference: Continuous (6)	0.105 (< 0.001)	0.095 (< 0.001)	0.028 (0.122)	0.214 (< 0.001)	0.210 (< 0.001)	4.49 (1.15)		
Numeracy - Preference: Dichotomous Low vs. Other (7)	0.081 (< 0.001)	0.067 (< 0.001)	0.028 (0.022)	0.172 (< 0.001)	0.188 (< 0.001)	0.748 (< 0.001)	--	
Numeracy - Ability: Continuous (8)	0.097 (< 0.001)	0.090 (< 0.001)	-0.003 (0.795)	0.279 (< 0.001)	0.268 (< 0.001)	0.516 (< 0.001)	0.403 (< 0.001)	4.76 (1.19)
Numeracy - Ability: Dichotomous Low vs. Other (9)	0.080 (< 0.001)	0.070 (< 0.001)	-0.014 (0.266)	0.216 (< 0.001)	0.210 (< 0.001)	0.408 (< 0.001)	0.346 (< 0.001)	0.816 (< 0.001)

Note: Mean and standard deviations (in parentheses) are along the diagonal. Pearson correlations and p-values are in cells below the diagonal.

Table 6. Correlations of 12-week follow-up exercise self-efficacy (continuous, dependent variable) with baseline exercise self-efficacy (continuous), treatment assignment (independent variable), health literacy and health numeracy (moderator variables; N = 6591).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
12-week Follow-up Exercise Self-Efficacy: Continuous (1)	7.74 (2.15)							
Baseline up Exercise Self-Efficacy: Continuous (2)	0.629 (< 0.001)	7.80 (2.15)						
Letter Intervention vs. Control (3)	-0.001 (0.911)	-0.007 (0.586)	--					
Literacy: Continuous (4)	0.107 (< 0.001)	0.120 (< 0.001)	-0.005 (0.668)	4.33 (0.85)				
Literacy: Dichotomous Low vs. Other (5)	0.096 (<0.001)	0.101 (< 0.001)	0.001 (0.916)	0.809 (< 0.001)	--			
Numeracy - Preference: Continuous (6)	0.153 (< 0.001)	0.207 (< 0.001)	0.018 (0.137)	0.214 (< 0.001)	0.210 (< 0.001)	4.48 (1.15)		
Numeracy - Preference: Dichotomous Low vs. Other (7)	0.117 (< 0.001)	0.161 (< 0.001)	0.028 (0.022)	0.172 (< 0.001)	0.188 (< 0.001)	0.748 (< 0.001)	--	
Numeracy - Ability: Continuous (8)	0.168 (<0.001)	0.210 (< 0.001)	-0.003 (0.795)	0.279 (< 0.001)	0.268 (< 0.001)	0.516 (< 0.001)	0.403 (< 0.001)	4.76 (1.19)
Numeracy - Ability: Dichotomous Low vs. Other (9)	0.132 (< 0.001)	0.175 (< 0.001)	-0.014 (0.266)	0.216 (< 0.001)	0.210 (< 0.001)	0.409 (< 0.001)	0.346 (< 0.001)	0.816 (< 0.001)

Note: Mean and standard deviations (in parentheses) are along the diagonal. Pearson correlations and p-values are in cells below the diagonal.

Table 7. Pooled and site-specific regression models (with and without covariates) assessing heterogeneity of treatment effects by literacy level (dichotomous) on 12-week follow-up exercise behavior (N = 6591).

	Overall		U. Iowa		UAB		KPGA	
	β (se)	p-value	β	p-value	β	p-value	β	p-value
Continuous Exercise Behavior Measure								
Unadjusted								
Treatment	0.181 (0.151)	0.23	0.396 (0.333)	0.23	0.178 (0.219)	0.42	0.082 (0.268)	0.76
Literacy	0.261 (0.117)	0.03	0.532 (0.257)	0.04	0.446 (0.175)	0.01	-0.188 (0.200)	0.35
Treatment*Literacy	-0.130 (0.165)	0.43	-0.419 (0.361)	0.25	-0.064 (0.241)	0.79	-0.041 (0.288)	0.89
Adjusted								
Treatment	0.214 (0.185)	0.25	0.428 (0.327)	0.19	0.186 (0.217)	0.39	0.027 (0.263)	0.92
Literacy	0.042 (0.117)	0.72	0.192 (0.254)	0.45	0.271 (0.175)	0.12	-0.357 (0.201)	0.08
Treatment*Literacy	-0.113 (0.163)	0.49	-0.348 (0.354)	0.33	-0.085 (0.239)	0.72	-0.008 (0.284)	0.98
Dichotomous Exercise Behavior Measure								
Unadjusted								
Treatment	0.159 (0.141)	0.26	0.337 (0.304)	0.27	0.055 (0.209)	0.79	0.203 (0.251)	0.42
Literacy	0.139 (0.109)	0.20	0.359 (0.234)	0.12	0.180 (0.166)	0.28	-0.122 (0.187)	0.51
Treatment*Literacy	-0.094 (0.152)	0.54	-0.348 (0.327)	0.29	0.103 (0.229)	0.65	-0.171 (0.268)	0.52
Adjusted								
Treatment	0.178 (0.177)	0.31	0.436 (0.308)	0.16	0.064 (0.214)	0.76	0.143 (0.258)	0.58
Literacy	-0.068 (0.113)	0.55	0.088 (0.239)	0.71	0.029 (0.171)	0.87	-0.295 (0.197)	0.14
Treatment*Literacy	-0.063 (0.156)	0.68	-0.353 (0.332)	0.29	0.086 (0.235)	0.71	-0.117 (0.277)	0.67

Table 8a. Pooled and site-specific regression models (with and without covariates) assessing heterogeneity of treatment effects by level of preference for numerical display (dichotomous) on 12-week follow-up exercise behavior (N = 6591).

	Overall		U. Iowa		UAB		KPGA	
	β	p-value	β	p-value	β	p-value	β	p-value
Continuous Exercise Behavior Measure								
Unadjusted								
Treatment	-0.002 (0.146)	0.99	-0.144 (0.314)	0.65	0.163 (0.238)	0.49	-0.070 (0.228)	0.76
Numeracy - Preference	0.384 (0.111)	< 0.001	0.335 (0.243)	0.17	0.509 (0.184)	0.006	0.296 (0.172	0.08
Treatment*Numeracy - Preference	0.078 (0.160)	0.63	0.217 (0.343)	0.53	-0.051 (0.258)	0.84	0.124 (0.254)	0.62
Adjusted								
Treatment	0.043 (0.181)	0.81	-0.099 (0.312)	0.75	0.143 (0.236)	0.54	-0.076 (0.224)	0.73
Numeracy - Preference	0.131 (0.113)	0.22	-0.001 (0.246)	0.99	0.235 (0.187)	0.21	0.100 (0.172)	0.56
Treatment*Numeracy - Preference	0.088 (0.158)	0.58	0.275 (0.341)	0.42	-0.035 (0.256)	0.89	0.114 (0.250)	0.65
Dichotomous Exercise Behavior Measure								
Unadjusted								
Treatment	0.010 (0.138)	0.94	-0.108 (0.289)	0.71	-0.011 (0.230)	0.96	0.089 (0.220)	0.69
Numeracy - Preference	0.308 (0.105)	0.003	0.302 (0.219)	0.17	0.249 (0.177)	0.16	0.347 (0.164)	0.03
Treatment*Numeracy - Preference	0.075 (0.150)	0.62	0.177 (0.313)	0.57	0.172 (0.249)	0.49	-0.057 (0.241)	0.81
Adjusted								
Treatment	0.017 (0.139)	0.90	-0.052 (0.293)	0.86	-0.027 (0.235)	0.91	0.099 (0.222)	0.66
Numeracy - Preference	0.071 (0.108)	0.51	0.049 (0.231)	0.83	0.022 (0.185)	0.91	0.139 (0.169)	0.41
Treatment*Numeracy - Preference	0.099 (0.152)	0.52	0.222 (0.320)	0.49	0.187 (0.253)	0.46	-0.074 (0.246)	0.76

Table 8b. Pooled and site-specific regression models (with and without covariates) assessing heterogeneity of treatment effects by level of numeracy ability (dichotomous) on 12-week follow-up exercise behavior (N = 6591).

	Overall		U. Iowa		UAB		KPGA	
	β	p-value	β	p-value	β	p-value	β	p-value
Continuous Exercise Behavior Measure								
Unadjusted								
Treatment	0.124 (0.136)	0.36	0.204 (0.275)	0.46	0.107 (0.214)	0.62	0.071 (0.230)	0.76
Numeracy - Ability	0.392 (0.109)	< 0.001	0.559 (0.223)	0.01	0.256 (0.178)	0.14	0.446 (0.179)	0.01
Treatment*Numeracy - Ability	-0.061 (0.152)	0.69	-0.199 (0.311)	0.52	0.031 (0.237)	0.90	-0.035 (0.256)	0.89
Adjusted								
Treatment	0.105 (0.134)	0.43	0.226 (0.275)	0.41	0.087 (0.212)	0.68	0.014 (0.225)	0.95
Numeracy - Ability	0.104 (0.111)	0.35	0.173 (0.232)	0.46	-0.054 (0.180)	0.77	0.216 (0.180)	0.23
Treatment*Numeracy - Ability	-0.022 (0.150)	0.88	-0.118 (0.309)	0.70	0.037 (0.235)	0.875	0.009 (0.251)	0.97
Dichotomous Exercise Behavior Measure								
Unadjusted								
Treatment	0.113 (0.129)	0.38	0.177 (0.254)	0.49	0.206 (0.205)	0.32	-0.052 (0.224)	0.82
Numeracy - Ability	0.333 (0.103)	0.001	0.425 (0.206)	0.04	0.279 (0.169)	0.10	0.324 (0.171)	0.06
Treatment*Numeracy - Ability	-0.037 (0.142)	0.79	-0.164 (0.283)	0.56	-0.074 (0.226)	0.74	0.131 (0.245)	0.59
Adjusted								
Treatment	0.090 (0.130)	0.49	0.213 (0.259)	0.41	0.197 (0.208)	0.34	-0.129 (0.226)	0.57
Numeracy - Ability	0.048 (0.108)	0.66	0.102 (0.222)	0.65	0.011 (0.177)	0.95	0.065 (0.178)	0.72
Treatment*Numeracy - Ability	0.016 (0.144)	0.91	-0.097 (0.290)	0.74	-0.075 (0.230)	0.74	0.212 (0.250)	0.40

Table 9. Pooled and site-specific regression models (with and without covariates) assessing heterogeneity of treatment effects by literacy level (dichotomous) on 12-week follow-up exercise self-efficacy (N = 6591).

	Overall		U. Iowa		UAB		KPGA	
	β	p-value	β	p-value	β	p-value	β	p-value
Continuous Self-Efficacy Measure								
Unadjusted								
Treatment	0.069 (0.104)	0.50	-0.091 (0.210)	0.67	-0.079 (0.164)	0.63	0.413 (0.172)	0.02
Literacy	0.229 (0.080)	0.005	-0.022 (0.165)	0.89	0.268 (0.131)	0.04	0.324 (0.129)	0.01
Treatment*Literacy	-0.068 (0.113)	0.55	0.126 (0.228)	0.58	0.107 (0.181)	0.55	-0.452 (0.186)	0.01
Adjusted								
Treatment	0.100 (0.103)	0.33	-0.063 (0.205)	0.76	-0.031 (0.162)	0.85	0.364 (0.171)	0.03
Literacy	0.122 (0.081)	0.13	-0.208 (0.159)	0.19	0.207 (0.131)	0.11	0.181 (0.131)	0.17
Treatment*Literacy	-0.101 (0.112)	0.37	0.125 (0.221)	0.57	0.042 (0.179)	0.81	-0.407 (0.185)	0.03

Table 10. Pooled and site-specific regression models (with and without covariates) assessing heterogeneity of treatment effects by level of numeracy ability (dichotomous) and level of preference for numerical display (dichotomous) on 12-week follow-up exercise self-efficacy (N = 6591).

	Overall		U. Iowa		UAB		KPGA	
	β	p-value	β	p-value	β	p-value	β	p-value
Continuous Self-Efficacy Measure								
Unadjusted								
Treatment	0.160 (0.093)	0.09	0.022 (0.175)	0.90	0.119 (0.161)	0.46	0.320 (0.148)	0.03
Numeracy - ability	0.217 (0.075)	0.004	0.157 (0.144)	0.28	0.162 (0.132)	0.22	0.308 (0.115)	0.008
Treatment*Numeracy - ability	-0.182 (0.104)	0.08	-0.010 (0.198)	0.96	-0.132 (0.178)	0.46	-0.365 (0.165)	0.03
Adjusted								
Treatment	0.133 (0.092)	0.15	-0.023 (0.171)	0.90	0.117 (0.158)	0.46	0.274 (0.146)	0.06
Numeracy - ability	0.065 (0.078)	0.39	-0.144 (0.145)	0.32	0.050 (0.135)	0.71	0.203 (0.117)	0.08
Treatment*Numeracy - ability	-0.146 (0.103)	0.16	0.082 (0.193)	0.67	-0.138 (0.175)	0.43	-0.323 (0.163)	0.05
Continuous Self-Efficacy Measure								
Unadjusted								
Treatment	-0.107 (0.100)	0.28	0.095 (0.199)	0.63	-0.382 (0.178)	0.03	0.045 (0.147)	0.76
Numeracy - preference	0.022 (0.077)	0.77	0.198 (0.155)	0.20	-0.132 (0.139)	0.34	0.069 (0.111)	0.54
Treatment*Numeracy - preference	0.142 (0.110)	0.20	-0.103 (0.217)	0.64	0.459 (0.193)	0.02	-0.027 (0.164)	0.87
Adjusted								
Treatment	-0.103 (0.100)	0.30	0.112 (0.195)	0.56	-0.373 (0.177)	0.03	0.049 (0.146)	0.73
Numeracy - preference	-0.067 (0.078)	0.39	0.044 (0.154)	0.77	-0.214 (0.141)	0.13	-0.007 (0.112)	0.95
Treatment*Numeracy - preference	0.143 (0.109)	0.19	-0.084 (0.213)	0.69	0.445 (0.191)	0.02	-0.043 (0.162)	0.79

Table 11. Summary of study hypotheses and findings.

Research Objective/Hypothesis	Result
Research Objective 1: To determine whether there is variation in exercise behavior across high and low levels of health literacy/numeracy. For each of these hypotheses, the dependent variable (exercise behavior) will be assessed using both a continuous and categorical variable.	
Hypothesis H1.A.1. Compared to the high health literacy group, the low health literacy group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low health literacy group is reporting significantly lower levels of exercise behavior at baseline and 12-week follow-up timepoints (Table 3).
Hypothesis H1.A.2. Compared to the high health literacy group, the low health literacy group will have a lower proportion of individuals meeting National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low health literacy group has a significantly lower percentage of individuals meeting exercise recommendations at baseline and 12-week follow-up timepoints (Table 3).
Hypothesis H1.B.1. Compared to the high numeracy ability group, the low numeracy ability group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low numeracy ability group is reporting significantly lower levels of exercise behavior at baseline and 12-week follow-up timepoints (Table 3).
Hypothesis H1.B.2. Compared to the high numeracy ability group, the low health numeracy ability group will have a lower proportion of individuals meeting National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low numeracy ability group has a significantly lower percentage of individuals meeting exercise recommendations at baseline and 12-week follow-up timepoints (Table 3).
Hypothesis H1.C.1. Compared to the high preference for numerical display group, the low preference for numerical display group will have a lower levels of exercise behavior at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low preference for numerical display group is reporting significantly lower levels of exercise behavior at baseline and 12-week follow-up timepoints (Table 3).
Hypothesis H1.C.2. Compared to the high preference for numerical display group, the low preference for numerical display group will have a lower proportion of individuals meeting National Osteoporosis Foundation exercise guidelines at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low preference for numerical display group has a significantly lower percentage of individuals meeting exercise recommendations at baseline and 12-week follow-up timepoints (Table 3).

Table 11 (continued). Summary of study hypotheses and findings.

Research Objective/Hypothesis	**Result**
Research Objective 2: To determine whether there is variation in exercise self-efficacy across high and low levels of health literacy/numeracy. For each of these hypotheses, the dependent variable (exercise self-efficacy) will be assessed using a continuous variable.	
Hypothesis H2.A. Compared to the high health literacy group, the low health literacy group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low health literacy group is reporting significantly lower levels of exercise self-efficacy at baseline and 12-week follow-up timepoints (Table 4).
Hypothesis H2.B. Compared to the high numeracy ability group, the low numeracy ability group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low numeracy ability group is reporting significantly lower levels of exercise self-efficacy at baseline and 12-week follow-up timepoints (Table 4).
Hypothesis H2.C. Compared to the high preference for numeric display group, the low preference for numeric display group will have a lower levels of exercise self-efficacy at baseline and at 12-week follow-up time-points.	There is sufficient statistical evidence to reject the null hypothesis and conclude that the low preference for numerical display group is reporting significantly lower levels of exercise self-efficacy at baseline and 12-week follow-up timepoints (Table 4).

Table 11 (continued). Summary of study hypotheses and findings.

Research Objective/Hypothesis	**Result**
Research Objective 3: Pending the identification of the hypothesized difference for exercise behavior across low and high levels of health literacy/numeracy (Research Objective 1), research objective three is to test for heterogeneity of treatment effects due to health literacy and numeracy on 12-week exercise behavior outcomes.	
Hypothesis H3.A.1. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for 12-week exercise behavior. Specifically, individuals with low health literacy in the control group will have the lowest levels of exercise behavior, compared to individuals with low health literacy in the treatment group, and compared to individuals with high health literacy (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of health literacy for 12-week exercise behavior (Table 7).
Hypothesis H3.A.2. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for meeting exercise recommendations at 12-week follow-up. Specifically, individuals with low health literacy in the control group will have the lowest proportion of individuals meeting National Osteoporosis Foundation guidelines for exercise behavior, compared to individuals with low health literacy in the treatment group, and compared to individuals with high health literacy (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of health literacy for percentage of individuals meeting exercise recommendations at 12-weeks (Table 7).
Hypothesis H3.B.1. Heterogeneity of treatment effects will be evident across high and low levels of numeracy ability for 12-week exercise behavior. Specifically, individuals with low numeracy ability in the control group will have the lowest levels of exercise behavior, compared to individuals with low numeracy ability in the treatment group, and compared to individuals with high numeracy ability (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of numeracy ability for 12-week exercise behavior (Table 8b).
Hypothesis H3.B.2. Heterogeneity of treatment effects will be evident across high and low levels of numeracy ability for meeting exercise recommendations at 12-week follow-up. Specifically, individuals with low numeracy ability in the control group will have the lowest proportion of individuals meeting National Osteoporosis Foundation guidelines for exercise behavior, compared to individuals with low numeracy ability in the treatment group, and compared to individuals with high numeracy ability (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of numeracy ability for percentage of individuals meeting exercise recommendations at 12-weeks (Table 8b).

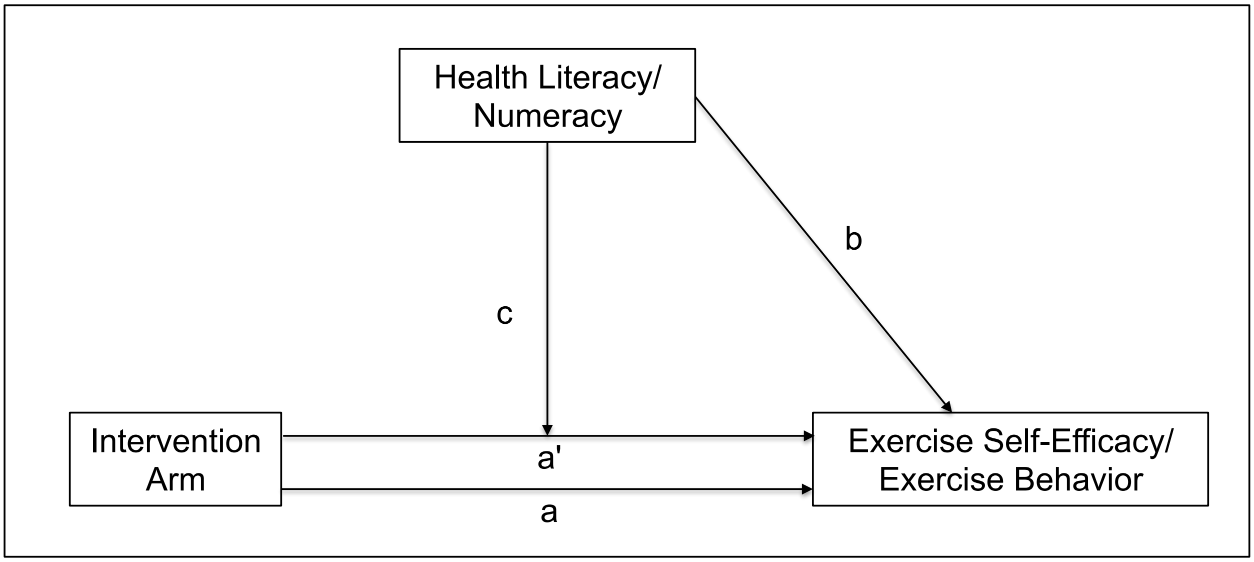
Table 11 (continued). Summary of study hypotheses and findings.

Research Objective/Hypothesis	Result
Hypothesis H3.C.1. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for 12-week exercise behavior. Specifically, individuals with low preference for numerical display in the control group will have the lowest levels of exercise behavior, compared to individuals with low preference for numerical display in the treatment group, and compared to individuals with high preference for numerical display (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of preference for numerical display for 12-week exercise behavior (Table 8a).
Hypothesis H3.C.2. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for meeting exercise recommendations at 12-week follow-up. Specifically, individuals with low preference for numerical display in the control group will have the lowest proportion of individuals meeting National Osteoporosis Foundation guidelines for exercise behavior, compared to individuals with low preference for numerical display in the treatment group, and compared to individuals with high preference for numerical display (regardless of treatment condition).	There is insufficient statistical evidence to reject the null hypothesis that there is no heterogeneity of treatment effects across high and low levels of preference for numerical display for percentage of individuals meeting exercise recommendations at 12-weeks (Table 8a).

Table 11 (continued). Summary of study hypotheses and findings.

Research Objective/Hypothesis	**Result**
Research Objective 4: Pending the identification of the hypothesized difference for exercise self-efficacy across low and high levels of health literacy/numeracy (Research Objective 2), research objective four is to test for heterogeneity of treatment effects due to health literacy and numeracy on 12-week self-efficacy outcomes.	
Hypothesis H4.A. Heterogeneity of treatment effects will be evident across high and low levels of health literacy for 12-week exercise self-efficacy. Specifically, individuals in the control group with low health literacy will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high health literacy and compared to individuals in the treatment group (regardless of health literacy level).	Fail to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of health literacy for 12-week exercise self-efficacy for the pooled effect or the site-specific effect at University of Alabama, Birmingham and Iowa. There is sufficient statistical evidence to reject the null hypothesis for the site-specific effect at the Kaiser Permanente, Georgia site, indicating the presence of heterogeneous treatment effects for health literacy on self-efficacy outcomes (See Table 9).
Hypothesis H4.B. Heterogeneity of treatment effecst will be evident across high and low levels of numeric ability for 12-week exercise self-efficacy. Specifically, individuals in the control group with low numeracy ability will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high numeracy ability and compared to individuals in the treatment group (regardless of numeracy ability level).	Fail to reject the null hypothesis that there is no evidence for heterogeneity of treatment effects across high and low levels of numeric ability for 12-week exercise self-efficacy for the pooled effect, or the site-specific effect at University of Alabama, Birmingham and Iowa. There is sufficient statistical evidence to reject the null hypothesis for the site-specific effect at the Kaiser Permanente, Georgia site, indicating the presence of heterogeneous treatment effects for health literacy on self-efficacy outcomes (See Table 10).
Hypothesis H4.C. Heterogeneity of treatment effects will be evident across high and low levels of preference for numerical display for 12-week exercise self-efficacy. Specifically, individuals in the control group with low preference for numerical display will report the lowest levels of exercise self-efficacy, compared to individuals in the control group with high preference for numerical display and compared to individuals in the treatment group (regardless of preference of numerical display).	Fail to reject the null hypothesis that preference for numerical display does not result in heterogeneity of treatment effects for 12-week exercise self-efficacy for the pooled sample, the University of Iowa site and the Kaiser Permanente, Georgia site. For the University of Alabama, Birmingham site, however, there is sufficient statistical evidence to reject the null hypothesis and conclude heterogeneity of treatment effects across high and low levels of preference for numerical display on 12-week exercise self-efficacy (See Table 10).

Figure 1. Conceptual model for moderation based on Baron and Kenny[95]



APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults.

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Kelley, et al., 2013				
Study Type: RCT's Only	Sex: Men Only	Intervention Length: 32-72 weeks	FN BMD: Hedge's g = 0.583 [95% CI: 0.031, 1.135, p = 0.04]	Small number of studies, of marginal quality. Signficant heterogeneity in study design and exercise modality
Number of Studies: 3	Age Range: 41-79 years	Compliance: 63% to 96% (mean: 72.4% ± 14.5%)	LS BMD: Hedge's g = 0.190 [95% CI: -0.036, 0.416, p = 0.10]	
Number of Participants: 275 (152 exercise, 123 control)		Load Ratings: 10 - 1375 (mean: 556.0 ± 747.6)	Total Hip BMD: Hedge's g = -0.035 [95% CI: -0.27, 0.199, p = 0.77]	
		Mode: Resistance Training and/or weight-bearing aerobic exercise		
Bolam, et al. 2013				
Study Type: RCT and Controlled Trials	Sex: Men Only	Intervention Length: 3 months to 4 years (mean = 13 months)	FN BMD: reported in 5 studies. 3/5 reported significant benefit of exercise over control	Walking ALONE, not supported for increasing BMD. Study quality is marginal. Significant heterogeneity in study design and exercise modality.
Number of Studies: 8 (9 publications)	Age Range: 50-79 years	Mode: walking only (2), resistance training only (3), both walking & resistance training (1), resistance training & impact loading activities (1), resistance training & tai chi (1)	LS BMD: reported in 7 studies. 1/7 showed significant benefit of exercise over control	
Number of Participants: 541 (289 exercise, 252 control)		Drop out: 2.1% to 6.8%		
Ma, et al. 2013				
Study Type: RCT and Controlled Trials	Sex: Women Only (Post-menopausal)	Intervention Length: 12 weeks to 2 years	FN BMD: reported in 7/10 studies; weighted mean difference in BMD was 0.01 g/cm^2 [95% CI: -0.00, 0.01, p = 0.07]	Study quality is marginal.
Number of Studies: 10	Age Range: 40-75 years	Compliance: 57.1% to 97.6%; 9/10 studies reported compliance > 70%	LS BMD: Reported in 10/10 studies; weighted mean difference in BMD was 0.01 g/cm^2 [95% CI: -0.00, 0.02, p = 0.05]	
Number of Participants: 622 (332 exercise, 290 control)		Mode: Walking only	Radius BMD: Reported in 4/10 studies; weighted mean difference in BMD was -0.01 g/cm^2 [95% CI: -0.06, 0.04, p = 0.71]	
		Duration: 40-60 min reported in 8/10 trials	Whole-body BMD: Reported in 2/10 studies; weighted mean difference in BMD was 0.04 g/cm^2 [95% CI: -0.00, 0.08, p = 0.06]	
		Frequency: 3-4 times/week in 9/10 trials		

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Kelley et al., 2012				
Study Type: RCT's Only	Sex: Women Only	Intervention Length: 24 weeks to 104 weeks	FN BMD: Hedge's g = 0.288 [95% CI: 0.102, 0.474, p = 0.002]	Changes in FN and LS BMD would reduce 20-year RR of fracture by 11% and 10%, respectively. Most studies low risk of bias due to blinding and sequence generation; risk of bias due to blinding is unclear
Number of Studies: 25	Age Range: 53 - 80 years	Mode: Resistance Training and/or weight-bearing aerobic exercise	LS BMD: Hedge's g = 0.179 [95% CI: -0.003, 0.361, p = 0.05]	
Number of Participants: 1775 (991 exercise, 826 control)		Frequency: 2-7 days/week		
		Compliance: 39% to 95%		
Gomez-Cabello et al., 2012				
Study Type: Controlled Trials, Meta-Analyses, Reviews	Sex: Men & Women	Aerobic Only: (14 studies) Intervention Length: 16 weeks to 24 months Intensity: Moderate Mode: walking/jogging/stair climbing	Aerobic Only: (14 studies) Most do not report increases in BMD; May maintain/slow the lost of BMD, with high impact activities (e.g., stepping, jogging) showing greater effects	Little data for very old adults (> 70 years); Exercise characteristics extremely heterogenous
Number of Studies: 74 (59 controlled trials, 7 meta-analyses, 8 reviews	Age Range: Mean age 55-70 years	Strength Training Only: (14 studies) Intervention Length: 16 weeks to 24 months	Strength Training Only: (14 studies) Most studies show improvement in BMD of FN, LS, and radius, but not whole-body BMD. Best improvements occur with high-loading intensities (at least 3 sessions/week, 2-3 sets/session)	
		Multi-Component Training: (16 studies) 12 weeks to 30 months	Multi-Component Training: (16 studies) Most studies show improvement or maintenance for BMD.	

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Howe et al., 2011 Study Type: RCT's Only Number of Studies: 37 Number of Participants: 4320	Sex: Women Only (Post-menopausal)	Weight Bearing, Low Force: 9 studies, 705 participants, compliance range 39% to 79.2% Weight Bearing, High Force: 10 studies, 568 participants, compliance range 82.6% to 86.2% Strength Training, Low Force: 6 studies, 231 participants, compliance range 65.0% to 90.0% Strength Training, High Force: 9 studies, 292 participants, compliance range 65.0% to 92.0% Combination Weight Bearing & Strength Training: 10 studies, 823 participants, compliance range 62.0% to 95.0%	Weight Bearing, Low Force: Vertebral BMD: mean difference = 0.87 [95% CI: 0.26, 1.48]; FN BMD:mean difference = -1.20 [95% CI: -4.45, 2.05]; Trochanter BMD: mean difference = 0.39 [95% CI: -0.59, 1.38]; Weight Bearing, High Force: Vertebral BMD: mean difference = -1.20 [95% CI: -4.45, 2.05]; FN BMD: mean difference = 1.06 [95% CI:-0.32, 2.45]; Trochanter BMD: mean difference = 1.23 [95% CI: -0.01, 2.47]; Total Hip BMD: mean difference = 1.55 [95% CI: 1.41, 1.69]; Strength Training, Low Force: No statistically signficant differences were observed for any outcome. Strength Training, High Force: Vertebral BMD: mean difference = 0.86 [95% CI: 0.58, 1.13]; FN BMD: mean difference = 1.03 [95% CI: 0.24, 1.82]; Combination Weight Bearing & Strength Training: Vertebral BMD: mean difference = 3.22 [95% CI: 1.80, 4.64]; Trochanter BMD: mean difference = 1.31 [95% CI: 0.69, 1.92]; FN BMD: mean difference = 0.45 [95% CI: 0.08, 0.82]; Overall Vertebral BMD change: exercise group = 0.85% higher [95% CI: 0.62, 1.07]; control group = range -4.38 to 1.05% Overall FN BMD change: exercise group = 0.08% lower [95% CI: 0.92, 1.08]; control group = range -3.19% to 3.12% Overall Total Hip BMD change: exercise group = 1.03% higher [95% CI: 0.56, 1.49]; control group = range-1.62% to 2.94%	Quality of study reporting is low, with sequence generation, allocation concealment, blinding, and loss to follow-up being problematic

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Martyn-St James & Carroll, Study Type: RCT and Controlled Trials Number of Studies: 9 Number of Participants: 521 (281 exercise, 240 control)	Sex: Women Only (Premenopausal) Age Range: 19 - 49 years	Intervention Length: 6 months to 24 months Frequency: 8/9 studies were 3 days/week Compliance: 43.7% to 91% Mode: Impact exercises (e.g., jumping, skipping, bench steping, hurdle bounding)	FN BMD: 0.012 g/cm^2 [95% CI: 0.005, 0.020, p = 0.001] LS BMD: 0.006g/cm^2 [95% CI: 0.002, 0.010, p = 0.003]	Study quality is low, citing failure to adequately blind participants, heterogeneity of exercise protocols, highly selected samples, failure to ascertain use of oral contraception, and attrition.
Martyn-St James & Carroll, Study Type: RCT and Controlled Trials Number of Studies: 15 Number of Participants: 692 (442 exercise, 250 control)	Sex: Women Only (Post-menopausal) Age Range: 50 - 75 years	Intervention Length: 6 months to 5 years Mode: Impact exercises (e.g., jumping, skipping, bench steping, hurdle bounding) Frequency: 2-5 days/week Compliance: 50% - 91%	FN BMD: 0.008 g/cm^2 [95% CI: 0.004, 0.013, p < 0.001] LS BMD: 0.015 g/cm^2 [95% CI: 0.005, 0.025, p = 0.004] Total Hip BMD: 0.013 g/cm^2 [95% CI: 0.001, 0.024, p = 0.004]	Study quality is low, with most studies scoring 0-3 points out of a possible 5 points. Most studies fail to adequately blind participants.
Martyn-St James & Carroll, Study Type: RCT and Controlled Trials Number of Studies: 8 Number of Participants: 427 (247 exercise, 180 control)	Sex: Women Only (Post-menopausal) Age Range: 45 - 75 years	Intervention Length: 6 to 24 months Mode: Walking only Frequency: Duration: 8/8 trials > 20 minutes Intensity: moderate/brisk pace Compliance: 77% to 85% Drop out: 8.5% to 41%	FN BMD: 0.008 g/cm^2 [95% CI: 0.004, 0.013, p < 0.001] LS BMD: 0.014 g/cm^2 [95% CI: 0.000, 0.028, p = 0.05]	Walking ALONE, not supported for increasing BMD at either FN or LS. Relatively small number of studies, with only 5 RCTs. Several studies are underpowered.

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Zehnacker & Bemis-Dougherty, 2007				
Study Type: RCT and Controlled Trials	Sex: Women Only (Post-menopausal)	Intervention Length: 4 months to 60 months		Characteristics of studies yeilding positive effect on BMD: (1) over 11 months in duration, (2) ≥ 70% of 1 maximal repitition, (3) 3-5 times/week, (4) 2-3 sets of 8-12 repititions, (5) lasted 45-70 minutes each session, (6) included leg exercises, stair-climbing/step boxes/jumping wearing weighed vests, chest, back, and arm exercises
Number of Studies: 20	Age Range: 40 - 75 years	Mode: Resistance Training and/or weighted exercises (e.g., weighted vest worn during calisthenics)	Only 3/20 studies showed no effect of exercise on BMD; Positive effect on BMD was found for FN (5 studies), LS (4 studies)	
Number of Participants: 1626		Frequency: 3-5 sessions/week		
Kelley & Kelley, 2006				
Study Type: RCT and Controlled Trials	Sex: Women Only (Post-menopausal)	Intervention Length: 24 weeks to 104 weeks	FN BMD: No effect of exercise on FN BMD	Pooled individual patient data from 43% of eligible studies; several studies were under-powered to detect an effect
Number of Studies: 10	Age Range: 42 - 92 years	Mode: Resistance Training and/or weight-bearing aerobic exercise		
Number of Participants: 595 (295 exercise, 300 control)		Frequency: 2-3 days/week		

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Martyn-St James & Carroll, 2006				
Study Type: RCT and Controlled Trials Number of Studies: 19 Number of Participants: n/a	Sex: Women Only (Post-menopausal) Age Range: 41 - 87 years	Intervention Length: 4.5 months to 24 months Mode: Resistance Training Only Intensity: 60-70% of one repitition maximum Frequency: 1-5 days/week Compliance: 61% to 99%	FN BMD: 0.010 g/cm^2 [95% CI:-0.002, 0.021, p = 0.11] LS BMD: 0.006 g/cm^2 [95% CI: 0.002, 0.011, p = 0.006] Total Hip BMD: 0.002 g/cm^2 [95% CI: -0.001, 0.005, p = 0.20]	Heterogeneity of exercise protocols; Overall, poor reporting of trial information; Concludes that high intensity training likely benefits the BMD of the Lumbar Spine
Palombaro, 2005				
Study Type: RCT and Controlled Trials Number of Studies: 10 Number of Participants: 768	Sex: Men and Women Age Range: 45 - 76 years	Intervention Length: 6 months to 45 months Mode: Walking only Frequency: 3-4 times/week Duration: 90 - 280 min/week Intensity: 40 to 85% maximum (moderate/vigorous)	FN BMD: Cohen's d = 0.00, p = 1.00 LS BMD: Cohen's d = 0.32, p < 0.03 Calcaneal BMD: Cohen's d = 0.32, p = 0.56	Majority of studies are underpowered to detect an effect;
Kelley & Kelley, 2004				
Study Type: Controlled Trials Number of Studies: 3 Number of Participants: 143 (74 exercise, 69 control)	Sex: Women Only (Premenopausal) Age Range: 18 - 40 years	Intervention Length: 18 weeks to 52 weeks Mode: Resistance Exercise Only Frequency: 3-5 times/week Intensity: 60%-80% one repetition maximum Volume: 2-6 sets, 6-20 repititions/set	FN BMD: F (1, 73) = 0.10, p = 0.75 LS BMD: F (1, 140) = 0.003, p = 0.96	Only 3 studies examining the effect of resistance training only; Most (if not all) were underpowered; heterogeneity of exercise protocol.

APPENDIX A

Reviews Examining the Effect of Exercise on Bone Mineral Density Among Adults (continued).

Characteristics of Included Studies	Participant Characteristics	Exercise Characteristics	Overall Result/Finding	Comments
Kelley et al., 2002				
Study Type: RCT and Controlled Trials	Sex: Women Only (Post-menopausal)	Intervention Length: 24 - 104 weeks	LS BMD: F (1, 697) = 15.23, p < 0.001	
		Mode: Resistance Training and/or weight-bearing aerobic exercise		
Number of Studies: 13	Age Range: ≥ 45 years			
Number of Participants: 699 (355 exercise, 344 control)		Frequency: 2-7 times/week		
		Compliance: mean 75% ± 17%		
Kelley et al., 2001				
Study Type: RCT and Controlled Trials	Sex: Women Only	Intervention Length: 18 weeks to 208 weeks	Femur BMD: 0.07 ± 0.36 (-0.02, 0.15)	
Number of Studies: 29	Age Range: ≥ 18 years	Mode: Resistance Training Only	LS BMD: 0.24 ± 0.36 (0.11, 0.38)	
Number of Participants: 1097 (555 exercise, 524 control)		Frequency: 2-7 times/week	Radius: 0.30 ± 0.33 (0.13, 0.48)	
		Intensity: 30%-85% one repetition maximum		
		Volume: 1-6 sets, 3-15 repititions/set		
Kelley et al., 2000				
Study Type: RCT and Controlled Trials	Sex: Men Only	Intervention Length: 4 months to 5 years	Femur BMD: Cohen's d = 0.482 (95% CI: 0.27, 0.71)	
		Mode: Resistance Training and/or weight-bearing aerobic exercise	LS BMD: Cohen's d = 0.749 (95% CI: 0.099, 1.33)	
Number of Studies: 8	Age Range: ≥ 17 years			
Number of Participants: 239		Compliance: > 90%		

Appendix B

Summary of Relevent Physical Activity Recommendations

Organization/Statement	Purpose/Population	Aerobic/Weight Bearing Exercise	Strength Training	Flexibility/Balance
Organization: National Osteoporosis Foundation (NOF; 2014) **Statement:** Clinician's Guide to Prevention and Treatment of Osteoporosis	**Purpose:** Offer recommendations regarding prevention, risk assessment, diagnosis, ad treatment of osteoporosis **Population:** postmenopausal women and men age 50 and older.	Regular weight-Bearing and muscle-strengthening exercise to improve agility, strength posture, balance, reduce the risk of falls and fractures, and modestly improve bone density. Weight-bearing exercise (in which bones and muscles work against gravity as the feet and legs bear the body's weight) include walking, jogging, Tai Chi, stair climbing, dancing and tennis. Muscle-strengthing exercise includes weight training and other resistive exercises, such as yoga, Pilates, and boot camp programs. Before and individual with osteoporosis initiates a new vigorous exercie program (e.g., running or heavy weight-lifting, a clinician's evaluation is appropriate.		Balance training exercise to reduce risk of falls.
Organization: *National Osteoporosis Foundation (NOF; 2015) **Statement:** Exercise for strong bones. Source: www.nof.org/exercise. Accessed 5/26/2015.	none stated	30 minutes on most days of the week. A single 30-minute session or multiple sessionsthroughout the day will result in the same bone health benefits.	2-3 days per week. Exercising just one body part each day is sufficient, if under time constraints. For example, exercise arms one day, legs the next day, and trunk the next. These exercises can also be spread throughout your normal day.	Every day or as often as needed. You can focus on one area more than others. For example, if you have fallen or lose your balance, do balance exercises. If you are getting rounded shoulders, do more posture exercises. If climbing stairs or getting up from the couch are difficult, do functional exercises. These exercises can be spread out throughout your day.
Organization: American College of Sports Medicine (ACSM; 2004) **Statement:** Physical Activity and Bone Health	**Purpose:** Promote bone health across the lifespan; Includes the role of physical activity in (1) increasing peak bone mass, (2) minimizing age-related bone loss and (3) preventing injurious falls and fractures. **Population:** Healthy, age 18-65 years	Weight bearing endurance activities (e.g., tennis, stair climbing, jogging), activities that involve jumping (e.g., volleyball, basketball), and resistance exercise (weight lifting) that are moderate to high, in terms of bone-loading forces. Weight bearing activities should be performed 3-5 times per week, and strength training activities 2-3 times per week. Activities should be 30-60 minutes in duration for a combination of weight bearing endurance activities, jumping activities, and resistance exercise.		Exercise programs for elderly men and women should include activities designed to improve balance and prevent falls.
Organization: **American College of Sports Medicine (ACSM) & American Heart Association (AHA; 2007) **Statement:** Physical Activity and Public Health: Updated Recommendation for Adults from ACSM and AHA	**Purpose:** Promote and maintain health, and reduce risk of chronic disease and premature mortality **Population:** Healthy, age 18-65 years	Minimum of 30 minutes of moderate-intensity aerobic (endurance) physical activity 5 days per week **OR** minimum of 20 minutes of vigorous-intensity aerobic (endurance) physical activity 3 days a week **OR** a combination of Moderate to Vigorous (endurance) physical activity (e.g., 2 days/week moderate + 2 days/week vigorous)	Activities that maintanin or increase muscular strength and endurance a minimum of 2 days per week, 8-10 exercises involving the major muscle groups, 8-12 repetitions for each exercise.	n/a

Note: *Website intended for physical activity promotion within the patient population;while an official communication from NOF, it is not a scientific position stand/statement. **Increased health benefits occur by exceending this minimum dose. Activity can be accumulated in 10 minute bouts.

Appendix B

Summary of Relevent Physical Activity Recommendations (continued)

Organization/Statement	Purpose/Population	Aerobic/Weight Bearing Exercise	Strength Training	Flexibility/Balance
Organization: **American College of Sports Medicine (ACSM) & American Heart Association (AHA) **Statement:** (1) Physical Activity and Public Health in Older Adults: Recommendation from ACSM and AHA (2009); (2) Exercise and Physical Activity for Older Adults (2009)	**Purpose:** Promote and maintain health & prevent disease & maintain physical independence **Population:** Adults, age 65 years and over and adults aged 50-64 years with clinically significant chronic conditions or functional limitations that affect movement ability, fitness, or physical activity	*Same as recommendation for healthy adults aged 18-65 years;* On a 10-point scale, where sitting is 0 and all-out effort is 10, moderate-intensity activity is a 5 or 6 and produces noticable increase in heart rate and breathing. Vigorous-intensity activity is a 7 or 8 and produces large increases in heart rate and breathing. This recommended level of activity is in addition to regular activities of daily living.	It is recommended that 8-10 exercises be performed on two or more nonconsecutive days per week using the major muscle groups. 10-15 repetitions for each exercise are recommended at a moderate to high level of effort. On a 10-point scale, where no movement is 0 and maximal effort of a muscle group is 10, moderate-intensity effort is a 5 or 6 and high-intenshity effort is a 7 or 8. Muscle strengthening activities can include progressive-weight training program, weight bearing calisthenics, and similar risistance exercise that use the major muscle groups.	To maintain the flexibility necessary for regular physical actvity and daily life, older adults should perform activities that maintain or increase flexibility on at least two days each week for at least 10 min each day. Major muscle and tendon groups should have moderate intensity (5-6 on 10-point scale) static streches for 10-30 seconds each, with 3-4 repitiions for each stretch. Perform flexibility exercises on all days where strength and aerobic exercises are performed. **AND** To reduce risk of injury from falls, community-dwelling older adults with substantial risk of falls (e.g., frequent falls or mobility problems) should perform exercises that maintain or improve balance. Balance exercises are recommended 3 days a week. Specific activities include: progressively difficult postures that reduce the base of support, dynamic movements that perturb the center of gravity, stressing postural muscle groups, and reducing sensory input.
Organization: United States Department of Health and Human Services (2008) **Statement:** Physical Activity & Health: A Report of the Surgeon General	**Purpose:** Promote overall health	People of all ages should engage in a minimum of 30 minutes of moderate-intensity physical activity on most, if not all days of the week. Greater health benefits can be obtained by engaging in physical activity of longer duration, and/or of vigorous intensity.	Cardiorespiratory endurance physical activity should be supplemented with strength-developing exercises at least 2 times per week, in order to improve musculoskeletal health, maintain independence in performing activities of daily life, and reduce the risk of falling.	n/a

Note: *Website intended for physical activity promotion within the patient population;while an official communication from NOF, it is not a scientific position stand/statement. **Increased health benefits occur by exceending this minimum dose. Activity can be accumulated in 10 minute bouts.

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior.

Study Design	Characteristics of the Sample	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA	Mediation Analysis
Cross-Sectional Studies						
Wolf et al. (2007); USA						
Cross-Sectional	N = 2923 US adults (54.2% of eligible participants) Mean Age = 71 years Convenience sample of new Medicare enrollees across four US metropolitan areas	Health Literacy: Short version of the Test of Functional Health Literacy in Adults (S-TOFHLA) Physical Activity: Adapted from National Health Interview Survey, the number of times/week exercised 20 minutes or longer (< 1, 1-2, 3, ≥4).	Adequate HL: n = 1944 (66.5%) Marginal HL: n = 330 (11.29%) Inadequate HL: n = 649 (22.2%)	Multinomial logistic regression, adjusted for age, gender, race/ethnicity, language (English/Spanish), income, occupation, education, research site, and physical functioning; comparison group is "adequate health literacy" compared to "marginal health literacy" and "inadequate health literacy."	Marginal (vs. Adequate): 1-2 times (vs < 1): OR = 1.3 (95% CI: 0.9, 1.8) 3 times (vs < 1): OR = 1.0 (95% CI: 0.7, 1.5) ≥4 times (vs < 1): OR = 1.0 (95% CI: 0.7, 1.4) Inadequate (vs. Adequate): 1-2 times (vs < 1): OR = 1.0 (95% CI: 0.7, 1.4) 3 times (vs < 1): OR = 0.9 (95% CI: 0.7, 1.3) ≥4 times (vs < 1): OR = 1.3 (95% CI: 0.9, 1.7)	n/a
von Wagner et al. (2007); UK						
Cross-Sectional	N = 719 UK adults Mean Age = 47.6 years (sd = 18.3) Two stage random sampling	Health Literacy: Test of Functional Health Literacy in Adults (S-TOFHLA); modified for the UK Physical Activity: specific measurement tool not stated; "whether or not they had undertaken any form of physical exercise in the past 7 days"	Adequate HL: n = 637 (88.6%) Marginal HL: n = 41 (5.7%) Inadequate HL: n = 41 (5.7%)	Multivariate logisitc regression (inadequate/marginal health literacy vs. adequate health literacy), adjusted for age, education, gender, ethnicity and income.	OR = 1.00 (95% CI: 0.98, 1.02, p = 0.88)	n/a
Cavanough et al. (2008); USA						
Cross-Sectional	N = 398 adults with type 1 or type 2 diabetes Age: 18-95 years Convenience sample from primary care and endocrinology clinics	Diabetes-related Numeracy: Diabetes Numeracy Test (DNT) Exercise: Summary of Diabetes Self-Care Activities Scale	Highest Numeracy Scores: Quartile 4: n = 90 Quartile 3: n = 102 Quartile 2: n = 99 Lowest Numeracy Scores: Quartile 1: n = 107	Wilcoxon rank-sum tests, unadjusted	Quartile 1: 3.5 (1-4.5) Quartile 2: 3 (1.5-5) Quartile 3: 2.5 (0.5-5) Quartile 4: 2.75 (1-4.5) (p = 0.25)	n/a

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA	Mediation Analysis
			Cross-Sectional Studies			
Bains & Egede (2011); USA						
Cross-Sectional	N = 125 men and women Age: *range/mean for total sample not specified* Convenience sample of diabetes mellitus patients attending appointments at a university medical hospital	Health Literacy: Rapid Estimate of Adult Literacy in Medicine - Revised (REALM-R) Exercise: Summary of Diabetes Self-Care Activities Questionnaire (SDSCA)	n/a	Linear regression, controlling for age, sex, race/ethnicity, education, income, and health status	β = -0.06 (95% CI: -0.22, 0.11) p > 0.05	n/a
McCleary-Jones (2011); USA						
Cross-Sectional	N = 50 African American Adults with Diabetes Mellitus Mean Age = 58.6 years (sd = 11.5) Female: 76% Convenience sample from a community health clinic	Health Literacy: Rapid Estimate of Adult Literacy in Medicine (REALM) Physical Activity: Summary of Diabetes Self-Care Activities Questionnaire (SDSCA)	n/a	Multiple regression, controlling for diabetes knowledge, and diabetes self-efficacy	β = -0.47 (SE = 0.45); t = 1.05, p = 0.30	n/a
Osborn et al. (2011); USA						
Cross-Sectional	N = 330 adults Mean Age = 53.6 years (sd = 12.0) Convenience sample from a community health clinics in three metropolitain areas	Health Literacy: Short version of the Test of Functional Health Literacy in Adults (S-TOFHLA) Physical Activity: 1-item masure of physical activity over the past 4 weeks (response options: never, only once or twice, at least once a week, 3-4 times a week, and every day)	Adequate HL: n = 230 (69.6%) Inadequate HL: n = 100 (30.3%)	Path analysis with observed variables ("all potential paths between variables were included to test both those hypothesized to be significant and those hypothesized to be non-significant. Model…omitted all nonsignificant paths…")	Health literacy to physical activity pathway assumed to be nonsignificant, based on its omission from the final model.	Relationship between HL and PA mediated by three intervening pathways: (1) HL to Knowledge: path coefficient = 0.22, p < 0.05 (2) Knowledge to Disease Management Self-efficacy: path coefficient = 0.13, p < 0.05 (3) Disease Management Self-efficacy to PA: path coefficient = 0.17, p < 0.05

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA	Mediation Analysis
			Cross-Sectional Studies			
Adams et al. (2013); Australia						
Cross-Sectional	N = 2825 Age: ≥ 15 years Randomly selected sample	Health Literacy/Numeracy: Newest Vital Sign (NVS) Physical Activity: self-reported PA from the previous week; *no specific tool/source cited.*	Adequate HL: 54.9% Marginal HL: 24.1% Inadequate HL: 21%	Multivariate logisitc regression (inadequate health literacy vs. adequate health literacy), adjusted for age, gender, income, education, occupation category, country of birth, and area of residence.	Sufficient PA vs. Sedentary: OR = 2.2 (95% CI: 1.5, 3.2) Some PA vs. Sedentary: OR =1.2 (95% CI: 0.8, 1.7)	Although SEM was employed, PA was grouped with other lifestyle health behaviors, and therefore can not be utilized for this review.
Lai et al. (2014); Japan						
Cross-Sectional	N = 63 Mean Age: 57.7 years (sd = 10.1), Range 32-78 Convenience sample from those undergoing hemodialysis at National Kidney Foundation	Health Literacy: Functional, Communicative and Critical Health Literacy (FCCHL) Scale Exercise: Summary of Diabetes Self-Care Activities Questionnaire (SDSCA), revised	n/a	Pearson Correlation; not adjusted	Functional HL and PA: r = 0.05 (p = 0.708) Communicative HL and PA: r = 0.28 (p = 0.027) Critical HL and PA: r = 0.27 (p = 0.033) Total HL and PA: r = 0.27 (p = 0.038)	n/a
Brega et al. (2012); USA						
Cross-Sectional	N = 2594 Age: ≥18 years Baseline sample from tne Special Diabetes Program for Indians Healthy Heart Project	Health Literacy: Special Diabetes Program for Indians Healthy Heart Literacy Survey Exercise: number of minutes/week in the month prior to enrollment (no specific measure identified)		Structural Equation Modeling, testing three nested mediation models, controlling for known Health Literacy covariates (i.e., age, education, income,& sex)	Final model demonstrates a direct negative effect of Health Literacy on Physical Activity (path coefficient = -0.97, p < 0.05)	Despite a direct effect of Health Literacy on Diabetes Knowledge (path coefficient = 0.695, p < 0.05), diabetes knowledge did not mediate the relationship between health literacy and physical activity.

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA	Mediation Analysis
			Cross-Sectional Studies			
van der Heide et al. (2014); Netherlands						
Cross-Sectional	N = 1714 adults with diabetes Age: ≥ 18 years Randomly selected sample of diabetes patients, clustered within a random sample of general practitioner offices	Health Literacy: three validated questions on a 5 point Likert scale, with higher scores indicating greater health literacy (Chew et al, 2004, 2008) Physical Activity: two-item assessment of self-reported non-sport related PA and sport participation from a typical previous week (Sluiter et al., 2012)	Adequate HL: n = 1547 (90.3%) Inadequate HL: n = 167 (9.7%)	Multiple regression, controlling for diabetes duration, insulin use, perceived health status, presence of comorbidities, age, education, and sex.	$\beta = 0.14$ (SE = 0.06) $p < 0.05$	Using multiple regression methods, also tested for mediational effect of diabetes knowledge (HL -> knowledge -> PA); no mediational pathway was detected c' path: $\beta = 0.13$ (SE = 0.06)
Husson et al. (2015); Netherlands						
Cross-Sectional	N = 1643 adult survivors of colorectal cancer Age: *range/mean for total sample not specified* Convenience sample from national cancer registry (RR = 83%)	Health Literacy: three validated questions on a 5 point Likert scale, with higher scores indicating greater health literacy (Chew et al, 2004, 2008) Physical Activity: European Prospective Investigation into Cancer PA Questionnaire; average number of weekly hours spent in various activities and dichotomized into meeting Dutch PA guideline of 150 min/week of MVPA	Adequate HL: n = 677 (42%) Marginal HL: n = 725 (45%) Inadequate HL: n = 224 (14%)	Chi-Square; not adujsted	Compared to patients categorized as moderate or high HL, patients categorized as having low HL were less likely to meet MVPA recommendations ($p < 0.01$).	n/a

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA	Mediation Analysis
			Longitudinal Studies			
Kim et al. (2004); USA						
Longitudinal (Prospective Observational Study)	N = 77 Age: ≥ 18 years Convenience sample from existing diabetes education classes at a University hospital	Health Literacy: Short version of the Test of Functional Health Literacy in Adults (S-TOFHLA) Exercise: Summary of Diabetes Self-Care Activities Questionnaire (SDSCA)	Adequate HL: n = 61 (79.2%) Inadequate HL: n = 16 (20.8%)	ANCOVA comparing 3 month change, adjusting for age, years of education, income, and baseline values	Adequate HL: Mean = 2.8 ± 0.31 Limited HL: Mean = 2.1 ± 0.67; $p = 0.022$; effect size = 0.13	n/a
Geboers et al. (2014); Netherlands						
Longitudinal (pre-post quasi-experimental controlled design)	N = 538 older adults participating in a 9-month community-based social ecological intervention designed to increase physical activity and healthy eating among seniors Age: ≥ 55 years Randomly selected sample from intervention and control communities	Health Literacy: three validated questions on a 5 point Likert scale, with higher scores indicating greater health literacy (Chew et al, 2004, 2008) Physical Activity: SQUASH, a validated Dutch questionnaire; estimates compliance with guidelines	Adequate HL: n = 345 (64.1%) Inadequate HL: n = 193 (35.9%)	Multivariate logisitc regression (inadequate health literacy vs. adequate health literacy), adjusted for age, gender, and intervention condition.	OR = 1.52 (95% CI: 1.00, 2.31, $p = 0.053$)	Relationship between HL and PA mediated by self-efficacy: OR = 2.36 (95% CI: 1.60, 3.49, $p < 0.001$)

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Treatment Description	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA
			Randomized Control Trials			
Cavanaugh et al. (2009); USA						
RCT	N = 198 adults with Type I or Type II diabetes, with A1C ≥7.0, and referred by physician Age: 31-60 years (median = 52 years) Computer randomized in blocks of 4, 6, or 8; 1:1 intervention-to-control ratio	Control: "usual care" treatment including up to 6 face-to-facediabetes care visits over 3 months Intervention: "usual care" program, plus program staff received special training to deliver materials at a 6th grade reading level (e.g., bulleting for key points, color coding, pictures, step-by-step instructions, speaking in clear, simple sentences)	Health Literacy: Rapid Estimate of Adult Literacy in Medicine (REALM) Diabetes-related Numeracy: Diabetes Numeracy Test (DNT) Exercise: Summary of Diabetes Self-Care Activities Scale	Total Sample: 40% of patients had below 9th grade literacy level Control Arm: 35% of patients had below 9th grade literacy level Intervention Arm: 40% of patients had below 9th grade literacy level	Wilcoxon rank-sum tests examining between group differences in self-management behaviors (Physical activity)	ns (no specific statistics provided)
Eckman et al. (2012); USA						
RCT	N = 170 adult patients with a diagnosis of coronary artery disease, angina, or prior myocardial infarction Age: ≥ 18 years Sequentially assigned to intervention arms at first visit; based on assumption that patient scheduling and appointment time was a random process; 1:1 intervention-to-control ratio	Control: "usual care" treatment including a 20-page printed "Living with Corolnary Heart Disease - Doing Your Part" booklet written at the 5th grade reading level (Foundation for Informed Medical Decision Making) Intervention: "usual care" booklet plus 30 minute DVD/VCR video	Health Literacy: Rapid Estimate of Adult Literacy in Medicine (REALM) Exercise: Physical Activity Scale for the Elderly (PACE)	Total Sample: 40% (n = 68) of participants had low Health Literacy	Only within treatment arm changes physical using t-tests are provided; Subgroup analysis (by high/low health literacy) examinig the moderating influence of health literacy on changes to physical activity was also examined *(specific analysis details not provided)*	No between group treatment effect for physical activity (specific statistical information not reported); Health Literacy (high versus low) did not moderate the change in physical activity [low health literacy M = 2.32, sd = 20.52; high health literacy M = 5.13, sd = 19.16; p = 0.37] Note: subgroup analysis (high vs. low health literacy) was not planned apriori, and therefore the sample size calculation did not take this factor into account and participant randomization was not stratified by this characteristic

Appendix C

Studies examining the relationship between health literacy/numeracy and physical activity behavior (continued).

Study Design	Characteristics of the Sample	Treatment Description	Measurement Tools	Health Literacy Prevalence	Analysis	Association between HL and PA
Randomized Control Trials						
Dominick et al. (2013); USA						
RCT	N = 89 healthy, Spanish-speaking Latinas enrolled in a 6-month RCT testing the efficacy of a print-based culturally and linguistically adapted Mean Age = 41.37 years (sd = 11.18) Individual Randomized, stratified by stage of change	Control: mailed general wellness information on topics other than physical activity at identical intervals as intervention arm Seamos Activas: 6-month theory-based cognitive-behavioral print-based individually tailored intervention	Health Literacy: Short version of the Test of Functional Health Literacy in Adults (S-TOFHLA); Spanish Version Physical Activity: 7-day Physical Activity Recall (7-day PAR)	Adequate HL: 81% Marginal HL: 9% Inadequate HL: 10%	Between group differences in changes in moderateto-vigorous physical activity were determined using using ANOVA Linear regression, controlling for treatment assignment, baseline physical activity, and key confounders *(not specified)*	No significant between-group differences for changes in physical activity were found (F [1, 91] 1.37, p 0.25; Pekmezi et al., 2009); No association ($p > 0.05$) between health literacy and physical activity changes were found at 6 months; *(no other statistics provided)*
Glasgow et al. (2012); USA						
RCT	N = 463 adult patients with a diagnosis of Type II Diabetes Age: 25 - 70 years (M = 58.4 ± 9.2 years) Individually randomized by computer;	Control: computer-based health risk appraisal and recommendations for preventive care behaviors using the same contact schedule as CASM/CASM+, but did not include the key intervention content CASM: Given access and instruction on web-based, cognitive- behavioral intervention. Website allowed self-monitoring, personalized feedback, goal-setting. Automated telephonic system provided reminder calls and collected data. CASM+: Received CASM intervention plus two interventionst telephone follow-up calls and invitation to attend three 120 minute group-based meetings.	Health Literacy: three validated questions on a 5 point Likert scale, with higher scores indicating greater health literacy (Chew et al, 2004) Health Numeracy: Subjective Numeracy Scale Exercise: Total weekly caloric expenditure using the CHAMPS	Total Sample: 5.9% of participants had low to moderate health literacy	Intent-to-treat analysis assessing control vs. CASM/CASM+ groups using treatment x time chi-square Because there was a significant effect of the intervention on physical activity,the moderating effect of Health Literacy on physical activity changes was assessed using the interaction of treatment and health literacy (alpha level set to 0.01 for moderator analyses)	Health Literacy did not moderate the intervention changes in physical activity. Therefore, the intervention is robust across all health literacy levels