Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Di H. Cross Date

Effects of Individual and Environmental Factors
on Intervention Efficacy and Participant Retention
in the American Cancer Society's Study of Nutrition and Physical Activity

By

Di H. Cross
Doctor of Philosophy
Epidemiology

_____
Carol J. Rowland Hogue
Advisor

_____
Roberd M. Bostick
Committee Member

_____
Youngmee Kim
Committee Member

_____
K. M. Venkat Narayan
Committee Member

_____
Lance A. Waller
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D. Dean of the Graduate School

_____
Date

Effects of Individual and Environmental Factors on Intervention Efficacy and Participant Retention in the American Cancer Society's Study of Nutrition and Physical Activity

By

Di H. Cross
B.S.E., Case Western Reserve University, 2004

Advisor: Carol J. Rowland Hogue, Ph.D., MPH

An abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Epidemiology
2009

Abstract

Effects of Individual and Environmental Factors
on Intervention Efficacy and Participant Retention
in the American Cancer Society's Study of Nutrition and Physical Activity

By Di H. Cross

Observational studies have demonstrated positive associations between individual physical activity (PA) and environmental factors like increased access to parks, safer streets, and dense, mixed use of surrounding land. Similarly, greater fruit and vegetable consumption (FVC) has been observed to be associated with increased availability or greater variety of healthy foods. However, a question of temporality remains unanswered as studies documenting such associations have largely employed cross-sectional designs. The few longitudinal studies that have been performed may still leave that question unanswered, as the same associations would arise if individuals selected their environments based on desired behaviors rather than if the environment caused the behavior.

This dissertation aimed to address the question of temporality by examining environmental-level characteristics with respect to loss to follow-up and differential efficacy in a study of an intervention designed to modify PA and FVC behaviors. This was accomplished using baseline, and 3-, 6-, and 12-month follow-up data from the American Cancer Society's Nutrition and Physical Activity (NuPA) Study – a randomized controlled trial of a low-intensity, telephone-administered counseling intervention – and commercially-available databases of grocer and park locations, and street connectivity.

The greatest intervention effect on FVC was observed among participants with the least access to grocers. The greatest intervention effect on PA was observed among participants with the least access to parks. No consistent difference in intervention effect was found across levels of street connectivity. No characteristics of the nutrition or physical activity environment were associated with loss to follow-up.

Our findings suggest that participants with the least access to resources may benefit the most from this low-intensity counseling intervention. This reveals a more complex picture of the relationship between environment and behavior than previously thought and suggests that even if access is poor, well-structured interventions can overcome disparities in health behaviors. Furthermore, this dissertation demonstrates a non-traditional use of randomized trial data in understanding problematic exposure-outcome relationships. Replication of these analyses should be performed using data from randomized trials of different interventions with emphasis placed on examining patterns of intervention effects across levels of potential effect modifiers rather than on evaluating statistical significance.

Effects of Individual and Environmental Factors on Intervention Efficacy and Participant Retention in the American Cancer Society's Study of Nutrition and Physical Activity

By

Di H. Cross
B.S.E., Case Western Reserve University, 2004

Advisor: Carol J. Rowland Hogue, Ph.D.

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Epidemiology
2009

Acknowledgements

The author would like to acknowledge

The contributions of Heather Adams and Joanne Pike
from the American Cancer Society
in their dedication to the NuPA Study

And

The support of her friends and family

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Chronic Disease: Burden and Preventability

Among the 6 leading causes of death, four conditions – heart disease, malignant neoplasms, cerebrovascular disease and diabetes mellitus – are related to obesity or overweight or share the same cardio-metabolic risk factors of physical inactivity and poor diet. Together, these four conditions were the primary cause for 57.7% of all deaths in the United States in 2006.[1] However, more than just contributing to mortality, these conditions had a combined cost of $80 billion in health care expenditures and an estimated $17.3 billion dollars in lost productivity in the U.S. based on data from 2000 to 2002. These figures excluded cardiovascular disease which, by itself, was estimated to have incurred $448.5 billion in direct and indirect costs in 2008.[2-5]

Low Physical Activity and Fruit and Vegetable Consumption as Risk Factors

Many of these health conditions are preventable and treatable with evidence consistently pointing to low fruit and vegetable consumption (FVC) and physical activity (PA) as risk factors.[2,6-9] For the prevention of chronic disease, current dietary guidelines recommend 9 servings of fruits and vegetable (2,000 calorie reference).[10] Physical activity guidelines currently recommend at least 150 minutes/week of moderate-intensity, or 75 minutes/week of vigorous-intensity aerobic exercise for substantial health benefits including lowered risk of early death, heart disease, diabetes, colon and breast cancers, and prevention of weight gain.[11] However, based on 2005 BRFSS data, only 32.6% of adults consumed fruit on two occasions or more per day and 27.2% ate vegetables three or more times per day.[12] Similarly, only 26.2% of US adults engaged in 30 minutes or

more of moderate intensity activity on 5 or more days per week or 20 minutes or more of vigorous intensity activity per day on 3 or more days per week, the levels recommended for prevention of chronic disease at the time.[13] Examined in combination, only 14.6% of Americans engage in recommended amounts of both physical activity and diet behaviors.[14] Moreover, meaningful decreases in disease burden can be achieved with only small changes in behavior, even at levels below the published guidelines. Observational studies report a significant 4% decrease in the risk of coronary heart disease with every 1 serving/day increase in FVC and a 6% decrease in cancer incidence.[15,16] Similarly for PA, an increase in 500 kcal/week energy expenditure – or the equivalent of 1 hour of moderate-intensity aerobic activity per week for a 70 kg (154 lb) individual – resulted in 6% decrease in the risk of non-insulin-dependent diabetes.[17,18]

Efforts to change behavior, however, have been only moderately successful at best. Mass campaigns designed to educate Americans about recommended levels of FVC and PA for chronic disease prevention have not yielded meaningful changes in the prevalence of FVC and PA behaviors, with some indications that PA may in fact be decreasing.[19-21] Individual-level interventions fare much the same. Despite the introduction of behavioral interventions, which have included components such as social support, guidance, and motivational techniques, randomized trials to increase PA to predetermined levels have generally been unsuccessful, despite intense intervention periods and high costs.[22-24] Randomized trials for FVC behavior change have been moderately successful, yielding increases of up to 1 serving/day of FVC in intervention arms compared to control.[25-28] However, recidivism to pre-intervention levels remains a problem based on findings from studies with follow-up extending past the intervention

period and because limited evidence has been provided of sustained intervention effects past 9 months.[28,29]

Experts in the field have suggested that such individual-oriented approaches may not be adequate to address behavior change.[30-33] This has partly been informed by observational studies demonstrating associations between environmental factors and individual levels of PA[34-41] and diet or FVC.[42-44] Specifically, neighborhood qualities, such as increased access to parks and recreational facilities, safer streets, and dense and mixed use of surrounding land have been associated with greater PA, even after controlling for individual-level characteristics[45-50], and access to healthful foods through lower prices, increased availability, or greater variety has been shown to be associated with increased FVC.

Although these studies demonstrate an association between environmental factors and individual behaviors, it is not clear whether changing the environment would be adequate for changing behavior. Studies documenting these associations have almost exclusively employed cross-sectional designs, leaving the temporal relationship between environment and behavior to be ambiguous.[51,52] The few longitudinal studies that have been performed may still leave the question of temporality unanswered, as individuals might select their environments based on desired behaviors rather than vice versa.[53] Furthermore, studies conducted with participants that are more racially and/or ethnically homogeneous have not found associations between the physical environment and individual behavior.[54-56]

Research Objectives

Given the nature of environmental exposures, it would be difficult to conduct a randomized trial in order to provide a more satisfactory level of evidence.[57,58] However, it is possible to address whether environmental exposures are causal precursors to PA and FVC behaviors by examining them as potential barriers to or facilitators of success within a trial. The proposed research aimed to address limitations in the observational literature implicating the environment as a determinant of individual behavior by examining the built, physical environment as potential barriers to or facilitators of success among participants in an intervention trial. As an additional advantage, this approach may also identify defined populations where the examined intervention is differentially effective, informing decisions regarding the allocation of resources for maximum impact on population health.[59]

Potential explanations for the moderate to null effects observed in the randomized trial literature are threefold:

1. Truly null effects
2. Unidentified modifiers of intervention effects
3. Informative or non-informative loss to follow-up

The first reason requires little explanation – estimates of intervention effects may approach zero simply because the intervention, in truth, does not have an effect.

The second reason presents the possibility that there exist characteristics that modify the intervention effect. When these characteristics are ignored, the apparent intervention effect is averaged over the characteristics and may be attenuated and appear to be null. This is supported by evidence from randomized trials which, despite demonstrating limited overall effects, have identified more pronounced successes within

specific subpopulations usually identified by demographic variables. However, these variables – occupation[60], psychological predictors[61,62], gender[62,63], baseline behaviors[63,64], age[62,63,65], education[62], race/ethnicity[66], [67]marital status – are generally un-modifiable. Although they can be used to identify sub-groups among whom the intervention may be more successful they cannot be changed as a means of making the intervention more effective and address the underlying public health problem. In contrast, environmental factors may be changed at the population level through policy-change or legal measures.[68,69]

However, the same variables examined as potential effect modifiers are often predictors of the health outcomes that are under study and are also predictors of loss to follow-up, leading into reason number 3: both informative and non-informative loss to follow-up. In the case of behavioral modification trials, this is particularly alarming because some trials of PA change have reported up to 87% loss to follow-up.[22] Substantial losses may result in significantly decreased power, but also call into question the validity of the assumption of non-informative loss.[70] To the authors' knowledge, there have been no studies that have examined whether environmental characteristics potentially related to individual behavior are also predictors of loss to follow-up. If the environment is a causal precursor for behaviors which are the outcomes under study, it would be prudent to assess whether they are also predictors of loss from the study and address the magnitude and direction of biases that would result from these losses.

This research will be conducted using data from a subset of participants enrolled in the American Cancer Society's Nutrition and Physical Activity (NuPA) Study – a randomized trial examining a telephone-administered counseling intervention designed to

improve lifestyle behaviors including, but not limited to, physical activity and fruit and vegetable consumption.[71] Complete data from enrollment, 3-, 6-, and 12-month follow-up evaluations will be used to address the following research questions:

1. *Do physical characteristics of the participant residential environment modify the effect of the intervention in increasing the proportion of participants meeting the 30 minutes/day recommendation for physical activity?* The physical characteristics examined will include street and intersection density, and availability of and proximity to green space. The residence will be used as the focal point for ascertaining environment. Guided by the literature, it is hypothesized that increased street and intersection density, and greater access to green space will be associated with greater increases in the odds of individuals meeting PA recommendations among those in the intervention arm compared to the control arm at follow-up after accounting for differences in PA at baseline. [37,39,45,46,72]
2. *Do physical characteristics of the participant residential environment modify the average change in FVC?* Studies have reported that major supermarkets offer more than 4 times the number of healthful foods including fresh produce when compared to other types of food retailers such as neighborhood grocers and convenience stores.[73] The physical characteristics under study are proximity to, variety of and diversity of supermarkets surrounding participant residences. While the intervention has been shown to significantly increase FVC among those randomized to the intervention arm in comparison to controls, this intervention effect is hypothesized to be greater among those with greater access to supermarkets.[71,74,75]

3. *Can the intervention effect or lack thereof, be explained by selective loss to follow-up among participants living in areas with specific physical environmental characteristics related to the intended behaviors?* This question will be addressed by examining environmental measures from both previous research questions and other potential predictors of loss to follow-up. The outcomes of interest will be time until the participant is not contacted for follow-up evaluation or time until refusal to participate. Individuals residing in environments providing more opportunities for increased PA and/or greater FVC are hypothesized to be contacted at a rate no different than others and to refuse to participate at a lower rate than those with less opportunities.

In addition, the existing literature on randomized trials of interventions designed to increase FVC will be examined addressing the fourth and last research question:

4. *Do published reports of randomized trials of FVC interventions report on effect modification and, if so, what are the challenges faced in its estimation, interpretation, and application of findings?* Non-pharmaceutical behavioral trials designed to increase FVC with no component involving provision of food, and no environmental- or policy-level change will be examined to determine the extent to which researchers investigate and report on potential effect modification, noting the type of effect modifier (ie., demographic, social-contextual, cognitive variable, etc.), and specific challenges faced in the application of study findings.

References

1. Heron M, Hoyert DL, Murphy SL, Xu J, Kochanek KD, Tejada-Vera B. Deaths: Final Data for 2006. National Vital Statistics Reports 2009;57(14).
2. Heart Disease and Stroke Statistics: 2008 Update. Dallas, TX: American Heart Association, 2008.
3. Sullivan PW, Ghushchyan V, Wyatt HR, Hill JO. The medical cost of cardiometabolic risk factor clusters in the United States. Obesity (Silver Spring) 2007;15(12):3150-8.
4. Sullivan PW, Ghushchyan V, Wyatt HR, Wu EQ, Hill JO. Productivity costs associated with cardiometabolic risk factor clusters in the United States. Value Health 2007;10(6):443-50.
5. Heart Disease and Stroke Statistics: 2005 Update. Dallas, TX: American Heart Association, 2005.
6. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. Journal of the National Cancer Institute 1981;66(6):1191-1308.
7. Espey DK, Wu XC, Swan J, Wiggins C, Jim MA, Ward E, Wingo PA, Howe HL, Ries LA, Miller BA, Jemal A, Ahmed F, Cobb N, Kaur JS, Edwards BK. Annual report to the nation on the status of cancer, 1975-2004, featuring cancer in American Indians and Alaska Natives. Cancer 2007;110(10):2119-52.
8. Kushi LH, Byers T, Doyle C, Bandera EV, McCullough M, McTiernan A, Gansler T, Andrews KS, Thun MJ. American Cancer Society Guidelines on Nutrition and Physical Activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2006;56(5):254-81; quiz 313-4.
9. Diabetes in America. 2nd Edition ed National Institutes of Health, 2008.
10. Dietary Guidelines for Americans. In: Services USDoHaH, Agriculture USDo, eds, 2005.

11. 2008 Physical Activity Guidelines for Americans. In: Services USDoHaH, ed. Washington, DC, 2008.

12. Fruit and vegetable consumption among adults--United States, 2005. MMWR Morb Mortal Wkly Rep 2007;56(10):213-7.

13. Prevalence of physical activity, including lifestyle activities among adults--United States, 2000-2001. MMWR Morb Mortal Wkly Rep 2003;52(32):764-9.

14. Prevalence of fruit and vegetable consumption and physical activity by race/ethnicity--United States, 2005. MMWR Morb Mortal Wkly Rep 2007;56(13):301-4.

15. Joshipura KJ, Hu FB, Manson JE, Stampfer MJ, Rimm EB, Speizer FE, Colditz G, Ascherio A, Rosner B, Spiegelman D, Willett WC. The Effect of Fruit and Vegetable Intake on Risk for Coronary Heart Disease. Ann Intern Med 2001;134(12):1106-1114.

16. van't Veer P, Jansen MCJF, Klerk M, Kok FJ. Fruits and vegetables in the prevention of cancer and cardiovascular disease. Public Health Nutrition 2000;3(01):103-107.

17. Helmrich S, Ragland D, Leung R, Paffenbarger R. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. N Engl J Med 1991;325(3):147-152.

18. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum. Nutr. Clin. Nutr. 1985;39C(Suppl 1):5-41.

19. Brownson RC, Boehmer TK, Luke DA. Declining rates of physical activity in the United States: what are the contributors? Annu Rev Public Health 2005;26:421-43.

20. Casagrande SS, Wang Y, Anderson C, Gary TL. Have Americans increased their fruit and vegetable intake? The trends between 1988 and 2002. Am J Prev Med 2007;32(4):257-63.

21. Marcus BH, Owen N, Forsyth LH, Cavill NA, Fridinger F. Physical activity interventions using mass media, print media, and information technology. Am J Prev Med 1998;15(4):362-78.

22. Marcus BH, Williams DM, Dubbert PM, Sallis JF, King AC, Yancey AK, Franklin BA, Buchner D, Daniels SR, Claytor RP. Physical activity intervention studies: what we know and what we need to know: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity); Council on Cardiovascular Disease in the Young; and the Interdisciplinary Working Group on Quality of Care and Outcomes Research. Circulation 2006;114(24):2739-52.

23. Marcus BH, Stanton AL. Evaluation of relapse prevention and reinforcement interventions to promote exercise adherence in sedentary females. Res Q Exerc Sport 1993;64(4):447-52.

24. Foster C, Hillsdon M, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180. pub2.

25. Coates RJ, Bowen DJ, Kristal AR, Feng Z, Oberman A, Hall WD, George V, Lewis CE, Kestin M, Davis M, Evans M, Grizzle JE, Clifford CK. The Women's Health Trial Feasibility Study in Minority Populations: changes in dietary intakes. American Journal of Epidemiology 1999;149(12):1104-12.

26. Heimendinger J, O'Neill C, Marcus AC, Wolfe P, Julesburg K, Morra M, Allen A, Davis S, Mowad L, Perocchia RS, Ward JD, Strecher V, Warnecke R, Nowak M, Graf I, Fairclough D, Bryant L, Lipkus I. Multiple tailored messages are effective in increasing fruit and vegetable consumption among callers to the Cancer Information Service. Journal of Health Communication 2005;10 Suppl 1:65-82.

27. Kristal AR, Curry SJ, Shattuck AL, Feng Z, Li S. A randomized trial of a tailored, self-help dietary intervention: the Puget Sound Eating Patterns study. Preventive Medicine 2000;31(4):380-9.

28. Marcus AC, Heimendinger J, Wolfe P, Rimer BK, Morra M, Cox D, Lang PJ, Stengle W, Van Herle MP, Wagner D, Fairclough D, Hamilton L. Increasing fruit and vegetable

consumption among callers to the CIS: results from a randomized trial. Preventive Medicine 1998;27(5 Pt 1):S16-28.

29. Brunner EJ, Rees K, Ward K, Burke M, Thorogood M. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2007(4):CD002128.

30. Garner D, Wooley S. Confronting the failure of behavioral and dietary treatments for obesity. Clin Psychol Rev 1991;11:729-80.

31. Brownson RC, Haire-Joshu D, Luke DA. Shaping the context of health: a review of environmental and policy approaches in the prevention of chronic diseases. Annu Rev Public Health 2006;27:341-70.

32. Economos C, Brownson R, DeAngelis M, Novelli P, Foerster S, Foreman C, Gregson J, Kumanyika S, Pate R. What lessons have been learned from other attempts to guide social change? Nutr Rev 2001;59(3 pt 2):S40-56, discussion S57-65.

33. Cummins S, Macintyre S. Food environments and obesity - neighbourhood or nation? International Journal of Epidemiology 2006;35:100-104.

34. Bedimo-Rung AL, Mowen AJ, Cohen DA. The Significance of Parks to Physical Activity and Public Health: A Conceptual Model. Am J Prev Med 2005;28(2S2):159-168.

35. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Ann Behav Med 2003;25(2):80-91.

36. Vernex Moudon A. Active Living Research and the Urban Design, Planning, and Transportation Disciplines. Am J Prev Med 2005;28(2S2):214-215.

37. Targa F, Clifton KJ. Built Environment and Nonmotorized Travel: Evidence from Baltimore City Using the NHTS. Journal of Transportation and Statistics 2005;8(3):55-70.

38. Powell KE. Land Use, the Build Environment, and Physical Activity. Am J Prev Med 2005;28(2S2):216-217.

39. Wendel-Vos GC, Schuit AJ, de Niet R, Boshuizen HC, Saris WH, Kromhout D. Factors of the physical environment associated with walking and bicycling. Med Sci Sports Exerc 2004;36(4):725-30.

40. Owen N, Humpel N, Leslie E, Bauman A, Sallis JF. Understanding environmental influences on walking; Review and research agenda. Am J Prev Med 2004;27(1):67-76.

41. McCormack G, Giles-Corti B, Lange A, Smith T, Martin K, Pikora TJ. An update of recent evidence of the relationship between objective and self-report measures of the physical environment and physical activity behaviours. J Sci Med Sport 2004;7(1 Suppl):81-92.

42. Rose D, Richards R. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. Public Health Nutrition 2004;7(8):1081-1088.

43. Shohaimi S, Welch A, Bingham S, Luben R, Day N, Wareham N, Khaw KT. Residential area deprivation predicts fruit and vegetable consumption independently of individual education level and occupational social class: a cross sectional population study in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). Journal of Epidemiology and Community Health 2004;58:686-691.

44. Diez-Roux AV, Nieto FJ, Caulfield L, Tyroler HA, Watson RL, Szklo M. Neighbourhood differences in diet: the Atherosclerosis Risk in Communities (ARIC) Study. J Epidemiol Community Health 1999;53(1):55-63.

45. Frank LD, Saelens BE, Powell KE, Chapman JE. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc Sci Med 2007;0(0):0.

46. Boer R, Zheng Y, Overton A, Ridgeway GK, Cohen DA. Neighborhood design and walking trips in ten U.S. metropolitan areas. Am J Prev Med 2007;32(4):298-304.

47. Pikora TJ, Giles-Corti B, Knuiman MW, Bull FC, Jamrozik K, Donovan RJ. Neighborhood environmental factors correlated with walking near home: Using SPACES. Med Sci Sports Exerc 2006;38(4):708-14.

48. McNeill LH, Wyrwich KW, Brownson RC, Clark EM, Kreuter MW. Individual, social environmental, and physical environmental influences on physical activity among black and white adults: a structural equation analysis. Ann Behav Med 2006;31(1):36-44.

49. Zimring C, Joseph A, Nicoll GL, Tsepas S. Influences of building design and site design on physical activity: research and intervention opportunities. Am J Prev Med 2005;28(2 Suppl 2):186-93.

50. van Lenthe FJ, Brug J, Mackenbach JP. Neighbourhood inequalities in physical inactivity: the role of neighbourhood attractiveness, proximity to local facilities and safety in the Netherlands. Soc Sci Med 2005;60(4):763-75.

51. Kamphuis CB, Giskes K, de Bruijn GJ, Wendel-Vos W, Brug J, van Lenthe FJ. Environmental determinants of fruit and vegetable consumption among adults: a systematic review. Br J Nutr 2006;96(4):620-35.

52. Hanson S, Berkowitz B, Ainsworth B, Blair S, Cervero R, Chen D, Crane R, Thompson Fullilove M, Giuliano G, Lawton T, Mokhtarian P, Powerll K, Stutts JV, RP. Does the Built Environment Influence Physical Activity? Examining the Evidence. Washington, D.C.: Transportation Research Board and the Institute of Medicine of the National Academics, 2005.

53. Weiss DR, O'Loughlin J L, Platt RW, Paradis G. Five-year predictors of physical activity decline among adults in low-income communities: a prospective study. Int J Behav Nutr Phys Act 2007;4:2.

54. Rohm Young D, Voorhees CC. Personal, social, and environmental correlates of physical activity in urban African-American women. Am J Prev Med 2003;25(3 Suppl 1):38-44.

55. Voorhees CC, Rohm Young D. Personal, social, and physical environmental correlates of physical activity levels in urban Latinas. Am J Prev Med 2003;25(3 Suppl 1):61-8.

56. Evenson KR, Sarmiento OL, Tawney KW, Macon ML, Ammerman AS. Personal, social, and environmental correlates of physical activity in North Carolina Latina immigrants. Am J Prev Med 2003;25(3 Suppl 1):77-85.

57. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. Social Science and Medicine 2004;58:1929-1952.

58. Diez Roux AV. Estimating neighborhood health effects: the challenges of causal inference in a complex world. Social Science and Medicine 2004;58(10):1953-1960.

59. Kraemer HC, Frank E, Kupfer DJ. Moderators of Treatment Outcomes: Clinical, Research, and Policy Importance. Jama 2006;296(10):1286-1289.

60. Sorensen G, Barbeau E, Stoddard AM, Hunt MK, Kaphingst K, Wallace L. Promoting behavior change among working-class, multiethnic workers: results of the healthy directions--small business study. Am J Public Health 2005;95(8):1389-95.

61. Pinto BM, Friedman R, Marcus BH, Kelley H, Tennstedt S, Gillman MW. Effects of a computer-based, telephone-counseling system on physical activity. Am J Prev Med 2002;23(2):113-20.

62. Stewart AL, Verboncoeur CJ, McLellan BY, Gillis DE, Rush S, Mills KM, King AC, Ritter P, Brown BW, Jr., Bortz WM, 2nd. Physical activity outcomes of CHAMPS II: a physical activity promotion program for older adults. J Gerontol A Biol Sci Med Sci 2001;56(8):M465-70.

63. Calfas KJ, Sallis JF, Nichols JF, Sarkin JA, Johnson MF, Caparosa S, Thompson S, Gehrman CA, Alcaraz JE. Project GRAD: two-year outcomes of a randomized controlled

physical activity intervention among young adults. Graduate Ready for Activity Daily. Am J Prev Med 2000;18(1):28-37.

64. Effects of physical activity counseling in primary care: the Activity Counseling Trial: a randomized controlled trial. Jama 2001;286(6):677-87.

65. Tinker LF, Rosal MC, Young AF, Perri MG, Patterson RE, Van Horn L, Assaf AR, Bowen DJ, Ockene J, Hays J, Wu L. Predictors of dietary change and maintenance in the Women's Health Initiative Dietary Modification Trial. J Am Diet Assoc 2007;107(7):1155-66.

66. Havas S, Anliker J, Damron D, Langenberg P, Ballesteros M, Feldman R. Final results of the Maryland WIC 5-A-Day Promotion Program. American Journal of Public Health 1998;88(8):1161-7.

67. Campbell MK, Demark-Wahnefried W, Symons M, Kalsbeek WD, Dodds J, Cowan A, Jackson B, Motsinger B, Hoben K, Lashley J, Demissie S, McClelland JW. Fruit and vegetable consumption and prevention of cancer: the Black Churches United for Better Health project. American Journal of Public Health 1999;89(9):1390-6.

68. Ashe M, Feldstein LM, Graff S, Kline R, Pinkas D, Zellers L. Local venues for change: legal strategies for healthy environments. J Law Med Ethics 2007;35(1):138-47.

69. Reversing the obesity epidemic: policy strategies for health funders. Issue brief (Grantmakers Health) 2007(28):i-v, 1-49.

70. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. New York, NY: Lippincott Williams & Wilkins, 1998.

71. Kim Y, Pike J, Adams H, Cross D, Doyle C, Foreyt J. Worksite Telephone Intervention Promoting Weight-Related Health Behaviors. American Journal of Preventive MEdicine *Under Review*.

72. Foster C, Hillsdon M, Thorogood M. Environmental perceptions and walking in English adults. J Epidemiol Community Health 2004;58(11):924-8.

73. Morland K, Wing S, Diez Roux A, Poole C. Neighborhood characteristics associated with the location of food stores and food service places. Am J Prev Med 2002;22(1):23-9.

74. Morland K, Filomena S. Disparities in the availability of fruits and vegetables between racially segregated urban neighbourhoods. Public Health Nutr 2007:1-9.

75. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. Am J Public Health 2002;92(11):1761-7.

Chapter 2: Literature Review

Physical activity behavior and environment

Observational studies on physical activity and environment

The field of urban planning has long been concerned with how the design of neighborhoods and cities affect the lives of residents.[1,2] Only in the late 1990s, however, have collaborative efforts between urban planners and epidemiologists emerged to specifically address the impact of the built environment on physical activity (PA) and health.[3] This is partly explained by the emergence of evidence suggesting that moderate physical activity may be enough to effect changes in chronic disease risk. Research previous to this had focused on chronic disease risk and vigorous physical activity such as running, tennis, etc.

A wide variety of environmental-level exposures have been investigated in relation to individual PA behavior. Perceived social norms such as seeing neighbors more being more physically active[4], perceived neighborhood safety[5], greater residential density[6], higher business diversity[7], more conducive overall environment[6], access to green space that is larger, more attractive, and closer in proximity[8], greater street connectivity[6], and higher 4-way intersection density[7] have all been shown to be related to increased physical activity.

The multi-disciplinary nature of this field, the number of aspects of the environment and the variety of definitions by which each may be measured necessitate a framework by which environmental factors can be grouped and systematically studied.[9-11] First, all environmental factors are, by definition, outside of the individual. From there,

characteristics may be grouped into the physical, socio-cultural, economic, and political domains.[10] Policies and regulations on the location and procedures for creation of parks, and the zoning and density of land use would be included in the political domain. Examples of socio-cultural characteristics may include social norms, social support, and crime while the economic domain would include characteristics such as access to for-pay facilities for physical activity.[4,6,7,10,12-21] Examples of characteristics in the physical domain include street structure, availability of parks and other green space, presence of free and pay facilities for PA, sidewalks, and adequate neighborhood lighting. [4,6,15,21-23] Still other characteristics outside the individual include weather or climate, which change across time and may be aggregated across broader geographies.[4,12,22,24-27] However, even with this framework, there remain several significant challenges in studying the environment as it relates to PA behavior.

First, many environmental variables studied are not strictly affected by factors within only one domain. The location of parks or green space may be considered a factor in the physical domain. However, location is affected by rules and regulations in the political domain. In turn, the rules and regulations regarding the location of green space may be affected by factors in the socio-cultural domain as community members may (or may not) mobilize to change the political factors, thereby changing characteristics of the physical domain. Because of the complexity of this interplay, and particularly the difficulty in measuring characteristics of the social and political domains, many researchers have instead focused on factors in the physical domain.[28] This approach also has the advantage of identifying specific, observable factors in the environment upon which changes may be implemented.

Second, investigators must distinguish between participant-perceived environment and objectively measured environment. Undeniably related to objectively measurable characteristics of the environment, participant-reported perceived environment has also been shown to be influenced by individual characteristics and experiences.[29,30] However, even after making the distinction between perceived and objectively measured environment, there are some inconsistencies in findings. Two studies in African-American women and Latino women reported no statistically significant association between perceived PA environment and individual-level PA while one study in a more heterogeneous population demonstrated a positive association between perceived street connectivity, overall environment, residential density and PA behavior.[6,31,32] The difference in findings may be explained by residual confounding that occurred in the more heterogeneous population. Alternatively, the explanation may lie in the difference between perceived environment and objectively measured environment. An analysis of national data suggests that communities with low socio-economic status or a higher proportion of minorities are less likely to have access to facilities for physical activity.[33] In a nationally representative study comparing a sample of 19% of all US census block groups, block groups with only 5% minorities had up to a 20% greater odds of having access to 1 or more facilities for PA in comparison to block groups with 95% minorities. This difference was less pronounced in block groups whose residents were more highly educated (55% or more college educated, OR=0.93). In combination, those living in the lowest-educated and highest minority BGs were about 50% as likely to have access to 1 or more facility than BGs with the highest educated and lowest minority populations (OR=0.54; 95% CI: [0.51–0.58]).

Another challenge stems from the use of traditional epidemiologic study designs. The majority of studies examining physical activity environment and individual behavior have been cross-sectional – perhaps necessitated by the nature of the exposures studied. This calls into question the temporal relationship between environment and behavior and therefore the causality of observed associations.[6,34] A few longitudinal studies have attempted to address this concern. However, even in these cases, the question of temporality remains since adults may choose their place of residence based on their desired physical activity behaviors rather than vice versa. Since this decision is likely to have occurred before enrollment into the study, participant responses will still yield an association between behavior and environment because they share a common cause. If this is the case, a change in an individual's environment would not be sufficient to effect changes in individual behavior.

Although few and far between, studies of environmental interventions to change PA environment would more directly test the hypothesis.[35] These studies have been limited to those examining physical activity before and after the building of a multi-use trail.[36-38] Findings, however, have not been conclusive, with some studies showing no change in physical activity behavior despite a change in the environment.[36]

Randomized trials for increasing physical activity

Interventions to increase PA have generally been unsuccessful despite, in some cases, high participation and theory-driven interventions[39], involvement of primary care physicians[40], and intensive efforts on the part of both researchers and participants. Furthermore, substantial losses have been reported in other studies with up to 87% loss to follow-up over the study period.[41,42]

However, interventions studies that have been informed by findings from the observational literatures have been more successful. These include individual-level trials restricted to disadvantaged populations such as participants in federally funded job programs[43], and participants from low-income and multi-ethnic communities.[44-46] While the components of these interventions generally did not differ from other less successful interventions, they do have the advantage of having a starting population where PA prevalence was low.

Community-level trials which engage the community, employ a multitude of strategies and change the environment in a combination of the socio-cultural, political, economic, and physical domains have also been more successful. Increases in PA were observed among work-site studies with interventions after offering employees incentives, increasing knowledge about recommendations and the relationship between PA and chronic disease and mortality risk, and changing the environment.[47] However, these studies are limited in that they engage only those who are employed and only address the context of a work environment. Other studies that have attempted similar social and physical environmental changes in church or community health center studies have not been as successful in changing physical activity behavior.[44]

Fruit and vegetable consumption behavior and environment

Observational studies on fruit and vegetable consumption and environment

Although not as extensive as the literature on PA and built environment, a number of studies have shown that access to fruits and vegetables is associated with increased individual-level consumption of fruits and vegetables.[48,49] Similar to the studies of PA and environment, these studies are limited in that most are cross-sectional in design.

Using the same methods, other environmental factors that have been studied include the presence of small grocers[50], fast food outlets[51], and perceived availability of healthful products.[52] However, findings suggest that the most relevant food outlet for obtaining a variety of affordable fruit and vegetables is the supermarket which provides lower cost and a greater variety of produce in comparison to other outlets such as the corner store, convenience stores, and privately owned stores.[53-55] Complicating the situation however, is the evidence suggesting that supermarket location are associated with area-level socio-economics and race.[51,56,57] This has led some researchers to suggest that perhaps the pricing of product and the location of supermarkets is a simple issue of supply and demand: residents of neighborhoods with low SES and high minority populations may be less likely to purchase fruits and vegetables, regardless of whether or not they are available. Because of low demand for such products, supermarkets are less likely to stock them and may be less profitable in these areas.[58] Although this hypothesis cannot be tested using current study designs, data from some other studies seem to support this hypothesis, showing lower consumption of fruits and vegetables among non-Hispanic black and Hispanic women than among non-Hispanic white women, and among those with lower income.[59-61] However, at least one study has shown that once income-related barriers to consuming fruit and vegetables are removed, these individuals are likely to increase their consumption fruits and vegetables.[62]

Another challenge in studying the relationship between grocer access and individual behavior is in the measurement of grocer access. Investigators in this field have employed a variety of methods to determine the location of grocers with respect to their participants. These methods have ranged from using commercially available

databases of business information such as InfoUSA[7] and Dun and Bradstreet[63], to use of regional data maintained by local authorities[17,64], to – more rarely – investigator-collected data. However, each of these solutions presents its own advantages and limitations.

While commercially available databases provide national coverage, little research has been done to verify their quality in terms of misclassification of existing business and erroneous omission or inclusion of business. However, they offer the additional advantage of access to historic data to examine time-periods relevant to when participant data were collected. In contrast, investigator-collected data may be more flexible with respect to amount of information collected and features definitions, but they may reflect only current businesses. Furthermore, the amount of time and resources required to collect such data may be prohibitive in comparison to commercially available databases. Even with such high costs, misclassification, and erroneous omission or inclusion of businesses may still remain a problem. Furthermore, data cannot be collected after the fact to reflect time-periods relevant to historical participant data. Although investigation into historic records may be possible, this again increases the cost of creating such a database.

Regional data may be available to academic researchers in some regions. However, national coverage is far from complete and tends to focused on metropolitan areas. Furthermore, data may not be comparable across different geographies due to differences in resources, definitions, etc.

Randomized trials of fruit and vegetable consumption

As with the literature on PA a number of studies have been designed with recruitment restricted to high-risk populations and with interventions designed to fit

population-specific expectations and needs.[46,65-68] Unfortunately, these studies have also only had limited success. A number of studies have been designed with recruitment restricted to African American[69-71] or multi-ethnic[44,70] populations, rural communities[65,66], and low-income[44,69,72-74] or lower educational attainment[75] communities with some limited success.[46,65-68]

Loss to follow-up

Randomized Trials of Behavior Modification.

Many studies focusing on continued participation in a study as the outcome have been conducted in studies of high-risk populations such as HIV/AIDS patients[76-78], high-risk youth[79,80], and those with problems of substance abuse.[81-88] In these studies, variables most commonly associated with drop out were demographic variables – minority race/ethnicity[85], male gender[83], and young age[88] were found to be associated with increased likelihood of dropping out of the study.

In the behavioral intervention literature, these demographic variables are also commonly associated with outcomes under measure. However, the possibility for systematic loss based on these variables and therefore also stratified by the outcome has not been consistently assessed. The studies that do report results of analyses contrasting those who drop out of the study with those who remain in the study report similar findings as observational studies – again minority race/ethnicity[89,90], gender[89,91], and the extremes of the age distribution[89,91-93] have been shown to be associated with increased likelihood of loss to follow-up.[94] The direction of association with gender, however, is not consistently in one direction and may depend on the intervention condition. In addition, depression or psychological distress[95,96], baseline health behaviors[90,91,97,98], and

baseline health or severity of disease[99-101] have also been shown to be associated with discontinued follow-up. Beyond being an important factor in assessing the generalizability or external validity of a study, loss to follow-up is also a concern in determining whether a study is internally valid.[34] This is of particular concern when loss to follow-up is high. In the case of telephone-based studies of fruit and vegetable consumption and physical activity losses have ranged between 0% for one or two studies, to the more typical 15-30%, to upwards of 62% and 81% loss to follow-up over the study period.[102,103]

Rationale and Significance of Studies

The current study aims to address research questions one through four as outlined in Chapter 1 using data from the American Cancer Society (ACS) Nutrition and Physical Activity (NuPA) Study, a randomized trial of a telephone-based counseling intervention designed to increase the prevalence of lifestyle behaviors.

The NuPA Intervention

Preliminary analysis of participants from the main NuPA study who had completed 6 month follow-up evaluation (N=421) indicate that there was no statistically significant change in PA from enrollment to 6 month follow-up overall or by intervention arm (p=0.84). This was true even after controlling for whether or not participants already met the recommended 30 min/day of PA at baseline (48.2% of participants met recommendations at baseline compared to 56.5% at 5 months, OR=1.12 for meeting recommendation at 6 months contrasting intervention with self-help, p=0.59).[104]

Among 814 individuals who had completed 6 month evaluation and who had both enrollment and 6-month data, those in the intervention group (N=423) averaged 3.69

srv/day of fruits and vegetables while those in the self-help group (N=391) averaged 3.76 srv/day at baseline (difference, p=0.61). Both groups increased FVC by about 1 srv/day overall at the 6-month evaluation (p<0.0001). However, the amount of increase in FVC was not different across study groups (self-help increased by 0.88 srv/day, counseling increased by 1.12 srv/day, p=0.0710).

These findings are consistent with null to moderate findings from other randomized trials. However, it is also important to note that the intervention does not include a component of environmental (either physical or social) change. As the proposed analysis already requires complex analysis of interaction between environment and intervention, inclusion of environmental components in the intervention itself would further confuse analysis.

The proposed studies also examine loss to follow-up in relation to the demographic factors discussed as well as in relation to potential environmental-level predictors of fruit and vegetable consumption and physical activity behaviors. Characterizing loss to follow-up with respect to the exposures of interest in the main analysis will allow us to then predict the magnitude and direction of biases that may result in the estimation of the main intervention effect, as well as in the potential modification of the main intervention effect by environmental-level variables.

In order to best address the range of populations in the US, the intervention includes a wide range of demographic characteristics among participants from a number of geographic locations throughout the US. This is enabled through a low-intensity, broadly applicable intervention delivered through the phone that can be delivered to a wide range of participants throughout the US. Furthermore, multiple follow-ups facilitate

the examination of behavior change, examining the issue of temporality that has been so difficult to address in previous studies.

Environmental Variables

Physical Activity. In the case of physical activity, the environmental exposures that will be examined will include street structure (4-way intersection density, street density) and access to parks (distance to and size of nearest, total number and area accessible within a given distance). These variables were chosen as they have been extensively studied in the epidemiologic literature. Furthermore, they are concrete characteristics that can be reliably measured and may be solid metrics by useful metrics for subsequent interventions for changing the environment. Both groups of environment exposures address only physical activity that takes place outside of the home and does not examine personal ownership of physical activity equipment, or for-pay or free facilities for PA outside of parks.

Fruit and Vegetable Consumption. Access to grocery stores will be examined in relation to fruit and vegetable consumption as measured by proximity to the closest grocer, number of stores within a given buffer area (diversity), and the average distance to three differently named grocers (variety).[105] Although these three measures are highly correlated, conceptually they reflect different aspects of how accessible grocers may be in a given neighborhood. Proximity reflects the minimal effort required to reach any grocer, regardless of differential availability or pricing that may occur across different stores. Diversity reflects the range of choices a participant may have given a fixed distance that he or she may travel, regardless of whether specific stores in the area may be more

similar. Variety takes into account the similarity between stores of the same chain, and addresses how far a participant must travel to choose between chains.

References

1. Scott AJ, Roweis ST. Urban planning in theory and practice: a reappraisal. Environment and Planning 1977;9(10):1097-1119.
2. Myers D. Building Knowledge about Quality of Life for Urban Planning. American Planning Association 1988;54(Summer 1988):347-358.
3. Frank LD, Engelke P. The Impacts of the Built Environment on Physical Activity: Active Community Environments Initiative Working Paper (ACES)Working Paper #1. 1999.
4. Addy CL, Wilson DK, Kirtland KA, Ainsworth BE, Sharpe P, Kimsey D. Associations of perceived social and physical environmental supports with physical activity and walking behavior. Am J Public Health 2004;94(3):440-3.
5. Neighborhood safety and the prevalence of physical inactivity--selected states, 1996. MMWR Morb Mortal Wkly Rep. Vol. 48, 1999;143-6.
6. Atkinson JL, Sallis JF, Saelens BE, Cain KL, Black JB. The association of neighborhood design and recreational environments with physical activity. Am J Health Promot 2005;19(4):304-9.
7. Boer R, Zheng Y, Overton A, Ridgeway GK, Cohen DA. Neighborhood design and walking trips in ten U.S. metropolitan areas. Am J Prev Med 2007;32(4):298-304.
8. Giles-Corti B, Broomhall MH, Knuiman M, Collins C, Douglas K, Ng K, Lange A, Donovan RJ. Increasing Walking: How Important is Distance To, Attractiveness, and Size of Public Open Space? Am J Prev Med 2005;28(2S2):169-176.

9. Ewing R, Kreutzer R. Understanding the relationship between public health and the built environment: A report prepared for the LEED-ND Core Committee. 2006.

10. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev Med 1999;29(6 Pt 1):563-70.

11. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: a review. Am J Prev Med 2002;22(3):188-99.

12. Wendel-Vos GC, Schuit AJ, de Niet R, Boshuizen HC, Saris WH, Kromhout D. Factors of the physical environment associated with walking and bicycling. Med Sci Sports Exerc 2004;36(4):725-30.

13. Targa F, Clifton KJ. Built Environment and Nonmotorized Travel: Evidence from Baltimore City Using the NHTS. Journal of Transportation and Statistics 2005;8(3):55-70.

14. Suminski RR, Poston WS, Petosa RL, Stevens E, Katzenmoyer LM. Features of the neighborhood environment and walking by U.S. adults. Am J Prev Med 2005;28(2):149-55.

15. Brownson RC, Baker EA, Housemann RA, Brennan LK, Bacak SJ. Environmental and policy determinants of physical activity in the United States. Am J Public Health 2001;91(12):1995-2003.

16. Frank LD, Saelens BE, Powell KE, Chapman JE. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc Sci Med 2007;0(0):0.

17. Frank LD, Schmid TL, Sallis JF, Chapman J, Saelens BE. Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ. Am J Prev Med 2005;28(2 Suppl 2):117-25.

18. Michael Y, Beard T, Choi D, Farquhar S, Carlson N. Measuring the influence of built neighborhood environments on walking in older adults. J Aging Phys Act 2006;14(3):302-12.

19. De Bourdeaudhuij I, Sallis JF, Saelens BE. Environmental correlates of physical activity in a sample of Belgian adults. Am J Health Promot 2003;18(1):83-92.

20. Kamphuis CB, Giskes K, de Bruijn GJ, Wendel-Vos W, Brug J, van Lenthe FJ. Environmental determinants of fruit and vegetable consumption among adults: a systematic review. Br J Nutr 2006;96(4):620-35.

21. Kamphuis CB, van Lenthe FJ, Giskes K, Brug J, Mackenbach JP. Perceived environmental determinants of physical activity and fruit and vegetable consumption among high and low socioeconomic groups in the Netherlands. Health Place 2007;13(2):493-503.

22. Foster C, Hillsdon M, Thorogood M. Environmental perceptions and walking in English adults. J Epidemiol Community Health 2004;58(11):924-8.

23. Pikora TJ, Giles-Corti B, Knuiman MW, Bull FC, Jamrozik K, Donovan RJ. Neighborhood environmental factors correlated with walking near home: Using SPACES. Med Sci Sports Exerc 2006;38(4):708-14.

24. Hoehner CM, Brennan Ramirez LK, Elliott MB, Handy SL, Brownson RC. Perceived and objective environmental measures and physical activity among urban adults. Am J Prev Med 2005;28(2 Suppl 2):105-16.

25. Wendel-Vos W, Droomers M, Kremers S, Brug J, van Lenthe F. Potential environmental determinants of physical activity in adults: a systematic review. Obes Rev 2007;8(5):425-40.

26. Humpel N, Owen N, Iverson D, Leslie E, Bauman A. Perceived environment attributes, residential location, and walking for particular purposes. Am J Prev Med 2004;26(2):119-25.

27. Bauman A, Smith B, Stoker L, Bellew B, Booth M. Geographical influences upon physical activity participation: evidence of a 'coastal effect'. Aust N Z J Public Health 1999;23(3):322-4.

28. Brownson RC, Chang JJ, Eyler AA, Ainsworth BE, Kirtland KA, Saelens BE, Sallis JF. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. Am J Public Health 2004;94(3):473-83.

29. Boslaugh SE, Luke DA, Brownson RC, Naleid KS, Kreuter MW. Perceptions of neighborhood environment for physical activity: is it "who you are" or "where you live"? J Urban Health 2004;81(4):671-81.

30. Giles-Corti B, Donovan RJ. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. Prev Med 2002;35(6):601-11.

31. Rohm Young D, Voorhees CC. Personal, social, and environmental correlates of physical activity in urban African-American women. Am J Prev Med 2003;25(3 Suppl 1):38-44.

32. Voorhees CC, Rohm Young D. Personal, social, and physical environmental correlates of physical activity levels in urban Latinas. Am J Prev Med 2003;25(3 Suppl 1):61-8.

33. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 2006;117(2):417-24.

34. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. New York, NY: Lippincott Williams & Wilkins, 1998.

35. Bauman A. The physical environment and physical activity: moving from ecological associations to intervention evidence. J Epidemiol Community Health 2005;59(7):535-6.

36. Evenson KR, Herring AH, Huston SL. Evaluating Change in Physical Activity with the Building of a Multi-Use Trail. Am J Prev Med 2005;28(2S2):177-185.

37. Merom D, Bauman A, Vita P, Close G. An environmental intervention to promote walking and cycling--the impact of a newly constructed Rail Trail in Western Sydney. Preventive Medicine 2003;36(2):235-242.

38. Brownson R, Housemann R, Brown D, Jackson-Thompson J, King A, Malone B, Sallis J. Promoting physical activity in rural communities: walking trail access, use, and effects. American Journal of Preventive Medicine 2000;18(3):235-241.

39. Calfas KJ, Sallis JF, Nichols JF, Sarkin JA, Johnson MF, Caparosa S, Thompson S, Gehrman CA, Alcaraz JE. Project GRAD: two-year outcomes of a randomized controlled physical activity intervention among young adults. Graduate Ready for Activity Daily. Am J Prev Med 2000;18(1):28-37.

40. Effects of physical activity counseling in primary care: the Activity Counseling Trial: a randomized controlled trial. Jama 2001;286(6):677-87.

41. Marcus BH, Williams DM, Dubbert PM, Sallis JF, King AC, Yancey AK, Franklin BA, Buchner D, Daniels SR, Claytor RP. Physical activity intervention studies: what we know and what we need to know: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity); Council on Cardiovascular Disease in the Young; and the Interdisciplinary Working Group on Quality of Care and Outcomes Research. Circulation 2006;114(24):2739-52.

42. Abrams DB, Follick MJ. Behavioral weight-loss intervention at the worksite: feasibility and maintenance. J Consult Clin Psychol 1983;51(2):226-33.

43. Albright CL, Pruitt L, Castro C, Gonzalez A, Woo S, King AC. Modifying physical activity in a multiethnic sample of low-income women: One-year results from the IMPACT (increasing motivation for physical activity) project. Annals of Behavioral Medicine 2005;30(3):191-200.

44. Emmons KM, Stoddard AM, Fletcher R, Gutheil C, Suarez EG, Lobb R, Weeks J, Bigby JA. Cancer prevention among working class, multiethnic adults: results of the Healthy Directions-Health Centers Study. American Journal of Public Health 2005;95(7):1200-1205.

45. Collins R, Lee RE, Albright CL, King AC. Ready to be Physically Active? The Effects of a Course Preparing Low-Income Multiethnic Women to be more Physically Active. Health Educ Behav 2004;31(1):47-64.

46. Banks-Wallace J, Conn V. Interventions to promote physical activity among African American women. Public Health Nurs 2002;19(5):321-35.

47. The Guide to Community Preventive Services: What Works to Promote Health? Oxford University Press, 2005.

48. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. Am J Public Health 2002;92(11):1761-7.

49. Rose D, Richards R. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. Public Health Nutrition 2004;7(8):1081-1088.

50. Bodor JN, Rose D, Farley TA, Swalm C, Scott SK. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. Public Health Nutr 2007:1-8.

51. Baker EA, Schootman M, Barnidge E, Kelly C. The role of race and poverty in access to foods that enable individuals to adhere to dietary guidelines. Prev Chronic Dis 2006;3(3):A76.

52. Cheadle A, Psaty BM, Curry S, Wagner E, Diehr P, Koepsell T, Kristal A. Community-level comparisons between the grocery store environment and individual dietary practices. Prev Med 1991;20(2):250-61.

53. Morland K, Wing S, Diez Roux A, Poole C. Neighborhood characteristics associated with the location of food stores and food service places. Am J Prev Med 2002;22(1):23-9.

54. Glanz K, Sallis JF, Saelens BE, Frank LD. Nutrition Environment Measures Survey in stores (NEMS-S): development and evaluation. Am J Prev Med 2007;32(4):282-9.

55. Zenk S, Schulz A, Israel B, James S, Bao S, Wilson M. Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit. Am J Public Health 2005;95:660-667.

56. Berg N, Murdoch J. Access to grocery stores in Dallas. International Journal of Behavioural and Healthcare Research 2008;1(1):22-37.

57. Burns CM, Inglis AD. Measuring food access in Melbourne: Access to healthy and fast foods by car, bus and foot in an urban municipality in Melbourne. Health Place 2007;13(4):877-85.

58. Cummins S, Macintyre S. Food environments and obesity - neighbourhood or nation? International Journal of Epidemiology 2006;35:100-104.

59. Diez-Roux AV, Nieto FJ, Caulfield L, Tyroler HA, Watson RL, Szklo M. Neighbourhood differences in diet: the Atherosclerosis Risk in Communities (ARIC) Study. J Epidemiol Community Health 1999;53(1):55-63.

60. Larson N, Story M, Nelson M. Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S. American Journal of Preventive Medicine 2009;36(1):74-81.

61. Prevalence of fruit and vegetable consumption and physical activity by race/ethnicity--United States, 2005. MMWR Morb Mortal Wkly Rep 2007;56(13):301-4.

62. Anderson JV, Bybee DI, Brown RM, McLean DF, Garcia EM, Breer ML, Schillo BA. 5 a day fruit and vegetable intervention improves consumption in a low income population. J Am Diet Assoc 2001;101(2):195-202.

63. Powell L, Slater S, Mirtcheva D, Bao Y, Chaloupa F. Food store availability and neighborhood characteristics in the United States. Prev Med 2007;44:189-95.

64. Li F, Fisher J, Brownson RC. A multilevel analysis of change in neighborhood walking activity in older adults. J Aging Phys Act 2005;13(2):145-59.

65. Carcaise-Edinboro P, McClish D, Kracen AC, Bowen D, Fries E. Fruit and vegetable dietary behavior in response to a low-intensity dietary intervention: the rural physician cancer prevention project. Journal of Rural Health 2008;24(3):299-305.

66. Campbell MK, Demark-Wahnefried W, Symons M, Kalsbeek WD, Dodds J, Cowan A, Jackson B, Motsinger B, Hoben K, Lashley J, Demissie S, McClelland JW. Fruit and vegetable consumption and prevention of cancer: the Black Churches United for Better Health project. American Journal of Public Health 1999;89(9):1390-6.

67. Resnicow K, Jackson A, Wang T, De AK, McCarty F, Dudley WN, Baranowski T. A motivational interviewing intervention to increase fruit and vegetable intake through black churches: results of the Eat for Life Trial. American Journal of Public Health 2001;91(10):1686-1693.

68. Block G, Wakimoto P, Metz D, Fujii ML, Feldman N, Mandel R, Sutherland B. A randomized trial of the Little by Little CD-ROM: demonstrated effectiveness in increasing fruit and vegetable intake in a low-income population. Preventing Chronic Disease 2004;1(3):A08.

69. Dubowitz T, Heron M, Bird CE, Lurie N, Finch BK, Basurto-Davila R, Hale L, Escarce JJ. Neighborhood socioeconomic status and fruit and vegetable intake among whites, blacks, and Mexican Americans in the United States. Am J Clin Nutr 2008;87(6):1883-1891.

70. Gans KM, Burkholder GJ, Risica PM, Lasater TM. Baseline fat-related dietary behaviors of white, Hispanic, and black participants in a cholesterol screening and education project in New England. J Am Diet Assoc 2003;103(6):699-706; discussion 706.

71. Gary TL, Baptiste-Roberts K, Gregg EW, Williams DE, Beckles GL, Miller EJ, 3rd, Engelgau MM. Fruit, vegetable and fat intake in a population-based sample of African Americans. J Natl Med Assoc 2004;96(12):1599-605.

72. Drewnowski A. Obesity and the food environment: dietary energy density and diet costs. Am J Prev Med 2004;27(3 Suppl):154-62.

73. Drewnowski A, Darmon N. The economics of obesity: dietary energy density and energy cost. Am J Clin Nutr 2005;82(Suppl):265S-73S.

74. Drewnowski A, Darmon N, Briend A. Replacing fats and sweets with vegetables and fruits--a question of cost. Am J Public Health 2004;94(9):1555-9.

75. Agudo A, Amiano P, Barcos A, Barricarte A, Beguiristain JM, Chirlaque MD, Dorronsoro M, Gonzalez CA, Lasheras C, Martinez C, Navarro C, Pera G, Quiros JR, Rodriguez M, Tormo MJ. Dietary intake of vegetables and fruits among adults in five regions of Spain. European Journal of Clinical Nutrition 1999;53:174-180.

76. Brown DM, Thorne JE, Foster GL, Duncan JL, Brune LM, Munana A, Meinert CL, Jabs DA. Factors affecting attrition in a longitudinal study of patients with AIDS. AIDS Care 2006;18(7):821-9.

77. Cozzi Lepri A, Smith GD, Mocroft A, Sabin CA, Morris RW, Phillips AN. A practical approach to adjusting for attrition bias in HIV clinical trials with serial marker responses. Aids 1998;12(10):1155-61.

78. Davis L, Evans S, Fishman B, Haley A, Spielman LA. Predictors of attrition in HIV-positive subjects with peripheral neuropathic pain. AIDS Care 2004;16(3):395-402.

79. Esbensen FA, Miller MH, Taylor TJ, He N, Freng A. Differential attrition rates and active parental consent. Eval Rev 1999;23(3):316-35.

80. Letourneau N. Attrition among adolescents and infants involved in a parenting intervention. Child Care Health Dev 2001;27(2):183-6.

81. Condelli WS, Koch MA, Fletcher B. Treatment refusal/attrition among adults randomly assigned to programs at a drug treatment campus: The New Jersey Substance Abuse Treatment Campus, Seacaucus, NJ. J Subst Abuse Treat 2000;18(4):395-407.

82. Condelli WS. Domains of variables for understanding and improving retention in therapeutic communities. Int J Addict 1994;29(5):593-607.

83. Claus RE, Kindleberger LR, Dugan MC. Predictors of attrition in a longitudinal study of substance abusers. J Psychoactive Drugs 2002;34(1):69-74.

84. Gottheil E, Sterling RC, Weinstein SP. Pretreatment dropouts: characteristics and outcomes. J Addict Dis 1997;16(2):1-14.

85. Agosti V, Nunes E, Ocepeck-Welikson K. Patient factors related to early attrition from an outpatient cocaine research clinic. Am J Drug Alcohol Abuse 1996;22(1):29-39.

86. Alterman AI, Bedrick J, Howden D, Maany I. Reducing waiting time for substance abuse treatment does not reduce attrition. J Subst Abuse 1994;6(3):325-32.

87. Nishimoto RH. Who drops out of drug-user treatment research on women? Subst Use Misuse 1998;33(6):1291-313.

88. Roffman RA, Klepsch R, Wertz JS, Simpson EE, Stephens RS. Predictors of attrition from an outpatient marijuana-dependence counseling program. Addict Behav 1993;18(5):553-66.

89. Janson SL, Alioto ME, Boushey HA. Attrition and retention of ethnically diverse subjects in a multicenter randomized controlled research trial. Control Clin Trials 2001;22(6 Suppl):236S-43S.

90. McCann TJ, Criqui MH, Kashani IA, Sallis JF, Calfas KJ, Langer RD, Rupp JW. A randomized trial of cardiovascular risk factor reduction: patterns of attrition after randomization and during follow-up. J Cardiovasc Risk 1997;4(1):41-6.

91. Snow WM, Connett JE, Sharma S, Murray RP. Predictors of attendance and dropout at the Lung Health Study 11-year follow-up. Contemp Clin Trials 2007;28(1):25-32.

92. Dennis BP, Neese JB. Recruitment and retention of African American elders into community-based research: lessons learned. Arch Psychiatr Nurs 2000;14(1):3-11.

93. Grilo CM, Money R, Barlow DH, Goddard AW, Gorman JM, Hofmann SG, Papp LA, Shear MK, Woods SW. Pretreatment patient factors predicting attrition from a multicenter randomized controlled treatment study for panic disorder. Compr Psychiatry 1998;39(6):323-32.

94. Nitzke S, Kritsch K, Boeckner L, Greene G, Hoerr S, Horacek T, Kattelmann K, Lohse B, Oakland MJ, Beatrice P, White A. A stage-tailored multi-modal intervention increases fruit and vegetable intakes of low-income young adults. Am J Health Promot 2007;22(1):6-14.

95. Cahill J, Barkham M, Hardy G, Rees A, Shapiro DA, Stiles WB, Macaskill N. Outcomes of patients completing and not completing cognitive therapy for depression. Br J Clin Psychol 2003;42(Pt 2):133-43.

96. Shaw WS, Cronan TA, Christie MD. Predictors of attrition in health intervention research among older subjects with osteoarthritis. Health Psychol 1994;13(5):421-31.

97. Meyers K, Webb A, Frantz J, Randall M. What does it take to retain substance-abusing adolescents in research protocols? Delineation of effort required, strategies undertaken, costs incurred, and 6-month post-treatment differences by retention difficulty. Drug Alcohol Depend 2003;69(1):73-85.

98. Morrison TC, Wahlgren DR, Hovell MF, Zakarian J, Burkham-Kreitner S, Hofstetter CR, Slymen DJ, Keating K, Russos S, Jones JA. Tracking and follow-up of 16,915 adolescents: minimizing attrition bias. Control Clin Trials 1997;18(5):383-96.

99. Cooley ME, Sarna L, Brown JK, Williams RD, Chernecky C, Padilla G, Danao LL. Challenges of recruitment and retention in multisite clinical research. Cancer Nurs 2003;26(5):376-84; quiz 385-6.

100. Given CW, Given BA, Coyle BW. Prediction of patient attrition from experimental behavioral interventions. Nurs Res 1985;34(5):293-8.

101. Schmidt JA, Gruman C, King MB, Wolfson LI. Attrition in an exercise intervention: a comparison of early and later dropouts. J Am Geriatr Soc 2000;48(8):952-60.

102. Eakin EG, Lawler SP, Vandelanotte C, Owen N. Telephone interventions for physical activity and dietary behavior change: a systematic review. Am J Prev Med 2007;32(5):419-34.

103. Grossi E, Dalle Grave R, Mannucci E, Molinari E, Compare A, Cuzzolaro M, Marchesini G. Complexity of attrition in the treatment of obesity: clues from a structured telephone interview. Int J Obes (Lond) 2006;30(7):1132-7.

104. ACS Nutrition and Physical Activity (NuPA) Study: Progress Report and Discussion. Atlanta, GA: American Cancer Society, 2007;49.

105. Apparicio P, Cloutier MS, Shearmur R. The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. Int J Health Geogr 2007;6:4.

Chapter 3: Methods

Study Population

The NuPA Study

Participant survey data were obtained from the American Cancer Society's Nutrition and Physical Activity (NuPA) Study. This study was initiated at the American Cancer Society (ACS) in response to a growing demand for information, guidance, and support in attaining healthy lifestyle behaviors and was modeled after the ACS's Smoking Cessation Quitline. Conducted from the National Cancer Information Center (NCIC) located in Austin, Texas, counselors administered the intervention and separate evaluators administered follow-up surveys over the phone.

Recruitment

Enrollment occurred from September 2005 through June 2007. Most participants were informed of the study through their employers although participants were also recruited by email newsletters, websites, media outlets (TV, newspaper, magazines, and radio) and friends or family. Upon hearing about the intervention, interested individuals called the NCIC and underwent pre-enrollment screening. Non-institutionalized callers who had no diagnosed eating disorder, were not pregnant if female, were over the age of 18, and who were willing to make a change within the next 2 weeks were eligible for enrollment. All participants were asked to complete the baseline survey (**Appendix 1**: Health for Life Intake Questionnaire (March 2005)) which assessed demographic information and included measures of health knowledge and beliefs, self-reported current weight and height, consumption of various foods (fruits, vegetables, milk, meat, saturated

fat, etc.), physical activity, and pre-existing co-morbid conditions (diabetes, hypertension, hypercholesterolemia, and previous cancer).

Randomization

After completing the survey, participants were randomized to either the control arm or the intervention arm using a computer algorithm. Both groups receive educational materials in the mail which addressed appropriate amounts of physical activity, provided specific ideas to improve diet, and discussed issues of goal-setting (**Appendix 2**: Get on Track, Stay on Track Outline (March 2005)). Those randomized to the counseling group also received 6 counseling sessions during a period of 4 months delivered over the telephone by trained counselors.[1] These counseling sessions were designed to teach stress management techniques, exercises to improve self-awareness and goal-setting and raise awareness of other resources or specific methods for identifying, setting, and achieving behavioral goals and lasted up to 30 minutes (**Appendix 3**: Health for Life Counseling Outline (March 2005)). Intervention participants also received up to 3 booster sessions, lasting 10 minutes each, during months 5 and 6.

A total of 2,550 individual responded to the baseline survey and were randomized, of whom 51.5% (N= 1,313) were assigned to the intervention, compared to 1,237 assigned to control. This is not too unexpected given a binomial random variable with 2,550 trials and probability of randomization to intervention equal to 0.5 (2-tailed, p=0.13).

Follow-up Contact

Evaluators at the NCIC then attempted to contact each individual at 3, 6, and 12 months after enrollment. The surveys used at these follow-up points (**Appendix 4**: Health

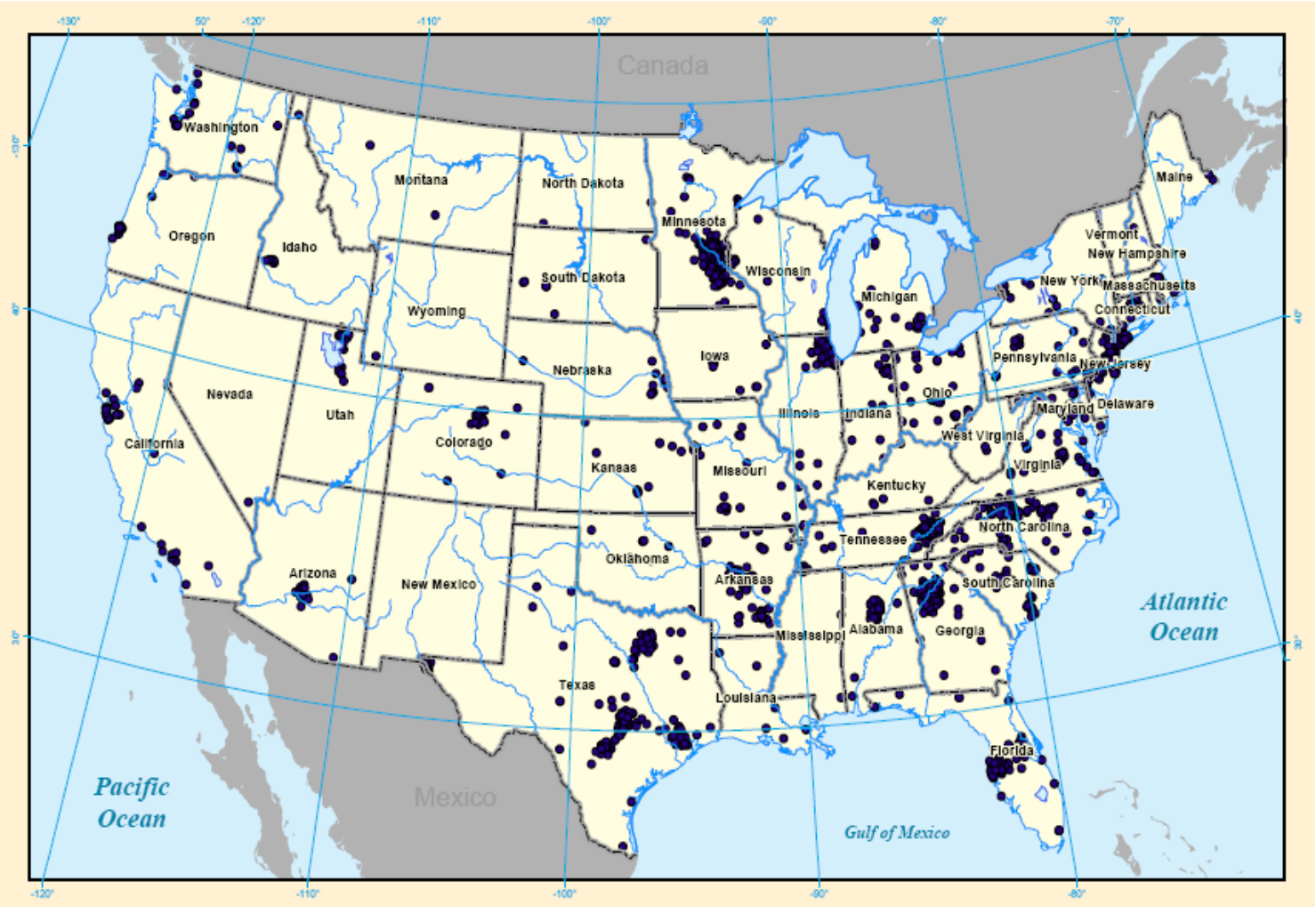


Figure 3.1. Map of participants in the Nutrition and Physical Activity Study (N=2,550). Each participant represented with (•). Residents of Alaska and Hawaii not shown.

for Life Evaluation 3, 6, and 12 months (March 2005)) were identical to one another but different from the one used at baseline. Evaluators were distinct from counselors involved in delivering the intervention to those in the intervention arm.

At each follow-up time, evaluators attempted to contact each participant, making up to 15 calls over a 4 week period while varying the time of day and day of the week during which calls were made. Since NuPA counseling arm participants may defer their intervention initiation by a period of up to 30 days after the initial contact with the NCIC, all participants were placed on the call list for the 3 month follow-up evaluation 120 days after their initial contact with NCIC. Six-month evaluation calls started at 210 days (7 months), and 12-month calls start at 390 days (13 months) after initial contact.

Demographics

Participants were recruited from 49 states (all states but New Mexico), the District of Columbia and Puerto Rico (**Figure 3.1**). Demographic characteristics were similar across study arms (**Table 3.1**). Typical of a volunteer population, participants were more likely to be female (80.6% vs. 50.8%) in comparison to the general US population and also more highly educated (55.6% with a bachelors degree or more vs. 27.0%).[2] However, the distribution of race/ethnicity was fairly representative of the general US population where 74.1% are white (72.4% in NuPA), and 12.4% are black or African-American (14.0% in NuPA).[2] Unemployment (5.9%) was approximately equal to that reported for the 2005-2007 period which ranged between 4.4 and 5.6%.[3]

Many of the participants were of normal/underweight (BMI in kg/m^2 below 25, N=511, 20.2%) with two participants who were underweight (BMI under 18.5 kg/m^2)

Table 3.1. Demographic Characteristics of All Participants in the Nutrition and Physical Activity (NuPA) study.

Variable	Counseling		Self-Help		p
Total N=	**1313**		**1237**		
Age					
mean (sd)	43.53	10.46	43.66	10.80	0.77
missing	*49*		*42*		
Sex					0.10
Male	238	18.2%	256	20.7%	
Female	1073	81.8%	981	79.3%	
missing	2		0		
Employment Status					0.30
Employed	1240	94.6%	1158	93.6%	
Unemployed	71	5.4%	79	6.4%	
missing	2		0		
Marital Status					0.26
Married	807	61.5%	786	63.6%	
Not Married	506	38.5%	449	36.4%	
missing	0		2		
Education					0.79
High school of less	188	14.3%	168	13.6%	
Some college or technical school	404	30.8%	372	30.1%	
College graduate	493	37.5%	488	39.5%	
Post-College	228	17.4%	208	16.8%	
missing	0		1		
Ethnicity					0.61
White	935	71.8%	903	73.1%	
Black	190	14.6%	165	13.4%	
Hispanic	118	9.1%	119	9.6%	
Other	60	4.6%	48	3.9%	
missing	10		2		
Medical Conditions (Dx by doctor)					
High Blood Pressure*	308	23.5%	261	21.1%	0.27
Cholesterol*	373	28.4%	321	25.9%	0.24
Cancer*	102	7.8%	103	8.3%	0.66
Diabetes2					0.35
Type I	7	0.5%	13	1.1%	
Type II	75	5.7%	79	6.4%	

Satisfied with Life					0.96
Yes	950	72.4%	894	72.3%	
No	363	27.6%	343	27.7%	
Sad/Blue					0.71
Yes	101	7.7%	100	8.1%	
No	1212	92.3%	1137	91.9%	
BMI Category					0.63
Normal/Underweight (<25 kg/m^2)	253	19.5%	258	21.0%	
Overweight (25, 30 kg/m^2)	448	34.5%	418	34.0%	
Obese (>30 kg/m^2)	598	46.0%	552	45.0%	
missing	*14*		*9*		
Physical Activity					0.02
0 to 29 minutes/day	735	56.1%	633	51.3%	
30 or more minutes/day	576	43.9%	601	48.7%	
missing	*2*		*3*		
Fruit and Vegetable Consumption					
mean (sd)	3.44	2.00	3.56	2.04	0.14
missing	*21*		*19*		

*** Fisher's exact test**

based on self-reported height and weight at baseline. The 79.8% of participants who were overweight or obese at baseline is slightly higher than the national average at 66.0% reported for 2001-2004[4]. Consistent with national figures reporting a 10.2% prevalence of diabetes, 0.8% of NuPA participants reported having type I diabetes while 6.0% reported type II diabetes. About a quarter of participants (N=569, 22.3%) had been diagnosed with hypercholesterolemia compared with the national figure of 16.5%, and 27.2% (N=694) of participants reported hypertension which is approximately equal to the national age-adjusted prevalence of 26.7%.[4]

Baseline Physical Activity and Fruit and Vegetable Consumption Behaviors

Physical Activity. PA was self-reported in total minutes/day using the past three months as the reference period. At baseline, a higher percentage of NuPA participants (**Table 3.1**) reported more than 30 minutes/day of physical activity (enrollment survey, item 17, **Appendix 1**, N=1177, 46.2%) then the general US population. Unreasonable values of exceeding 720 minutes/day (12 hours of PA/day) were excluded. Among those randomized to receive the intervention, a statistically significant 4.8% less reported exceeding this threshold than those randomized to receive control.

Fruit and Vegetable Consumption. Fruit and vegetable consumption ranged between 0 and 14 servings/day in the NuPA population and appeared to be approximately normally distributed (**Figure 3.2a**). Log-transformation did not improve the shape of the curve (**Figure 3.2a**). There was no significant difference in average servings of fruits and vegetables consumed

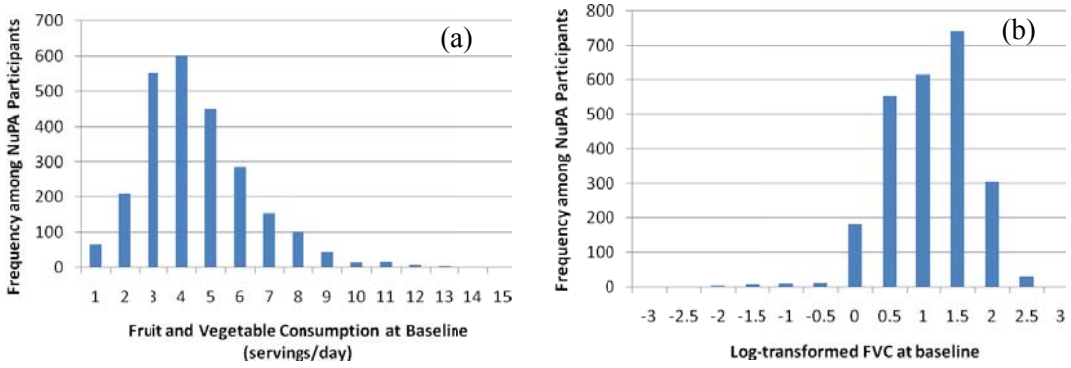


Figure 3.2. Frequency distribution of **(a)** fruit and vegetable consumption (FVC) and **(b)** log-transformed FVC from baseline survey of 2,550 eligible participants in the Nutrition and Physical Activity (NuPA) Study.

across study arm (**Table 3.1**), although consumption was above the national average at 3.49 servings/day with standard deviation (SD) 1.92. Unreasonable values of more than 20 servings/day were excluded.

Exclusion to 5 Metropolitan Areas

Although participants were recruited from throughout the United States, there were several regions where large portions of participants resided. Among the 2,550 participants randomized 43.2% (N=1104) of participants resided in 5 metropolitan statistical areas (MSAs) with more than 100 participants each. These MSAs, herein referred to as cities, were: Minneapolis/St. Paul, MN; Atlanta, GA; Austin, TX; Chicago, IL; and Knoxville, TN (**Table 3.2**).[5]

Given the geographic distribution of participants, the amount of data extraction and calculation involved in obtaining the exposure measures, and the type of statistical modeling that would be required to obtain valid estimates, only participants residing in these five cities were included for the current studies. By restricting to individuals

Table 3.2. Demographic Characteristics of Participants in the Nutrition and Physical Activity (NuPA) Study Residing in Atlanta, GA; Austin, TX; Chicago, IL; Knoxville, TN; and Minneapolis/St. Paul, MN.

	Counseling		Self-Help		p
Total N=	**558**		**546**		
Variable	**N**	**%**	**N**	**%**	
Age *mean(sd)*[1]	42.7	(10.4)	42.5	(10.9)	0.79
missing	*1*		*7*		
Sex					0.63
Male	104	18.6%	108	19.8%	
Female	454	81.4%	438	80.2%	
Metropolitan Area					0.83
Atlanta	100	17.9%	86	15.8%	
Austin	56	10.0%	59	10.8%	
Chicago	54	9.7%	51	9.3%	
Knoxville	47	8.4%	53	9.7%	
Minneapolis	301	53.9%	297	54.4%	
Employment Status					0.05
Yes	538	96.6%	514	94.1%	
No	19	3.4%	32	5.9%	
missing	*1*		*0*		
Marital Status					0.88
Yes	335	60.0%	329	60.5%	
No	223	40.0%	215	39.5%	
missing	*0*		*2*		
Education					0.31
HS or less	71	12.7%	55	10.1%	
Some college/Tech	160	28.7%	178	32.6%	
College grad	222	39.8%	220	40.3%	
Post college	105	18.8%	93	17.0%	
Ethnicity					0.70
White	402	72.7%	412	75.6%	
Black	95	17.2%	87	16.0%	
Hispanic	32	5.8%	27	5.0%	
Other	24	4.3%	19	3.5%	
Missing	*5*		*1*		
Medical History[2]					
Hypertension	115	20.6%	95	17.4%	0.17
Cancer	43	7.7%	36	6.6%	0.47
Hypercholesterolemia	160	28.7%	124	22.7%	0.02

Diabetes[3]					0.20
Type I	1	0.2%	4	0.7%	
Type II	24	4.3%	32	5.9%	
Satisfaction with Life					0.10
Yes	401	71.9%	416	76.2%	
No	157	28.1%	130	23.8%	
Feeling Sad/Blue					0.29
Yes	45	8.1%	35	6.4%	
No	513	91.9%	511	93.6%	
BMI Category					0.91
Normal/Underweight (<25 kg/m^2)	121	22.0%	121	22.4%	
Overweight (25, 30 kg/m^2)	185	33.6%	187	34.6%	
Obese (>30 kg/m^2)	224	40.7%	233	43.1%	
missing	*8*		*5*		
Physical Activity					0.00
0 to 29 minutes/day	319	57.4%	267	48.9%	
30 or more minutes/day	237	42.6%	279	51.1%	
missing	*2*		*0*		
Fruit and Vegetable Consumption					
mean (sd)	3.63	2.03	3.71	2.19	0.56
missing	*9*		*8*		

P-values reported for $\chi 2$ test except where indicated.

1. 2-sample t-test, pooled variance
2. Self-reported physician diagnosis or treatment of condition
3. Fisher exact test

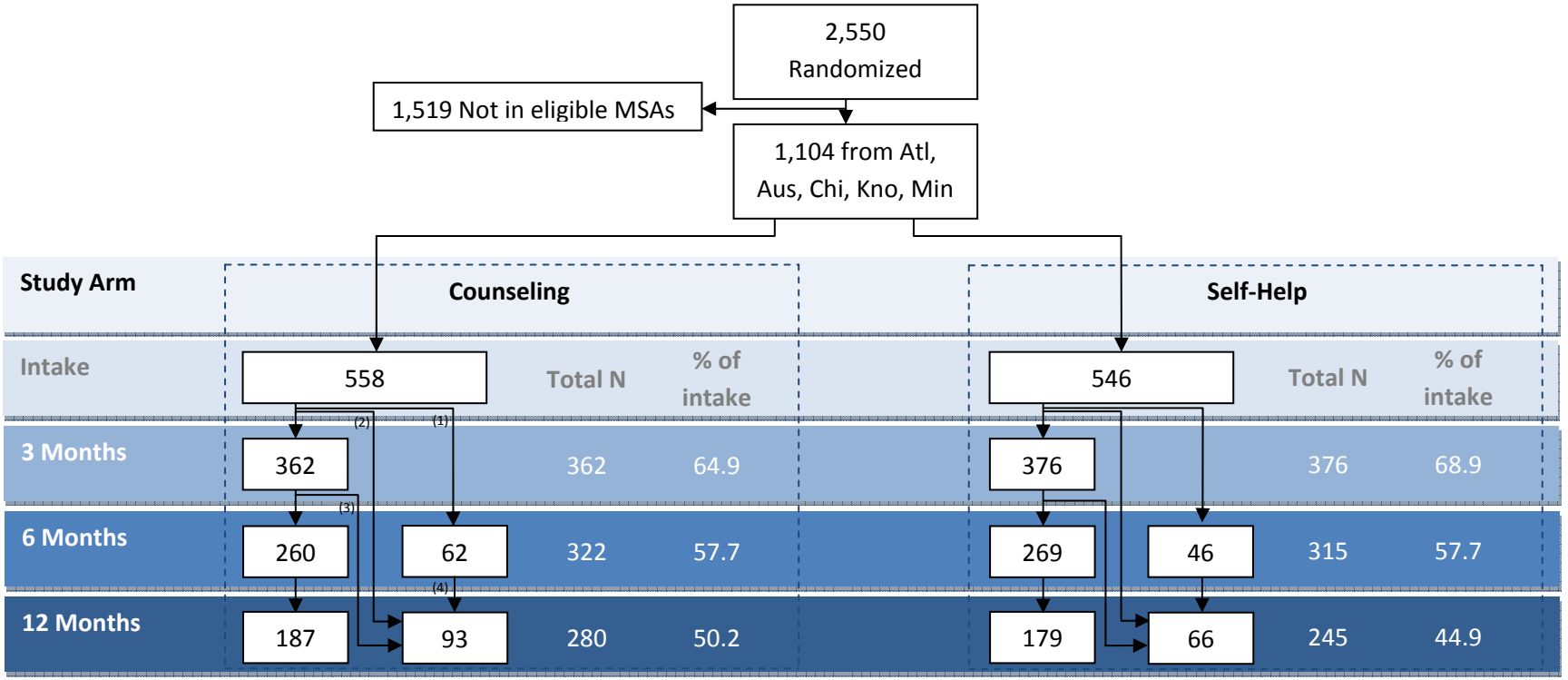


Figure 3.3. Flow-chart of participant follow-up among 1,104 individuals residing in Atlanta, Georgia (Atl); Austin, Texas (Aus); Chicago, Illinois (Chi); Knoxville, Tennessee (Kno); and Minneapolis/St. Paul, Minnesota (Min).

Notes:

- All numbers indicated at each follow-up are those who were contacted (record reported from SIEBEL) and consented to the follow-up (responded: "Yes" or "Yes, but call back later").
- Arrows with numbers among those randomized to receive Counseling indicate patterns of skipped follow-ups and are similarly indicated, without numbers, among those randomized to Self-Help.
  (1) Participants who were not contacted or did not consent to the 3-month follow-up but were contacted and consented to 6-month follow-up (N=62 for Counseling, 46 for Self-Help, as indicated)
  (2) Participants who were not contacted or did not consent to the 3- or 6-month follow-ups but were contacted and consented to the 12-month follow-up (N=23 for Counseling, 10 for Self-Help)
  (3) Participants who consented to intake, 3-, and 12- month follow-up, but not 6-month (N=35 for Counseling, 26 for Self-Help)
  (4) Participants who consented to intake, 6-, and 12-month but not 3-month (N=35 for Counseling, 30 for Self-Help).

enrolled from the five MSAs listed above, more focus could be placed on exposure ascertainment resulting in more precise and accurate assessment of exposure. From here, participants in the current studies will be referred to merely as participants while those from the larger study will be referred to as NuPA participants.

The distribution of demographic characteristics within the subgroup included in subsequent analysis did not substantially differ in comparison to the overall NuPA population (**Table 3.2**). At 3, 6, and 12 months respectively 66.8% (362 intervention, 376 comparison), 57.7% (322 intervention, 315 comparison), and 47.6% (280 intervention, 245 comparison) remained in the study (**Figure 3.3**).

The distribution and range of baseline behaviors across study arms did not differ in comparison to the overall NuPA population with baseline fruit and vegetable consumption ranging between 0 and 14 servings/day with mean 3.62 and standard deviation 1.91 (**Figure 3.4a** and **3.4b**) and a significantly greater portion of participants in the control arm meeting the physical activity recommendation at baseline (**Table 3.2**).

Institutional Review Board Review

This study was approved by the Emory IRB on November 22, 2007 (study IRB00005116). The NuPA study (IRB 235-2005) was approved by the Emory IRB on April 13, 2005.

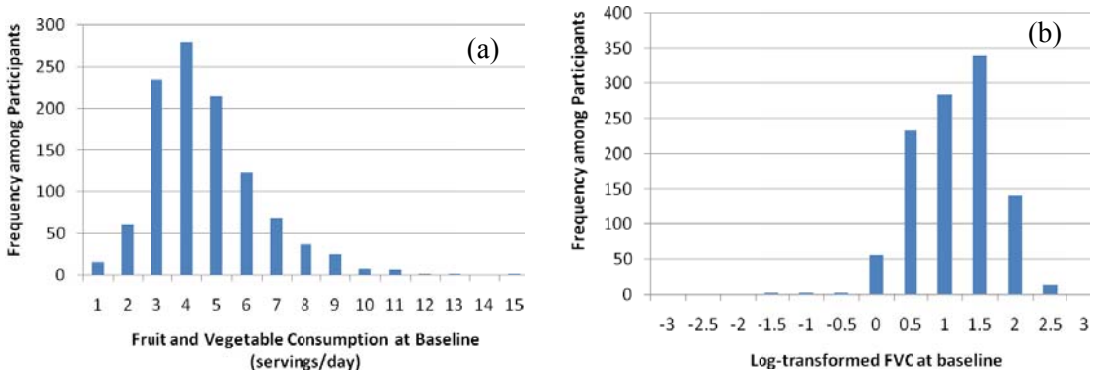


Figure 3.4. Frequency distribution of **(a)** fruit and vegetable consumption (FVC) and **(b)** log-transformed FVC from baseline survey of 1,104 eligible participants in the Nutrition and Physical Activity (NuPA) Study residing in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota.

Assessment of the Environment

All spatial calculations were performed using ArcView 9.2 (ESRI, Inc., Redlands, California), and were dichotomized, treated in quartiles, or, where quartiles were not possible, treated in tertiles. Quartiles are referred to as highest, high, low, and lowest, and tertiles are referred to as high, medium, and low.

Geocoding Procedure and Validation. Participants residences were geocoded to a block group using the Qualitative Marketing Software (QMS) Geocoder (QMSoft, Clearwater, Florida) based on mailing address reported for delivery of educational materials. A random sample of 10% of all addresses linked at the precision level of the residence were examined for accurate pairing with the block group code using the address search function on the United States (U.S.) Census website and data from the 2000 Census.[6] Addresses that could not be found in this manner were assumed to be new and were located using Google Earth and the U.S. Census Block Groups shapefile (2007, Environmental Systems Research Institute, Inc. (ESRI), Redlands, California).[7] All addresses that were approximated at the block group or zip code level were examined using the same procedure. Where codes obtained from different data sources disagreed,

preference was given to the code obtained from, in order, the census address tool, the block group shapefile, and the QMS-approximated block group.

The 1,104 participants included in the analyses were geocoded to 905 separate block groups with the majority of participants (N=947, 85.8%) geocoded during the first round. All addresses sampled from this group for further verification produced codes that were consistent between the Census address tool and the QMS Geocoder. The majority of participants whose addresses were approximated at the block group or zip code level during the first round (N=157), were linked to a block group using one of the other 2 methods (29 using the Census, 96 using the shapefile). Approximated block groups were used for 28 participants who reported a post office box as their mailing address.

The unweighted geographic centroid of the block group was used as the focal point for assessing environmental factors. Although the use of the BG centroid rather than a full address introduces misclassification into the assessment of exposure, this is required to maintain confidentiality. Furthermore, little or no bias should result because the degree of misclassification is limited by the size of the BGs and the variability of the exposure across space. The size of block groups are relatively small in metropolitan areas and ranged from 0.02 to 61.0 mi^2 for BGs identified to have NuPA participants.

By design, individuals who were assigned the same block group will share focal points from which exposures will be measured and therefore have identical values for the exposures. The shared nature of these exposures will be accounted for in a multi-level model with observations nested in three levels: individual time points nested within individuals, individuals nested within a block group identified with the residence, and block groups further nested within a city.

Geographic Exposure Assessment of Exposure

Fruit and vegetable environment. The Dun & Bradstreet (D&B) Million Dollar Database provides data on businesses meeting one or more of the following criteria: (1) conducted $1 million or more in sales, (2) have 20 or more employees, or (3) have branches with 50 or more employees.[8] Smaller businesses are excluded from this database. Although this will exclude smaller businesses, these exclusions are consistent with the observational literature suggesting that it is the larger chains and businesses that contribute to greater consumer access to fruits and vegetables.[9-11] Businesses with primary North American Industry Classification System (NAICS) codes of 44523 (fruit and vegetable specialty markets) or 44511 (supermarkets and other grocery stores, except convenience stores) were included in the analysis. A total of 1,537 grocers were included in this database in the 5 cities examined. Although the majority of these grocers were chain supermarkets, identified as either branch or headquarter locations for the above-mentioned NAICS coded businesses (N=1,024, 66.6%), a substantial portion were single-location businesses (N=513, 33.4%). Thus, these businesses are referred to as grocers rather than supermarkets.

The Dun and Bradstreet Million Dollar Database was used in this study because it is continuously updated and has nation-wide coverage. This insured greater comparability of data quality between the 5 different cities studied and is in contrast to feasible alternatives which would have required combining data obtained from several different sources and across different time periods.

Each grocer was mapped using the longitude and latitude from the D&B database which provides the business location with precision within 11 meters (4 decimal places). Combining that information with block group centroids, 3 measures of grocer access were adapted from the work of Apparicio et al: (1) proximity: the distance from the focal point to the nearest grocer; (2) diversity: the number of grocers within a buffer distance from the focal point; and (3) variety: the average distance from the focal point to the 3 nearest differently named grocers.[12] Buffer regions

with 8.05 km (5-mile) and 0.805-km (0.5-mile) radii will be used, assuming non-motorized and motorized transport, respectively.[13-17]

Because studies comparing commercially available business listings to other data sources highlight the potential for misclassification of businesses across different sources and suggest closer examination to determine which sources are most appropriate for the purposes of research, an additional validation study was designed to assess the impact of misclassification by other individual- and block group-level predictors of participant behavior (**Chapter 5, Section 2:** Additional Analyses).[18,19]

Physical activity environment. The physical activity environment was measured as street structure and access to parks. Street structure was operationalized using two separate components – 3- and 4-way intersection density and total street length contained within a buffer region surrounding the participant focal point. Both were assessed using the buffer regions defined above. The participant's surrounding intersection density and total street length were calculated using data from Environmental Systems Research Institute (ESRI) StreetMap data, which are based on TIGER/Line Files (2000) maintained by the US Census Bureau. These data are used to support conducting the decennial census and are updated annually by commercial firms such as NavTeq and Tele Atlas. Parks data were obtained from the ESRI U.S. Geographic Data Technology (GDT) Park Landmarks shapefile. This dataset provides the name, location, boundaries, square mileage and general usage information for each site.

Statistical Analysis

All analyses were performed using SAS 9.1 and 9.2.

Fruit and Vegetable Consumption and Physical Activity Studies

Fruit and Vegetable Consumption. The continuous measure of servings of fruits and vegetable per day was used as the dependent variable for analysis. As such, a linear mixed model using restricted maximum likelihood was chosen for regression using SAS PROC MIXED.

Conceptually relevant random effects included individual (for repeated measures at baseline, 3-, 6-, and 12 months), block group, and city. Analysis of intra-class correlations indicated that 0.3% of total variation was explained by city, 21.8% by block group, and 11.3% by individual with 66.6% of total variation unaccounted for. Although not all sources of variation reached pre-established criteria for inclusion in a mixed effect model, they remained in the final model due to their conceptual importance.

Physical Activity. Physical activity was analyzed as a dichotomous outcome. Although it was measured as a continuous outcome (total minutes/day), there were some indications in the data that there may have been some coding errors. Specifically, during the start of the enrollment period in 2005, many small values such as 2, 3, and 4 were reported for minutes/day of physical activity. Because such small and specific values were unlikely to reflect actual minutes of physical activity engaged in per day and because instances of such small values decreased after additional evaluator training, it was hypothesized that these values reflected hours instead of minutes. As such, the categorized outcome was chosen to minimize the impact of these errors and facilitate the assessment of misclassification on final analyses. Both logistic regression and binomial regression with random effects were used for regression analysis using the SAS GLIMMIX Macro.[20] The GLIMMIX macro was used rather than PROC GLIMMIX because it provided increased flexibility in specifying multiple random effects and correlation structures. PROC GENMOD was not used as multiple random effects could not be specified.

In analysis of intra-class correlations, city explained 2.1% of the variation in the dichotomous physical activity outcome, block group explained 41.0% and individual 12.3% of variation leaving 44.6% unexplained. Again, each of these levels remained in the final model as random effects because of their conceptual importance, not based on an external criterion.

Regression model for FVC and PA outcomes. Although specific estimation techniques and statistical assumptions for the FVC and PA analyses differ, the general

approach for creating the regression models were the same. The following model served as the basis for statistical analysis:

$$Outcome(t, I, E, C) = f(\beta_0 + \beta_1 \times t) \quad [3.1]$$

$$\beta_0 = (\gamma_{00} + b_{0k} + b_{00}) + \gamma_{01} \times E + \gamma_{02} \times C_{indiv} + \gamma_{03} \times C_{BG} \quad [3.2]$$

$$\beta_1 = \gamma_{10} + \gamma_{11} \times E + \gamma_{12} \times I + \gamma_{13} \times E \times I + \gamma_{14} \times C_{indiv} + \gamma_{15} \times C_{BG} \quad [3.3]$$

where, t indicates the follow-up timepoint, C_{indiv} is a vector containing all individual-level covariates and C_{BG} is a vector containing all block group-level covariates. E indicates the environment exposure of interest, and I represents the intervention arm.

This model assumes behavior, either the continuous outcome for FVC or the logit of the dichotomous outcome for PA, to be a function of two components (**equation 3.1**): a baseline behavior, β_0; and a change from baseline associated with each follow-up time point, β_1. The baseline behavior (**equation 3.2)**, β_0, is assumed to be explained by individual- and block group-level covariates (C_{indiv}, and C_{BG}) as well as with the environmental exposures of interest (E), essentially replicating analysis from cross-sectional studies (reviewed in Chapter 2). Furthermore, as evidenced by the omission of a term associated with I in equation 3.2, baseline behavior is assumed to be independent of the individual's intervention status. Change in behavior, given by β_1 (**equation 3.3**), is a function of the environment, intervention arm, an interaction between environment and intervention, and same individual- and BG-level covariates in equation 3.2. The effects of interest are given in equation 3.3 as γ_{11} and γ_{13}, the terms for the effect of the two-way interaction between time and the environmental exposure and the three-way interaction between time, environment, and exposure.

Random effects for BG (b_{00}) and for MSA (b_{0k}) were also included in the analysis to allow for correlation between individual sharing those variables. Furthermore, observations within a given individual were assumed to be correlated following an auto-regressive order-1 structure. Additional analysis was performed using a compound symmetric correlation structure to address the potential for bias in conclusions drawn from the model due to mis-specification of the correlation structure. Empirical estimation of the correlation was also attempted.

Power for FVC and PA outcomes. Using a simulation technique (**Appendix 5:** Power Simulation), power was shown to be dependent only on the proportion that is exposed to both the environmental exposure and the intervention, and the effective sample size. A range for the effective sample size (ESS) was estimated from equation 3.4, assuming balanced distribution of observations within clusters given a range of 0.02 to 0.60 for the intraclass correlation (ICC), and 57.7% losses by 6 months in both arms (**Figure 3.3**)[21]:

$$ESS = \frac{mk}{1 + ICC \times (m - 1)} \qquad [3.4]$$

where m is the number of clusters, and k is the number of observations within each cluster. A range of actual sample size (here m*k) was obtained and taken to be approximately 320 for each arm. Taking m to be approximately 1104/905 =1.22 (the number of participants / the number of block groups), the effective sample size was estimated to be between 283 and 318 for each arm.[22]

Assuming that exposure to the environmental factor of interest is independent of randomization and keeping power constant at 80%, the detectable effect size of the interaction between environment and intervention arm is dependent only on the proportion that is exposed to both the environmental exposure and the intervention, and the effective sample size. In fact, it is independent of the effect sizes of the first-order terms for environment and intervention. Assuming a dichotomous measure of environment (proportion expose=0.50), the smallest detectable effect size of the interaction parameters as γ_{11} and γ_{13} is estimated to be 0.4 srv/day for fruit and vegetable consumption. For the dichotomous outcome for physical activity, the smallest detectable odds ratio is approximately 1.94 or, assuming 60% meeting the PA recommendation in the reference category, 74.4% meeting the PA recommendation in the exposed group.

Model Assessment for FVC and PA Outcomes. Each exposure was first examined using an unadjusted model with random effects for both BGs and MSA and no covariates in equation 3.2 and 3.3. Those that were thus statistically significant in γ_{11} or γ_{13} of equation 3.3 were then examined adjusting for individual-level covariates and then both individual- and BG-level covariates.

Individual-level covariate data were obtained from the NuPA baseline survey. Identified *a priori* based on the literature, potential confounders of associations between environment both FVC and PA behaviors included age, sex, race/ethnicity, employment, marital status, educational attainment, hypercholesterolemia, diabetes types 1 or 2, hypertension, previous diagnosis of cancer, being sad/blue, and satisfaction with life.[23]

Environmental-level covariates were assessed on the block group level and included educational attainment (% of block group residents with a bachelors degree or

more, % with a high school degree or equivalent), neighborhood median income, racial composition (% minority including black, Asian, Hispanic, other, and mixed), median age, and unemployment rate (% unemployed). These data were obtained from the Claritas Demographic Update and were categorized into quartiles.

To determine which covariates should be included in the final model as potential confounders of behavior change, all individual-level covariates were entered into equation 3.2. Using the fully adjusted equation 3.2 and introducing the intervention term into equation 3.3, each individual-level covariate was entered separately into equation 3.3 and assessed for statistical significance at the α 0.05 level since a confounder must also be an individual cause or correlate of the outcome. All those that were statistically significant in equation 3.3 were added into the model together and assessed for independent contribution to β_1. Together these individual-level covariates were introduced into the model and those that were not statistically significant after adjusting for other significant covariates were removed from the model. This was then used as the final individual-level model.

A different approach was used to examine BG-level covariates as independent confounders of environment and behavior change since BG-level covariate may act as proxy measures of an individual-level analog. All individual-level covariates were entered into equation 3.2. Using this model, each BG-level covariate was separately entered into equation 3.2 to determine whether it contributed to the baseline outcome. Those that were statistically significant after adjusting for individual-level covariates were entered into equation 3.3 for further assessment. Finally, all BG-level covariates that were statistically significantly associated with behavior, whether in equation 3.2 or in

both equation 3.2 and 3.3, were entered into a model with all other significant BG-level covariates and all individual-level covariates. Those that maintained an independent effect were kept in the model, and those that did not were removed.

Additional Analysis for FVC and PA Outcomes. Adjusted analyses were repeated excluding data from participants whose block group codes were approximated at the block group or zip code level and participants who reported a PO Box mailing address to determine the effect potential geocoding errors on the final analysis outcome.

In addition, imputation analyses were undertaken to determine the effect of loss to follow-up on findings. Among the 1,104 participants in the analysis, 18 were missing on demographic factors, which were not imputed. As such, only participants with complete baseline data (N=1,086) were analyzed in imputation analysis, requiring imputation of missing PA and FVC behavior variables only. As participants could skip follow-ups and return for evaluation at later time-points, the pattern of missingness was not monotonic. The data were made to be monotonic by omitting subsequent observations recorded from participants who had skipped a previous evaluation.

From there, baseline values for PA (met recommendation: Yes/No) and FVC (servings/day) were imputed based on all individual-level predictors and block group-level characteristics that were associated with the outcome (as described above) or with loss to follow-up (as described below). PA and FVC for subsequent follow-up time-points (3-, 6-, and 12-month) were imputed based on all previous measures of the behavior in addition to baseline demographics using the regression method for monotonic data. For example, 6-month values for FVC were imputed based on 3- month FVC, baseline FVC, individual-level baseline variables, and block-group level characteristics.

Five imputed datasets were created and analyzed using models described for the main analyses. FVC data were imputed using a linear regression method[24] and PA data were imputed using a logistic regression method.[25]

The results from each of the analyses of the imputed datasets were then combined using standard procedures.[26] Estimates of parameters from each of the D=5 datasets, θ_d(hat), were combined to obtain a summary measure, θ_D(bar), using **equation 3.5**.

$$\overline{\theta}_D = \frac{1}{D}\sum_{d=1}^{D}\hat{\theta}_d \qquad [3.5]$$

From there, total variance was estimated from two components: within imputation variance, W_D(bar) (**equation 3.6**); and between imputation variance, B_D (**equation 3.7**):

$$\overline{W}_D = \frac{1}{D}\sum_{d=1}^{D} W_d \qquad [3.6]$$

$$B_D = \frac{1}{D-1}\sum_{d=1}^{D}\left(\hat{\theta}_d - \overline{\theta}_D\right)^2 \qquad [3.7]$$

These components were then combined to estimate total variance, T_D (**equation 3.8**):

$$T_D = \overline{W}_D + \frac{D+1}{D} B_D \qquad [3.8]$$

This was estimated for mean change in FVC at each time-point by environment and study arm and for the linearized parameters by time-point, environment, and study arm in logistic and binomial regression for the PA outcome.

Strengths and Limitations for FVC and PA Analyses. The approach described above uses all data available in one model rather than separating outcomes measured at different follow-ups into different models. The advantage to this approach is that baseline measurement of the behavior is treated in the same way as other follow-ups, as

they should be. This is in contrast to other approaches where baseline measures may be treated as an adjustment variable, or subtracted from the follow-up measure of the behavior to look at the change from baseline.

Furthermore, the analysis accounts for several conceptual sources of correlation. While this may yield small numbers in some clusters (average 1.22 participants/block group), there is some indication in the literature that correlated analysis is still appropriate and yields valid estimates in contrast with uncorrelated analyses.[27]

And lastly, results from additional analyses were present in order to address potential limitations. Misclassification of block group code due to errors in geocoding was addressed by repeating analysis in only those participants for whom geocoding was successful during the first round. Potential bias due to loss to follow-up is addressed in the imputation analyses.

Survival Analysis

Outcomes. Survival analysis was performed using two different events. For individuals to remain in this study, they must first be contacted by the telephone evaluators and then consent to continue to participate in the study. These two steps allow for two different types of losses – those that are due to inability to contact the participant, and those that are due to explicit refusal of participants to continue to participate in the study.

In the first analysis, the outcome was defined as time from enrollment until the first missed contact between NuPA evaluators and participants. Since evaluations only occurred at 3, 6, and 12 months, these were taken to be the time points at which a participant could be censored or experience the event. Although it is possible that

participants skip a follow-up evaluation between two other evaluations – completing, for example, baseline and 3-month evaluations, skipping 6-month evaluation, and returning for 12-month evaluation – the participant would be recorded as having experienced the outcome at the first instance of a skipped follow-up. Participants who refused a follow-up evaluation were censored and their follow-up time recorded to be when they first refused. Those that were evaluated through the 12 month follow-up were censored at 12 months. Overall, 456 participants (41.3%) experienced the event and 553 (50.1%) completed the 12-month follow-up.

In the second analysis, refusal to participate was analyzed as the outcome and those who were lost to follow-up were censored. Of 1,104 participants, 145 (13.1%) refused to participate at some point during the trial.

Statistical Analysis. For both outcomes, Kaplan-Meier curves and the -2logR statistic were examined to determine whether there were significant differences in the rate of either event by study arm, metropolitan statistical area (MSA), or any individual-level demographic, health and psycho-social information, BG-level demographics, or environmental exposures. All variables that were statistically significant at the $\alpha=0.05$ level were examined for meeting the proportional hazards assumption by examining Log-Log curves.[28] After identifying variables that were significant when unadjusted, all significant variables were entered into a Cox model using SAS PROC TPHREG and PROC LIFEREG. Each variable was assessed for meeting the PH assumption using Schoenfield residuals. Those that met the PH assumption were kept in the model and examined for an independent association with the outcome after adjusting for other variables. Those that were not significant after adjustment for other variables were

removed in a stepwise fashion, eliminating the least significant ones first and checking for the PH assumption after each variable is removed. Once a reduced model containing independent predictors of refusal was identified, mixed models were examined using PROC NLMIXED adjusting for random effects for MSA or block group. These were compared with uncorrelated models estimated using PROC GENMOD.

Model Statement. The statistical models used allow for main effects of both the intervention arm (*I*), environment exposures of interest (*E*), and individual- and neighborhood-level demographics (combined in the vector *C*). Furthermore, they also allow for statistical interaction between the environment and the intervention arm to which the participant was randomized at baseline. Furthermore, each analysis will also allow for a potential random effect due to geographic clustering (b_0).

$$h(t,I,E,C) = h_o(t) \times \exp(\gamma_0 + b_0 + \gamma_1 \times I + \gamma_2 \times E + \gamma_3 \times C + \gamma_4 \times E \times I) \quad [3.5]$$

$$Logit\left[P(I,E,C)\right] = \gamma_0 + b_0 + \gamma_1 \times I + \gamma_2 \times E + \gamma_3 \times C + \gamma_4 \times E \times I \quad [3.6]$$

Power. Using the models described above, PASS (Power Analysis and Sample Size Software, NCCS, Kaysville, Utah) was used to determine power for the proposed study using the same approach for estimating effective sample size as described for the PA and FVC analyses. Fixing power at 80%, alpha at 0.05, the baseline event rate at 50%, exposure rate ranging from 5 to 95%, actual sample size of 300 to 900, the smallest detectable hazard ratio ranged from 1.3 to 2.5.

Strengths and Limitations. Survival analysis was used in the analysis of both refusal to participate and loss of contact. Although logistic regression analysis would

have been a feasible alternative to examine refusal at each follow-up, this current method was used instead in order to incorporate data from all participants across the entire follow-up time in one analysis.

The analysis also controlled for correlation between participants residing in the same block group and/or city.

Systematic Review of Fruit and Vegetable Consumption Interventions

Search Strategy

A search strategy was developed to identify controlled trials of behavioral interventions aimed at increasing fruit and vegetable consumption (**Appendix 6:** Search Strategy for Systematic Review). The same search strategy was applied to the CINAHL, MedLine, EMBASE, PsychInfo, databases and the Cochrane Central Register of Controlled Trials for reports published between January 1, 1990 and December 31, 2008. Behavioral interventions were identified using search terms adapted from previous systematic reviews.[29-32] These terms were then combined with other terms identifying increase in fruit and vegetable consumption as a study aim.[29,31] Results were restricted to studies performed among adult, non-institutionalized populations.

Assessment of Eligibility

The titles and abstracts of all articles identified from the search strategy were evaluated by two reviewers (DHC and REP) determining eligibility based on study aims, population, and intervention type. Studies were excluded if the intervention was strictly a pharmaceutical intervention, if there was a component of the intervention involving environmental modification or policy change, if the intervention was performed exclusively among smokers, and if the study was performed only among participants with

specific chronic disease conditions (hypercholesterolemia, diabetes or impaired glucose tolerance, hypertension, cancer survivors, etc.) except for overweight/obese participants. Studies that recruited from the general population and included some participants with chronic diseases were not excluded from the analysis.

Study settings were expected to include the workplace, church, university/college, primary care, or population-based studies, although no restrictions were made on this basis. Furthermore, intervention delivery was expected to be through the mail, through online mechanisms (email, website, or chatroom), through the telephone, or through individual- or group sessions delivered in person. Although most studies were randomized trials (either individual- or cluster-randomized), interventions with a pre- and post-test and no control group were not excluded.

Full texts of eligible studies were retrieved and examined to determine whether fruit and vegetable consumption was reported for baseline and follow-up evaluations in servings/day, grams/day, energy-adjusted daily intake, etc. and whether any modification of the intervention effect was examined or reported. References from studies that reported investigating effect modification were examined to identify additional studies.

Studies that reported only on baseline demographic information, baseline FVC, or described study design and motivations of otherwise eligible intervention were entered into ISI Web of Science (Thomson Reuters, New York, NY) to determine whether they had been cited by subsequent articles reporting on follow-up data. Follow-up articles were identified through author lists and/or study name. Again, these studies were assessed for eligibility to be included in the review.

Data Extraction

Once all eligible studies were identified, citation information, intervention design and setting, study population and sample size, a description of the intervention and control (if applicable), evaluation time-points, intervention effect, effect modifier examined, and the direction of the effect modification were extracted. The direction of effect modification was denoted as (0) for no effect modification, (+) for a greater intervention effect associated with a greater value of the effect modifier (or a greater intervention effect in the non-reference category in comparison to the reference category), or (-) for a smaller intervention effect associated with a greater value of the effect modifier (or a smaller intervention effect in the non-reference category of the effect modifier in comparison to the reference category). Information from the text identifying different reports originating from the same study (study name, sample size, intervention description, etc.) were grouped together. The authors' reasons for examining effect modification were noted.

No attempt was made to examine un-published studies or studies published in conference proceedings. Because of the diversity of study characteristics, and the variability in the effect modifiers measured, no attempt was made to arrive at any summary measures.

References

1. Kim Y, Pike J, Adams H, Cross D, Doyle C, Foreyt J. Worksite Telephone Intervention Promoting Weight-Related Health Behaviors. American Journal of Preventive MEdicine *Under Review*.
2. American FactFinder: 2005-2007 American Community Survey 3-Year Estimates. Vol. 2009.

3. Databases, Tables & Calculators by Subject: Labor Force Statistics from the Current Population Survey. Vol. 2009 Bureau of Labor Statistics.

4. Health, United States, 2007: With Chartbook on Trends in the Health of American. Hyattsville, MD: National Center for Health Statistics, 2007.

5. Metropolitan and micropolitan statistical areas and components, December 2006, with codes. Population Division, U.S. Census Bureau, 2007.

6. American FactFinder. In: Bureau USC, ed. Vol. 2009.

7. Goodchild MF. GIScience Ten Years After *Ground Truth*. Transactions in GIS 2006;10(5):687-692.

8. Dun and Bradstreet Million Dollar Directory. Cross Reference Geography, 2007.

9. Glanz K, Sallis JF, Saelens BE, Frank LD. Nutrition Environment Measures Survey in stores (NEMS-S): development and evaluation. Am J Prev Med 2007;32(4):282-9.

10. Morland K, Wing S, Diez Roux A, Poole C. Neighborhood characteristics associated with the location of food stores and food service places. Am J Prev Med 2002;22(1):23-9.

11. Zenk S, Schulz A, Israel B, James S, Bao S, Wilson M. Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit. Am J Public Health 2005;95:660-667.

12. Apparicio P, Cloutier MS, Shearmur R. The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. Int J Health Geogr 2007;6:4.

13. Addy CL, Wilson DK, Kirtland KA, Ainsworth BE, Sharpe P, Kimsey D. Associations of perceived social and physical environmental supports with physical activity and walking behavior. Am J Public Health 2004;94(3):440-3.

14. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 2006;117(2):417-24.

15. Antonakos C. Nonmotor travel in the 1990 nationwide personal transportation survey, 1995.

16. Sallis JF, Hovell MF, Hofstetter CR, Elder JP, Hackley M, Caspersen CJ, Powell KE. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. Public Health Rep 1990;105(2):179-85.

17. Li F, Fisher KJ, Brownson RC, Bosworth M. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. Journal of Epidemiology and Community Health 2005;59(7):558-564.

18. Sharkey JR, Horel S. Neighborhood Socioeconomic Deprivation and Minority Composition Are Associated with Better Potential Spatial Access to the Ground-Truthed Food Environment in a Large Rural Area. Journal of Nutrition 2008;138(3):620-627.

19. Larson N, Story M, Nelson M. Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S. American Journal of Preventive Medicine 2009;36(1):74-81.

20. Witte J, Greenland S, Kim L, Arab L. Multilevel Modeling in Epidemiology with GLIMMIX. Epidemiology 2000;11(6):684-688.

21. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. Ann Fam Med 2004;2(3):204-8.

22. Hannan PJ, Murray DM, Jacobs DR, Jr., McGovern PG. Parameters to aid in the design and analysis of community trials: intraclass correlations from the Minnesota Heart Health Program. Epidemiology 1994;5(1):88-95.

23. Prevalence of fruit and vegetable consumption and physical activity by race/ethnicity--United States, 2005. MMWR Morb Mortal Wkly Rep 2007;56(13):301-4.

24. Yuan YC. Multiple imputation for missing data: Concepts and new development. Twenty-Fifth Annual SAS Users Group International Conference. Cary, NC, 2000.

25. Allison PD. Imputation of Categorical Variables with PROC MI. SAS Users Group International, 30th Meeting (SUGI 30). . Philadelphia, PA, 2005.

26. Little RJA, Rubin DB. Statistical Analysis with Missing Data. Wiley Series in Probability and Statistics. 2nd edition ed. Hoboken, NJ: John Wiley & Sons, Inc., 2002.

27. Clarke P. When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. Journal of Epidemiology & Community Health 2008;62:752-758.

28. Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text. Statistics for Biology and Health. 2nd ed. New York, NY: Springer Science+Business Media, Inc., 2005.

29. Brunner EJ, Rees K, Ward K, Burke M, Thorogood M. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2007(4):CD002128.

30. Foster C, Hillsdon M, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180. pub2.

31. Kremers SP, de Bruijn GJ, Droomers M, van Lenthe F, Brug J. Moderators of environmental intervention effects on diet and activity in youth. Am J Prev Med 2007;32(2):163-72.

32. Katz DL, O'Connell M, Yeh MC, Nawaz H, Njike V, Anderson LM, Cory S, Dietz W. Public health strategies for preventing and controlling overweight and obesity in school and worksite settings: a report on recommendations of the Task Force on Community Preventive Services. MMWR Recomm Rep 2005;54(RR-10):1-12.

33. ACS Nutrition and Physical Activity (NuPA) Study: Progress Report and Discussion. Atlanta, GA: American Cancer Society, 2007;49.

Chapter 4:

Effects of park access and street structure on randomized trial efficacy

Do characteristics of the physical activity environment affect response to an intervention to increase physical activity? Results from a randomized controlled trial

Di H. Cross[1], Roberd M. Bostick[1,2], K.M. Venkat Narayan[1,2,3], Youngmee Kim[4], Lance Waller[5], Carol J. Rowland Hogue[1,6,7]

Affiliations

1. Department of Epidemiology, Emory University, Atlanta, GA, USA
2. Hubert Department of Global Health, Emory University, Atlanta, GA, USA
3. Department of Medicine, Emory University, Atlanta, GA, USA
4. Department of Psychology, University of Miami, Coral Gables, FL, USA
5. Department of Biostatistics, Emory University, Atlanta, GA, USA
6. Women's and Children's Center, Emory University, Atlanta, GA, USA
7. Department of Behavioral and Health Sciences, Emory University, Atlanta, GA, USA

Corresponding Author: Di H. Cross

Mailing address:

4408 Thompson Farm

Bedford, MA 01730

Email: dhe3@sph.emory.edu

Telephone: 404-290-1519

Fax: 404-727-8737

ABSTRACT

Background. Many studies report a positive association between physical activity (PA) and park access or street connectivity. However, these studies have almost exclusively been cross-sectional in design. The present study examines these measures of the physical activity environment as a potential modifier of a low-intensity counseling intervention, taking advantage of the established temporal sequence between environment and PA behavior and of the changes in behavior elicited from the study design.

Methods. Addresses from 1,104 participants in the Nutrition and Physical Activity Trial – a nationwide, randomized study of a telephone-based counseling intervention – were linked with geographic data to obtain density of streets and intersections near participants, proximity to parks, and the number of and total area of parks accessible within 8.05 km and 0.805 km radii. Self-reported PA was recorded at baseline and 3, 6, and 12 months post-randomization and investigated using generalized linear mixed models.

Results. Approximately 10% more intervention participants increased PA to ≥30 minutes/day at 6 months than control participants (risk difference (RD): 10.4%, 95% CI: [1.6%, 19.2%]). The intervention effect was not statistically significantly different by measures of the PA environment. However, the magnitude of the intervention effect increased with decreasing number of parks (0.805-km) and decreasing area of parks accessible (0.805-km), with 5-10% increases among those with the least number or area of parks accessible. Furthermore, maintenance of the intervention effect, measured at 12 months follow-up, increased with decreasing number of parks (8.05-km).

Conclusion. These findings suggest that the PA environment may encourage PA behavior. Unexpectedly, participants with fewer resources may have the most to gain

from a low-level counseling intervention. Future replications may take advantage of the established temporal relationship between behavior and environment in this design to understand the association between behavior and environment.

INTRODUCTION

Physical inactivity is linked to risk for chronic diseases such as obesity, cardiovascular and cerebrovascular disease, diabetes, and hypercholesterolemia, and other health conditions [1, 2]. In 2005, healthcare expenditures for these and other conditions totaled more than $2 trillion in the United States [3]. Stemming from its wide influence as a risk factor for disease and morbidity, health economists estimate that were physical inactivity and overweight/obesity to be prevented, total healthcare expenditures could be reduced 9-27% [4, 5], or by at least $180 billion.

However, efforts to increase physical activity to levels recommended for preventing chronic disease have been met with a number of challenges. At best, randomized counseling and educational interventions have reported only moderate effects on self-reported physical activity [6, 7] despite the application of strategies to address a multitude of barriers to behavior change such as knowledge, limited time, and lack of social support [8, 9]. Oftentimes, no changes in physical activity behavior are observed despite increases in intermediate outcomes such as self-efficacy [8].

Observational studies have reported that physical activity behavior is linked to the built, physical activity environment, a factor relatively unaddressed in the randomized trial literature. Even after accounting for individual-level characteristics, greater access to parks [10-12] and greater street connectivity [13-15] have been associated with increased physical activity. However, these findings cannot establish a causal link between the physical activity environment and physical activity levels as they have almost exclusively employed a cross-sectional study design. Furthermore, results have not always been consistent, particularly in studies restricted to homogeneous groups such

as Latinas [16, 17] and urban-dwelling, African-American women [18] where no such associations were found. Furthermore, studies attempting to study the direct effect of changes in the physical activity environment have been inconsistent [19, 20].

To address limitations in the literatures on randomized trials and observational studies, the present analysis combines objective measures of the physical activity environment with data from a randomized controlled trial. While this analysis aids in identifying subgroups defined by physical activity environment for whom a physical activity counseling intervention may be more effective, it also presents an opportunity to address the question of whether the built environment may be a barrier to physical activity improvement, examining the question of temporality in the association between behavior and environment from a different perspective. We hypothesized that intervention effects would be more pronounced among participants who have greater access to parks and greater street connectivity.

METHODS

Participants

The Nutrition and Physical Activity (NuPA) Trial, conducted by the American Cancer Society between September 2005 and August 2008, was a randomized controlled trial of an educational intervention to increase vegetable and fruit intake and physical activity. Participants enrolled by contacting the National Cancer Information Center in Austin, Texas by phone. Eligibility criteria included not having been diagnosed with an eating disorder, over 18, not pregnant or nursing, and willing to make a change in their daily habits within the next 2 weeks. Both randomized groups received three self-help

guidebooks that discussed behavior change strategies and provided educational material and specific suggestions. Those randomized to the intervention arm also received counseling by telephone over six months. All participants were evaluated at baseline, and with follow-up attempts at three-, six-, and 12-months post-randomization. Details of recruitment and trial procedures have been reported elsewhere (Youngmee Kim, 2009, "Worksite Telephone Intervention Promoting Weight-Related Health Behaviors", unpublished manuscript).

Although participants were recruited from throughout the United States, only participants residing in cities from which 100 or more participants were recruited were chosen for the present analysis to improve the detail with which environmental exposures could be calculated. From 2,623 eligible participants in the NuPA Trial, 1,104 residing in the metropolitan statistical areas of Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota were included in the analysis.

This study and its parent study were approved by the Emory University Institutional Review Board.

Survey Data

Individual-level demographic information, psychosocial variables, and relevant medical information were ascertained at baseline. Covariates included in the analysis were age, sex, race/ethnicity, employment, marital status, educational attainment, hypercholesterolemia, diabetes types 1 or 2, hypertension, previous diagnosis of cancer, being sad/blue, and satisfaction with life. Self-reported physical activity in minutes/day

during the previous three months was ascertained at baseline, three, six, and 12 months follow-up and categorized as ≥30 minutes/day or <30 minutes/day.

Geocoding

Participants' reported mailing addresses were used with the Qualitative Marketing Software (QMS) Geocoder (QMSoft, Clearwater, Florida) to map participants to census-defined block groups. Following this first round of geocoding, addresses that had not been mapped were matched to a block group using one of three alternative methods. Addresses were located: via the United States Census Bureau American FactFinder Address function, using the block group shape file in combination with continuously updated street information from TeleAtlas (TeleAtlas North America, Incoporated, Lebanon, New Hampshire), or approximated using the QMS Geocoder at the block group or zip code level. Twenty-eight participants reported post office box mailing addresses. A detailed description of the geocoding procedures was described elsewhere (Cross, Di H. "The effects of access to grocers on fruit and vegetable consumption in a randomized controlled trial". 2009. Unpublished manuscript). To maintain confidentiality, the unweighted geographic centroid of the block group was used as the focal point from which exposures were ascertained.

Environmental-level covariates were obtained from the Claritas Demographic Update (2007 Update, Claritas Nielson, San Diego, California), and measured within each block group. These included educational attainment (percentage of block group residents with a bachelors degree or more, percentage with a high school degree or equivalent), neighborhood median income, racial composition percentage minority

including black, Asian, Hispanic, other, and mixed), median age, and unemployment rate (percentage unemployed) and were categorized into quartiles for analysis.

Exposure ascertainment

Park Access

Participant access to parks (local park/recreation area, national park service land, and state/local park/forest) was calculated using the United States Geographic Data Technology, Incorporated (GDT), Park Landmarks shapefile (2007, Environmental Systems Research Institute, Inc. (ESRI), Redlands, California). Measures of park access were park diversity (total number of parks within buffer regions with 0.805, and 8.05 km radii), proximity to parks (distance to the nearest park), and park space (size of the nearest park as well as the sum of the size of all parks within a buffer region).

Street Connectivity

Data from StreetMap USA (2007, ESRI, Inc., Redlands, California) were used to calculate street and intersection density within the above-defined proximity buffers. Measures of street connectivity included the number of 4-way intersections within the buffer area, number of high-speed 4-way intersections (1 or more tributary streets with a 25 mile/hour (40 km/hour) or greater speed limit), total length of roads within the buffer, and the combined length of any high-speed (speed limit of 25 miles/hour or more) roads.

All continuous or count measures were categorized into quartiles; where quartiles were not possible, variables were dichotomized or categorized into tertiles as appropriate.

Statistical analysis

Unadjusted associations

Baseline individual-level characteristics were examined across study arms using χ^2 tests for categorical variables and two sample t-tests with pooled variances for continuous variables. The overall intervention effect on meeting or exceeding the threshold of 30 minutes/day of physical activity was modeled as a linear change in risk across all timepoints using correlated generalized linear models, allowing for an autoregressive order one correlation between observations from the same participant.

Increase in proportion of participants meeting the recommendation was estimated for each arm comparing three-, six-, and 12-month follow-up to baseline data across study arms using a two-way interaction with time and study arm, and across study arm by physical activity environment categories using a three-way interaction variable consisting of time, study arm, and characteristics of their physical activity environment. Categories of the physical activity environment were treated as nominal as well as ordinal variables in order to test for trend. Allowing for the possibility of no intervention effect, a two-way interaction between environment and time was also examined. Models were fit via generalized linear mixed models with autoregressive order one correlation between observations within individuals and, where appropriate, random effects shared between individuals residing in the same block group and/or city.

Analyses were repeated adjusting for potential confounders identified as individual- and environmental-level variables associated with the outcome at the $p \leq 0.05$ level. City of residence and month of enrollment were also assessed as potential confounders.

Sensitivity Analyses

The above analyses were repeated using only data collected from participants who had been linked to their block group during the first round of geocoding. In addition, the analyses were repeated in which only participants who reported <30 minutes/day of physical activity at baseline were included.

Using a proportional hazards survival model with random effects, measures of the physical activity environment were evaluated to determine whether they were associated with participant retention in the trial. Random effects for block group and/or city, and individual- and block-group level covariates were included in the models as appropriate.

All spatial calculations were performed using ArcView 9.2 and statistical analyses were performed using SAS versions 9.1 and 9.2.

RESULTS

Of 1,104 eligible participants, 1,086 residing in 890 different block groups were included in the analysis after excluding 17 participants on whom baseline demographic data were missing, and 1 participant who could not be matched to a block group. From these, self-reported physical activity information was obtained on 1,085 at the baseline evaluation, 653 from the 3-month, 540 from the 6-month, and 439 from the 12-month follow-ups.

The intervention groups did not significantly differ in selected characteristics at baseline except for a slightly higher frequency of persons who were employed or who had hypercholesterolemia in the active intervention arm (**Table 4.1**). Furthermore, less than 1% of the variation in physical activity levels was explained by either shared city of

residence or shared block group. Thus, random effects for these variables were not included in subsequent analyses.

The two intervention groups were balanced on indicators of physical activity environment at baseline. However, there were some significant differences in the characteristics of the physical environment between persons who did or did not engage in ≥30 minutes/day of physical activity at baseline (**Table 4.2**). Of note, park diversity at 0.805 and 8.05 km were significantly positively associated with greater physical activity. This was true for park space within 0.805 km as well. Higher density of high-speed 4-way intersections (over 25 miles per hour (mph) tributary streets), and high-speed road lengths (combined total lengths of road with speed limits ≥25 mph) were also positively associated with more physical activity.

The associations between baseline physical activity levels and environment remained significant after controlling for individual- and environmental-level covariates. Participants with a park within 0.805 km were 10.3% (95% CI: [2.8, 17.7], p=0.004) more likely to exceed 30 minutes/day of physical activity at baseline than those with more parks. Similarly, the trend by park diversity (8.05-km) remained significant (p=0.05) as did the association between baseline behavior and park space accessible (0.805-km, p=0.01, trend p=0.05). In contrast, the associations with high-speed intersection (p=0.40, trend p=0.09) and high-speed lengths (p=0.67, trend p=0.33) were not significant after controlling for other covariates.

A substantial portion of participants in both arms increased their physical activity to ≥30 minutes/day during the 12 months of follow-up. Among those randomized to receive the intervention, 42.6% reported meeting this threshold at baseline. At 3 months

this increased to 55.0%, at 6 months 57.7% and at 12 months 52.8%. Among those randomized to receive the control, 51.2% exceeded the threshold at baseline. An increase was also observed in this group, increasing to 62.8% at 3 months, 55.9% at 6 months, and 56.6% at 12 months. The intervention effect, defined as the difference between the two arms in the change in percentage of participants meeting the recommendation from baseline to follow-up, was statistically significant at 6 months with 10.4% more intervention than control participants increasing to meet the recommended 30 minutes/day (95% confidence interval (CI): [1.6, 19.2]). At 12 month follow-up, a significantly greater proportion of intervention participants reported ≥30 minutes/day of physical activity than at baseline (risk difference (RD): 10.2%, 95% CI: [3.4, 17.0]). In contrast, the increase from baseline to 12 months among controls was not significant (RD: 5.4%, 95% CI: [-1.6, 12.4]).

The magnitude of the intervention effect varied by some characteristics of the physical activity environment found to be associated with physical activity behavior at baseline, although these interactions and the two-way interactions between physical activity environment and time were not statistically significant (ps>0.05).

First, by quartiles of numbers of parks within 8.05 km, the highest 12-month intervention effect was among those with the fewest parks (0 to 12 parks), with an intervention effect of 9.4% (95% CI: [-10.5, 29.2], p=0.36). In the second and third quartiles the intervention effects were, respectively, 3.0% (95% CI: [-17.9, 23.9], p=0.78) and 9.3% (95% CI: [-14.2, 22.1], p=0.67). In the highest quartile, the intervention effect was 0.3% (95% CI: [19.5, 20.2], p= 0.97).

By tertiles of park area accessible within 0.805 km, participants with no park space accessible showed a consistent 5-10% intervention effect at 3, 6, and 12 months (5.2, 8.2, and 5.9% respectively) while the intervention effect was less consistent in the other two tertiles (second tertile: -3.2, 13.5, -3.6%; third tertile: -3.3, 10.1, 9.3%). A similar pattern was seen among participants with no parks within 0.805 km. The intervention effects at 3, 6, and 12 months were 5.2, 8.2, and 5.9% respectively among those with no parks and -2.9, 12.2, and 3.3% at 3, 6, and 12 months among those with ≥1 park within 0.805 km.

Adjustment for potential confounders (all individual-level covariates, block group-level % non-Hispanic white population, block-group level median age, and time-by-block group-level median age interaction) or for month of response and city of residence (either as fixed or random effects) did not yield substantively different conclusions.

Our findings did not materially differ when the analyses were restricted to only individuals who were successfully mapped during the first round of geocoding, or to participants who reported <30 minutes/day at baseline. Furthermore, no measures of the physical activity environment were statistically significantly associated with participant retention whether loss to follow-up was due to explicit refusal by the participant to continue in the trial at any time after baseline, or because study staff was unable to contact the participant.

DISCUSSION

This is the first study known to the authors to examine objective measures of the physical activity environment within a randomized controlled trial of an intervention to increase physical activity. The findings suggest that the effect of the counseling intervention investigated may differ across measures of the physical activity environment. However, contrary to the *a priori* hypothesis – formed based on the observational literature on the built environment and individual behavior – the greatest intervention effect was observed among participants with the least access to parks. Furthermore, there was no evidence suggesting differences in intervention effect across measures of street connectivity such as intersection density and street density.

The observed patterns were identified despite several limitations. First, a high proportion of participants reported ≥30 minutes/day of physical activity at baseline. This may be explained by the survey measure used which does not distinguish between leisure- and work-time physical activity, and grouped mild physical activity together with moderate and vigorous activity. This would decrease the total amount of change in physical activity behaviors observable in this group of individuals. Second, the physical activity environmental measures examined were limited to those that may affect outdoor physical activity. Access to private or community gyms and ownership of equipment at home were not accounted for in this analysis. However, this association was not confounded by time as adjustment for seasonality, operationalized as the month of response, did not change the association between environment and behavior. Third, exposures were likely misclassified due to the shift from the focal point of interest (the participant residence) and the focal point used (the block group centroid). This was limited by the size of block groups and the variability of the exposure across space.

There is little information quantifying the variability of this exposure across space. However, the size of block groups ranged between 0.05 km^2 and 158.0 km^2 in this study. All but 3 block groups in the analysis were smaller in total area than the buffer area enclosed by the 8.05-km radius used for assessing exposure. Additional misclassification may have occurred because of changes in the environment across time unaccounted for in this analysis. Participant exposures were assessed using data from 2007 although follow-up occurred between 2005 and 2008. Both sources of misclassification of the exposures were likely non-differential with respect to the outcome under study and resulted in bias towards the null.

This design presents several benefits over directly examining environment and behavior in longitudinal cohort studies or cross-sectional surveys. First, it is important to note that behavior change, rather than behavior, was the outcome under study. In contrast to cross-sectional studies, the current study establishes that the physical activity environment experienced by participants preceded any observed changes in behavior. Second, the administration of an intervention is likely to elicit more behavior change in a population than observational data from a longitudinal study. This leads to greater power to detect differences in behavior change across different levels of environmental exposures. And third, contrasting intervention participants with control participants allows the researcher to factor out the effect of increased knowledge, desirability effects, or secular trends in behavior change and behavior reporting which may remain unaccounted for in a longitudinal study.

Overall, the findings are suggestive of an inverse relationship between intervention effect and access to parks, despite several sources of potential bias towards

the null. Furthermore, it is striking that the relationship is limited to parks and that null associations were found with street connectivity. These findings are in contrast to at least one study conducted examining perceived characteristics of the physical activity environment [21]. King et al. found that less perceived speeding among drivers, greater perceived safety, and greater presence of crosswalks in the neighborhood to be associated with greater intervention effects on physical activity in three different randomized trials. The difference between the findings reported here and reported by King may be accounted for by the use of perceived physical activity environment rather than objectively measured environment, and differences in the intervention administered. Three different interventions were examined by King et al: one involving a specific aerobic exercise regimen administered over the phone, another investigating computer-based community resource support, and another examining telephone-based motivational interviewing techniques.

Taken together, this study finds evidence suggestive of a causal relationship between physical activity environment and physical activity behavior. Furthermore, it demonstrates a novel approach to examining an association that has proven to be problematic to examine in previous studies. Although the complex social, physical, and biologic aspects of behavior remain difficult to disentangle, this study design demonstrates the use of additional tool that may be used in the future. Larger studies and studies of different interventions should be conducted to match population subgroups to more appropriate and more effective interventions to address the challenges in attaining and maintaining population levels of healthy physical activity behaviors.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions of Joanne Pike and Heather Adams in the design and execution of the NuPA study. Furthermore, they would like to acknowledge the work of employees of the National Cancer Information Center in data collection.

COMPETING INTERESTS

None declared.

FUNDING

This study was funded by the Intramural Research and Health Promotions Departments of the American Cancer Society.

REFERENCES

1 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. *Med Sci Sports Exerc* 2007;**39**:1423-34.

2 Kesaniemi YA, Danforth E, Jensen MD*, et al.* Dose-response issues concerning physical activity and health: an evidence-based symposium. *Symposium on Dose-Response Issues Concerning Physical Activity and health*. Ontario, Canada: Lippincott Williams & Wilkins 2000:S351-S8.

3 Health, United States, 2007. Hyattsville, Maryland: National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services 2007.

4 Anderson L, Martinson B, Crain A*, et al.* Health care charges associated with physical inactivity, overweight, and obesity. *Prev Chronic Dis* 2005;**2**:A09.

5 Colditz G. Economic costs of obesity and inactivity. *Med Sci Sports Exerc* 1999;**31**:S663-7.

6 Foster C, Hillsdon M, Thorogood M. Interventions for promoting physical activity. *Cochrane Database Syst Rev* 2005:CD003180. pub2.

7 Marcus BH, Williams DM, Dubbert PM*, et al.* Physical activity intervention studies: what we know and what we need to know: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity); Council on Cardiovascular Disease in the Young; and the Interdisciplinary Working Group on Quality of Care and Outcomes Research. *Circulation* 2006;**114**:2739-52.

8 Collins R, Lee RE, Albright CL*, et al.* Ready to be Physically Active? The Effects of a Course Preparing Low-Income Multiethnic Women to be more Physically Active. *Health Educ Behav* 2004;**31**:47-64.

9 Albright CL, Pruitt L, Castro C*, et al.* Modifying physical activity in a multiethnic sample of low-income women: One-year results from the IMPACT (increasing motivation for physical activity) project. *Ann Behav Med* 2005;**30**:191-200.

10 Foster C, Hillsdon M, Thorogood M. Environmental perceptions and walking in English adults. *J Epidemiol Community Health* 2004;**58**:924-8.

11 Kamphuis CB, van Lenthe FJ, Giskes K*, et al.* Perceived environmental determinants of physical activity and fruit and vegetable consumption among high and low socioeconomic groups in the Netherlands. *Health Place* 2007;**13**:493-503.

12 Wendel-Vos GC, Schuit AJ, de Niet R*, et al.* Factors of the physical environment associated with walking and bicycling. *Med Sci Sports Exerc* 2004;**36**:725-30.

13 Atkinson JL, Sallis JF, Saelens BE*, et al.* The association of neighborhood design and recreational environments with physical activity. *Am J Health Promot* 2005;**19**:304-9.

14 Boer R, Zheng Y, Overton A*, et al.* Neighborhood design and walking trips in ten U.S. metropolitan areas. *Am J Prev Med* 2007;**32**:298-304.

15 Targa F, Clifton KJ. Built Environment and Nonmotorized Travel: Evidence from Baltimore City Using the NHTS. *Journal of Transportation and Statistics* 2005;**8**:55-70.

16 Evenson KR, Sarmiento OL, Tawney KW*, et al.* Personal, social, and environmental correlates of physical activity in North Carolina Latina immigrants. *Am J Prev Med* 2003;**25**:77-85.

17 Voorhees CC, Rohm Young D. Personal, social, and physical environmental correlates of physical activity levels in urban Latinas. *Am J Prev Med* 2003;**25**:61-8.

18 Rohm Young D, Voorhees CC. Personal, social, and environmental correlates of physical activity in urban African-American women. *Am J Prev Med* 2003;**25**:38-44.

19 Evenson KR, Herring AH, Huston SL. Evaluating Change in Physical Activity with the Building of a Multi-Use Trail. *Am J Prev Med* 2005;**28**:177-85.

20 Merom D, Bauman A, Vita P*, et al.* An environmental intervention to promote walking and cycling--the impact of a newly constructed Rail Trail in Western Sydney. *Preventive Medicine* 2003;**36**:235-42.

21 King AC, Toobert D, Ahn D*, et al.* Perceived environments as physical activity correlates and moderators of intervention in five studies. *Am J Health Promot* 2006;**21**:24-35.

Table 4.1. Selected characteristics of participants.

Variable	Counseling (n=558)		Self-Help (n=546)	
Age (*years)*[1]	42.7	(10.4)	42.5	(10.9)
Sex *(%)*				
Male	104	18.6	108	19.8
Female	454	81.4	438	80.2
Metropolitan Area *(%)*				
Atlanta	100	17.9	86	15.8
Austin	56	10.0	59	10.8
Chicago	54	9.7	51	9.3
Knoxville	47	8.4	53	9.7
Minneapolis	301	53.9	297	54.4
Employed *(%)*	538	96.6	514	94.1
Married *(%)*	335	60.0	329	60.5
Education *(%)*				
HS or less	71	12.7	55	10.1
Some college/Tech	160	28.7	178	32.6
College grad	222	39.8	220	40.3
Post college	105	18.8	93	17.0
Ethnicity *(%)*				
White	402	72.7	412	75.6
Black	95	17.2	87	16.0
Hispanic	32	5.8	27	5.0
Other	24	4.3	19	3.5
missing	*5*		*1*	
Medical History *(%)*[2]				
Hypertension	115	20.6	95	17.4
Cancer	43	7.7	36	6.6
Hypercholesterolemia	160	28.7	124	22.7
Diabetes				
Type I	1	0.2	4	0.7
Type II	24	4.3	32	5.9
Satisfied with Life *(%)*	401	71.9	416	76.2
Feel Sad/Blue *(%)*	45	8.1	35	6.4

1. mean (standard deviation)
2. Self-reported physician diagnosis or treatment of condition

Table 4.2. Baseline participant physical activity levels according to selected physical activity environment characteristics

Variable		≥30 min/day %	p	trend
Number of Parks				
Within 0.805 km			**0.004**	--
None		41.8		
One or more		50.8		
Within 8.05 km			0.09	**0.01**
Q1:	0 to 12	41.5		
Q2:	13 to 47	46.0		
Q3:	48 to 105	47.4		
Q4:	106 o 192	52.5		
Area of parks accessible (km^2)				
Within 0.805 km			**0.04**	**0.01**
T1:	None	41.6		
T2:	(0, 0.028]	51.5		
T3:	≥ 0.028	49.8		
High-speed 4-way intersections				
Within 0.805 km			0.13	**0.03**
Q1:	0 to 24	43.8		
Q2:	25 to 55	44.4		
Q3:	56 to 86	46.8		
Q4:	87 or more	52.9		
Road length *(over 25 mph, km)*				
Within 0.805 km			0.18	**0.05**
Lowest	[0, 11.3)	44.4		
Low	[11.3, 18.85)	44.2		
High	[18.85, 26.17)	46.4		
Highest	≥ 26.15	52.7		

p-value for equality of proportions using linear risk difference
Trend test from above model

Chapter 5:

Effects of grocer access on randomized trial efficacy

The effects of access to grocers on fruit and vegetable consumption in a randomized controlled trial

Running title: Fruit and vegetable consumption by grocer access

Keywords: access, diet, effect modification, fruits, grocers, randomized controlled trial, vegetables

Word Count: 3,464

Acronyms: American Cancer Society (ACS), CI (confidence interval), NuPA (Nutrition and Physical Activity), SE (standard error)

ABSTRACT

Although fruit and vegetable consumption decreases the risk of chronic disease, interventions to increase it have demonstrated only modest effects. In addition, cross-sectional studies suggest a positive association between fruit and vegetable consumption and grocer access but cannot establish temporality. The randomized trial environment presents an opportunity to study the temporal relationship between environment and behavior and identify subgroups for whom counseling interventions may be more effective. Addresses from 1,104 participants in the Nutrition and Physical Activity Trial – a nationwide, randomized study of a telephone-based counseling intervention – were linked with grocer addresses to obtain 3 measures of access: proximity, diversity, and variety. Self-reported fruit and vegetable consumption was recorded at baseline, 3, 6, and 12 months and investigated using mixed models. The intervention effect was significantly greater among participants with the lowest access compared to those with low access (by diversity of grocers, 1.46 servings/day greater (95% confidence interval (CI): 0.09, 2.00); by variety, 1.21 servings/day (95% CI: 0.66, 2.49)). These findings support the hypothesis that the built environment is a precursor to individual behavior. Participants with the least access to grocers may have the most to gain from individual-level counseling interventions to increase fruit and vegetable consumption.

INTRODUCTION

Although an estimated 30% of cardiovascular disease and 35% of cancer is linked to poor diet, few Americans consume diets consistent with patterns which reduce the risk of these and other chronic conditions (1-3). Current guidelines recommend 4 servings of fruits and 5 servings of vegetables per day, but data from 2005 show that only 32.6% of Americans consume fruit 2 or more times a day and 27.2% eat vegetables 3 or more times per day (4, 5).

Attempts to change dietary patterns have been met with only limited success. Randomized trials of counseling interventions consistently report limited to moderate changes in fruit and vegetable consumption with little compelling evidence that these effects are sustained (6, 7). These moderate effects may be explained by a truly moderate or null intervention effect or they may be observed due to reporting overall results that are averaged over heterogeneous effects across participant subgroups (8, 9).

At a population level, identification of such subgroups may be critical to appropriate allocation of resources (8). Given finite resources, assigning an intervention to sub-populations where it will be effective would produce a greater population health benefit. Traditionally, subgroup analysis has been applied to estimate intervention effects within participant groups characterized by individual-level demographic variables measured at baseline. However, the literature suggests that examination of intervention heterogeneity across participant characteristics, rather than within homogeneous groups, may be more appropriate (10). This has proven fruitful as meaningful differences in intervention efficacy have been found in participant groups identified by socio-economic status, and ethnicity (11). However, the observational literature on fruit and vegetable

consumption suggests that environmental-level characteristics may also be important predictors of individual behavior, even after controlling for individual-level characteristics (12-15).

Findings from a number of cross-sectional studies suggest that access to grocers where one can purchase produce is associated with individual fruit and vegetable consumption, although study findings have not always been consistent (16-18). However, it is difficult to determine whether access is a precursor to behavior or vice versa in cross-sectional studies (19). While individual consumption patterns may be influenced by the local food environment, it is equally plausible that the local food environment is influenced by individual consumption, if only in aggregate. This limitation is not strictly due to study design, but is rooted in the nature of the exposure, making it difficult to execute a study that would provide a higher level of evidence such as a randomized trial of high access versus low access to grocers. Even longitudinal studies of behavior cannot overcome this challenge because a participant with a given behavioral intent may choose a residence based in part on the local environment, rather than vice versa. In either case, analysis of study data would result in the same observed association.

This study demonstrates an alternative solution, addressing the question of temporality by examining grocer access as a modifier of the effect of a counseling intervention designed to improve lifestyle behaviors. Using the framework of VanderWeele and Robins, heterogeneity in the intervention effect may arise if both the intervention and the exposure are causes of the outcome (20). Overall, the intervention was associated with increased change in fruit and vegetable consumption (University of

Miami, unpublished manuscript: "Worksite Telephone Intervention Promoting Weight-Related Health Behaviors"). Based on this finding and the observational literature, the authors hypothesize that participants who were randomized to the intervention arm will increase their consumption more than those in the comparison arm, and that this group difference will be larger among participants with greater access to grocers than those with less access.

MATERIALS AND METHODS

Participants

This study is derived from the Nutrition and Physical Activity (NuPA) Trial conducted by the American Cancer Society between September 2005 and August 2008. Of 2,623 eligible participants (not having been diagnosed with an eating disorder, over 18, not pregnant or nursing, and willing to make a change in their daily habits within the next 2 weeks), 1,353 were randomized to the intervention arm and 1,270 to the comparison arm. The intervention arm received 3 guidebooks discussing changes in lifestyle behaviors and 6 telephone counseling sessions conducted over 6 months. Participants in the comparison arm received only the 3 guidebooks. Detailed description of recruitment procedures and characteristics of the entire study population are reported elsewhere (University of Miami, unpublished manuscript: "Worksite Telephone Intervention Promoting Weight-Related Health Behaviors").

As participants were recruited from throughout the United States, only cities, defined by boundaries of metropolitan statistical area, with 100 or more participants were included in the present analysis in order to maximize the detail with which grocer access could be calculated. Residents of Atlanta, Georgia; Austin, Texas; Chicago, Illinois;

Knoxville, Tennessee, and Minneapolis/St. Paul, Minnesota were included and totaled 1,104 at baseline.

This study and its parent study were approved by the Emory University Institutional Review Board.

Geocoding

Participants were geocoded to a block group using the Qualitative Marketing Software (QMS) Geocoder (QMSoft, Clearwater, Florida) based on mailing address reported for delivery of guidebooks. The location of the full address was not used for analysis to maintain confidentiality. A random sample of 10% of all addresses linked at the precision level of the residence were examined for accurate pairing with the block group code using the address search function on the United States (U.S.) Census website and data from the 2000 Census (21). Addresses that could not be found in this manner were assumed to be new and were located using Google Earth and the U.S. Census Block Groups shapefile (2007, Environmental Systems Research Institute, Inc. (ESRI), Redlands, California). All addresses that were approximated at the block group or zip code level were examined using the same procedure. Where codes obtained from different data sources disagreed, preference was given to the code obtained from, in order, the census address tool, the block group shapefile, and the QMS-approximated block group.

The majority of participants (N=947, 85.8%) were geocoded during the first round. All addresses sampled from this group for further verification produced codes that were consistent between the Census address tool and the QMS Geocoder. The majority of participants whose addresses were approximated at the block group or zip code level

during the first round (N=157), were linked to a block group using one of the other 2 methods (29 using the Census, 96 using the shapefile). Approximated block groups were used for 28 participants who reported a post office box as their mailing address.

The unweighted geographic centroid of the block group was used as the focal point for assessing access to grocers. Block group-specific covariates were obtained from the Claritas Demographic Update (2007 Update, Claritas Nielson, San Diego, California), and locations of grocers were obtained from the Dun & Bradstreet Million Dollar Database (December 2007, Dun & Bradstreet, Short Hills, New Jersey).

Exposure Ascertainment

The Dun & Bradstreet (D&B) Million Dollar Database provides data on businesses meeting one or more of the following criteria: (1) conducted $1 million or more in sales, (2) have 20 or more employees, or (3) have branches with 50 or more employees (22, 23). Smaller businesses were excluded from exposure assessment, consistent with the literature suggesting that it is the larger chains and businesses that contribute to greater consumer access to fruits and vegetables (24-26). Businesses with primary North American Industry Classification System codes of 44523 (fruit and vegetable specialty markets) or 44511 (supermarkets and other grocery stores, except convenience stores) were included in the analysis.

Each business was mapped using the longitude and latitude from the D&B database. Combining that information with block group centroids, 3 measures of grocer access were adapted from the work of Apparicio et al: (1) proximity: the distance from the focal point to the nearest grocer; (2) diversity: the number of grocers within a buffer distance from the focal point; and (3) variety: the average distance from the focal point to

the 3 nearest differently named grocers (27). Both 0.805-km (0.5-mile) and 8.05 km (5-mile) buffers were used for calculating diversity, assuming non-motorized and motorized transport, respectively (28).

All spatial calculations were performed using ArcView 9.2 (ESRI, Inc., Redlands, California), and were dichotomized or, where possible, treated in quartiles. Quartiles are referred to as highest, high, low, and lowest.

Outcomes

Fruit and vegetable consumption was measured at baseline and at each follow-up as the sum of self-reported servings/day of fruit and vegetable consumption, using the past 3 months as the reference period. Unreasonable values of 30 combined servings per day or greater were excluded from the analysis.

Covariates

Individual-level variables that may influence behavior were assessed at baseline. Covariates included age, sex, race/ethnicity, employment, marital status, educational attainment, hypercholesterolemia, diabetes types 1 or 2, hypertension, previous diagnosis of cancer, being sad/blue, and satisfaction with life. Environmental-level covariates were assessed on the block group level and included educational attainment (% of block group residents with a bachelors degree or more, % with a high school degree or equivalent), neighborhood median income, racial composition (% minority including black, Asian, Hispanic, other, and mixed), median age, and unemployment rate (% unemployed). These data were obtained from the Claritas Demographic Update (2007) and were categorized into quartiles.

Statistical Analysis

Of 1,104 participants, 17 were excluded from analysis due to missing baseline covariate data. One additional participant was excluded because the reported mailing address could not be geocoded to any block group. The 1,086 remaining participants resided in 890 different block groups and reported 2,708 assessments, of which 1,065 were at baseline, 655 at 3 months, 547 at 6 months, and 441 at 12 month follow-up evaluations.

Unadjusted associations. Individual-level covariates and exposure variables were examined across study arms using χ^2 tests. The associations between fruit and vegetable consumption and exposure variables were examined at baseline using linear mixed models with random effects for block group and city. Average fruit and vegetable consumption at baseline and each follow-up, and average change in consumption from baseline were compared across study arms using separate mixed models for each time-point, again including random effects for block group and city. Linear mixed models were used to examine all follow-ups together including at baseline. Models were fit via restricted maximum likelihood estimation allowing autoregressive order one correlation between observations within a given individual and between individuals sharing a block group and/or city. Average change in fruit and vegetable consumption was estimated across exposure-by-study arm groups to identify exposure variables with statistically significant heterogeneity in the intervention effect at the 0.05 level.

Adjusted associations. All individual-level covariates were included in adjusted models as potential predictors of baseline fruit and vegetable consumption. Each individual-level covariate was then assessed separately as a potential predictor of behavior change using a covariate-by-time interaction. Block group-level covariates

were only examined in models already adjusted for analogous individual-level covariates. Interactions and block-group level covariates that were significantly associated with behavior or behavior change at the 0.05 level were included in the fully adjusted model along with all individual-level covariates. The main effects of interest were the estimated changes in fruit and vegetable consumption associated with exposure-by-time-by-study arm interactions.

Sensitivity Analysis. Sensitivity analysis was performed excluding participants whose block group codes were approximated at the block group or zip code level and participants who reported a post office box mailing address in order to assess the potential effect of misclassification occurring from block group codes obtained using different methods.

Survival analysis was performed to determine whether access to grocers was associated with loss to follow-up, adjusting for correlation between participants from the same block group and/or city, and individual- and block group-level covariates.

All statistical analyses were performed using SAS 9.1 and 9.2.

RESULTS

Unadjusted analysis

There were no significant imbalances in individual-level covariates (**Table 5.1**) or block group-level exposures (**Table 5.2**) by study arm, with 2 exceptions: hypercholesterolemia was prevalent in 6% more of those randomized to intervention than to the comparison arm and employment was slightly higher among those in the comparison arm than those in the intervention arm. A substantially larger portion of participants were recruited from Minneapolis/St. Paul, Minnesota (54.2%) than from

other cities (University of Miami, unpublished manuscript: "Worksite Telephone Intervention Promoting Weight-Related Health Behaviors"). At baseline, fruit and vegetable consumption ranged between 0 and 12 servings/day and did not differ by diversity or proximity. However, participants with a high variety of grocers at baseline (3 grocers within 1.848 to 2.59 km average distance, 3.92 servings/day, 95% confidence interval (CI): 3.62, 4.21) reported a greater average fruit and vegetable consumption than other quartiles (**Table 5.3**). There was no detectable trend in baseline fruit and vegetable consumption by any measure of grocer access.

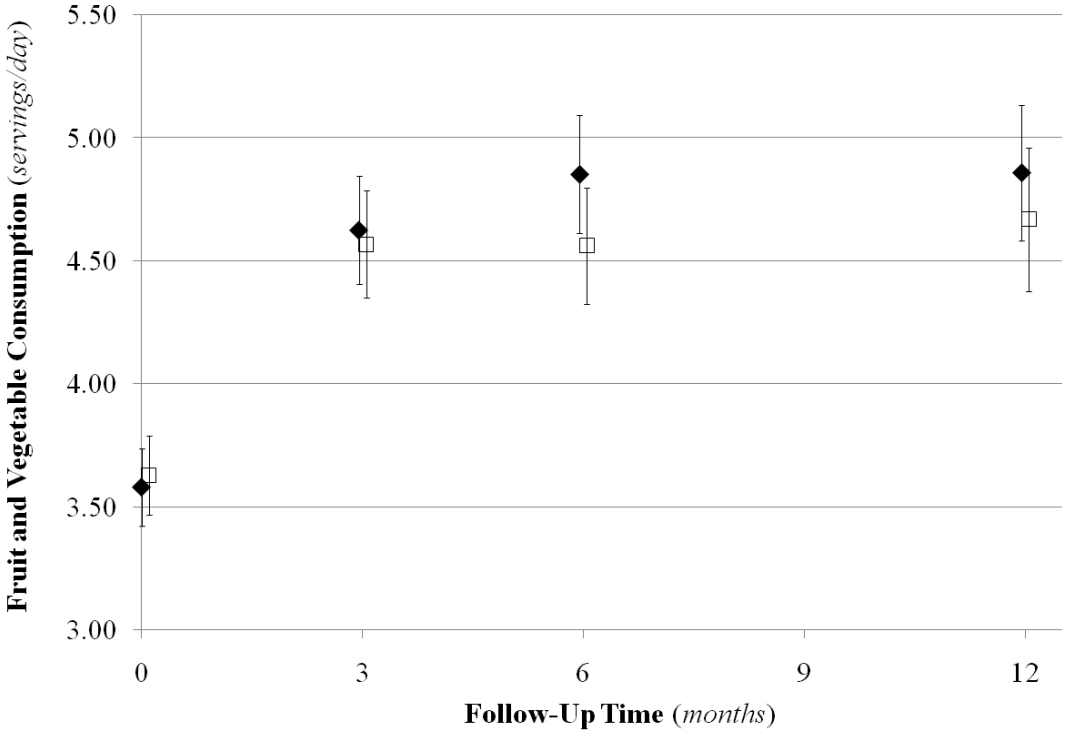
Although average consumption did not significantly differ by study arm at any time-point (**Figure 5.1**), average consumption did significantly increase from baseline to follow-up. Participants from both arms increased consumption by an average of 0.92 servings/day at 3 months (N=648, 95% CI: 0.78, 1.06), 0.97 servings/day at 6 months (N=541, 95% CI: 0.82, 1.13), and 1.13 servings/day at 12 months (N=432, 95% CI: 0.96, 1.31). Despite similarities between study arms at each follow-up, there was a cumulative intervention effect. Based on a linear mixed model examining change from baseline, participants in the intervention arm increased fruit and vegetable consumption by an average of 1.32 servings/day compared to baseline at 12 months (N=228, 95% CI: 1.07, 1.55), which was significantly more than participants in the comparison arm who increased consumption by an average of 0.94 servings/day (N=204, 95% CI: 0.68, 1.19, P=0.04).

Fruit and vegetable consumption was significantly associated with 2 exposure-by-time-by-study arm interactions: diversity of grocers within 8.05 km (P =0.04) and variety of grocers (P =0.02).

Adjusted analysis

The magnitude of changes associated with the interaction of diversity (8.05-km) and variety with time and study arm did not substantially change after adjusting for significant individual- and block group-level covariates. However, p-values were

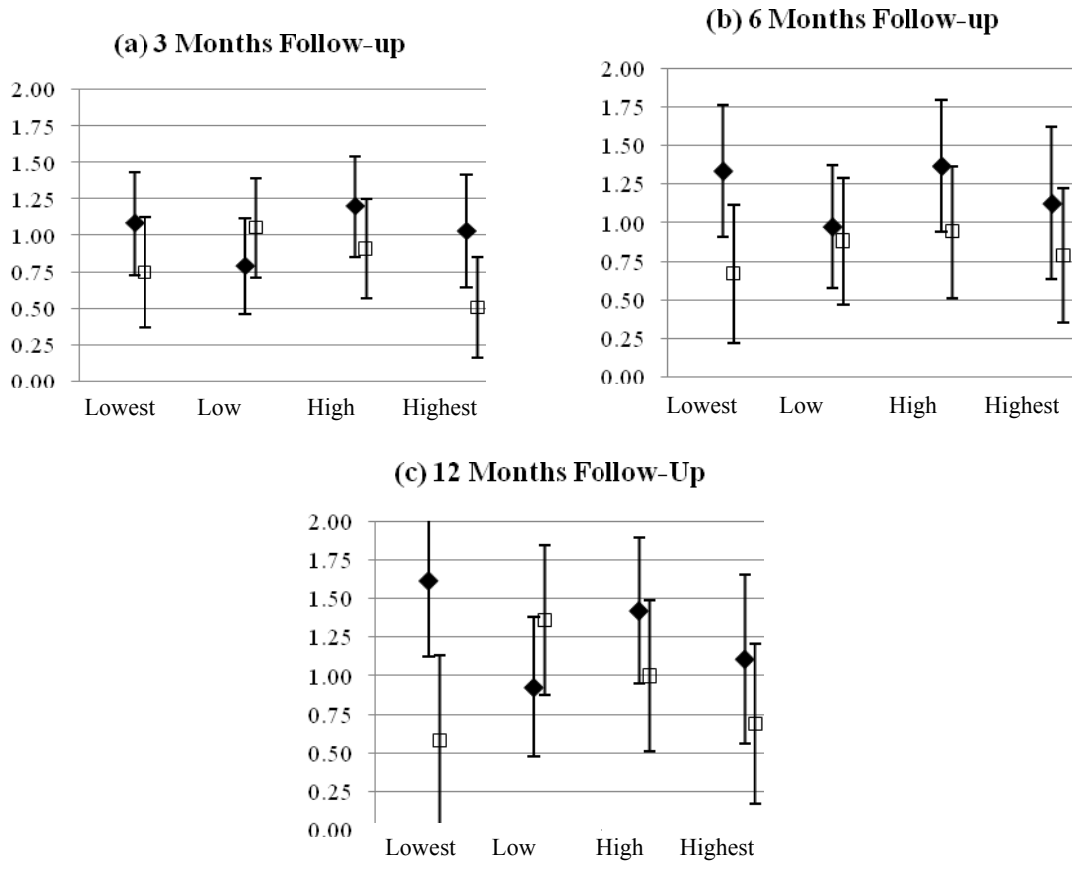
Figure 5.1. Mean fruit and vegetable consumption by Study at Each Follow-up Among Participants in the Nutrition and Physical Activity Trial Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008



Intervention arm (◆) and comparison arm (□)
Bars shown for 95% confidence intervals accounting for random effects of block group and city
Statistically significant increase from intake to 3-, 6- and 12-month follow-ups (*P*s<0.0001).
N=541, 325, 270, and 233 for intervention arm at intake, 3, 6, and 12 months and 524, 330, 277, and 208 for comparison arm

attenuated to 0.08 and 0.05, respectively. No individual- or block group-level covariate-by-time interaction was significantly associated with fruit and vegetable consumption. Thus, fully adjusted models included only first-order terms for individual- and block group-level covariates. Among block group-level covariates, proportion of residents with a bachelors' degree, proportion of black non-Hispanic residents, and proportion Hispanic

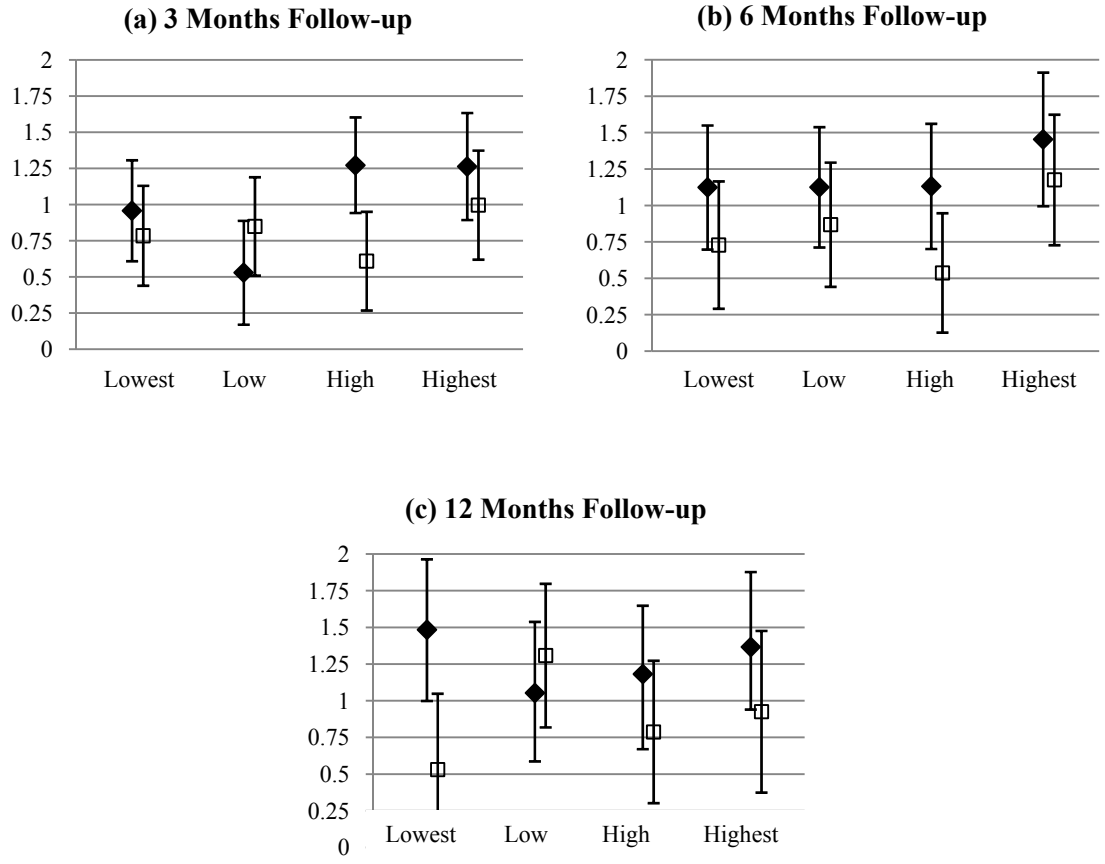
Figure 5.2. Mean Change in Fruit and Vegetable Consumption from Baseline By Diversity of Grocers and Study Arm Among Participants in the Nutrition and Physical Activity Trial Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008



Intervention arm (◆) and comparison arm (□)
Bars shown indicate 95% confidence interval for adjusted mean. Includes random effects for block group and city.
Adjusted for individual-level age, sex, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and block group-level racial composition and education

residents were associated with fruit and vegetable consumption after adjusting for individual-level covariates. City of residence did not explain any differences in average consumption.

Figure 5.3. Mean Change in Fruit and Vegetable Consumption from Baseline By Variety of Grocers and Study Arm Among Participants in the Nutrition and Physical Activity Trial in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008



Intervention arm (◆) and comparison arm (□)
Bars shown indicate 95% confidence interval for adjusted mean. Includes random effects for block group and city.
Adjusted for individual-level age, sex, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and block group-level racial composition and education.

Significant differences in intervention effect were found in the fully adjusted models. At 12 months, participants with the lowest diversity (0 to 6 grocers within 8.05-km) increased fruit and vegetable consumption by 1.61 servings/day (95% CI: 1.12, 2.10) among those in the intervention arm in contrast to 0.58 servings/day (95% CI: 0.09, 1.07) in the comparison arm (**Figure 5.2**). This intervention effect was 1.46 servings/day (95% CI: 0.50, 2.42) greater than the smallest intervention effect, observed among participants with low diversity (7 to 12 grocers within 8.05-km, intervention arm: 0.93 servings/day, 95% CI: 0.48, 1.38; comparison arm: 1.36 servings/day increase, 95% CI: 0.85, 1.87).

Similarly, participants residing in block groups with the lowest variety of grocers (> 4.10 km average to reach 3 differently named grocers) reported significantly greater changes in fruit and vegetable consumption than other participants (**Figure 5. 3**). These participants increased consumption by 1.48 servings/day (95% CI: 1.00, 1.96) in the intervention arm compared to 0.53 servings/day (95% CI: 0.01, 1.05) in the comparison arm. This intervention effect was 1.21 servings/day (95% CI: 0.22, 2.19) greater than participants with a low variety of grocers (2.59 to 4.10 km, intervention arm: 1.05 servings/day increase, 95% CI: 0.57, 1.54; comparison arm: 1.31 servings/day increase, 95% CI: 0.82, 1.80).

Sensitivity analysis

Analysis using a compound symmetric correlation structure for observations from the same participant and analysis restricted to only those who were geocoded to their residence during the first round did not yield materially different results. Empirical estimates could not be obtained due to unequal distribution and small numbers of

observations in each block group. Independent effects of variety and diversity could not be obtained because of the high correlation between these measures (ρ>0.5).

No exposures were significantly associated with loss to follow-up in survival analysis.

DISCUSSION

This is the first study known to us to examine grocer access as a modifier of the intervention effect on fruit and vegetable consumption in a randomized trial. In contrast to the stated hypothesis, the largest intervention effect was observed among participants with the lowest diversity and variety of grocers.

The heterogeneity in the intervention effect cannot be explained by differences in baseline behavior. Based on the observational literature, greater access to grocers should have been associated with higher consumption of fruits and vegetables. Intuitively, these participants are also less likely to further increase consumption at follow-up. This would explain a greater intervention effect among participants with less access to grocers. However, this was not the case in this study. Although participants with high variety of grocers had significantly higher fruit and vegetable consumption at baseline, the greatest difference in intervention effect was observed between those with low and lowest variety of grocers. Furthermore, there was no significant difference in baseline consumption across quartiles of grocer diversity. This lack of a cross-sectional association between consumption and access has been reported in some populations and may be under-reported in others (17).

The magnitude of the differences observed here are meaningful in terms of chronic disease risk reduction and comparable to findings from other randomized trials.

Observational studies report a significant 4% decrease in the risk of coronary heart disease with every 1 serving/day increase in fruit and vegetable consumption and a 6% decrease across all cancers (29, 30). The difference in intervention effect here exceeds that magnitude and is similar to findings from a meta-analysis of dietary advice performed by Brunner et al, who reported an average intervention effect of 1.25 servings/day increase in fruit and vegetable consumption (6). The authors concluded that there was little evidence to support maintenance of intervention effects past 10 months based on the reviewed studies. In contrast, the greatest intervention effect in this study was observed at 12 months follow-up or 6 months after the intervention was completed. This differential maintenance effect cannot be explained by loss to follow-up as there was no difference in participant retention across levels of grocer access.

There are several limitations to this study. First, variety and diversity could not be examined simultaneously as they were highly correlated. Future work may recruit participants stratified on these variables in order to determine the independent effects of variety and diversity which may reflect different aspects of the local food environment. Second, studies comparing commercially available business listings to other data sources highlight the potential for misclassification of businesses across different sources and suggest closer examination to determine which sources are most appropriate for the purposes of research (31, 32). The Dun and Bradstreet Million Dollar Database was used in this study because it is continuously updated and has nation-wide coverage, insuring greater comparability of data quality between the 5 different cities studied. This is in contrast to feasible alternatives which would have required combining data obtained from several different sources such as local health departments and across different time-

points. Third, participants could not be linked to their exact residence due to concerns for confidentiality. The use of block group centroids rather than participant addresses introduces misclassification of the exposure because the focal point from which participant exposure is calculated is shifted. However, the amount of shift is limited by the size of the block groups which, in this study, ranged in size between 0.05 and 158.0 km^2. In contrast, the area contained within an 8.05-km buffer is a much greater 203.6 km^2. The resulting misclassification is likely to be limited and non-differential with respect to the outcome. This would result in a bias toward the null. Fourth, this study does not account for participant re-location during the study. While difficult to quantify since the parent study was not designed to investigate residential exposures, this is also likely to result in a bias towards the null as re-location would likely be non-differential with respect to the outcome. Last, exposures were calculated for 2007 while follow-up occurred between September 2005 and August 2008. This is also likely to result in non-differential misclassification and result in a bias towards the null.

Despite limitations, this study presents a novel method for examining a problematic exposure-behavior relationship. Furthermore, the findings were robust to changes in adjustment variables and correlation structure despite several sources of potential bias towards the null. These findings suggest that there are significant differences in the effect of a fruit and vegetable consumption intervention by access to grocers. Participants who had the lowest access to grocers experienced the greatest benefit from the individual-level counseling intervention. These findings are counter-intuitive given the observational study literature but they are not necessarily contradictory. Together, they suggest a more complex relationship between access and

behavior than previously thought but support the hypothesis that access to outlets to purchase fruits and vegetables is a causal factor in determining individual fruit and vegetable consumption.

REFERENCES

1. Casagrande SS, Wang Y, Anderson C, et al. Have Americans increased their fruit and vegetable intake? The trends between 1988 and 2002. *Am J Prev Med*. 2007;32:257-63.
2. Prevalence of fruit and vegetable consumption and physical activity by race/ethnicity--United States, 2005. *MMWR Morb Mortal Wkly Rep.* 2007;56:301-4.
3. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *JNCI.* 1981;66:1191-308.
4. Dietary Guidelines for Americans. In: Services USDoHaH, Agriculture USDo, eds, 2005.
5. Fruit and vegetable consumption among adults--United States, 2005. *MMWR Morb Mort Wkly Rep*. 2007;56:213-7.
6. Brunner EJ, Rees K, Ward K, et al. Dietary advice for reducing cardiovascular risk. *Cochrane Database of Syst Rev.* 2007:CD002128.
7. Pignone MP, Ammerman A, Fernandez L, et al. Counseling to promote a healthy diet in adults: A summary of the evidence for the U.S. Preventive Services Task Force. *Am J Prev Med*. 2003;24:75-92.
8. Kravitz RL, Duan N, Braslow J. Evidence-Based Medicine, Heterogeneity of Treatment Effects, and the Trouble with Averages. *The Milbank Quarterly.* 2004;82:661-87.
9. Kraemer HC, Frank E, Kupfer DJ. Moderators of Treatment Outcomes: Clinical, Research, and Policy Importance. *JAMA.* 2006;296:1286-9.
10. Kent DM, Hayward RA. Limitations of Applying Summary Results of Clinical Trials to Individual Patients: The Need for Risk Stratification. *JAMA.* 2007;298:1209-12.

11. Oldroyd J, Burns C, Lucas P, et al. The effectiveness of nutrition interventions on dietary outcomes by relative social disadvantage: a systematic review. *J Epidem Comm Health*. 2008;62:573-9.

12. Moore LV, Diez Roux AV, Nettleton JA, et al. Associations of the Local Food Environment with Diet Quality--A Comparison of Assessments based on Surveys and Geographic Information Systems: The Multi-Ethnic Study of Atherosclerosis. *Am J Epidem.* 2008;167:917-24.

13. Rose D, Richards R. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutr.* 2004;7:1081-8.

14. Bodor JN, Rose D, Farley TA, et al. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutr*. 2007:1-8.

15. Caldwell EM, Miller Kobayashi M, DuBow WM, et al. Perceived access to fruits and vegetables associated with increased consumption. *Public Health Nutr.* 2009; In press.

16. Kamphuis CB, Giskes K, de Bruijn GJ, et al. Environmental determinants of fruit and vegetable consumption among adults: a systematic review. *Br J Nutr.* 2006;96:620-35.

17. Pearce J, Hiscock R, Blakely T, et al. The contextual effects of neighbourhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. *J Epidem Comm Health* 2008;62:198-201.

18. Pearson T, Russell J, Campbell MJ, et al. Do 'food deserts' influence fruit and vegetable consumption?--A cross-sectional study. *Appetite*. 2005;45:195-7.

19. Hennekens CH, Buring JE. *Epidemiology in Medicine*. 1st ed. New York, NY: Lippincott Williams & Wilkins, 1987.

20. Vanderweele TJ, Robins JM. Four Types of Effect Modification: A Classification Based on Directed Acyclic Graphs. *Epidemiology.* 2007;18:561-8.

21. United States Census Bureau. American FactFinder. Washington, DC: United States Census Bureau; 2008. (http://factfinder.census.gov/servlet/AGSGeoAddressServlet?_lang=en&_programYear=50&_treeId=420&_sse=on). (Accessed January 2008).

22. Dun and Bradstreet. "D&B's Million Dollar Databases". Short Hills, NJ: Dun and Bradstreet; 2008. (http://www.dnbmdd.com/mddi/last_update.aspx#_mdds). (Accessed March 11, 2009).

23. Dun and Bradstreet. "Million Dollar Directory 2007: America's Leading Public and Private Companies". Short Hills, NJ: 2007.

24. Morland K, Wing S, Diez Roux A, et al. Neighborhood characteristics associated with the location of food stores and food service places. *Am J Prev Med.* 2002;22:23-9.

25. Glanz K, Sallis JF, Saelens BE, et al. Nutrition Environment Measures Survey in stores (NEMS-S): development and evaluation. *Am J Prev Med.* 2007;32:282-9.

26. Zenk S, Schulz A, Israel B, et al. Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit. *Am J Public Health.* 2005;95:660-7.

27. Apparicio P, Cloutier MS, Shearmur R. The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. *Int J Health Geogr.* 2007;6:4.

28. Antonakos C. Nonmotor travel in the 1990 nationwide personal transportation survey. 1995.

29. Joshipura KJ, Hu FB, Manson JE, et al. The Effect of Fruit and Vegetable Intake on Risk for Coronary Heart Disease. *Ann Intern Med.* 2001;134:1106-14.

30. van't Veer P, Jansen MCJF, Klerk M, et al. Fruits and vegetables in the prevention of cancer and cardiovascular disease. *Public Health Nutr.* 2000;3:103-7.

31. Sharkey JR, Horel S. Neighborhood Socioeconomic Deprivation and Minority Composition Are Associated with Better Potential Spatial Access to the Ground-Truthed Food Environment in a Large Rural Area. *J Nutr*. 2008;138:620-7.

32. Larson N, Story M, Nelson M. Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S. *Am J Prev Med.* 2009;36:74-81.

Table 5.1. Characteristics of Participants in the NuPA Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008

	Intervention (N=558)		**Comparison** (N=546)		
Variable	**No.**	**%**	**No.**	**%**	***P***
Age *mean(SD)*[1]	42.7	(10.4)	42.5	(10.9)	0.79
missing	*1*		*7*		
Sex					0.63
Male	104	18.6	108	19.8	
Female	454	81.4	438	80.2	
City					0.83
Atlanta, Georgia	100	17.9	86	15.8	
Austin, Texas	56	10.0	59	10.8	
Chicago, Illinois	54	9.7	51	9.3	
Knoxville, Tennessee	47	8.4	53	9.7	
Minneapolis, Minnesota	301	53.9	297	54.4	
Employed					0.05
Yes	538	96.6	514	94.1	
No	19	3.4	32	5.9	
missing	*1*		*0*		
Marital Status					0.88
Yes	335	60.0	329	60.5	
No	223	40.0	215	39.5	
missing	*0*		*2*		
Education					0.31
HS or less	71	12.7	55	10.1	
Some college/Tech	160	28.7	178	32.6	
College grad	222	39.8	220	40.3	
Post college	105	18.8	93	17.0	
Ethnicity					0.70
White	402	72.7	412	75.6	
Black	95	17.2	87	16.0	
Hispanic	32	5.8	27	5.0	
Other	24	4.3	19	3.5	
missing	*5*		*1*		
Medical History[2]					
Hypertension	115	20.6	95	17.4	0.17
Cancer	43	7.7	36	6.6	0.47
Hypercholesterolemia	160	28.7	124	22.7	0.02

Diabetes[3]					0.20
Type I	1	0.2	4	0.7	
Type II	24	4.3	32	5.9	
Satisfaction with Life					0.10
Yes	401	71.9	416	76.2	
No	157	28.1	130	23.8	
Feeling Sad/Blue					0.29
Yes	45	8.1	35	6.4	
No	513	91.9	511	93.6	

Abbreviations: SD, standard deviation
P-values reported for $\chi 2$ test except where indicated.

1. 2-sample t-test, pooled variance
2. Self-reported physician diagnosis or treatment of condition
3. Fisher exact test

Table 5.2. Access to Grocers By Study Arm Among Participants in the NuPA Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008

		Intervention (N=558)		**Comparison** (N=546)		***P***
Exposure		**No.**	**%**	**No.**	**%**	
Diversity						
Grocers within 0.805 km						0.88
Low	None	479	85.8	467	85.5	
High	One or more	79	14.2	79	14.5	
Grocers within 8.05 km						0.40
Lowest	0 to 6	139	24.9	143	26.2	
Low	7 to 12	153	27.4	142	26.0	
High	13 to 20	148	26.5	127	23.3	
Highest	21 and over	118	21.1	134	24.5	
Variety						
Average distance to 3 different *(km)*						0.96
Lowest	over 4.10	141	25.3	136	24.9	
Low	[2.59, 4.10)	141	25.3	140	25.6	
High	[1.848, 2.59)	150	26.9	141	25.8	
Highest	[0, 1.848)	126	22.6	129	23.6	
Proximity						
Distance to nearest grocer*(km)*						0.82
Lowest	over 2.82	134	24.0	138	25.3	
Low	(1.74, 2.82]	143	25.6	143	26.2	
High	(1.60, 1.74]	152	27.2	135	24.7	
Highest	[0, 1.60]	129	23.1	130	23.8	

P-values reported for $\chi 2$ test

Table 5.3. Fruit and vegetable consumption by access to grocers at baseline Among Participants in the NuPA Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008

Exposure		**Baseline** (N=1,065)		***P***	**Trend**
		Mean	**95% CI**		
Diversity					
Grocers within 0.805 km				0.23	
Low	None	3.67	3.47, 3.88		
High	One or more	3.47	3.13, 3.82		
Grocers within 8.05 km				0.48	0.44
Lowest	0 to 6	3.51	3.24, 3.78		
Low	7 to 12	3.61	3.35, 3.88		
High	13 to 20	3.73	3.46, 4.00		
Higher	21 and over	3.73	3.45, 4.02		
Variety					
Average distance to 3 different *(km)*				0.02	0.84
Lowest	over 4.10	3.66	3.36, 3.96		
Low	[2.59, 4.10)	3.60	3.30, 3.89		
High	[1.848, 2.59)	3.92	3.62, 4.21		
Highest	[0, 1.848)	3.40	3.09, 3.70		
Proximity					
Distance to nearest grocer *(km)*				0.32	0.58
Lowest	over 2.82	3.69	3.41, 3.97		
Low	(1.74, 2.82]	3.76	3.48, 4.03		
High	(1.60, 1.74]	3.66	3.38, 3.94		
Highest	[0, 1.60]	3.46	3.17, 3.75		

Abbreviations: CI, Confidence Interval, km, kilometers
Notes:
Mean consumption at baseline by exposure group using mixed linear models with for random effects for block group and city
P-value shown for difference in mean using mixed linear models, including random effects for block group and city.
Test for trend using mixed linear models, as above, restricting to a linear relationship between quartiles
Conclusions were not substantively different using the original continuous variable.

Restricted to those Geocoded During the First Round

Results. The main analyses for fruit and vegetable consumption were repeated restricting to only those participants who were geocoded to their residence during the first round of geocoding. This included 947 participants (**Table 5.4**). The characteristics of participants who were excluded from the analysis did not greatly differ from those who were included. However, they were more likely to be residents of Atlanta, Georgia, more likely to be black and less likely to be white.

Among those participants included in this analysis, 932 reported fruit and vegetable consumption (FVC) in at least one follow-up. In total 2,342 observations were recorded from these participants. Analyses were adjusted for individual-level age, gender, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and BG-level racial composition and education.

Thus restricted, the adjusted mean changes in FVC by study arm and grocer diversity (**Figure 5.4**) did not substantially differ from the means reported in the standard analysis. For example, at three months in the standard analysis, the mean change in FVC in the comparison arm from lowest to highest quartile of grocer diversity among those in the comparison arm were 0.74, 1.05, 0.91, 0.50 srv/day (**Figure 5.2**a). In the restricted analysis, 3-month changes in FVC in the comparison arm were 0.82, 1.02, 0.73, and 0.58 srv/day (**Figure 5.4a**), only nominally different in magnitude from the standard analysis. The same appeared to be true of the intervention arm at 3, 6, and 12 months, and of the comparison arm at 12 months. However, at 6-months, the restricted analysis seems to indicate a linear increase in FVC with increasing quartile of grocer diversity through the

Table 5.4. Demographic Characteristics of Participants in the Nutrition and Physical Activity (NuPA) Study Residing in Atlanta, GA; Austin, TX; Chicago, IL; Knoxville, TN; and Minneapolis/St. Paul, MN. Contrasting those who were geocoded during the first round [Geolevel Indicator (GLI) = Residence (R)] and those who were not [GLI=block group (B), ZIP code (Z)].

	Included N=947		Excluded N=157		p
Variable	**N**	**%**	**N**	**%**	
Age *mean(sd)*[1]	42.79	10.73	41.27	9.98	0.10
missing	*7*		*1*		
Sex					0.40
Male	178	18.8	34	21.7	
Female	769	81.2	123	78.3	
Metropolitan Area					0.04
Atlanta	147	15.5	39	24.8	
Austin	101	10.7	14	8.9	
Chicago	95	10.0	10	6.4	
Knoxville	88	9.3	12	7.6	
Minneapolis	516	54.5	82	52.2	
Employment Status					0.61
Yes	901	95.2	151	96.2	
No	45	4.8	6	3.8	
missing	*1*		*0*		
Marital Status					0.55
Yes	566	59.9	98	62.4	
No	379	40.1	59	37.6	
missing	*2*		*0*		
Education					0.74
HS or less	109	11.5	17	10.8	
Some college/Tech	292	30.8	46	29.3	
College grad	373	39.4	69	43.9	
Post college	173	18.3	25	15.9	
Ethnicity					0.00
White	713	75.7	101	64.7	
Black	142	15.1	40	25.6	
Hispanic	53	5.6	6	3.8	
Other	34	3.6	9	5.8	
missing	*5*		*1*		
Medical History[2]					

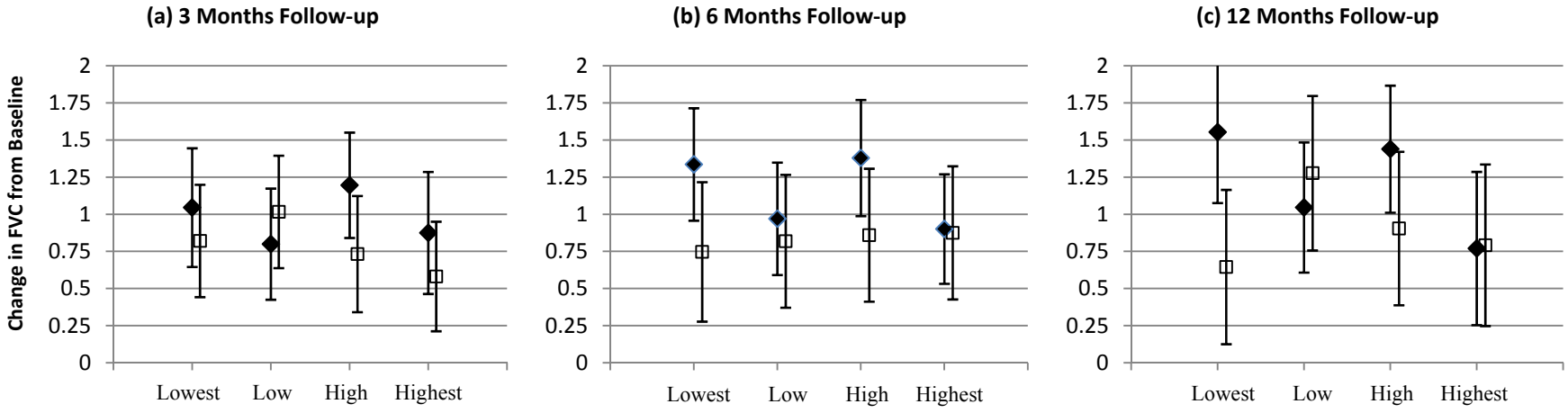
Hypertension	186	19.6	24	15.3	0.20
Cancer	70	7.4	9	5.7	0.46
Hypercholesterolemia	245	25.9	39	24.8	0.78
Diabetes[3]					0.28
Type I	3	0.3	2	1.3	
Type II	49	5.2	7	4.5	
Satisfaction with Life					0.67
Yes	703	74.2	114	72.6	
No	244	25.8	43	27.4	
Feeling Sad/Blue					
Yes	66	7.0	14	8.9	
No	881	93.0	143	91.1	

p-values reported for $\chi 2$ test except where indicated
1. 2-sample t-test, pooled variance
2. Self-reported physician diagnosis or treatment of condition
3. Fisher exact test

high quartile, at which point the increase in FVC becomes level (**Figure 5.4b**). Although this pattern was not observed in the standard analysis, the difference lay in only one data point. While the restricted analysis suggested a steady increase from the lowest through the highest quartile, the standard analysis suggested an increase from the lowest through the high quartile, and then a drop in the highest quartile.

The restricted analysis of grocer variety (**Figure 5.5**) also suggested only minimal differences from the standard analysis (**Figure 5.3**). Relative patterns remained the same among those randomized to the comparison arm and mean change in FVC remained below 1 srv/day with the exception of the highest quartile at 3-month and 6-month follow-ups, and the low quartile at 12 months. This was true of both the standard and restricted analyses. Among those randomized to the intervention, patterns of FVC change at 3- and 12-month remained similar between the two analyses. However, the

Figure 5.4. Change in fruit and vegetable consumption (FVC, in servings/day) from intake by grocer diversity within 8.05-km. Restricted to 2,342 observations where participants were geocoded to their residence during the first round.

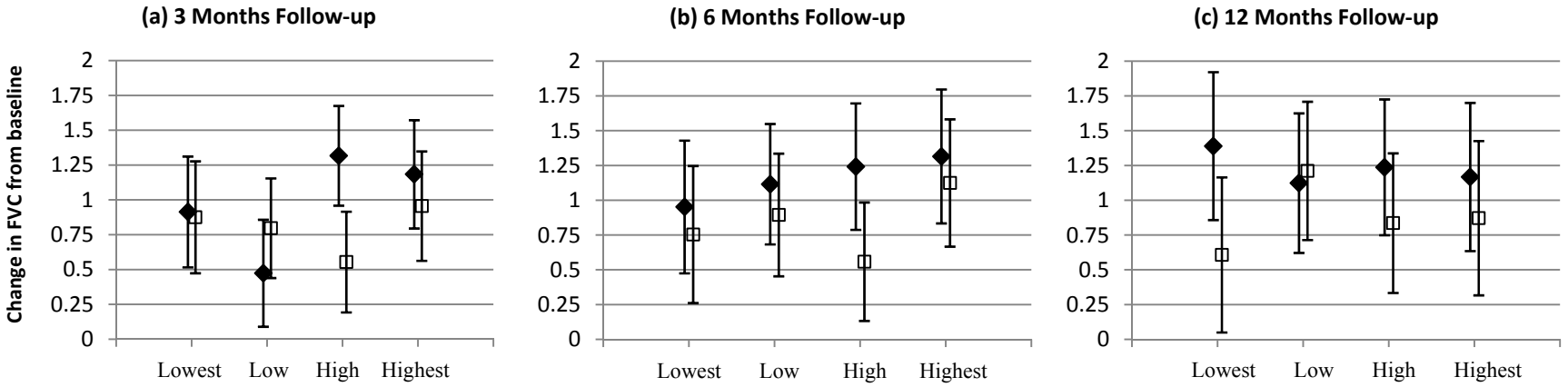


Intervention arm (◆) and comparison arm (□)

Bars shown indicate 95% confidence interval for predicted mean

Adjusted for individual-level age, gender, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and BG-level racial composition and education

Figure 5.5. Change in fruit and vegetable consumption (FVC, in servings/day) from intake by grocer variety (p=0.07). Restricted to 2,342 observations where participants were geocoded to their residence at the first round.



Intervention arm (◆) and comparison arm (□)

Bars shown indicate 95% confidence interval for predicted mean

Adjusted for individual-level age, gender, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and BG-level racial composition and education

steady increase in change in FVC with increasing grocer variety from the restricted analysis (lowest to highest quartile: 0.95, 1.12, 1.24, 1.32 srv/day) was not an obvious replication of the pattern observed in the standard analysis (lowest to highest quartile: 1.13, 1.13, 1.12, 1.45). The different is largely explained by the higher change in FVC in the lowest quartile of grocer diversity.

Taken together, the relative differences in change in FVC by study arm across levels of grocer diversity were not meaningfully changed when the analyses were restricted to those who were geocoded during the first round. Also of note, at 12 months, the largest difference in change in FVC across study arms, or intervention effect, was still observed in the lowest quartile of both grocer diversity and variety.

Conclusions. These analyses were undertaken in order to assess the potential effect of geocoding errors on the main research question. As such, these analyses included only participants for whom we are more certain of the correct geocoding linkage from the mailing address to the block group . While this serves to decrease the effect of potential misclassification, it also serves to decrease the sample size with which the analysis can be performed. Although patterns in FVC change in some study arm-by-grocer diversity or variety groups were slightly different in the restricted analysis, the conclusions from these analyses do not differ from those drawn from results in the standard analysis. These results are consistent with a greater intervention effect among those in the lowest quartile of grocer diversity and variety in comparison to participants in the low, high, and highest quartiles.

Imputation Analysis

Results. Missing values for fruit and vegetable consumption (FVC) were imputed using linear regression for each follow-up among participants with complete baseline demographic data (N=1,086). Five complete datasets with 4,344 observations each (4 evaluation points x 1,086 participants) were imputed (**Table 5.5**).

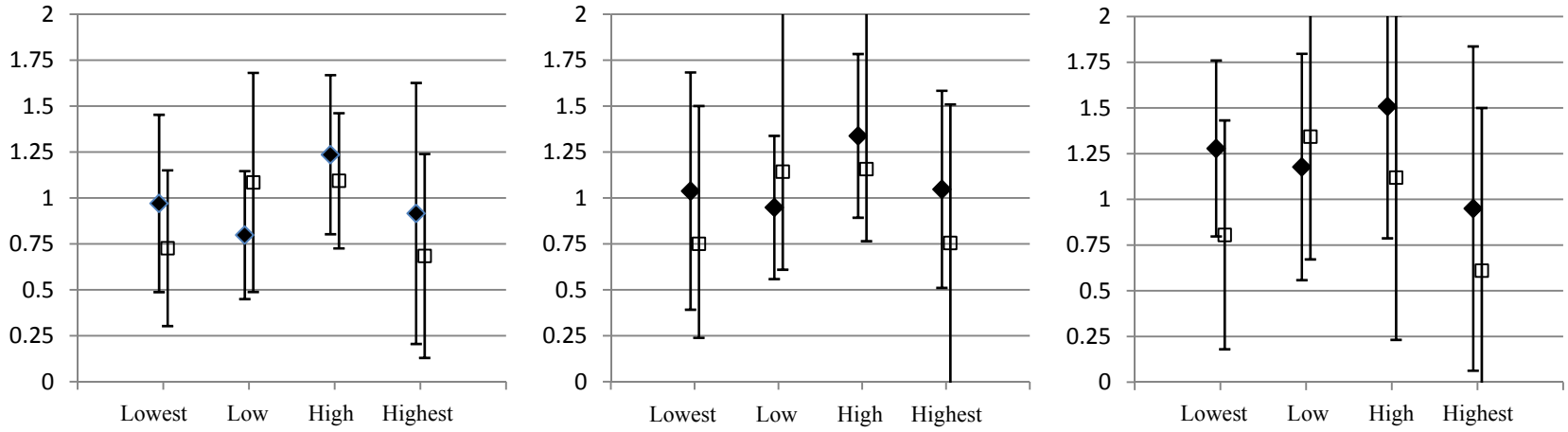
Since the data were not monotonically missing, collected data were made to be monotonic by initially omitting data that were collected after 1 or more missed follow-ups. Using this monotonic dataset containing 1, 360 observations, 1,919 missing data points were filled in regressing on all individual-level covariates (age, gender, education, ethnicity, satisfaction with life, city, employment status, marital status, hypercholesterolemia, diabetes, hypertension, cancer, whether the participant felt sad/blue) and select environmental- or block group-level covariates (proportion with a bachelors degree, proportion black non-Hispanic, proportion Hispanic) and grocer variety and diversity (8.05-km). For follow-up evaluations after baseline, all previous fruit and vegetable consumption, whether imputed or originally collected, were also included in the regression.

Table 5.5. Description of collected, monotonic, imputed, and replaced data points from the 1,086 participants with complete baseline demographic data. Total data points (per dataset) = 4,344 (4 follow-up time points x 1,086 participants).

Time	Collected	Monotonic	Used Imputed	Replaced
Baseline	1,065	1,065	21	0
3 Months	655	648	431	7
6 Months	547	434	539	113
12 Months	441	278	645	163
Total data points:	*2,708*	*1,360*	*1,636*	*283*

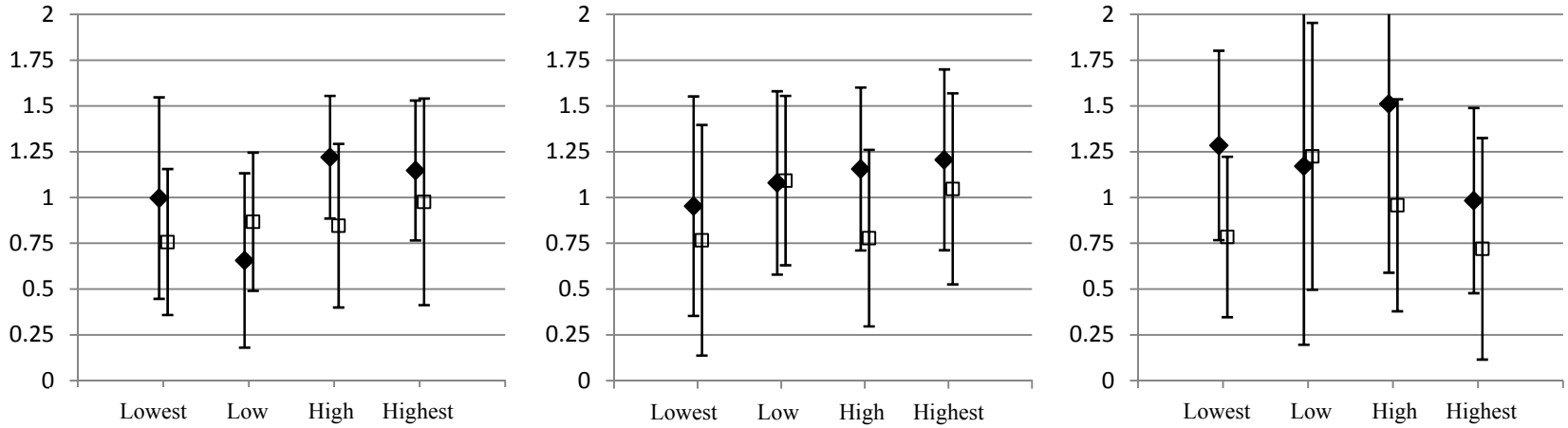
Adjusted analysis as described in the main study was repeated for grocer variety and diversity within 8.05 km (5 miles) to determine the potential effect of loss to follow-

Figure 5.6. Change in Fruit and Vegetable Consumption (FVC) from Baseline by Diversity of Grocers and Study Arm Among Participants in the Nutrition and Physical Activity (NuPA) Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008. Change from baseline at **(a)** three months, **(b)** six months, and **(c)** twelve months in servings per day.



- Intervention arm (◆) and comparison arm (□)
- Bars shown indicate 95% confidence intervals for adjusted mean
- Adjusted for individual-level age, gender, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and block group-level racial composition and education
- Missing values imputed regressing on all individual-level covariates (age, gender, education, ethnicity, satisfaction with life, city, employment status, marital status, hypercholesterolemia, diabetes, hypertension, cancer, whether the participant felt sad/blue) and select environmental- or block group-level covariates (proportion with a bachelors degree, proportion black non-Hispanic, proportion Hispanic) and grocer variety and diversity (8.05-km). For follow-up evaluations after baseline, all previous fruit and vegetable consumption, whether imputed or originally collected, were also included in the regression.

Figure 5.7. Change in Fruit and Vegetable Consumption (FVC) from Baseline by Variety of Grocers and Study Arm Among Participants in the Nutrition and Physical Activity (NuPA) Study in Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota; September 2005-August 2008. Change from baseline at **(a)** three months, **(b)** six months, and **(c)** twelve months in servings per day.



- Intervention arm (◆) and comparison arm (□)
- Bars shown indicate 95% confidence intervals for adjusted mean
- Adjusted for individual-level age, gender, ethnicity, education, marital status, employment status, chronic conditions, previous cancer diagnosis, satisfaction with life, feeling sad/blue, and block group-level racial composition and education
- Missing values imputed regressing on all individual-level covariates (age, gender, education, ethnicity, satisfaction with life, city, employment status, marital status, hypercholesterolemia, diabetes, hypertension, cancer, whether the participant felt sad/blue) and select environmental- or block group-level covariates (proportion with a bachelors degree, proportion black non-Hispanic, proportion Hispanic) and grocer variety and diversity (8.05-km). For follow-up evaluations after baseline, all previous fruit and vegetable consumption, whether imputed or originally collected, were also included in the regression.

up on the estimation of change in FVC from intake until follow-up by study arm-by-PA environment categories.

In contrast to the standard analysis, there was no statistically significant three-way interaction between study arm, time, and grocer diversity or variety (*P*s > 0.05) in imputed analyses.

In the analysis of grocer diversity, this is likely because the direction of change in fruit and vegetable consumption remained the same while some differences in FVC change became attenuated. For example, in the standard analysis for change in FVC by grocer diversity at 3 months (**Figure 5.2a**), the difference between intervention arm and comparison arm were approximately 0.34 srv/day and -0.26 srv/day in the quartile with lowest and low diversity of grocers, respectively. However, in the imputation analysis (**Figure 5.6a**) the differences in change in FVC in the same quartiles are 0.24 sr/vday and -0.29 srv/day, respectively. Similar patterns were observed at 6 and 12 months (**Figures 5.6b** and **5.6c**).

The same cannot be said about the analysis of grocer variety. While there was not a significant difference between the standard analysis and the imputed analysis in the direction of FVC change at 3 months, patterns were substantially different at 6 and 12 months. In the standard analysis change in FVC among those in the comparison arm were (from lowest to highest quartile) 0.78, 0.85, 0.61, and 1.00 srv/day (**Figure 5.3a**). Similarly, in the imputation analysis (**Figure 5.7a**), the changes were 0.76, 0.87, 0.85, and 0.98 srv/day. Among participants in the intervention arm, at 3 months results from standard analysis and imputed analysis were similar as well. However, at 6 months, the imputed analysis seems to indicate a steady increase in change in FVC with increasing

variety of grocers among those in the intervention arm. This pattern was not observed in the standard analysis. Furthermore, the changes in FVC among those in the comparison arm were almost identical in the imputed analysis (from lowest to highest quartiles) at 0.77, 1.10, 0.78, and 1.05 srv/day and in the standard analysis at 0.73, 1.10, 0.78, and 1.05 srv/day.

Conclusions. Formal statistical tests using the imputed data did not identify significantly different intervention effects across groups of participants with different diversity and variety of grocers in their local environments. However, the magnitude and direction of changes in FVC were similar to that found in the standard analysis. This, in combination with findings from analysis of survival time of loss of contact or refusal to participate (**Chapter 6**), suggests that findings from the main analysis are not strongly affected by missing values.

Misclassification of grocer businesses by Block Group-Level Demographic Characteristics

Rationale. On-site ascertainment of businesses, or ground-truth data, is the most direct method of determining the location of businesses. However, the majority of epidemiologic research involving such data has employed the use of commercially available databases.[1-4] This is likely due to limited resources and the need to obtain such data for a large geographic area.

Despite the widespread use of such databases, little research has been performed to quantify the magnitude and/or direction of potential misclassification of businesses either through omission or inclusion of businesses, or errors in geocoding. An alternative

to on-site assessment would be to use online data sources such as search engines, publicly available aerial photographs, etc. However, these would be redundant with the methods used in creating the commercial database.[5] Furthermore, the reliability and consistency of those data by themselves cannot be assured. Although ESRI updates their database annually and D&B updates continuously using information from online search engines, print resources, and business listings, a more direct, real-world comparison is needed to insure valid measurements for epidemiologic research.

This project assessed the degree of completeness and validity of data from commercial databases of grocers in comparison with data obtained from on-site ascertainment to determine whether any discrepancies are related to area-level demographic characteristics. This information is critical to determine whether adjustment is necessary in the analysis of the association between access to grocers and individuals behavior due to differential misclassification of the exposure. Previous research has found that the presence of parks and grocery stores is associated with not only individual behavior but also with area-level socio-economic and racial characteristics. Therefore, to tease out the effect of the built environment from that of socio-economic factors related to neighborhood, this research will include both socio-economic measures and built environment characteristics.

Methods. A sampling scheme was developed to identify block groups (BGs) to be assessed in each of the five metropolitan statistical areas (MSAs) or cities included in the main analysis (Atlanta, Georgia; Austin, Texas; Chicago, Illinois; Knoxville, Tennessee; and Minneapolis/St. Paul, Minnesota).

Table 5.6. Description of block groups in each Metropolitan Statistical Area included in the assessment.

MSA	Total BGs	BGs Sampled	Median Values		
			Median Income	Business Density	Minority Population
Atlanta, GA	1,923	68	$54,000	53.3	41.4%
Austin, TX	765	72	$50,000	90.3	36.2%
Chicago, IL	6,589	67	$55,000	151.4	38.4%
Knoxville, TN	422	62	$40,000	29.8	6.9%
Minneapolis, MN	2,172	65	$61,000	64.1	11.3%

Each of the 11,871 BGs across the five MSAs (**Table 5.6**) were categorized at the median of city-specific business density (from the US Economic Census 2002), proportion of minorities (black, Hispanic, other estimated in the Claritas Update Demographics for 2007 based on the US Census 2000), median income (estimated for 2007 from 2000 Census), and the presence of urban areas or clusters (referred to here as urbanicity, US Census 2000) – all of which are factors believed to be related to the presence of businesses. This created 16 strata, from each of which a maximum of eight block groups (BGs) were sampled in a given MSA. All BGs with missing values for demographic variables were also included as a separate category (**Table 5.7**).

A driving route was designed for each BG using MapSource Version 6.10.2 and uploaded into a Garmin eTrex GPS unit. Any parks or grocery stores that were found while driving the route were recorded with a waypoint using the GPS unit (long/lat, WGS 1984), and additional data such as time, waypoint accuracy (feet), type (park or grocery store), and name were recorded during on-site data collection (**Table 5.8**).

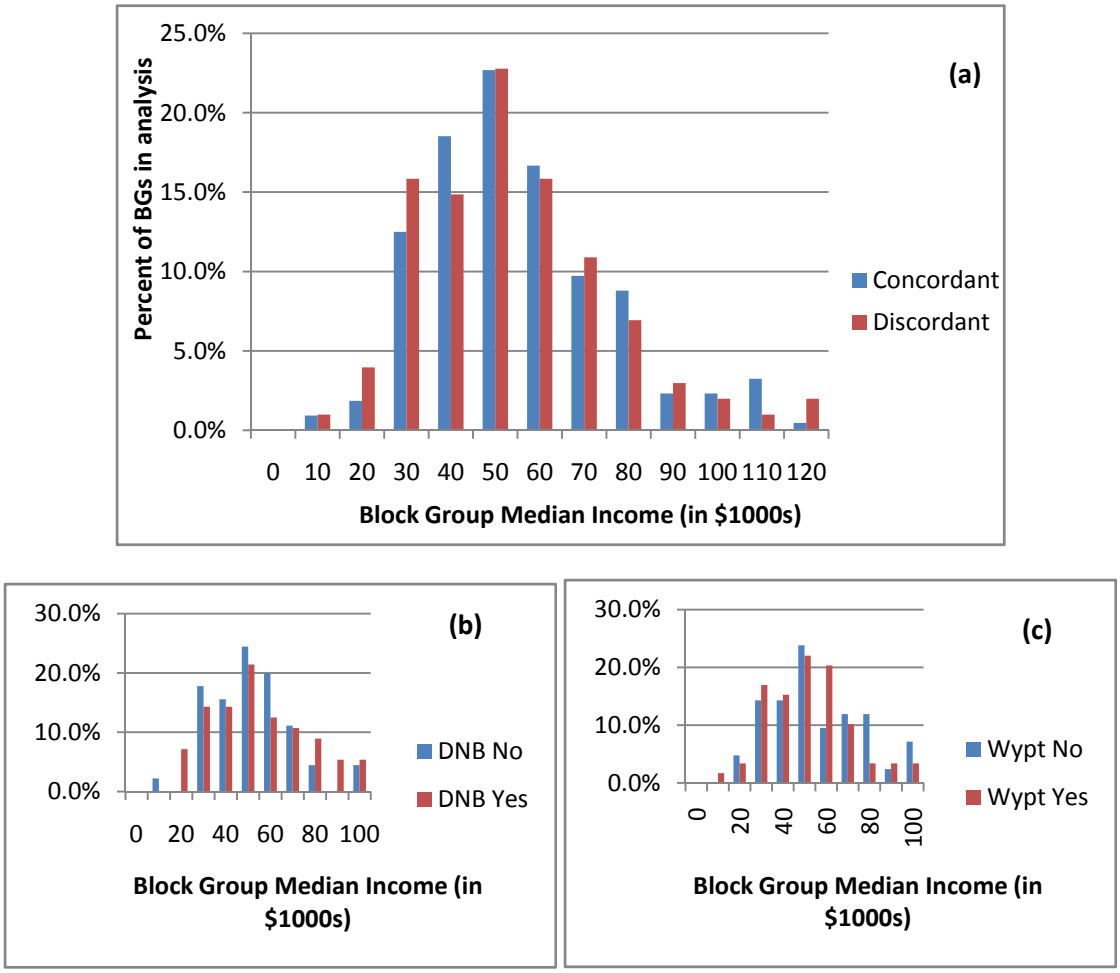
Table 5.7. Combinations and frequencies of urbanicity, and high or low income, business density, and minority population. All block groups with missing values for the group value had missing values for median income. Of these, 13 also had missing information on minority population. There were no missing values for urbanicity and business density. One BG in Atlanta, 4 in Austin, 6 in Chicago, 2 in Knoxville, and 4 in Minneapolis were included in the *missing* category.

Group	Urban	Median Income	Business Density	Minority Population	N
1	No	Low	Low	Low	30
2	No	Low	Low	High	13
3	No	Low	High	Low	3
4	No	High	Low	Low	30
5	No	High	Low	High	4
6	Yes	Low	Low	Low	29
7	Yes	Low	Low	High	29
8	Yes	Low	High	Low	29
9	Yes	Low	High	High	30
10	Yes	High	Low	Low	30
11	Yes	High	Low	High	30
12	Yes	High	High	Low	30
13	Yes	High	High	High	30
Missing					17
Total					**334**

Table 5.8. Dates of on-site assessments. Since the Dun and Bradstreet Million Dollar Database is updated continuously, updated versions of the database were obtained based on when data were collected.

MSA	Assesment Dates	D&B Download Date
Atlanta, GA	June 3, 4, 13-15; Nov. 2, 2008	June 30, 2008
Austin, TX	April 22-26, 2008	May 5, 2008
Chicago, IL	July 27-29, 2008	August 11, 2008
Knoxville, TN	June 5-7, 2008	June 30, 2008
Minneapolis, MN	July 22-25, 2008	August 11, 2008

Figure 5.8. (a) Concordance: Distribution of median income by concordance between D&B and waypoints data. **(b) Completeness:** Distribution of income among those discordant between D&B and waypoints data by presence of waypoint-only data. **(c) Validity:** Distribution of income among those discordant between D&B and waypoints data by presence of D&B-only data.



Each of the 334 sampled BGs was evaluated for the presence of grocers within 500 ft of the BG border using, separately, the D&B database and the way points collected.

Those BGs where either D&B or waypoint data indicated there was a grocer present were further examined to determined whether the business(es) listed in the two data sources were identical by name. BGs were then categorized as concordant on D&B and waypoint data or not concordant. Those that were not concordant were further

categorized into groups having D&B listings of businesses not found in the waypoint assessment (D&B-only businesses) and/or vice versa (waypoint-only businesses). Simple comparisons of means, medians, or proportions were made and both linear and logistic regressions were used to examine the independent associations between demographic or BG-level variables and the concordance, completeness (fewer waypoint-only businesses, presence indicated by wypt_pos) and validity (fewer D&B-only businesses, presence indicated by DNB_pos) of the D&B database.

Results. Of the 334 BGs sampled, 232 were completely concordant (N=232, 69.5%). However, this was largely due to BGs where neither data source indicated the presence of a grocer (N=219, 65.6% overall, 94.4% of BGs that were concordant). Of the remaining 102, there were 16 BGs where D&B and waypoints shared at least 1 business listing but had other unshared listings (partially concordant), 56 BGs had a business listed in D&B that was not found during on-site assessment (validity), and 60 BGs had a business not listed in D&B that was found during assessment (completeness). These categories were not mutually exclusive.

Overall, median income, urbanicity and minority population of the BG were not associated with concordance, completeness, or validity of the D&B data source. However, BGs with a high business density were more likely to contain businesses in D&B that were not found during data collection. This may be a function of both the limitations in collecting waypoint data as well as a true difference in the validity of those business listings in the DNB database.

Median income was normally distributed and was not statistically significantly different between those BGs that were concordant and those that were not (concordant

(N=216, mean: $55,693), discordant (N=101, $54,360); p=0.61, total N not equal to 334 due to missing values in median income, **Figure 5.8a**). Further subsetting to those who were not concordant and examining by completeness (complete (N=59, mean=$51,795) vs. incomplete (N=42, mean=$57,962, p=0.23) and validity (valid (N=56, mean=$55,858) vs. in-valid (N=45, mean=$52,495), p=0.47), there was still not a significant difference. Simultaneously examining the associations between completeness, validity, and concordance and further accounting for correlation between BGs sharing the same county and MSA, there were still no significant associations with median income.

Among 334 BGs, only 82 (24.6%) did not include an urban area or cluster. Although this was not statistically significant in unadjusted analysis, 10% more urban BGs than non-urban BGs were discordant (23.2% among non-urban, 32.9% among urban, χ^2=2.8, p= 0.10). There was no significant relationship between completeness and urbanicity. However, there was a significantly higher proportion of urban BGs that had businesses listed in DNB that were not found during data collection (60.2%) than among non-urban businesses (31.6%). These differences were not significant using logistic regression accounting for shared county and MSA. Similar analysis DNB_pos did not demonstrate a statistically significant relationship between urbanicity and validity accounting for correlation between BGs sharing county and MSA, likely due to non-overlap between groups and small numbers.

Business density ranged between 0 and 4638 businesses/mi^2. These densities were dichotomized for each city with values below the median. BGs with a higher business density were less likely to be concordant. However, subsetting to those that were discordant showed no significant relationship between business density completeness, but

a significant association between high business density and potentially decreased validity (among those not concordant, low business density (N=46, 35.7% invalid) vs. high business density (N=56, 64.3% invalid), $\chi 2=4.4$, p=0.04). This difference in validity remained of borderline significance in logistic regression (p=0.05) where BGs with a higher business density were more likely to contain potentially invalid business listings in the DNB database after accounting for concordance and completeness.

Minority population ranged from 0 to 100% of the BG population. Again, city-specific medians were used to categorize BGs as having a low or high minority population. BGs that had a higher minority population were slightly less likely to be concordant (high minority population (N=137, 51.5% concordant) vs low minority population (N=184, 61.4% concordant), $\chi 2=4.7$, p=0.03). Among BGs that were discordant, there was no association between minority population and validity (overall 55.5% invalid) or completeness (overall 41.6% incomplete). However, this association did not remain when using logistic regression with random effects for state and county, examining independent effects of concordance, completeness, and validity.

Conclusions and Implications. Although there is substantial misclassification of grocers, this misclassification is largely non-differential with respect to area-level demographic characteristics. Thus, adjustment for misclassification is probably unwarranted as any bias is likely to result towards the null.

However, these conclusions must be considered in light of some important limitations. Although the block groups assessed in this study were stratified by 4 area-level demographic characteristics (business density, median income, urbanicity, and racial composition) other factors (renters vs. owner-occupant residences, median age)

could also be related misclassification businesses. Due to limitations in time and financial resources, not all strata demographic characteristic could be examined. Furthermore, because not all block groups in the five cities were assessed, conclusions may not necessarily be generalizable to all block groups. However, because the sample was a simple random sample stratified by demographics, it is unlikely that the sampled block groups are substantially different from those that were not sampled. Furthermore, because the block groups included in this analysis are sampled from the five cities of interest, results may be generalizable to any study performed in these studies.

References

1. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 2006;117(2):417-24.
2. Moore LV, Diez Roux AV, Nettleton JA, Jacobs DR, Jr. Associations of the Local Food Environment with Diet Quality--A Comparison of Assessments based on Surveys and Geographic Information Systems: The Multi-Ethnic Study of Atherosclerosis. Am. J. Epidemiol. 2008;167(8):917-924.
3. Zenk S, Powell L. U.S. secondary schools and food outlets. Health & Place 2008;14:336-346.
4. Powell L, Slater S, Mirtcheva D, Bao Y, Chaloupa F. Food store availability and neighborhood characteristics in the United States. Prev Med 2007;44:189-95.
5. Dun and Bradstreet Million Dollar Directory. Cross Reference Geography, 2007.

Chapter 6:

Effects of Individual- and Block Group-Level Demographics and Built Environment on Loss to Follow-up

Results

As discussed in the methods, loss to follow-up was conceptualized as occurring through two mechanisms: missed contact (where the evaluator is unable to make contact with the participant), and refusal to participate (where the evaluator is able to make contact with the participant and the participant refuses further follow-up).

Missed Contact

Overall, 456 participants or 41.3% missed at least one follow-up evaluation and 553 completed all evaluations through the 12 month-follow-up.

The data were tested among all 1104 participants at the alpha=0.05 level to determine whether random effects would be necessary in modeling survival time data. Assessed separately, there was no significant random effect of metropolitan statistical area (MSA, $0.21<p<0.98$, 5 MSAs) or block group (BG, $0.99<p<1.00$, 905 BGs) in an otherwise unadjusted model using NLMIXED.

Among the 1104 participants in the analysis dataset, 14 were excluded from adjusted analyses because of missing covariate data. Each potential covariate was evaluated for association with follow-up time using Kaplan-Meier curves Log-Rank test (proc lifetest) and assessed for meeting the proportional hazards assumption using the graphical approach (parallelism of LLS curves). The significant findings are summarized in **Table 6.1**.

Other individual-level covariates that were assessed but that were not statistically significant included gender, hypercholesterolemia, diabetes (type I or type II), feeling sad/blue, and employment status. BG-level covariates included urbanicity, median income, median age, % Hispanic population, unemployment rate, and % with HS/GED equivalent.

Table 6.1. Summary of direction of association, p-value, and whether or not the variable met the PH assumption among variables that were found to be significantly associated with survival using unadjusted Kaplan-Meier curves (α=0.05). For the description of the direction of association, those categories that are shown together in parentheses showed approximately equivalent survival curves. KM curves are attached at the end of this file.

Level (Indiv- vs. BG) Variable	Direction	p	Met PH*
Individual			
Intervention Arm	Greater survival among counseling	0.006	Y
Age (Quartiles)	Increased age, increased survival	<0.0001	Y
Ethnicity	In order of increasing survival: Hispanic, Black, (Other, White)	0.002	N
Education	Increased education, increased survival	0.01	Y
Marital Status	Greater survival among those married	0.03	Y
Tobacco Use	Greater survival among non-smokers	0.05	Y
Cancer	Greater survival among those with cancer dx	0.004	Y
Blood Pressure	Greater survival among those with high BP	0.03	Y
Satisfied with Life	Greater survival among those satisfied	0.007	Y
Block Group			
MSA	In order of increasing survival: Knoxville, (Chicago, Austin), (Minn., Atlanta)§	0.006	N
% white residents	High %, increased survival	0.001	N
% black residents	High %, decreased survival	0.01	N
% w/bachelors degree	High %, increased survival	0.05	N

* proportional hazards assumption
§ Reported highest survival to lowest. Categories shown in parentheses are approximately equivalent in observed survival function.

Individual-level variables that met the PH assumption and were statistically significant were entered into a Cox model together. Thus examined, only age, education, intervention arm, and satisfaction with life were independently associated with time until missed contact. Marital status, tobacco use, previous cancer diagnosis, and high blood pressure were no longer statistically significant. Individual-level variables that had not met the PH assumption when unadjusted were put into the reduced adjusted model to assess for the PH assumption using Schonfield residuals after adjusting for other variables. Ethnicity was the only such variable, and it was determined to meet the PH assumption after adjustment for other variables (p=0.80), and was statistically significantly associated with time until event (p=0.03). The final individual-level

model included age, education, ethnicity, study arm, and satisfaction with life. This individual-level model was then used to examine BG-level covariates.

For BG-level covariates, each variable was first assessed in combination with its corresponding individual-level variable to determine whether the BG-level variable was likely to be serving as a proxy for the individual-level characteristic. For example, individual-level education was highly significantly associated with proportion of residents in BG with a bachelors degree (χ^2_9=83.0, p<0.0001) with participants with greater educational attainment residing in BGs where the residents are also more highly educated. Using a Cox model, both variables met the PH assumption when examined together, but neither was statistically significant. Thus, proportion of those in the BG with a bachelors degree was removed in favor of keeping the individual-level education variable.

Similarly, proportion white and black residents were both significantly associated with individual-level ethnicity (χ^2_9=312.0, 310.0, respectively, ps<0.0001). In a Cox model with individual-level ethnicity and both BG-level variables, all three satisfied the PH assumption but the BG-level covariates were not independently associated with survival time.

Thus, no BG-level covariates were entered into the model after adjusting for individual-level analogs. Only MSA remained which, when entered into the reduced individual-level model, did not violate the PH assumption (p=0.1371). Thus for the final model, age, ethnicity, education, study arm, satisfaction with life and MSA were examined in the linearized portion of the model. Including MSA in the analysis, neither MSA (p=0.1046) nor ethnicity (p=0.1573) were associated with survival time.

Using the adjusted model, all exposure variables were assessed for meeting the PH assumption and significance. All met the PH assumption and none was statistically significantly associated with time until missed contact. Since no exposures were significantly associated with survival time, the final model included only intervention arm, ethnicity, education, age, and satisfaction with life (**Table 6.2**).

Table 6.2. Multiple model estimates for missed contact with and without stratification by city (N=1090 individuals).

Model:	Uncorr[1]			LIFEREG[2]			NLMIXED no re†[3]			NLMIXED BG re†[4]			NLMIXED MSA re†[5]		
Variable	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI	
Intervention Arm															
Counseling	0.79	0.65	0.95	0.86	0.78	0.96	0.89	0.80	0.99	0.86	0.78	0.96	0.86	0.78	0.96
Self-Help	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
Education															
HS or less	0.95	0.76	1.20	1.24	1.05	1.47	1.22	1.02	1.46	1.24	1.05	1.47	1.24	1.05	1.47
Some college	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
College grad	1.40	1.03	1.89	0.97	0.86	1.11	1.01	0.89	1.15	0.97	0.86	1.10	0.97	0.86	1.10
Post college	0.88	0.65	1.18	0.93	0.79	1.09	1.03	0.87	1.21	0.93	0.79	1.09	0.93	0.79	1.09
Age															
Q1: [19, 33]	1.65	1.26	2.15	1.37	1.18	1.60	1.45	1.24	1.70	1.37	1.18	1.60	1.37	1.18	1.60
Q2: [34, 42]	1.30	1.00	1.70	1.18	1.02	1.38	1.20	1.02	1.40	1.18	1.02	1.38	1.18	1.02	1.38
Q3: [43, 50]	0.95	0.72	1.26	0.98	0.83	1.14	0.98	0.83	1.15	0.98	0.83	1.14	0.98	0.83	1.14
Q4: [51, 80]	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
Ethnicity															
Black	1.40	1.11	1.79	1.24	1.08	1.42	1.31	1.15	1.50	1.24	1.08	1.42	1.24	1.08	1.42
Hispanic	1.31	0.89	1.94	1.19	0.95	1.48	1.14	0.90	1.44	1.19	0.95	1.48	1.19	0.95	1.48
Other	1.12	0.72	1.76	1.05	0.82	1.35	1.10	0.86	1.42	1.05	0.82	1.35	1.05	0.82	1.35
White	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
Satisfied with life															
No	1	1.03	1.56	1	1.03	1.31	1	1.04	1.32	1	1.03	1.31	1	1.03	1.31
Yes	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
-2LogL	6047.83			1866.80			3528.40			3524.30			3524.3		

1. SAS proc phreg, no random effects

2. SAS proc lifereg, no random effects

3. SAS proc nlmixed, no random effect

4. SAS proc nlmixed, random effects for each of 905 block groups

5. SAS proc nlmixed with random effects for each of 5 MSAs

Regardless of the estimation method and inclusion of random effects, participants had a decreased hazard for having been lost to follow-up if they were in the intervention group, if they had any college education or more, were in the 3rd or 4th quartile of age (over 42), if they were White, Hispanic, or categorized themselves "Other" for race/ethnicity, or were satisfied with life. This was true for both correlated and uncorrelated models.

Refusal

Of 1104 participants, 145 (13.1%) refused to participate at some point during the trial. One hundred and sixty were censored after intake because they were not contacted for any other follow-ups. Only 5 variables were significantly associated with the rate of refusal (**Table 6.3**).

Table 6.3. Summary of direction of association, p-value, and whether or not the variable met the PH assumption among variables that were found to be significantly associated with survival using unadjusted Kaplan-Meier curves (α=0.05). For the description of the direction of association, those categories that are shown together in parentheses showed approximately equivalent survival curves. KM curves are attached at the end of this file.

Variable	Direction*	p-value	Met PH assumption?
MSA	(Chi ,Min), (Atl, Aus), Kno	0.0003	Y
Age (Quartiles)	1,4,3,2	0.04	Y
Gender	Male, Female	0.06	Y
Employment	Yes, No	0.004	Y
Chain grocer within ½ mi	Yes, No	0.04	Y

*Reported highest survival to lowest. Categories shown in parentheses are approximately equivalent in observed survival function.

When entered into the Cox model together, all variables met the PH assumption.

Adjusting for all variables, the number of chain grocers within 0.805 km and gender were no longer significant at the 0.05 level. The final reduced model contained fixed effects for MSA, employment status, and age. The results from standard uncorrelated analysis as well as correlated analysis adjusting for MSA or BG are presented in **Table 6.4**.

Table 6.4. Multiple model estimates for refusal to participate with and without stratification by city. N=1095 individuals.

Model:	Uncorr			LIFEREG			NLMIXED no re			NLMIXED BG re			NLMIXED MSA re		
		95% CI			95% CI			95% CI			95% CI			95% CI	
Variable	HR	Lower	Upper	HR	Lower	Upper	HR	Lower	Upper	HR	Lower	Upper	HR	Lower	Upper
CBSA															
Atlanta	1.37	1.10	1.72	1.31	1.11	1.54	1.30	1.11	1.53	1.26	1.07	1.50	1.30	1.11	1.53
Austin	1.11	0.82	1.50	1.12	0.90	1.39	1.10	0.89	1.37	1.16	0.93	1.44	1.10	0.89	1.37
Chicago	0.71	0.49	1.04	0.88	0.67	1.15	0.94	0.73	1.22	0.86	0.67	1.11	0.94	0.73	1.22
Knoxville	2.49	1.96	3.15	1.98	1.66	2.36	1.97	1.65	2.35	1.98	1.62	2.42	1.97	1.65	2.35
Minneapolis	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
Age															
Quartile 1	0.85	0.64	1.13	0.88	0.72	1.07	0.88	0.72	1.07	0.86	0.70	1.04	0.88	0.72	1.07
Quartile 2	1.62	1.30	2.03	1.44	1.23	1.70	1.43	1.22	1.68	1.49	1.26	1.77	1.43	1.22	1.68
Quartile 3	1.27	1.00	1.60	1.17	0.99	1.39	1.17	0.99	1.38	1.16	0.98	1.38	1.17	0.99	1.38
Quartile 4	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
Employment															
Yes	1	--	--	1	--	--	1	--	--	1	--	--	1	--	--
No	2.13	1.55	2.93	1.83	1.45	2.32	1.84	1.46	2.31	1.86	1.43	2.42	1.84	1.46	2.31
-2LogL		1892.08			945.81			1417.60			1399.80			1417.6	

The findings are similar regardless of model. Participants from Atlanta and Knoxville were significantly more likely to refuse to participate in comparison to those from Minneapolis while participants from Chicago and Austin were not significantly different from those from Minneapolis. Furthermore, those in the middle 2 quartile of age were more likely to refuse than those who were the youngest or oldest. And lastly, those who were not employed were more likely to refuse than those who were employed.

Notably, environmental exposures to opportunities for healthy choices and BG-level characteristics were not associated with the rate of refusal.

Conclusions

In conclusion, it appears that the potential effect of informative loss to follow-up in the main analyses should be performed based on imputation regressing on city of residence, age, education, study arm, ethnicity, employment status and satisfaction with life. No quantitative adjustment for environment seems necessary as there was no differential loss to follow-up by environmental factors. However, differential loss to follow-up across characteristics of the environment *within* the intervention arm should also be examined in future analyses.

Based on the findings here, missing values of fruit and vegetable consumption were imputed based on baseline values of all individual-level covariates (age, gender, education, ethnicity, satisfaction with life, city, employment status, marital status, hypercholesterolemia, diabetes, hypertension, cancer, whether the participant felt sad/blue) and select environmental- or block group-level covariates (proportion with a bachelors degree, proportion black non-Hispanic, proportion Hispanic) for fruit and vegetable consumption (described in Chapter 3 and 5). However, in the analysis of fruit and vegetable consumption, grocer variety and diversity (8.05-km) were also used in the imputation regression as they were associated with the outcome under study. No imputation analyses were undertaken for the analysis of physical activity.

Chapter 7:

Effect Modification in Randomized Trials of Behavioral Interventions for Increasing Fruit and Vegetable Consumption

A Systematic Review of Effect Modification in Randomized Trials of Behavioral Interventions for Increasing Fruit and Vegetable Consumption

Di H. Cross, K.M. Venkat Narayan, Roberd M. Bostick,

Lance A. Waller, Youngmee Kim, Rachel E. Patzer, Carol J. Rowland Hogue

Abstract

Introduction. Increased interest in differential treatment effects, also known as moderation or effect modification, is particularly poignant among randomized controlled trials (RCT) of behavioral interventions given the limited success of current methods in changing behavior. This review examines reported differential treatment effects in behavioral trials aimed at increasing fruit and vegetable consumption (FVC) with particular emphasis on motivations for assessing treatment effects and recommendations for future applications of this methodology.

Methods. The literature review included all behavioral, non-pharmaceutical interventions designed to increase fruit and vegetable consumption published from 1990-2008. Interventions involving policy or environmental modification were excluded. Study quality was evaluated by examining sample size, randomization success, loss to follow-up, statistical analysis, instrument and methods for evaluation of FVC, and blinding.

Results. After removing duplicates from 8,327 citations from MedLine, CINAHL, EMBASE, PsychINFO, and the Cochrane Registry of Clinical Trials, 468 full text articles were retrieved. Of 162 articles reporting fruit and vegetable consumption, 28 (17.9%) articles discussing 23 trials reported investigating potential modification of the intervention effect. Demographic variables [education (N=13), sex (N=12), age (N=11), race/ethnicity (N=8)] and psychosocial variables [stage of change (N=10)] were most commonly investigated. Studies varied in type of intervention, magnitude of effects, and duration of follow-up. Heavy reliance on statistical significance as a criterion for identifying effect modification led to few studies concluding any such modification occurred.

Discussion. Although calls for assessing effect modification have been increasing, its application is not reflected in current practice. In our review, trials that examined effect modification focused on demographic and/or psychosocial characteristics, often without explanation for this analysis. While this approach may serve to validate behavioral models or to identify sub-groups for subsequent targeting, as often reported by the authors, the approach has other important uses

including identification of potentially modifiable characteristics to improve intervention efficacy on a population level. One example of this is environmental factors, such as access to grocers. As a result of our review, we recommend that researchers examine and report effect modification in intervention trials and broaden the search for potentially modifying variables, using the observational literature as a guide. Furthermore, while important, statistical testing should not be the only means by which effect modification is assessed. Improved understanding of modifying variables will serve to inform the design of interventions and improve efficiency in the allocation of public health resources for both services delivery and research activities.

Introduction

There is little question that chronic disease risk is related to health behaviors such as fruit and vegetable consumption (FVC)[1]. However, poor health behaviors are still pervasive. Less than a quarter of the US population engages in FVC at levels recommended for chronic disease prevention [2]. Furthermore, intervention trials to increase FVC have had limited success [3], highlighting the need to identify methods of improving interventions.

One area that has received increasing attention is the identification of effect modification [4-6], also called moderation in the behavioral literature [6], and intervention- or treatment-effect heterogeneity in the randomized trial literature [7,8]. Effect modification, as it will be referred to in this article, is a difference in the association between a variable, E, and an outcome, O, across different levels of an effect modifier, X [9]. This is different from confounding of the relationship between E and O by X, where X is a predictor of the outcome, O, and also associated with the variable E under study [9,10]. In this latter case, the association between E and O across levels of X is the same. Confounding and effect modification are also distinct from mediation, studied in the behavioral and psychology fields, where E causes X, and X in turn causes the outcome, O [6].

Effect modification has typically been studied using two methods. The first is stratified analysis where the association between E and O is reported for each level of X [9]. Meaningful differences in this association yield a conclusion that effect modification is present. The second method is the examination of statistical interaction. Here, a statistical model is employed to study variation in the dependent variable O using E, X, and the product term between E and X as independent variables. In this method, effect

modification is defined as a statistically significant contribution of the product term between E and X [11].

While both methods have been used to investigate randomized trials, there has not been a systematic evaluation of the motivation for such analyses and the subsequent knowledge garnered from them. The purpose of this review is threefold: (1) to document published reporting of effect modification in behavioral trials with specific regard for interventions to increase fruit and vegetable consumption; (2) to examine ways in which the information obtained from analysis of effect modification has been used and (3) to make recommendations to take further advantage of such information.

Methods

Search Strategy

A search strategy (**Appendix 6**) was adapted from previous systematic reviews to identify controlled trials of behavioral interventions to increase fruit and vegetable consumption by [3,12-14]. Results were restricted to studies performed among adult, non-institutionalized populations from the CINAHL, MedLine, EMBASE, and PsychInfo databases and the Cochrane Central Register of Controlled Trials for reports published between January 1, 1990 and December 31, 2008.

Assessment of Eligibility

The titles and abstracts of all articles identified from the search strategy were evaluated by two reviewers (DHC, REP). Eligibility was determined based on study aims, study population, and intervention type. Studies were excluded if the intervention was a pharmaceutical intervention, if there was an environmental modification or policy

change component of the intervention, if the intervention was performed exclusively among smokers, and if the study was designed for participants with specific chronic disease conditions (hypercholesterolemia, diabetes or impaired glucose tolerance, hypertension, cancer survivors, etc.) except for overweight/obese. Studies that recruited from the general population and included some participants with chronic diseases were not excluded from the analysis. No restrictions were made based on participant recruitment or intervention delivery setting. Although most studies were randomized trials (either individual- or cluster-randomized), studies with only pre- and post-intervention measurements and no control group were not excluded.

Full texts of eligible studies were retrieved and examined to determine whether individual-level fruit and vegetable consumption were reported for baseline and follow-up evaluations in servings/day, grams/day, or energy-adjusted daily intake, and whether any modification of the intervention effect was examined. Articles reporting only on differences in intervention effect across levels of baseline behavior or levels of intervention dose were not included in the analyses. References from included articles were examined to identify additional eligible articles. Articles that reported only baseline demographic information, baseline FVC, or described study design and motivations of otherwise eligible intervention were entered into the ISI Web of Science database (Thomson Reuters, New York, NY) to determine whether they had been cited by subsequent articles that reported on follow-up data and met the inclusion criteria.

Data Extraction

Citation information, intervention design and setting, study population and sample size, a description of the intervention and control (if applicable), evaluation time-points,

intervention effect, effect modifier examined, and the direction of the effect modification were extracted from all eligible studies. The direction of effect modification was denoted as (0) for no evidence of effect modification, (+) for a greater intervention effect associated with a greater value of the effect modifier or a greater intervention effect in the non-reference category in comparison to the reference category, or (-) for a smaller intervention effect associated with a greater value of the effect modifier or a smaller intervention effect in the non-reference category of the effect modifier in comparison to the reference category. For articles using statistical interaction to assess effect modification, statistical significance as reported by the authors was used to determine the presence of effect modification. Among articles employing stratified methods, intervention effects with a difference greater than 2 standard errors of each other were recorded as meaningful effect modification.

Information reported in the text identifying different articles originating from the same study (study name, sample size, intervention description, etc.) were noted and study characteristics of such articles were reported together. Study quality or risk of bias was assessed for each article by examining six sources of potential bias: participant blinding, assessment tools for fruit and vegetable consumption, randomization success, sample size, loss to follow-up and the appropriateness of statistical analyses. Since participant blinding to the intervention was not possible, and because all assessments of FVC were self-reported, the evaluation of the potential risk of bias focused four sources of bias: randomization success, sample size, loss to follow-up, and statistical analysis. Studies likely to result in bias from 2 or more of the above sources were considered to have a high risk of bias, while those with 1 potential source were considered to have a medium

risk of bias. Articles unlikely to result in bias from any of the four sources were categorized as having a low risk of bias.

No attempt was made to examine un-published studies or studies published in conference proceedings. Because of the diversity of study characteristics, and the variability in the effect modifiers measured, no attempt was made to arrive at any summary measures of association.

Results

After removal of duplicates from 8,327 citations from MedLine, CINAHL, EMBASE, PsychINFO, and the Cochrane Registry of Clinical Trials, 468 full text articles were retrieved, of which 162 reported fruit and vegetable consumption (FVC). Among those reporting on FVC, 28 articles (17.3%) reported examining effect modification, with data coming from 23 individual trials (**Table 7.1**) [15-42]. Notable studies excluded were the Working Well Trial [43], the Healthy Directions – Small Business Study [44], and WellWorks-2 [45] which all included components of environmental or policy change; High-5 [46], Gimme 5 [47], CATCH [48], and Take-5 [49] which examined fruit and vegetable consumption behavior in children; and one study examining intervention effect across levels of rurality of study sites but which was performed among recipients of congregate meals [50]. Eight studies were rated to be at low risk of bias, while 16 were rated medium, and 4 were rated as having a high risk of bias (**Table 7.1**).

Among the 23 trials, 16 were individually-randomized trials while 7 were cluster-randomized trials. Five trials reported no significant intervention effect at the end of the trial [29,40-42,51]. Of 16 trials reporting a statistically significant intervention effect, six

reported an intervention effect of 0.5 srv/day or less. One trial did not compare the randomized groups and thus reported only change in fruit and vegetable consumption in the dietary intervention group [15], and one trial reported a smaller decline in fruit and vegetable consumption in the intervention group than in the control group [39].

A total of 39,515 participants were recruited for participation (N range: [32, 5,041], median: N=1,359) while loss to follow-up ranged between 0.0% [15] and 73.9% [27]. Study settings included churches, clinics or primary care facilities, health management organizations, the internet, telephone directories, community-based organizations, and worksites. The majority of studies were performed in clinical settings(N=12, 52.2%). Study evaluation ranged from under a month [42] to 24 months or two years [17,19,21,51] and intervention duration ranged from those involving only a one-time exposure [29,39,42] to those administered over a period of 24 months [51]. With only 2 exceptions [28,37], interventions with evaluations at or beyond 12 months had larger intervention effects. Eleven trials were conducted among at-risk populations (low-income, minority, etc.).

The majority of articles reporting on potential effect modifiers focused on demographic variables (**Table 7.2**) with 13 studies reporting examination by education, 12 by sex, 11 by age, and 8 by race/ethnicity. Despite the number of articles examining differences by demographic variables, few studies concluded that there was a meaningful difference by education [16,27,29], sex [15,30], age [16,19,23,27,40], or race/ethnicity [16]. All studies reporting such differences were rated for assessment of quality with a medium or low risk of bias, while none with high risk of bias reported effect modification. In addition, the few studies that used stratified analysis indicated differences in intervention effect more

often than studies employing statistical significance of a product term to evaluate effect modification.

Following demographic variables, psychosocial variables were also commonly examined. Stage of change was the most commonly examined effect modifier, evaluated in 8 articles [17,23,26-28,32,37,42] covering 10 separate studies with meaningful differences in intervention effect reported in 5 studies [23,26,27,32,37]. However, the direction of association was inconsistent. Three studies reported increase in intervention effect with increasing stage of change (from precontemplator to maintenance)[23,27,32] and two reported the reverse [26,37]. There were generally no difference reported by baseline intention to change FVC behavior (examined in two studies [18,23]), and only one [36] of three studies examining baseline motivation [24,36,39] reported a significant difference in intervention effect. In addition, one study examined intervention effect by need for cognition [38] and another examined intervention effect by self-efficacy [34]. Neither found differential intervention effects. One study examining autonomy showed a significantly greater intervention effect with greater baseline autonomy [40].

In addition to demographic and psychosocial variables, studies also examined social-contextual variables and health-related indicators, albeit more rarely. One study examined social networks, social norms and food security and reported no difference in intervention effect observed across levels of these variables [34]. A few studies examined smoking status [16,23,27,31] and baseline BMI status [23,39,51] of which only one study reported a significant difference with a smaller intervention effect among smokers [27]. One study found a greater intervention effect with increasing level of participant food responsibility, but no difference by number of restaurant meals consumed [23].

Investigator-cited justification for examining effect modification varied by the effect modifiers examined. Those studies examining psychosocial effect modifiers commonly cited validation of behavioral models as the motivation for the analysis [17,26,28,32,42]. Examination of demographic variables, when explicitly stated, was motivated in determining the generalizability of results [20,33,41], identifying subgroups for subsequent targeting of the intervention [27,41], determining public health importance or impact [30], or simply because such effect modification had been reported in previous studies [25]. One study, performed by Sorensen et al, cited findings from observational studies as the motivation for examining effect modification [34]. For a large proportion of articles examined (N=10, 35.7%), there was no explicitly stated motivation for examining effect modification [15,16,18,19,21,23,29,35,37,51].

Discussion

Despite increasing interest in studying heterogeneity in intervention effects [4,5,7], there are few articles that report on such findings in the behavior intervention trial literature. To the authors' knowledge, only one other review article, conducted by Oldroyd et al., has examined heterogeneity in the effect of interventions to increase FVC [52]. There was no overlap in the articles included by Oldroyd et al and in this study due to difference in inclusion criteria – particularly the inclusion of studies conducted among children [46,53,54]. The authors concluded that few articles reported examining intervention effect modification, consistent with our finding that fewer than one out of every five studies reported examining effect modification.

It is possible that these analyses are more commonly performed than they are reported, likely because the resulting associations do not reach statistical significance.

Many studies included in this analysis did indeed report no statistically significant effect modification. Barring the circumstances where a homogeneous intervention effect may be observed for an intervention designed for and administered within a specific sample, such a lack of heterogeneity may be counter-intuitive given the conventional wisdom that "one size does not fit all" [8,55]. This suggests that we are either wrong in our assumption that interventions should have heterogeneous effects, or wrong in the method by which heterogeneity is identified and quantified.

Indeed, this review suggests that because analysis of effect modification is generally conducted as part of a secondary aim rather than as a primary aim of an intervention study, less reliance on statistical significance and more careful examination of the magnitude of intervention effects is more likely to lead investigators to accurately conclude that there is meaningful modification of the intervention effect. Lack of statistical significance likely results as the number of participants included in intervention trials is determined based on statistical power to detect a significant overall intervention effect not power to detect interaction effects. As such, it is well known that stratified analysis is likely to be under powered [56]. Furthermore, stratified analysis introduces multiple testing and an increased likelihood of making obtaining false positive results [5]. These reasons are often stated as support for the alternative method: examining the statistical significance of the product term in a statistical model [5]. However, the power for detecting such an effect is dependent on the proportion of participants simultaneously exposed to both the intervention and the potential modifier, and simply including a product term in the model often is not a panacea for under-powered hypotheses. In addition, multiple testing remains a problem given the number of potential modifiers that

are commonly examined. As such, both methods – examining effect modification through a statistical interaction operationalized as a product term in the model, or through stratified analysis – offer similar disadvantages in traditional trial designs. However, while the direct statistical testing possible with the analysis of statistical interaction is often stated as an advantage of that method, the use of a single criterion of statistical significance – particularly in a situation where power is nearly guaranteed to be less than adequate – over-simplifies the picture. The same cannot be said for stratified analysis, which demands examination of magnitudes of the intervention effects across sub-groups since direct comparison is not made in the analysis. In light of limited statistical power, stratified analysis may be more interpretable, more accessible to broader audiences, and more likely to lead to conclusions useful for informing future studies.

However, beyond merely identifying effect modification, there is also a question of motivation. Why are such analyses important? What do they tell us about the intervention? What, if anything, do they tell us about the effect modifier?

Although demographic variables were the most commonly studied effect modifiers, few authors explicitly stated a motivation for examining them. In addition, it is likely that many more investigators have examined effect modification by demographic variables but have not reported their results due to apparent null findings. Thus, these results were not included in the literature and hence not addressed in this review. This lack of reporting may be explained by the motivation behind the investigation of effect modification. The most commonly cited reasons for investigating demographic variables were to identify subgroups among which the intervention had a differential effect, or to determine whether the estimated intervention effect was generalizable. If no such

differences are found, then there is no other reason to report effect modification. However, if differences are identified, efficient allocation of resources may create changes in behavior in particular subgroups. This may not address the public health problem of pervasive, poor health behaviors since sub-populations among whom interventions are less effective, which often includes underserved, or high-risk populations, would still be at high risk after resources are expended. Perhaps motivated by these findings, targeted interventions have been designed for such populations – often with moderate effects. However, beyond targeting sub-groups, little else has been done with the information garnered from effect modification across demographic characteristics. Indeed, this would appear to make sense since these characteristics are not modifiable. However, investigators should be careful, as demographic characteristics may merely serve as proxies for other modifiable characteristics in a more complex causal mechanism [57].

Among investigators of psychosocial modifiers, the most commonly cited motivation was testing or validating a behavioral model. In this case, modifying variables are of primary interest – not only to identify groups that will benefit from the intervention in its current form, but also to improve the design of interventions in the future by understanding the mechanisms by which the intervention under investigation did and did not work.

This latter approach is commendable, but has not been taken full advantage of in the literature since it has been restricted to psycho-social variables. While behavioral models may be a resource for identifying variables that may predict behavior, increasingly recognized is the importance of factors related to behavior *change* [58]. In

addition to, individuals' motivations and abilities to execute a behavior, external variables such as opportunities for behaviors are also important in changing behavior [58,59]. One source for identifying such factors is the observational literature. However, only one study included in this review [34] examined environmental-level characteristics, explicitly citing the observational literature as the motivator for examining these potentially modifying factors.

Investigators must make better use of the observational literature as a resource for identifying modifiable, environmental-level variables which are potentially modifiers the effect of behavioral interventions. This would aid in designing studies and allocating resources for addressing the public health problem of pervasive, poor health behaviors for entire populations. In much the same way that physicians may first address a modifying risk factor before prescribing a drug to an individual with elevated risks of adverse effects or attenuated benefits related to the risk factor, public health policy-makers may address modifiable, environmental risk factors on a population level before administering other individual- or policy-level.

We recommend that investigators more consistently report when effect modification is examined, even when findings are not statistically significant. Furthermore, we recommend that statistical significance should be only one criterion by which meaningful effect modification is identified and reported, and we recommend that patterns and trends in intervention effect or differences in magnitude be given more weight even in the absence of statistical significance. Such reporting will document apparent associations and allow for more focused follow-up with appropriately powered studies for studying interaction or subset effects. Most importantly, the choice of

variables for which effect modification is examined should be guided by specific causal hypotheses, making it possible not only to allocate resources to populations where interventions will be most effective, but also making it possible to change population-level characteristics in order to make interventions effective in all segments of the population. This can most effectively be accomplished by incorporating the knowledge obtained from the observational literature with data obtained from randomized trials.

Table 7.1. Description of included studies

Trial Name	Citation	Risk of Bias	Design	Setting	Population (Loss to f/up)	Intervention	Control	Evaluations	Intervention Effect
	Gambera et al, 1995 [15]	Medium	RCT	Air Force Base in Sacramento, California	32 active-duty Air Force members (0.0%)	90 days; dietary counseling, fitness program	Fitness only	0, 90 days	Significant 3.4 srv/day increase in FVC among those in diet+fitness at 90 days; control group consumption not reported
Maryland WIC 5-a-Day Promotion Program	Havas et al, 1998 [16] Campbell et al, 2008 [17]	Low Medium	cRT	16 WIC sites, Maryland	3,122 mothers over 18 years of age (24.5, 26.3%)	6 months; group sessions, print material, personalized mailings	Delayed intervention	0, 8, 20 months	Significant 0.4 srv/day greater increase in intervention compared to controls at 20-month follow-up
	Marcus et al, 1998 [18]	Medium	RCT	Callers to 6 NCIS Centers	2,126 callers to six regional call centers (39.5%)	Tailored educational message, concrete behavioral suggestions, print materials	No materials	0, 4 weeks, 4 months	Significant 0.7 srv/day greater increase among intervention than controls at 4 weeks, 0.4 srv/day greater increase at 4 months
Black Churches United for Better Health	Campbell et al, 1999 [19] Campbell et al, 2008 [17]	Medium Medium	cRT	50 churches in 10 rural, eastern North Carolina counties	3,737 African-American participants (32.6, 33.0%)	20 months; Tailored bulletins, lay health advisors, activities, pastor and community support, etc.	Delayed intervention	0, 1, 2 years	Significant 0.66 srv/day greater increase in intervention than controls at 2 year follow-up
Women's HealthTrial	Coates et al, 1999 [20]	High	RCT	Clinics in Atlanta, GA; Birmingham, AL; and Miami, FL	2,208 post-menopausal women, aged 50-79 (73.1%)	2 years; multiple group counseling sessions	Self-help material	0, 6, 12, 18 months	0.6 to 0.8 srv/day significantly greater increase among intervention than controls
Treatwell 5 a Day	Sorensen et al, 1999 [21]	Low	cRT	22 worksites in eastern	1,359 employees	19.5 months; worksite intervention:	Minimal: NCI cancer	0, 2 years	Significant 0.5 to 1.2 srv/day greater

	Campbell et al, 2008 [17]	Medium		Massachusetts	(3.9, 8.2%)	educational sessions, activities, plus minimal intervention; worksite plus family intervention	hotline, 1hr nutritional education session		increase in worksite intervention than control at 2 years
Next Step Trial	Tilley et al, 1999 [22]	Low	cRT	28 automobile manufacturing sites in 5 US states	5,042 employees (30.9%)	2 years; classes, print materials, personalized dietary feedback	No materials	0, 1, 2 years	No significant difference in change at 1 or 2 years
Puget Sound Healthy Eating Patterns Study	Kristal et al, 2000 [23] Satia et al, 2001 [24]	Medium Medium	RCT	Group Health Cooperative of Puget Sound HMO, Washington State	1,459 enrollees, age 18-69 (17.4, 17.4%)	Print materials, dietary analysis, motivational phone call	No materials	0,3,12 months	Significant 0.5 srv/day greater increase among intervention compared to controls at 3 and 12 months
Eat for Life Trial	Resnicow et al, 2001 [25] Resnicow et al, 2003 [26]	Medium Medium	cRT	14 black churches, Atlanta metropolitan area	1,011 participants (14.8, 14.8%)	Group 2: self-help intervention and 1 cue call; Group 3: self-help intervention, 1 cue call, 3 counseling calls	Delayed intervention	0, 1 year	1.0 to 1.4 srv/day greater increase in group 3 than controls; 1.0 to 1.1 srv/day greater increase in group 2 than controls
Maryland WIC Food for Life Program	Havas et al, 2003 [27]	Medium	cRT	10 WIC sites, Maryland	2,066 mothers (73.9%)	6 months; various educational media, activities, individualized feedback, incentives, and phone calls	Delayed intervention	0, 8, 20 months	Significant 0.4 srv/day greater increase in intervention compared to controls at 8 months
	John et al, 2003 [28]	Low	RCT	2 primary care centers, Oxfordshire, UK	729 participants aged 25-64 (5.3%)	6 months; negotiation to change behavior, print materials, 2-week self-monitoring diary	Delayed intervention	0, 6 months	Significant 1.4 srv/day greater increase in intervention than controls
	Block et al, 2004 [29]	Low	RCTl	Community-based organizations, California	491 African American or Non-Hispanic White, age 40-65, low income, female	One-time experience; Group One: Little-by-Little CD-ROM; Group Two: Little-by-Little CD-ROM plus two reminder phone calls	Stress management CD-ROM	0, 2 months	No significant difference in change intake

					(2.0%)				
Diabetes Prevention Program	Mayer-Davis et al, 2004 [30]	Low	RCT	27 clinical centers throughout the US	2,934 adults (9.3%)	Lifestyle modification; written materials, annual review with case manager, at least monthly contact with interventionist	Standard care, placebo	0, 1 year	Significant >1.5 srv/day greater increase in intervention compared to placebo
	Steptoe et al, 2004[31] Perkins-Porras et al, 2005 [32]	Medium Low	RCT	Primary care center, South London, UK	271 participants aged 18-70, low-income (19.6, 19.6%)	2 weeks, behavioral counseling	Nutrition education	0, 8 weeks, 12 months	Significant 0.6 srv/day greater increase in intervention compared to controls at 12 months
Healthy Directions – Health Centers Study	Emmons et al, 2005 [33] Sorensen et al, 2007 [34]	Low Medium	RCT	10 community health centers, greater Boston area	2,219 participants, aged 18-75 (11.9, 11.9%)	Tailored print materials, 1 in-person and 4 telephone counseling, links to local activities	Usual care	0, 8 months	Significant 0.3 srv/day greater increase in intervention than control at 8 months
	Heimendinger et al, 2005 [35]	High	RCT	6 NCIS regional call centers	3,402 callers (43.4%)	11 months; ST: control plus tailored material; MT: control plus multiple tailored material; MRT: control plus multiple re-tailored material	Untailored, brief educational message, print materials	0, 5, 12 months	Significant 0.6 srv/day greater increase in multiple re-tailored compared to untailored at 12 months
	Kreuter et al, 2005 [36]	Medium	RCT	10 urban health centers, St. Louis, MO	1,241 low-income African –American women, aged 18-65 (29.0%)	18 months; behavioral construct (BCT) and/or culturally relevant tailored (CRT) print materials	Delayed intervention/ Usual care	1,6,18 months	No significant difference at 6 months, 0.4 srv/day greater increase in BCT+CRT group compared to controls at 18 months
	Richards et al, 2006 [37]	Medium	RCT	College students, Midwest rural land grand	437 students, age 18-24 (28.1%)	4 months; stage-matched print material, motivational interviewing, email	No materials or contact	0, 4 months	Significant 0.9 srv/day greater increase in intervention compared to controls at 4

				university		contact			months
	Williams-Piehota et al, 2006 [38]	High	RCT	New England NCIS Call Center	517 adult callers (39.1%)	3 months; Complex phone message and tailored print material	Simple phone message and materials	0, 1 ,4 months	About 0.2 srv/day greater increase in complex message group than simple message group at 4 months
	Smeets et al, 2007 [39]	High	RCT	25 worksites in two towns, national phone directory, The Netherlands	2,821 adults (23.6%)	Tailored print materials	1 un-tailored letter	0, 3 months	Significantly smaller decline among intervention than controls at 3 months[1]
	Resnicow et al, 2008 [40]	Medium	RCT	2 integrated health-care delivery systems; Atlanta and Detroit metropolitan areas	512 African-Americans, aged 21 to 70 (17.4%)	Tailored newsletters based on self-determination theory and motivational interviewing	Newsletters tailored on demographic s and social cognitive variables	0, 3 months	No significant difference in change in intake
Rural Physician Cancer Prevention Project	Carcaise-Edinboro et al, 2008 [41]	Medium	RCT	3 primary care practices, rural Virginia	754 patients, age 18-72 (17.4%)	5-8 weeks; recruitment letter, baseline dietary analysis, personalized feedback, print materials, phone calls	Recruitment letter, baseline dietary analysis	0, 1, 6, 12 months	*Not presented overall*
	De Vet et al, 2008 [42]	Medium	RCT	Dutch internet research company	775 adult (26.1%)	One-time experience; stage-matched feedback	Non-stage-matched feedback	0, 2-9, 9-16 days	No significant difference in change intake

Loss to f/up: Loss to follow-up
RCT: Randomized controlled trial
cRT: cluster-randomized trial

WIC: Women, Infants and Children
NCIS: National Cancer Information Service
HMO: Health Maintenance Organization

Notes:
1. no change in fruit vs 0.2 pieces/day decline; 0.5 gm decline in vegetables vs. 10.4 gm/day decline in intervention and controls respectivel

Table 7.2. Findings from analysis of intervention effects across levels of potential demographic effect modifiers, by article.

Trial Name	Citation	Male (vs. Female)	Older Age	Higher Education	Married	Higher Income	White (vs. Non-White)	Immigrant/ Birth status	Employed	Living situation	Occupation	Car access
N/A	Gambera et al,1995 [15]	(-)[1]										
Maryland WIC 5-a-Day Promotion Program	Havas et al, 1998 [16]		(-)	(+)	(+)		(+)[2]		(-)			
	Campbell et al, 2008 [17]											
N/A	Marcus et al, 1998 [18]	(0)	(0)	(0)								
Black Churches United for Better Health Project	Campbell et al, 1999 [19]	(0)	(+)	(0)	(+)[3]	(0)						
	Campbell et al, 2008 [17]											
Women's Health Trial	Coates et al, 1999 [20]			(0)			(0)[4]					
Treatwell 5 a Day	Sorensen et al, 1999 [21]	(0)		(0)						(0)[5]	(0)	
	Campbell et al, 2008 [17]											
Next Step Trial	Tilley et al, 1999 [22]		(0)	(0)					(0)[6]			
Puget Sound Healthy Eating Patterns Study	Kristal et al, 2000 [23]	(0)	(+)							(-)[7]		
	Satia et al, 2001 [24]											
Eat for Life Trial	Resnicow et al,2001[25]	(0)	(0)	(0)		(0)						
	Resnicow et al, 2003 [26]											
fMaryland WIC Food for Life Program	Havas et al, 2003 [27]		(+)	(+)			(0)[8]					
N/A	John et al, 2003 [28]											
N/A	Block et al, 2004 [29]			(-)		(0)	(0)					
Diabetes Prevention Program	Mayer-Davis et al, 2004 [30]	(+)					(0)					
N/A	Steptoe et al, 2004 [31]	(0)				(0)	(0)					

	Perkins-Porras et al, 2005 [32]										
Health Directions – Health Centers Study	Emmons et al, 2005 [33]	(0)		(0)		(0)[9]	(0)	(0)[10]			
	Sorensen et al, 2007 [34]			(0)		(0)[9]		(0)	(0)[11]	(+)[12]	
N/A	Heimendinger et al, 2005[935]	(0)	(0)	(0)			(0)				
N/A	Krueter et al, 2005 [36]										
N/A	Richards et al, 2006 [37]										
N/A	Williams-Piehota et al, 2006 [38]										
N/A	Smeets et al, 2007 [39]	(0)	(0)	(0)							
N/A	Resnicow et al, 2008 [40]		(+)								
Rural Physician Cancer Prevention Project	Carcaise-Edinboro et al, 2008 [41]	(0)	(0)	(0)			(0)				
N/A	De Vet et al, 2008 [42]										

Where not explicitly tested, if means are within 2 standard errors of each other → not significant

(+) Intervention effect greater among those with higher values of the examined variable

(0) Intervention effect examined, but not significantly different by the stated variable

(-) Intervention effect inversely related to this variable

Notes:

1. 1.2 srv/day greater increase among women in intervention than men in intervention. No comparison with control group.
2. Most effective among white (0.73 srv/day greater in intervention) and other (1.72 srv/day greater increase). May be due to unstable estimates in the "other" group.
3. Intervention most effective among those married (+0.81 srv/day), widow, divorced, or other (0.92 srv/day). Less effective among those single (+0.28 srv/day)
4. Intervention most effective among whites (+0.71 srv/day at 6 months), less effective among Hispanics (+0.28 srv/day at 6 months) and blacks (+0.55 srv/day at 6 months)
5. Alone or with others
6. Retirement status
7. More effective among those living alone or with adults (0.64 srv/day) than those with children (0.15 srv/day)
8. Blacks increased -0.39 srv/day in controls, 0.09 in intervention. Whites increased -0.02 srv/day in controls, 0.27 srv/day in intervention
9. Poverty status
10. Respondent's and respondents' parents' country of birth
11. Crowding in household
12. 0.6 srv/day greater increase among those with access to a car than those without

References

1. Joshipura KJ, Hu FB, Manson JE, Stampfer MJ, Rimm EB, Speizer FE, Colditz G, Ascherio A, Rosner B, Spiegelman D, Willett WC. The Effect of Fruit and Vegetable Intake on Risk for Coronary Heart Disease. Ann Intern Med 2001;134(12):1106-1114.
2. Fruit and vegetable consumption among adults--United States, 2005. MMWR Morb Mortal Wkly Rep 2007;56(10):213-7.
3. Brunner EJ, Rees K, Ward K, Burke M, Thorogood M. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2007(4):CD002128.
4. King AC, Ahn DF, Atienza AA, Kraemer HC. Exploring Refinements in Targeted Behavioral Medicine Intervention to Advance Public Health. Annals of Behavioral Medicine 2008;35(3):251-260.
5. Kraemer HC, Frank E, Kupfer DJ. Moderators of Treatment Outcomes: Clinical, Research, and Policy Importance. Jama 2006;296(10):1286-1289.
6. Kraemer HC, Wilson T, Fairburn CG, Agras WS. Mediators and Moderators of Treatment Effects in Randomized Clinical Trials. Archives of General Psychiatry 2002;59(10):877-883.
7. Greenfield. Heterogeneity of Treatment Effects: Implications for Guidelines, Payment, and Quality Assessment. American Journal of Medicine 2007;120(4A):S3-S9.
8. Kravitz RL, Duan N, Braslow J. Evidence-Based Medicine, Heterogeneity of Treatment Effects, and the Trouble with Averages. The Milbank Quarterly 2004;82(4):661-687.
9. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. New York, NY: Lippincott Williams & Wilkins, 1998.
10. Miettinen O. Confounding and effect-modification. American Journal of Epidemiology 1974;100(5):350-353.
11. Greenland S. Modeling and Variable Selection in Epidemiologic Analysis. American Journal of Public Health 1989;79(3):340-349.

12. Foster C, Hillsdon M, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180. pub2.

13. Kremers SP, de Bruijn GJ, Droomers M, van Lenthe F, Brug J. Moderators of environmental intervention effects on diet and activity in youth. Am J Prev Med 2007;32(2):163-72.

14. Katz DL, O'Connell M, Yeh MC, Nawaz H, Njike V, Anderson LM, Cory S, Dietz W. Public health strategies for preventing and controlling overweight and obesity in school and worksite settings: a report on recommendations of the Task Force on Community Preventive Services. MMWR Recomm Rep 2005;54(RR-10):1-12.

15. Gambera PJ, Schneeman BO, Davis PA. Use of the Food Guide Pyramid and US Dietary Guidelines to improve dietary intake and reduce cardiovascular risk in active-duty Air Force members. Journal of the American Dietetic Association 1995;95(11):1268-73.

16. Havas S, Anliker J, Damron D, Langenberg P, Ballesteros M, Feldman R. Final results of the Maryland WIC 5-A-Day Promotion Program. American Journal of Public Health 1998;88(8):1161-7.

17. Campbell MK, McLerran D, Turner-McGrievy G, Feng Z, Havas S, Sorensen G, Buller D, Beresford SA, Nebeling L. Mediation of adult fruit and vegetable consumption in the National 5 A Day for Better Health community studies. Annals of Behavioral Medicine 2008;35(1):49-60.

18. Marcus AC, Heimendinger J, Wolfe P, Rimer BK, Morra M, Cox D, Lang PJ, Stengle W, Van Herle MP, Wagner D, Fairclough D, Hamilton L. Increasing fruit and vegetable consumption among callers to the CIS: results from a randomized trial. Preventive Medicine 1998;27(5 Pt 1):S16-28.

19. Campbell MK, Demark-Wahnefried W, Symons M, Kalsbeek WD, Dodds J, Cowan A, Jackson B, Motsinger B, Hoben K, Lashley J, Demissie S, McClelland JW. Fruit and

vegetable consumption and prevention of cancer: the Black Churches United for Better Health project. American Journal of Public Health 1999;89(9):1390-6.

20. Coates RJ, Bowen DJ, Kristal AR, Feng Z, Oberman A, Hall WD, George V, Lewis CE, Kestin M, Davis M, Evans M, Grizzle JE, Clifford CK. The Women's Health Trial Feasibility Study in Minority Populations: changes in dietary intakes. American Journal of Epidemiology 1999;149(12):1104-12.

21. Sorensen G, Stoddard A, Peterson K, Cohen N, Hunt MK, Stein E, Palombo R, Lederman R. Increasing fruit and vegetable consumption through worksites and families in the treatwell 5-a-day study. American Journal of Public Health 1999;89(1):54-60.

22. Tilley BC, Glanz K, Kristal AR, Hirst K, Li S, Vernon SW, Myers R. Nutrition intervention for high-risk auto workers: results of the Next Step Trial. Preventive Medicine 1999;28(3):284-92.

23. Kristal AR, Curry SJ, Shattuck AL, Feng Z, Li S. A randomized trial of a tailored, self-help dietary intervention: the Puget Sound Eating Patterns study. Preventive Medicine 2000;31(4):380-9.

24. Satia JA, Kristal AR, Curry S, Trudeau E. Motivations for healthful dietary change. Public Health Nutrition 2001;4(5):953-9.

25. Resnicow K, Jackson A, Wang T, De AK, McCarty F, Dudley WN, Baranowski T. A motivational interviewing intervention to increase fruit and vegetable intake through black churches: results of the Eat for Life Trial. American Journal of Public Health 2001;91(10):1686-1693.

26. Resnicow K, McCarty F, Baranowski T. Are precontemplators less likely to change their dietary behavior? A prospective analysis. Health Education Research 2003;18(6):693-705.

27. Havas S, Anliker J, Greenberg D, Block G, Block T, Blik C, Langenberg P, DiClemente C. Final results of the Maryland WIC Food for Life Program.[erratum appears in Prev Med. 2004 May;38(5):668]. Preventive Medicine 2003;37(5):406-16.

28. John JH, Yudkin PL, Neil HA, Ziebland S. Does stage of change predict outcome in a primary-care intervention to encourage an increase in fruit and vegetable consumption? Health Education Research 2003;18(4):429-38.

29. Block G, Wakimoto P, Metz D, Fujii ML, Feldman N, Mandel R, Sutherland B. A randomized trial of the Little by Little CD-ROM: demonstrated effectiveness in increasing fruit and vegetable intake in a low-income population. Preventing Chronic Disease 2004;1(3):A08.

30. Mayer-Davis EJ, Sparks KC, Hirst K, Costacou T, Lovejoy JC, Regensteiner JG, Hoskin MA, Kriska AM, Bray GA, Diabetes Prevention Program Research G. Dietary intake in the diabetes prevention program cohort: baseline and 1-year post randomization. Annals of Epidemiology 2004;14(10):763-72.

31. Steptoe A, Perkins-Porras L, Rink E, Hilton S, Cappuccio FP. Psychological and Social Predictors of Changes in Fruit and Vegetable Consumption Over 12 Months Following Behavioral and Nutrition Education Counseling. Health Psychology 2004;23(6):574-581.

32. Perkins-Porras L, Cappuccio FP, Rink E, Hilton S, McKay C, Steptoe A. Does the effect of behavioral counseling on fruit and vegetable intake vary with stage of readiness to change? Preventive Medicine 2005;40(3):314-20.

33. Emmons KM, Stoddard AM, Fletcher R, Gutheil C, Suarez EG, Lobb R, Weeks J, Bigby JA. Cancer prevention among working class, multiethnic adults: results of the Healthy Directions-Health Centers Study. American Journal of Public Health 2005;95(7):1200-1205.

34. Sorensen G, Stoddard AM, Dubowitz T, Barbeau EM, Bigby J, Emmons KM, Berkman LF, Peterson KE. The influence of social context on changes in fruit and vegetable

consumption: results of the healthy directions studies. American Journal of Public Health 2007;97(7):1216-1227.

35. Heimendinger J, O'Neill C, Marcus AC, Wolfe P, Julesburg K, Morra M, Allen A, Davis S, Mowad L, Perocchia RS, Ward JD, Strecher V, Warnecke R, Nowak M, Graf I, Fairclough D, Bryant L, Lipkus I. Multiple tailored messages are effective in increasing fruit and vegetable consumption among callers to the Cancer Information Service. Journal of Health Communication 2005;10 Suppl 1:65-82.

36. Kreuter MW, Sugg-Skinner C, Holt CL, Clark EM, Haire-Joshu D, Fu Q, Booker AC, Steger-May K, Bucholtz D. Cultural tailoring for mammography and fruit and vegetable intake among low-income African-American women in urban public health centers. Preventive Medicine 2005;41(1):53-62.

37. Richards A, Kattelmann KK, Ren C. Motivating 18- to 24-year-olds to increase their fruit and vegetable consumption. Journal of the American Dietetic Association 2006;106(9):1405-11.

38. Williams-Piehota P, Pizarro J, Navarro Silvera SA, Mowad L, Salovey P, Williams-Piehota P, Pizarro J, Navarro Silvera SA, Mowad L, Salovey P. Need for cognition and message complexity in motivating fruit and vegetable intake among callers to the cancer information service. Health Communication 2006;19(1):75-84.

39. Smeets T, Kremers SPJ, Brug J, de Vries H. Effects of tailored feedback on multiple health behaviors.[erratum appears in Ann Behav Med. 2007 Jul-Aug;34(1):104]. Annals of Behavioral Medicine 2007;33(2):117-23.

40. Resnicow K, Davis RE, Zhang G, Konkel J, Strecher VJ, Shaikh AR, Tolsma D, Calvi J, Alexander G, Anderson JP, Wiese C. Tailoring a fruit and vegetable intervention on novel motivational constructs: results of a randomized study. Annals of Behavioral Medicine 2008;35(2):159-69.

41. Carcaise-Edinboro P, McClish D, Kracen AC, Bowen D, Fries E. Fruit and vegetable dietary behavior in response to a low-intensity dietary intervention: the rural physician cancer prevention project. Journal of Rural Health 2008;24(3):299-305.

42. de Vet E, de Nooijer J, de Vries NK, Brug J. Testing the transtheoretical model for fruit intake: comparing web-based tailored stage-matched and stage-mismatched feedback. Health Education Research 2008;23(2):218-27.

43. Sorensen G, Thompson B, Glanz K, Feng ZD, Kinne S, DiClemente C, Emmons K, Heimendinger J, Probart C, Lichtenstein E. Work site-based cancer prevention: Primary results from the Working Well Trial. American Journal of Public Health 1996;86(7):939-947.

44. Sorensen G, Barbeau E, Stoddard AM, Hunt MK, Kaphingst K, Wallace L. Promoting behavior change among working-class, multiethnic workers: results of the healthy directions--small business study. Am J Public Health 2005;95(8):1389-95.

45. LaMontagne AD, Barbeau E, Youngstrom RA, Lewiton M, Stoddard AM, McLellan D, Wallace LM, Sorensen G. Assessing and intervening on OSH programmes: effectiveness evaluation of the Wellworks-2 intervention in 15 manufacturing worksites. Occupational and Environmental Medicine 2004;61(8):651-660.

46. Reynolds KD, Franklin FA, Binkley D, Raczynski JM, Harrington KF, Kirk KA, Person S. Increasing the fruit and vegetable consumption of fourth-graders: Results from the High 5 project. Preventive Medicine 2000;30(4):309-319.

47. Baranowski T, Davis M, Resnicow K, Baranowski J, Doyle C, Lin LS, Smith M, Wang DQT. Gimme 5 fruit, juice, and vegetables for fun and health: Outcome evaluation. Health Education & Behavior 2000;27(1):96-111.

48. Perry C, Lytle LA, Feldman H, Nicklas T, Stone E, Zive M, Garceau A, Kelder SH. Effects of the child and adolescent trial for cardiovascular health (CATCH) on fruit and vegetable intake. Journal of Nutrition Education 1998;30(6):354-360.

49. Cox DN, Anderson AS, Reynolds J, McKellar S, Lean MEJ, Mela DJ. Take Five, a nutrition education intervention to increase fruit and vegetable intakes: impact on consumer choice and nutrient intakes. British Journal of Nutrition 1998;80(2):123-131.

50. Hendrix SJ, Fischer JG, Reddy S, Lommel TS, Speer EM, Stephens H, Park S, Johnson MA. Fruit and vegetable intake and knowledge increased following a community-based intervention in older adults in Georgia senior centers. Journal of Nutrition for the Elderly 2008;27(1-2):155-178.

51. Tilley B, Vernon S, Myers R, al. e. The Next Step Trial: impact of a worksite colorectal cancer screening promotion program. Prev Med 1999;28:276-283.

52. Oldroyd J, Burns C, Lucas P, Haikerwal A, Waters E. The effectiveness of nutrition interventions on dietary outcomes by relative social disadvantage: a systematic review. Journal of Epidemiology and Community Health 2008;62:573-579.

53. Friel S, Kelleher C, Campbell P. Evaluation of the Nutrition Education at Primary School (NEAPS) programme. Public Health Nutrition 1999;2:549-555.

54. Reinaerts E, Nooijer Jd, Candel M. Increasing children's fruit and vegetable consumption: distribution or a multicomponent programme? Public Health Nutrition 2007;10:874-882.

55. Whitlock EP, Orleans CT, Pender N, Allan J. Evaluating Primary Care Behavioral Counseling Interventions: An Evidence-Based Approach. American Journal of Preventive Medicine 2002;22(4):267-284.

56. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariables risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Medical Research Methodology 2006;6(18).

57. Hogue CJR. Getting to the Why. Epidemiology 1997;8(3):230.

58. Brug J, Oenema A, Ferreira I. Theory, evidence and Intervention Mapping to improve behavior nutrition and physical activity interventions. International Journal of Behavioral Nutrition and Physical Activity 2005;2(2).

59. Rothschild ML. Carrots, Sticks, and Promises: A Conceptual Framework for the Management of Public Health and the Social Issue Behaviors. Journal of Marketing 1999;63:24-37.

60. Brunner EJ, Thorogood M, Rees K, Hewitt G. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2005(4):CD002128.

Chapter 8: Conclusions

Summary of Findings

This dissertation incorporated theory-based effect modification analysis with the review of the randomized literature on fruit and vegetable consumption interventions. The studies discussed in this dissertation examine possible determinants of health behaviors which have been hypothesized by Rothschild to be important in the determination of behavior and behavior change[1], but which have not received as much attention in the literature as more traditional determinants of behavior such as motivation and abilities[2]. Based on findings from the analyses of the effect of environmental exposures on response to the Nutrition and Physical Activity (NuPA) intervention on fruit and vegetable consumption (FVC) and physical activity (PA), much can be learned from the examination of the nutrition and physical activity environment as modifiers of the effect of a randomized, controlled trial.

In the first study of this dissertation we found the intervention effect on physical activity levels was consistently positive only among participants with the lowest proximity (more than 1.66 km to the nearest park) at baseline or the least total park space accessible within 0.805-km at baseline. We also found that the intervention maintenance effect measured at 12 months was linearly and inversely associated with the number of parks within 8.05-km. These findings were consistent with those from the second study examining modification of the intervention effect. In that study, we found that the intervention effect increased consistently with follow-up time only among participants who had the least diversity of grocers within 8.05-km or the least variety of grocers at baseline. Imputation analyses accounting for loss to follow-up did not change these

conclusions. Furthermore, analysis of loss to follow-up showed that there was no association between any measure of the environment (whether related to physical activity or fruit and vegetable consumption) and overall loss to follow-up (whether defined as loss of contact or refusal to participate). Additional analysis, however, should be undertaken to determine whether differential loss by environmental characteristics occurred *wthin* the intervention arm.

These analyses are prospective, with behavior change occurring within existing nutrition and physical activity environment exposure and after randomization. This is in contrast to previous studies of behavior and environment which have largely been cross-sectional, making it impossible to determine whether environment preceded and caused behavior, or vice versa. The processes governing individual behavior and environmental exposures are complex, and it is possible that behaviors (food purchasing patterns, for example) influenced the environment (location of grocers) rather than the more often-hypothesized idea that individual behaviors are determined by environment. In fact, results from at least one cross-sectional study support the possibility of self-selection into communities based on desired behavior[3]. In a cross-sectional study of residents in Atlanta, neighborhood walkability was associated with walking behavior only amongst participants who preferred a walkable neighborhood. Among residents who did not prefer a walkable neighborhood, there were few who engaged in walking behavior regardless of the walkability of the surrounding neighborhood.

This does not preclude the possibility that both processes occur, with mutual feedback between environment and behavior. While much research on the location of grocery stores has been conducted, the majority of the work has been performed in either

the United Kingdom or Australia[4,5]. Although the UK, Australia and the United States share similar economic structures, whether these findings from the other two countries are generalizable to the United States is unclear. A few studies have also been performed in North America[6]. However, information on grocery store distribution remains scarce in the public health literature, and information on processes governing park distribution appears to be almost non-existent.

Given the complexity in the relationship between behavior and environment, the potential mutual feedback between the two, and the nature of the environmental exposures under study, observational longitudinal studies would be faced with similar challenges as cross-sectional studies. In general, longitudinal studies offer the advantage of exposure measurement preceding the outcome of interest[7]. However, if the true state is reciprocity but the hypothesized direction of causality is that environment causes behavior, there is a built-in bias in favor of finding an association between environment and behavior. Under such circumstances, environment will have already been influenced by behavior at the time of study enrollment, making longitudinal studies little more useful in determining the temporal sequence than cross-sectional studies. Then only by changing the environment can investigators observe the effect of the environment on individual behavior.

Of course, community-level environmental change, poses many challenges and requires significant resources [8]. This author is unaware of any published studies in which communities have been randomized to receive opportunities for physical activity or outlets for purchasing fruits and vegetables. Moreover, few investigators have had the opportunity to examine the effect of introducing new physical activity opportunities to a

community. Observational studies examining natural experiments of environmental change have yielded inconclusive results. For instance, one study reported behavior in a community where a walking trail had been newly created[9]. However, it did not examine the behavior before and after the introduction of the trails. Rather, the study focused on the effects of campaigns to raise awareness of the walking trails. There have been at least two studies examining fruit and vegetable consumption in a community before and after the introduction of a large food outlet[4, 10]. While the first study[4], published in 2002, reported a significant increase in fruit and vegetable consumption after the introduction of a large-scale retailer, it did not control for secular trends or potential changes in reporting. The second study[10] included a control community. Both communities increased reported fruit and vegetable consumption over the study period, but the difference in the amount of increase was not significantly greater in the community where the new retail outlet had been introduced.

One potential explanation for the discrepancy between these natural experiments and the results of cross-sectional studies may lie in that fact that the different designs are measuring two different outcomes. There is growing awareness in the behavioral literature that the goal of behavioral interventions should be in predicting, eliciting, and explaining behavior *change*[2]. This is in contrast to previous studies explaining predictors of behavior itself which cannot inform investigators on how improve behaviors to levels recommended for health benefits. However, there is likely to be limited statistical power in longitudinal studies of this sort because change in behavior potentially elicited from change in the environment may not occur amongst enough participants to be statistically significant. In contrast, randomized trials including individual-level interventions, or a

mixture of individual- and environmental-level interventions have been shown to be more successful at eliciting behavior change[8].

In the analysis of data from the NuPA trial, our estimates suggest that any intervention effect was inversely proportional to access to resources. For FVC, the difference between the intervention arm and the control arm in the increase in daily servings of fruits and vegetables consumed was greatest among those with the least variety and diversity of grocers. Furthermore, the same relationship was seen with the intervention effect on PA being greatest among those with the least diversity of parks. These findings were observed despite multiple sources of non-differential misclassification. Additional analysis to determine the potential effect of geocoding errors (**Chapter 4** and **5**) and informative loss to follow-up (**Chapter 6**) revealed no substantial bias to explain the findings.

Given the observational literature – which has consistently reported a positive association between access to better nutrition and physical activity environments and healthy lifestyle behaviors – these findings were unexpected. They suggest that the relationship between behavior, behavior change, and the built environment may be more complex than previously thought. It is possible that the inverse relationship may be observed only when low intensity interventions are administered. Future studies involving more intense interventions – those involving supervision, more frequent counseling sessions, or face-to-face educational or counseling sessions – may yield different results.

Strengths and Limitations

There were several limitations to these findings.

First, the study may not have had adequate power to detect statistically significant differences in intervention effect across characteristics of the environment. This is a limitation that is faced by most analyses of intervention effect modification because randomized studies are generally designed to have adequate power to detect a statistically significant main intervention effect. This limited power was further exacerbated by non-differential misclassification of the exposures which may have biased estimates towards the null. However, these sources of bias could not be avoided given the design of the current study and may be common to secondary analyses of intervention trials. Concerns for confidentiality necessitated the use the block group (BG) centroid rather than actual address of participants for exposure ascertainment and the identification of businesses from a commercial database reduced the time, cost, and effort in ascertaining exposure to grocery stores in comparison to ground-truth data, which could not have been feasibly collected solely for the purpose of performing this study.

The use of the commercial databases, both for the grocer businesses examined in the FVC study and the parks and street data in the PA study, also provides several advantages. Time, personnel, financial resources, and computing resources that would have been spent to collect data on grocer locations, parks, and streets is not required, thus preventing the duplication of the efforts across the research team and the companies producing such databases and frees resources for use in other areas. Second, the variety of data collected (business size, location type, etc.) as well as the standard definitions and quality control measures applied to commercial resources provides the researcher with more data than can be collected solely for research purposes, and more uniformity in data quality across geographies. In this analysis, we were able to examine associations with

access to any and all grocery stores, as well as with just chain grocers. Of course, use of external data sources presents challenges in determining data quality, validity, and consistency across areas of interest to the researcher. However, findings from our validation study (**Chapter 5**) suggest that those concerns could not have strongly biased our findings, except for non-differential misclassification potentially biasing the observed associations towards the null

In addition, this study is limited in that exposures are highly correlated with one another. For example, the distance to the nearest grocery store (proximity) and the number of grocers within 8.05 km (5 miles) is correlated with a coefficient of -0.51 ($p<0.0001$). This makes it difficult to obtain estimates of the independent associations between behavior and these exposures. Further increasing the difficulty in obtaining stable estimates, participants were clustered within cities, but further adjustment for correlation at lower levels within city – such as random effects for block groups included in the analysis of fruit and vegetable consumption – is computationally intensive and models often do not converge.

There are also several strengths to this study. Specifically: (1) the main motivation for conducting these studies lay in the fact that the temporal relationship between environment and behavior change was established, (2) the use of heterogeneity in the intervention effect accounts for multiple sources of change in reporting of behavior that may be un-related to the questions of interest, (3) the original NuPA study elicits more behavior change than can be observed in a longitudinal study, and (4) the results may be generalizable. On this last point, the participants analyzed in the studies reported

for this dissertation were representative of the population in the United States based on race, education, employment and marital status.

The main advantage of this study, and the main motivation for conducting this dissertation, was that the design establishes the temporal sequence between environment and behavior change.

The use of the heterogeneity in the intervention effect across environment as a measure of the association between environment and behavior change addresses two limitations with self-reported behavior: (1) changes in reporting and desirability effects and (2) temporal trends in behavior previously mentioned. First, participants may report a change in behavior simply because they become increasingly knowledgeable about what constitutes a serving of fruits or vegetables once they are exposed to the survey instrument as well as to any educational element of an intervention or control. This has been a criticism of intervention studies that report change in behaviors only in one group of individuals who all received an intervention[11]. In the NuPA study, both the control arm and the intervention arm were evaluated using the same instrument and received the same educational materials. Thus, changes in reporting are accounted for in the analysis as they occur in both arms. Second, participants may report an increase in consumption because they want to report behaviors more consistent with those that are socially acceptable or desirable[12]. This is affected by the educational materials, but may also result from repeated questioning of participants regarding healthy behaviors by evaluators. However, participants in the control arm and in the intervention arm are both surveyed in the same way and by evaluators blinded to their randomization status. As such, increased reporting due to desirability effects are accounted for in the estimation of

the intervention effect. And last, secular trends in behavior change in the target population may also explain increases in healthy behaviors observed over the course of a randomized trial. Because participants are randomized as they are recruited, this trend should affect both randomized arms in the same way and thus does not affect the estimation of the intervention effect.

In addition, all participants in the NuPA study were given educational materials. This increases the possibility that participants will change their behavior as well as the magnitude of change that will result. This is in contrast to longitudinal studies which, being observational, may not elicit as much behavior change and therefore would be less likely to detect a statistically significant degree of behavior change, beit overall or across levels of the environment.

Overall, this dissertation demonstrates that this study design can be used effectively for studying a problematic and complex exposure-outcome relationship. Furthermore, it demonstrates how two bodies of literature can be combined to inform one another. In the case of this dissertation, we attempted to determine whether the moderate effects observed in this randomized trial were due to potential biases resulting from differential loss across levels of the environment. With respect to loss to follow-up, this is a relatively un-examined group of variables. We also attempted to address the limitations of observational studies by establishing temporality in the association between behavior and environment. We measured environment at enrollment and then examined behavior change.

Implications

On Future Research

As evidenced in the above discussion of limitations, there are several areas where improvements can be made to increase statistical efficiency, promote better understanding of the independent associations between behavior and exposures, and also improve the effectiveness of intervention studies.

First, refinement of definitions for assessing access to resources for healthy behaviors may be beneficial in efforts to improve the understanding of mechanisms by which these resources affect individual behavior. Few existing studies have carefully examined and contrasted different measures of access. Even fewer have examined different measures across different scales, as was done here in the use of two different buffer zones representing different preferences in commuting modality. Studies that have examined different definitions of access have not always found the relationships expected from the previous literature. For example, Apparicio et al. did not find the presence of food deserts in Montreal although previous research had suggested there should be areas with little or no access to grocers. The authors used definitions of access including variety, diversity, and proximity to grocery stores[6] within the scale of a city. In contrast, other studies have reported the existence of food deserts when examining national-level data[13, 14]. Differences in these findings may lay in multiple sources including the different geographies examined, the different scales examined, and differences in definitions of access. Further investigation is needed to determine whether small-scale difference as studied in Apparicio et al. or whether large-scale differences are more important for individual-level behavior.

The multitude of extant data regarding the built environment – from national-, state-, metropolitan-area level governments, commercial firms, or independent researchers – will make studies regarding the environment more feasible as data collection increases and data are made increasingly available. A challenge arising out of this wealth of data, however, is the correlated nature of the exposures[15]. This makes it difficult to determine the independent effects of different exposures. However, different methodologies have been adopted from other fields[16] and may prove to aid in addressing these challenges. In fact, one study performed by King et al examined the association between the perceived physical activity environment and intervention effect using a signal processing methodology designed to identify participant subgroups with significantly different intervention effects[17]. These subgroups can be defined using any combination of variables, including those that may be highly correlated.

King's study and the NuPA analyses represent a departure from methodologies used in the previous literature in which those conducting intervention studies have seldom examined intervention effects across sub-groups other than those defined by demographic variables. However, investigators are increasingly becoming aware that examining the potential for differential intervention effect is important. This is true not only for identifying subgroups for whom the intervention may be targeted for services delivery or future intervention studies, but also for understanding the mechanisms by which the intervention may or may not work and for identifying environment determinants that may not be a component of the intervention, but may serve as a modifier of the intervention effect. This less obvious role of the randomized trial design can be employed for other purposes as well, as demonstrated here. Future studies might

involve a widely-accepted, individual-level behavioral intervention with recruitment from populations geographically defined to examine the effect of the built environment on intervention effect. Since the intervention will be widely accepted, the design should involve a delayed intervention for a control group.

Effect modification is not as well understood or as easily reported as other phenomena, such as confounding[7]. Employing effect modification as a focus of research rather than as ancillary findings to a randomization trial requires better understanding of how effect modification may manifest. There is ample discussion in the literature regarding the effect of scale, model, and estimation methods on the interpretation of statistical findings and, specifically, the interpretation of effect modification. Although epidemiologists are aware of the issues in interpreting the multiplicative scale, a challenge remains in mobilizing researchers to use additive models[18]. This is particularly poignant as theoretical and methodologic discussion of effect modification has largely been limited to the additive scale[7, 19, 20].

Another challenge lies in the use of statistical significance as a criterion for meaningful effect modification. As discussed above, randomized trials are usually adequately powered only for detecting the main intervention effect[21]. As such, any examination of effect modification will likely be underpowered. This is not likely to change as funding structures will remain focused on attaining power for estimating the main intervention effect. However, the alternative (examining the magnitude of differences in intervention effect among subgroups) remains problematic. Judgment regarding what constitutes a meaningful difference in intervention is inherently subjective, and a lack of a summary measures across groups – particularly when there is

not a monotonic effect – may make interpretation difficult. As demonstrated in the **Chapters 4** and **5**, these challenges may be overcome by examining patterns in the intervention effect. Furthermore, results from observational studies may serve as effective guides for formulating a specific approach.

On Policies and Initiatives to Change Nutrition and Physical Activity Environments

Despite the multitude of challenges in studying the nutrition and physical activity environments and their association with their respective behaviors, there has already been much social and financial investment in environment-level changes, particularly in the creation of mixed use communities. The United States Environmental Protection Agency offers an award for communities promoting the principles of what has been termed "smart growth"[22]. Although composed of multiple components, the features distinguishing a "smart growth" community from others includes the promotion of mixed land use, the creation of walkable neighborhoods, and preserving open spaces. Furthermore, these initiatives attempts to address issues of quality of life, economics, and health[23].

However, evidence to support direct effects of these communities and initiatives on the issues of quality of life and health are still equivocal at best[24]. In particular, the claim of direct health and behavioral effects requires stronger evidence than what is currently available from cross-sectional studies. Some policy-makers claim that these initiatives are not a benign phenomenon because they further marginalize segments of the population who are at the most risk of adverse health behaviors and health outcomes[25]. By artificially increasing the cost of housing in the process of "smart growth" development, low-income, minority, under-served populations are the least likely to

benefit from the process and the most likely to feel the negative effects. Furthermore, these initiatives may draw attention away from truly causal relationships, not only detracting from resources and efforts required to address the existing public health problem, but also damaging the trust held between the public and public health institutions – be it researchers, policy-makers, or others[7].

Overall, we conclude that access to grocers and access to parks may be inversely associated with behavioral change from a modest, low-intensity, individual-level behavioral intervention. Since those who benefited the most from the intervention were participants with the least access to resources and opportunities for healthy behavior, future studies may benefit from similar analysis to determine whether this inverse relationship is maintained in populations exposed to a more rigorous or intense intervention.

References

1 Rothschild ML. Carrots, Sticks, and Promises: A Conceptual Framework for the Management of Public Health and the Social Issue Behaviors. *Journal of Marketing* 1999;**63**:24-37.

2 Brug J, Oenema A, Ferreira I. Theory, evidence and Intervention Mapping to improve behavior nutrition and physical activity interventions. *International Journal of Behavioral Nutrition and Physical Activity* 2005;**2**.

3 Frank LD, Saelens BE, Powell KE*, et al.* Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Soc Sci Med* 2007;**0**:0.

4 Wrigley N, Warm D, Margetts D*, et al.* Assessing the impact of improved retail access on diet in a 'food desert': a preliminary report. *Urban Studies* 2002;**39**:2061-82.

5 Smith RL. The Australian Grocery Industry: A Competition Perspective. *Australian Journal of Agricultural and Resource Economics* 2006;**50**:33-50.

6 Apparicio P, Cloutier MS, Shearmur R. The case of Montreal's missing food deserts: evaluation of accessibility to food supermarkets. *Int J Health Geogr* 2007;**6**:4.

7 Rothman K, Greenland S. *Modern Epidemiology*. New York, NY: Lippincott Williams & Wilkins 1998.

8 Brug J, van Lenthe F. Environmental determinants and interventions for physical activity, diet and smoking: a review. Rotterdam, The Netherlands: Erasmus MC 2005.

9 Merom D, Bauman A, Vita P*, et al.* An environmental intervention to promote walking and cycling--the impact of a newly constructed Rail Trail in Western Sydney. *Preventive Medicine* 2003;**36**:235-42.

10 Cummins S, Petticrew M, Higgins C*, et al.* Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. *J Epidemiol Community Health* 2005;**59**:1035-40.

11 Howard GS. Response-Shift Bias: A Problem in Evaluating Interventions with Pre/Post Self-Reports. *Eval Rev* 1980;**4**:93-106.

12 Kristal AR, Andrilla CHA, Koepsell TD*, et al.* Dietary Assessment Instruments are Susceptible to Intervention-associated Response Set Bias. *Journal of the American Dietetic Association* 1998;**98**:40-3.

13 Morton L, Blanchard T. Starved for access: life in rural America's food deserts. *Rural Realities* 2007;**1**.

14 Shaffer A. The persistence of L.A.'s grocery gap:the need for a new food policy and approach to market development.: Center for Food and Justice, Urban and Environmental Policy Institute (UEPI), Occidental College 2002.

15 Mullner RM, Chung K, Croke KG*, et al.* Geographic Information Systems in Public Health and Medicine. *Journal of Medical Systems* 2004;**28**:215-21.

16 Friedman JH. A Recursive Partitioning Decision Rule for Nonparametric Classification. *IEEE Transaction on Computers* 1977;**26**:404-8.

17 King AC, Toobert D, Ahn D*, et al.* Perceived environments as physical activity correlates and moderators of intervention in five studies. *Am J Health Promot* 2006;**21**:24-35.

18 Spiegelman D, Hertzmark E. Easy SAS Calculations for Risk or Prevalence Ratios and Differences. *Am J Epidemiol* 2005;**162**:199-200.

19 Vanderweele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol* 2007;**166**:1096-104.

20 Vanderweele TJ, Robins JM. Four Types of Effect Modification: A Classification Based on Directed Acyclic Graphs. *Epidemiology* 2007;**18**:561-8.

21 Kraemer HC, Frank E, Kupfer DJ. Moderators of Treatment Outcomes: Clinical, Research, and Policy Importance. *Jama* 2006;**296**:1286-9.

22 National Award for Smart Growth Achievement U. S. Environmental Protection Agency 2009.

23 About Smart Growth. Sustainable Communities Network 2009.

24 Frank LD. Land Use and Transportation Interaction: Implications on Public Health and Quality of Life. *Journal of Planning Education and Research* 2000;**20**:6-22.

25 Pozdena RJ. Smart Growth and Its Effects on Housing Markets: The New Segregation. Washington, DC: The National Center for Public Policy Research 2002.

Appendices

Appendix 1: Health for Life Intake Questionnaire (March 2005)

User Name _____Date _____Time _____

Client Name_____Constituent/ACS ID _____

1. Reason for calling
   ____Personally interested in Health for Life
   ____Family/Friend of a person interested in Health for Life
   ____Other (Drs. Office, Teachers, Community Org)

   If Family/Friends, go to #2, then 2a
   If "Other", go to #2, then 2b
   If caller is under 18, go to #2, then 2c
   If caller is personally interested in Health for Life, go to#2, then #3

2. How did you hear about this phone number?
   ☐Newspaper story ☐Newspaper ad ☐TV news story ☐TV ad
   ☐Radio news story ☐Radio ad ☐Doctor ☐ Phone Book
   ☐Pharmacist ☐Nurse ☐Other health care provider
   ☐Flyer from school ☐Flyer from work ☐Flyer from community
   ☐Friend ☐Relative ☐Internet/Website
   ☐ACS Office ☐Transferred from NCIC
   ☐County Health Dept.
   ☐Other/Describe Ad_____

2a. That's great that you want to help your friend or family member live a healthier life. I'm going to send you a booklet that talks about nutrition and physical activity. You can read it yourself and perhaps give it to the person you want to help. Also, please encourage your friend or family member to give us a call when they are ready to participate in the Health for Life program. (We will only enroll the participant if we speak with them over the phone, no third party participants will be enrolled without us talking with the participant directly and completing their individualized intake.)

Is there anything else I can do for you today? **(Answer questions/requests with applicable information from call handling resources)**
Thank you for calling your American Cancer Society.

Order "Living Smart" Fulfillment #2042.00

2b. What can I help you with today?

(Answer questions/requests with applicable information from call handling resources)

If from the media, e-mail "Health for Life Questions"

Thank you for calling your American Cancer Society.

2c. I see from your date of birth that you are not 18 years old. Unfortunately, we won't be able to include you in this program. However, I would be happy to send you some nutrition and physical activity information through the mail. Would you like to receive those?

☐Yes ☐ No **(Answer questions/requests with applicable information from call handling resources)**

Is there anything else I can do for you today?
Thank you for calling your American Cancer Society.

Order "Living Smart" Fulfillment # 2042.00

3. We have some services available to help you incorporate nutrition, physical activity and weight management into your lifestyle. But first I'd like to ask you a few demographic questions, will that be okay?

☐Yes **(go to # 4)** ☐No **(go to # 3a)**

3a. Okay, we have some nutrition and physical activity information available we can mail, if you are interested? Also, if you change your mind about sharing information with us feel free to call us back; we're here to help. Do you have our phone number?
(1-877-4R-LIVES /877-475-4837)

Is there anything else I can do for you today?
(If yes) Answer questions/requests with applicable information from call handling resources.
Thank you for calling your American Cancer Society, good bye.

If interested in receiving materials, Order "Living Smart" Fulfillment # 2042.00

4. Do you have access to the Internet for your personal use?

☐Yes ☐No

5. Do you have an email? ☐Yes ☐No **(If Yes, enter e-mail address in Siebel)**
_____

6. What is the highest grade or year of school completed?

☐ Never attend school or only Kindergarten
☐ Grades 1-5 (some Grade School) ☐ Some College or Technical School
☐ Grades 6-8 (some Jr. High School) ☐ College Graduate
☐ Grades 9-11 (some High School) ☐ Graduate School
☐ High School Graduate or GED ☐ Refused to answer

7. What is your current employment status?

☐ Paid full-time employment ☐ Student part-time
☐ Paid part-time employment ☐ Self-employed
☐ Homemaker ☐ Volunteer
☐ Not employed-disabled ☐ Not employed-retired
☐ Student full-time ☐ Unemployed
☐ Refused to answer

8. What is your marital status?

☐Single ☐Married ☐Widowed ☐Divorced
☐Separated ☐Other _____ ☐Refused to answer

9. What is your ethnic background or race?

☐White ☐Black ☐Hispanic
☐Asian American ☐Indian/Native American
☐Other _____
☐Refused to answer

Skip question # 10 for male callers

10. Are you currently pregnant or lactating?

☐ **Yes (go to #10a)** ☐**No (go to #11)**

10a. Unfortunately, this clinical trial does not meet the needs of your special population. However, I would be happy to send you some nutrition and physical activity information through the mail. Would you like to receive those?

☐Yes ☐ No **(Answer questions/requests with applicable information**

from call handling resources)

Is there anything else I can do for you today?

Thank you for calling your American Cancer Society.

Order "Living Smart" Fulfillment # 2042.00

11. Do you currently have an eating disorder? (Anorexia Nervosa, Bulimia Nervosa, Binge Eating Disorder) (self-reports suffering from an eating disorder and/or receiving treatment for an eating disorder)

☐ **Yes (go to #10a)** ☐ **No (go to #12)**

12. Have you felt sad or blue almost everyday for the last three months?

☐ Yes ☐ No

13. In the past three months, have you considered yourself satisfied with your life?

☐ Yes ☐ No

14. In general, would you say your health is:

☐ Excellent ☐ Fair
☐ Very Good ☐ Poor
☐ Good ☐ Refuse to answer

15. In the past three months, on a typical day, how many servings (1 serving = ½ cup) of fruits do you consume each day? Count any kind of fruit- fresh, frozen, canned or dried. (Please exclude fruit juice.)

16. In the past three months, on a typical day, how many servings (1 serving = ½ cup) of vegetables do you consume each day? Count raw, cooked, canned or frozen as well as cooked dried beans. (Please exclude vegetable juices, white potato products, fried or processed starchy vegetables like French fries, potato chips and hash browns.)

17. In the past three months, on a typical day, how many minutes of physical activity do you get each day? (physical activity to include mild, moderate and vigorous activity; i.e. minimal effort/no sweating, sweat lightly, or rapid heart beating/heavy sweating, respectively)

18. How often do you eat cheese or cheese spread (not low-fat)?

- ❑ Never
- ❑ <1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

19. How often do you eat beef, pork or lamb? Include bacon, sausage, salami and hot dogs.

- ❑ Never
- ❑ <1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

20. What kind of milk do you drink?

- ❑ Whole milk
- ❑ 2% fat milk
- ❑ 1% fat milk
- ❑ Skim, nonfat, or ½ % fat milk
- ❑ Soy milk
- ❑ Rice milk
- ❑ Other

21. How many days of the week do you eat breakfast?

22. How often do you eat regular potato chips, tortilla chips, corn chips, etc.?

- ❑ Never
- ❑ < 1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

23. How often do you eat sugary foods like cake, candy, soda, sugary cereals?

- ❑ Never
- ❑ < 1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

24. What do you think is the minimal number of daily servings of fruits and vegetables recommended to reduce health risks?

25. What do you think is the minimal amount of weekly physical activity recommended to reduce health risks?

26. What are ways you know of to reduce health risks?

- ❑ Eat fruits and vegetables
- ❑ Stay out of the sun/Use sunscreen
- ❑ Get plenty of sleep
- ❑ Reduce stress
- ❑ Reduce amount of fat consumed
- ❑ Lose weight
- ❑ Limit alcohol consumption
- ❑ Don't smoke
- ❑ Exercise
- ❑ Limit red meat consumption
- ❑ Minimize toxic environmental exposure
- ❑ Other
  - o Other has a free text box to allow additional response

27. What are ways you know of to reduce cancer risk?

- ❑ Eat fruits and vegetables
- ❑ Don't smoke
- ❑ Stay out of the sun/Use sunscreen
- ❑ Eat whole grains
- ❑ Limit red meat consumption
- ❑ Exercise
- ❑ Limit alcohol consumption
- ❑ Lose weight
- ❑ Minimize toxic environmental exposure
- ❑ Other

- Other has a free text box to allow additional response

28. Do you believe there is a link between being overweight and cancer risk?

Yes ☐ **No** ☐

29. Do you know what your Body Mass Index (BMI) is?

Yes ☐ **(go to #29a)** **No** ☐

29a. What is your BMI?

30. Do you know what the healthy BMI range is?

- ☐ 19-24.9
- ☐ 25-29.9
- ☐ Greater than 30
- ☐ Don't Know

31. What is your current weight?

32. What is your height?

33. Has your doctor diagnosed you with any of the following conditions?

- Diabetes ☐ **Yes** ☐**No** ☐ **Don't Know**
- High blood pressure (140/90mm Hg or higher) ☐ **Yes** ☐**No** ☐ **Don't Know**
- Cancer ☐ **Yes (choose the primary cancer site)** ☐**No**
  - If yes, use drop down menu to choose what type of cancer (ability to choose more than one choice)
    - Anal
    - Bile Duct
    - Bladder
    - Bone
    - Bone Metastasis
    - Brain (adult)
    - Brain (child)
    - Breast
    - Cancer of Unknown Primary
    - Cervical

- Colon and Rectum
- Endometrial
- Esophagus
- Hodgkin's Disease
- Leukemia
- Lung
- Lymphoma, Non-Hodgkin's
- Ovarian
- Pancreatic
- Prostate
- Skin, Melanoma
- Skin, Nonmelanoma
- Stomach
- Other
    - Other will have a free text box to allow additional response

- High cholesterol (total cholesterol 240 mg/dl or higher) ☐ **Yes** ☐**No** ☐ **Don't Know**

34. Has your doctor prescribed a special diet for you because of a medical condition such as diabetes, kidney problem, etc

    ☐ **Yes (go to 31a)** ☐**No**

    31a. What is the diet?

35. Do you use tobacco?

    ☐**Yes** ☐**No (go to 32a)**

    32a. Did you quit within the last 3 months?

    ☐**Yes** ☐**No**

36. Have you gained or lost 10 pounds or more in the last six months?

    ☐**Yes** ☐**No**

37. There are a variety of reasons why people change their lifestyle, please indicate if any of these reasons are true for why you want to engage in a healthier lifestyle.

- Because others like me better when I am in shape ☐ True ☐ False
- Because I would feel bad about myself if I didn't do it ☐ True ☐ False
- Because it is personally important to me to live a healthy lifestyle ☐ True ☐ False

- Because I simply enjoy living a healthier lifestyle ☐ True ☐ False

38. Are you willing to make a serious attempt at making changes in your food and activity choices within the next two weeks?

☐ **Yes (go to #39)** ☐**No (go to #41a)**

39. Because this is a study, I need to read you the following statement:

Your American Cancer Society is testing different services for incorporating better nutrition, increased physical activity and weight management into one's lifestyle. If you would like to join this research, we will randomly assign you to receive either our self-help materials or to participate in our telephone counseling program. We will check how well this service is working by calling you back up to three times within the next year to ask about your success in changing your lifestyle and how you feel about the help that your American Cancer Society has provided.

There are no known risks to you for taking part in our program, and if any part of the program makes you feel uncomfortable, let us know and we can skip that part. For quality assurance, your calls may be monitored, but any information you provide will be kept private and confidential.

Your taking part in this study is entirely voluntary and you may decide to quit at any time without penalty. We do hope that your taking part in our program will directly benefit you in your attempt to live a healthier lifestyle. If you have questions about this research study you may contact Dr. Youngmee Kim, the Principal Investigator at 404-329-7626 or Youngmee.Kim@cancer.org. If you have questions about your rights as a participant in this study, you may contact Dr. Karen Hegtvedt, Chair, Social, Humanist, and Behavioral Institutional Review Board, which oversees the protection of human participants. She can be reached at 404-727-7517 or khegtv@emory.edu. If you do not wish to take part, we can still provide materials and community resource information to help you learn more about nutrition, physical activity and weight management.

Are you willing to take part in this study?

☐Yes **(go to # 40)** ☐No (**go to #41)**

40. I'd like to ask you one last question. On a scale of 0-100%, where 0 is no chance and 100 is absolutely, what do you think your chances are for being able to make lasting changes in your food, activity and weight management choices? _____

Electronic randomization takes place to select caller for self-help materials or counseling
If chosen for counseling, go to #42
If chosen to receive self-help materials only, go to #43

41. Okay, we have some nutrition and physical activity information available we can mail, if you are interested? In addition, we may have some community resources in your area that could be beneficial to your goals of living healthier.

41a. Okay, we have some nutrition and physical activity information available we can mail, if you are interested? In addition, we may have some community resources in your area that could be beneficial to your goals of living healthier. Please feel free to call us back if you feel ready to make changes in your food and physical activity choices within the next 2 weeks.

Is there anything else I can do for you today?

(If yes) Answer questions/requests with applicable information from call handling resources

Thank you for calling your American Cancer Society, good bye.

If interested in receiving materials, Order "Living Smart" Fulfillment # 2042.00
If interested in community referrals, check CID or CRD for applicable resources.

42. We would like to offer you telephone counseling supplemented by self-help materials to help you make healthier choices. I will send you the self-help materials today. Also, we do have community resource information; there may be a support group in your area. Would you like me to look up that information for you? If you have time, would you like to speak to a counselor now?

Caller is then either transferred to a Counselor or an appointment is made.
If transferred, the Intake ends here, Intake Specialist sends the materials and the Counselor completes the first session or schedules the session.
If an appointment is made by the IS, go to #42a to end the call.

42a. Is there anything else I can do for you today?

(If yes) Answer questions/requests with applicable information from call handling resources.

Schedule the appointment and send the self-help materials, *Get on Track, Stay on Track*

Thank you for calling your American Cancer Society.

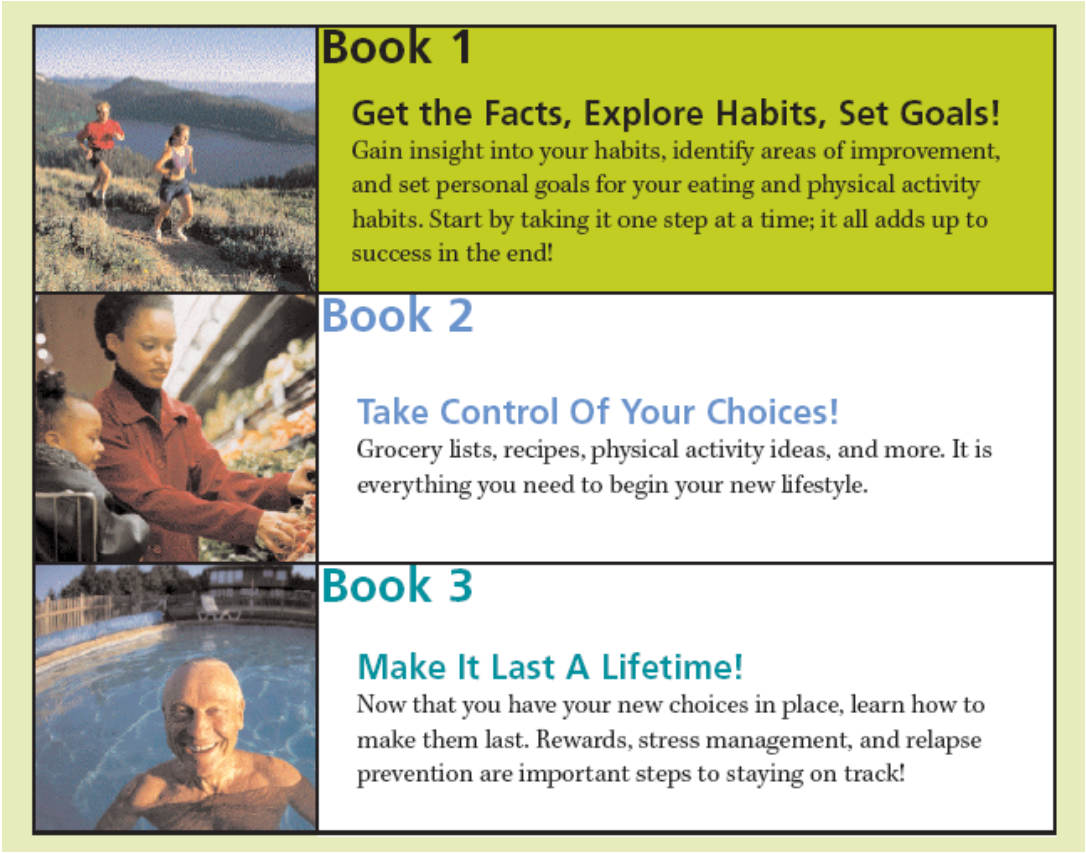
43. We would like to offer you our new self-help materials. This three-part guide provides information and helpful tips on incorporating good nutrition, physical activity and weight management into your lifestyle. In addition, we have access to community resource information. There may be a program or support group in your area. Would you like me to look up that information for you?

Send self-help materials, *Get on Track, Stay on Track*

Thank you for calling your American Cancer Society.

Appendix 2: Get on Track, Stay on Track Outline (March 2005)

Get on Track, Stay on Track is the three part guide designed to serve as the self-help materials provided to the participants in the Health for Life study.



Book 1

- Facts about benefits to overall health for good food choices and physical activity
- BMI chart and discussion of healthy weight ranges
- Goal Setting
- Gaining support/establishing a support system
- Keeping a food journal

Book 2

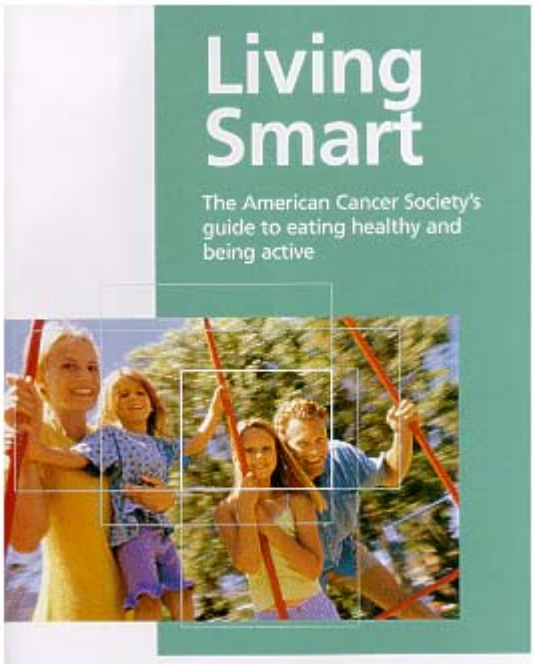
- Tips for reading a food label
- Standard portions and understanding portion sizes with visuals
- Healthy cooking/exchanging high fat ingredients with lower fat options
- Physical Activity ideas and suggestions for moderate versus vigorous
- Healthy recipes
- Shopping list with recommendation for basic ingredients in a healthy kitchen

Book 3

- Rewards and motivation
- Stress management
- Relaxation techniques
- Cognitive restructuring with thoughts and actions
- Lapse and Relapse prevention information
- Creating a supportive environment

Living Smart is the American Cancer Society's guide to eating healthy and being active. The pamphlet has some general suggestions for achieving success with good food choices and weight

management. This publication will be used to offer general educational information for callers who do not qualify or enroll in the Health for Life program.



Appendix 3: Health for Life Counseling Outline (March 2005)

Session One

- Confidentiality and format of sessions
- Review of lifestyle, hobbies, interests
- Reasons for wanting change
- Advantages and disadvantages of lifestyle change (decisional balance)
- Recommendation to read materials, use a food and physical activity journal to help decide what changes you can begin to make
- Identify where you need improvement and identify your strengths
- Recommendation to involve health care professional if on other medications
- Ask for guidance if needed, Drs., nutritionists, registered dieticians
- Introduction to key elements in plan:
  Food choices
  Physical activity
  Stress management
- Introduction to social support
- Scheduling of "change" date
- Commitment to create "change plan" in next session

Session Two

- Making a reasonable plan
- Reviewing using a food and activity journal
- Basic goal setting information:
  - Realistic goals means you are more likely to be successful
  - Moderation, make small changes over time
  - Set short term goals and set out to tackle them one at a time
  - Small changes eventually become regular behavior
- Review of reasons for changing
- Education about healthy eating (fruits, veggies, et)
- Elicit change strategy for eating
- Education about Physical Activity
- Elicit change strategy about activity

- Confirm selection of changes for Phase One of change plan
- Identifying and tracking triggers (stress) to eat
- Review of support system and how to use
- Creating written for plan for night before and days after

Session Three

- Review of Phase One change attempt
  Food Choices
  Physical Activity
- Teaching stress management strategies
- Teaching thinking and action strategies
- Discussion of challenges and evaluation of strategies
- Creating strategies for future tough situations
- Teach mental rehearsal
- Discussion of willpower vs. skill learning
- Discussion of slips/relapse prevention
- Review of support system
- Create timeline for Phase Two changes

Session Four

- Assessment of Phase One changes
- Identification of differences noted in energy and stress
- Assessment of changes in motivation and coping strategies
- Assess readiness to move to Phase Two
- Create written Phase Two plan (if appropriate)
- Evaluation of challenges in Phase Two
- Planning for future tough situations
- Review of successful strategies and mental rehearsal
- Create timeline for Phase Three implementation

Session Five

- Assessment of Phase Two changes
- Identification of differences noted in energy and stress

- Assessment of changes in motivation and coping strategies
- Assess readiness to move to Phase Three
- Create written Phase Three plan (if appropriate)
- Evaluation of challenges in Phase Three
- Review strategies with stress management, support, thoughts and rewards/motivation

Session Six

- Assessment of Phase Three changes
- Identification of differences noted in energy and stress
- Assessment of changes in motivation and coping strategies
- Establishing personal long-term goals with food, activity and weight
- Long-term planning for future tough situations
- Discussion of lifestyle changes and long term maintenance
- Review strategies with stress management, support, thoughts and rewards/motivation

Boosters (1, 2, & 3)

- Check-in on status of changes and success with long-term goals
- Review challenges and strategies in place to maintain goals
- Review strategies with stress management, support, thoughts and rewards/motivation

Appendix 4: Health for Life Evaluation 3, 6 and 12 months (March 2005)

PARTICIPANT INFORMATION	
STATE:	
CONSTITUENT ID:	
PHONE NUMBER:	
CREATION DATE:	/ /

FOLLOW-UP TYPE**: 3 MONTH**

6 MONTH

12 MONTH

Please circle one

Evaluator Name: _____

This is < *Evaluator Name*> calling from the Health for Life evaluation research unit at your American Cancer Society.

We are studying the effects of the assistance you recently received from the Health for Life study. This interview will take 5 to 10 minutes to complete. May I have your permission to continue?

The purpose of the interview is to find out about your food, activity and weight management choices after your call to your American Cancer Society.

It is important that you know that your participation is entirely voluntary. You may decide not to take part or to quit the interview at any time without penalty.

There will be no risk or discomfort to you in providing responses to the questions asked in this interview, however should you feel uncomfortable in providing a response to a specific question, you may skip that question.

Your participation will benefit the American Cancer Society by providing useful information on the effectiveness of the assistance that we've provided you.

Your answers will be kept confidential. Your name will not be known to anyone nor will it be used in any reports or publications from this study.

Are you willing to participate in this interview?
If yes, say "Good, thank you." and continue on the next page.
If no, thank the person for her or his time, probe for why not, and describe below.

I'd like to ask you a few questions about your food, activity and weight management choices since you contacted your American Cancer Society's Health for Life program.

1. In the past three months, on a typical day, how many servings (1 serving = ½ cup) of fruits do you consume each day? Count any kind of fruit- fresh, frozen, canned or dried. (Please exclude fruit juice.)

2. In the past three months, on a typical day, how many servings (1 serving = ½ cup) of vegetables do you consume each day? Count raw, cooked, canned or frozen as well as cooked dried beans. (Please exclude vegetable juices, white potato products, fried or processed starchy vegetables like french fries, potato chips and hash browns.)

3. In the past three months, on a typical day, how many minutes of physical activity do you get each day? (divide minutes of activity by 7 if they are active less than daily) (physical activity to include mild, moderate and vigorous activity; i.e. minimal effort/no sweating, sweat lightly, or rapid heart beating/heavy sweating, respectively)

4. How often do you eat cheese or cheese spread (not low-fat)?

- ❑ Never
- ❑ < 1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

5. How often do you eat beef, pork or lamb? Include bacon, sausage, salami and hot dogs.

- ❑ Never
- ❑ <1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

6. What kind of milk do you drink?

- ❑ Whole milk
- ❑ 2% fat milk
- ❑ 1% fat milk
- ❑ Skim, nonfat, or ½ % fat milk

- ❑ Soy milk
- ❑ Rice milk
- ❑ Other

7. How many days of the week do you eat breakfast?

8. How often do you eat regular potato chips, tortilla chips, corn chips, etc.?

- ❑ Never
- ❑ < 1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

9. How often do you eat sugary foods like cake, candy, soda, sugary cereals?

- ❑ Never
- ❑ <1 time per day
- ❑ 1 time per day
- ❑ 2 times per day
- ❑ 3 times per day
- ❑ 4 times per day
- ❑ 5 or more times per day

10. What is your current weight?

11. Have you felt sad or blue almost everyday for the last three months?

- ❑ Yes
- ❑ No

11a. In the past three months, do you consider yourself satisfied with your life?

- ❑ Yes
- ❑ No

12. What do you think is the minimal number of daily servings of fruits and vegetables recommended
to reduce health risks?

13. What do you think is the minimal amount of weekly physical activity recommended to reduce health
risks?

14. What are ways you know of to reduce health risks? (ability to choose up to 5 choices)

- ❑ Eat fruits and vegetables
- ❑ Stay out of the sun/Use sunscreen
- ❑ Get plenty of sleep
- ❑ Reduce stress

- ❑ Reduce amount of fat consumed
- ❑ Lose weight
- ❑ Limit alcohol consumption
- ❑ Don't smoke
- ❑ Exercise
- ❑ Limit red meat consumption
- ❑ Minimize toxic environmental exposure
- ❑ Other
  (Other has a free text box to allow additional response)

15. Do you believe there is a link between being overweight and cancer risk?

- ❑ Yes
- ❑ No

16. Do you know what your Body Mass Index (BMI) is?

- ❑ Yes (go to 16a)
- ❑ No

16a. What is your BMI?

17. Do you know what the healthy BMI range is?

- ❑ 19-24.9
- ❑ 25-29.9
- ❑ Greater than 30
- ❑ Don't Know

18. Has your doctor diagnosed you with any of the following conditions?

- Diabetes ☐ Yes ☐No ☐ Don't Know

- High blood pressure (140/90mm Hg or higher) ☐ Yes ☐No ☐ Don't Know

- Cancer ☐ Yes (choose the primary cancer site/s) ☐No
  - If yes, use drop down menu to choose what type of cancer
    - Anal
    - Bile Duct
    - Bladder
    - Bone
    - Bone Metastasis
    - Brain (adult)
    - Brain (child)
    - Breast
    - Cancer of Unknown Primary
    - Cervical
    - Colon and Rectum
    - Endometrial
    - Esophagus
    - Hodgkin's Disease
    - Leukemia
    - Lung
    - Lymphoma, Non-Hodgkin's
    - Ovarian
    - Pancreatic
    - Prostate
    - Skin, Melanoma
    - Skin, Nonmelanoma
    - Stomach
    - Other
      - Other will have a free text box to allow additional response

- High cholesterol (total cholesterol 240 mg/dl or higher) ☐ Yes ☐No ☐ Don't Know

18a. In general, would you say your health is:

☐ Excellent ☐ Fair
☐ Very Good ☐ Poor
☐ Good ☐ Refuse to answer

19. Has your doctor prescribed a special diet for you?

☐ Yes (go to 18a) ☐No

18a. What is the diet?

20. Do you use tobacco?

☐Yes ☐No (go to #20a)

20a. Did you quit within the last 3 months?

☐ Yes ☐No

21. Since contacting the Health for Life program, did you make changes to your food choices?

☐ Yes ☐No

21a. Did you make changes to your physical activity choices?

☐ Yes ☐No

22. Since contacting the Health for Life program have your family or close friends helped you with your
food and activity choices?

☐ Yes ☐No

22a. Do you think the Health for Life program you participated in has influenced the food and
activity choices among your family or close friends?

☐ Yes ☐No

23. Since you contacted the Health for Life program have you gotten other forms of assistance to help you with food, activity or weight management choices?

_____ 0. No

_____ 1. Other counseling

_____ 2. Bio-feedback

_____ 3. Hypnosis

_____ 4. Acupuncture

_____ 5. Internet

_____ 6. Support groups

_____ 7. Personal Trainer

_____ 8. Nutritionist

_____ 9. Other (specify) _____

24. On a scale of 0-100%, where 0 is no chance and 100 is absolutely, what do you think your chances are

for being able to make lasting changes in your food, activity and weight management choices? _____

Now I'd like to ask you about your satisfaction with our services. Could you please rate the next question for me on a scale from 1 to 10, where "1" represents "Poor"and"10" represents "Excellent"?

25. Please tell me how you would rate your overall experience with the American Cancer Society's Health for Life program?

_____ (If 1, 2, or 3 go to 25a)
(If 4 or more go to 26)

25a. Can you tell me why you feel that way?

PROBE: Are there any other reasons?

26. On a scale from 1 to 10, where "1" represents "did not meet expectations," and "10" represents "exceeded expectations", overall, how would you say the information you received from the American Cancer Society met your needs? _____

27. Did you receive our materials in the mail? (only ask at 3 month)

_____1 Yes (go to 28)
_____0 No (go to 29)
_____2 Don't recall (go to 29)

28. On a scale of 1 - 10, where "1" represents "not at all helpful" and "10" represents "extremely helpful", how helpful do you feel the self-help materials were in your ability to make changes?

_____

29. Did the American Cancer Society refer you to any community resources (support groups, etc.)?

_____1 Yes (go to 29a)
_____0 No (go to 30)
_____2 Don't recall (go to 30)

29a. Did you use them? _____1 Yes (go to 29b)
_____0 No (go to 30)

29b. Which ones? _____ (go to 30)

30. After you called the American Cancer Society were you set up to receive counseling?

____ 1 Yes & Received at least 1 session (go to 31)
____ 2 Yes – received zero sessions (go to 40)
____ 3 No – was not set up to receive counseling (go to 40)
____ 4 Don't recall (go to 40)
____ 5 Refused to answer (go to 40)

31. On a scale from 1 to 10, where "1" represents "Poor," and "10" represents "Excellent" please tell me how you would rate the overall quality of the service provided by the counselor you spoke with?

_____ (If 1, 2, or 3 go to 32)
(If 4 or more go to 33)

32. Ok is there anything in particular that you can tell us about why you feel that way?

PROBE: Are there any other reasons?

33. On a scale of 1 - 10, where "1" represents "not at all helpful" and "10" represents "extremely helpful", how helpful was your counselor in your efforts to make changes?

_____

34. On a scale of 1 - 10, where "1" represents "not at all satisfied" and "10" represents "extremely satisfied", please rate your satisfaction with the information you received from your counselor.

_____

On a scale from 1 to 10, where "1" represents "did not meet expectations," and "10" represents "exceeded expectations", please rate your counselor(s) on the following attributes:

35. Their sincerity_____

36. Having a positive and helpful attitude _____

37. Having compassion for your situation _____

38. Ability to understand you _____

39. Conveying confidence in your ability to do well at changing your lifestyle _____ (go to 42)

40. On a scale from 1 to 10, where "1" represents "Poor," and "10" represents "Excellent" please tell me how you would rate the overall quality of the service provided by the person you spoke with during your first call?

_____ (If 1, 2, or 3 go to 41)
(If 4 or more go to 42)

41. Ok is there anything in particular that you can tell us about why you feel that way?

PROBE: Are there any other reasons? (go to 42)

42. What parts of the telephone assistance or self-help materials were most useful to you?

43. How could this service be improved in the future?

44. This question is another way to help measure your satisfaction with the services that you received overall. If you had to assign a dollar value to the services you received, what would it be?

$_____

This concludes the interview.

<3 and 6 Month follow-up ONLY… Do not say this with 12 Month follow-up>

You will be contacted again in a few months to see how you are doing.

Please feel free to ask questions you may have about the interview or about your rights as a research subject. Do you have any questions now?

All**:** If other questions occur to you later, we have a number that you can call. Do you have a pen and paper? You may contact Dr. Youngmee Kim, the Principal Investigator at 404-329-7626. And for concerns specifically about your rights as a research subject, you can call Dr. Karen Hegtvedt, Chair of Social, Humanist, and Behavioral Institutional Review Board, which oversees the protection of human participants at 404-727-7517.
Thank you for your participation on behalf of your American Cancer Society.

Appendix 5: Power Simulations

$$Outcome(t) = \alpha + (\beta_0 + \beta_1 \times I + + \beta_2 \times E + + \beta_3 \times I \times E) \times t$$

Outcome(t): outcome measured at time t (srv/day of FVC, min/day of PA)

α: mean outcome at baseline

β0: time trend/Hawthorn effect in both intervention and control groups

β1: change in outcome from baseline to 6 months due to intervention (among those unexposed to environment)

β2: change in outcome from baseline to 6 months due to environmental exposure (among those in the control group)

β3: additional change in outcome from baseline to 6 months due to environmental exposure among those in the intervention group (or due to intervention among those in the exposed group).

Assume:

1. Alpha and the betas are fixed, but unknown quantities.
2. P(I=1)=P(I=0)=0.5 This is from the design.
3. P(E=1) taken from literature
4. I and E are independent due to randomization Good.
5. Ceffective sample size assuming independence of observations
    a. Take into account 60% losses by 6 months

i. Actual N in each arm is 750 (1250*0.6)

    b. Assuming ICC from 0.002 to 0.05

i. Changes effective sample size

6. No change in outcome across time among those not in self-help/control group

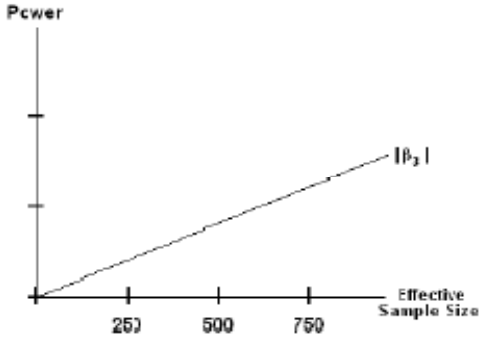
a. β0=0

b. β2=0

7. Estimates of β1 can be obtained from the clinical trial literatures

1. Parameter of interest is β3

Power to detect β3 perhaps best represented as a surface plot with X and Y axes being effective sample size and |β3|, and Z axis representing power



Keeping power constant at 80% and the alpha level constant at 5%, the minimum detectable $|\beta_3|$, given a range of the effective sample size (ESS) and exposure proportion, was estimated using a simulation method using 100 simulations of the data each time. The value of $|\beta_3|$ for which approximately 80% of trials resulted in a rejection at the 5% significance level, was taken to be the minimum detectable sample size. This was plotted on the Z axis against ESS and exposure proportion.

Simulation Coding

```
data one;
alpha=30;       sdalpha=0;      *mean and sd of baseline outcome;
beta0=0;        sdbeta0=0;      *mean and sd of effect of time/hawthorn effect;
```

```
delta1=10;      sddelta1=0;     *mean and sd of intervention effect among those
                                unexposed;
delta2=0;       sddelta2=0;     *mean and sd of exposure effect among those in
                                the control group;
gamma=5;        sdgamma=60;     *mean and sd of exposure effect among those in
                                the intervention group;

prope=0.25;                     *proportion of those who are exposed;

loss=0.6;
reduce=0.6;

keep=(1-loss)*(1-reduce);

intv=1273*keep*reduce;
sh=1199*keep*reduce;            *number in intervention and self-help arms, total
                                N's taken from sample size at the September 2007
                                update;
run;

%macro power();
%do i=1 %to 100;
```

```
data didi&i;                    *generates 100 datasets with properties given in
                                data set one;
set one;
dataset=&i*1;
do k=0 to (intv-1) by 1;
vari=1;
baseline=max(0,(rannor(0)*sdalpha+alpha));
vare=ranbin(0, 1, prope);
outcome=max(0, baseline
        +(beta0+rannor(0)*sdbeta0)
        +(delta1+rannor(0)*sddelta1)*vari
        +(delta2+rannor(0)*sddelta2)*vare
        +(gamma+rannor(0)*sdgamma)*vare*vari)
;
output;
end;
do k=0 to (sh-1) by 1;
vari=0;
baseline=max(0,(rannor(0)*sdalpha+alpha));
vare=ranbin(0, 1, prope);
outcome=max(0, baseline
        +(beta0+rannor(0)*sdbeta0)
        +(delta1+rannor(0)*sddelta1)*vari
```

```
                +(delta2+rannor(0)*sddelta2)*vare
                +(gamma+rannor(0)*sdgamma)*vare*vari)
        ;
        output;
        end;

        run;
%end;

data combined;                  *concatenates the 100 datasets generated by the
                                macro. Increases efficiency in running proc glm;
set
%do i=1 %to 100;
didi&i
%end;;
run;
%mend;

%power(); run;

ods output ParameterEstimates=pvalues;
proc glm data=combined;         *runs the model separately for each of the 100
                                datasets generated in the macro;
```

```
model outcome=baseline vare vari vare*vari;
by dataset;
run;
ods output close;

data three;
set pvalues;                    *extracts the p-values for the parameter of interest
                                in each of the 100 GLMs executed in the previous
                                procedure.;
where Dependent='outcome' and Parameter='vare*vari';
if probt<=0.05 then sig=1;
if probt>0.05 then sig=0;
run;

proc print data=three;
run;

proc freq data=three;           *produces a frequency table to indicate what
                                proportion of the 100 models resulted in a
                                significant result for the parameter of interest;
tables sig;
run;
```

Appendix 6: Search Strategy for Systematic Review of Effect Modification in Trials to Increase Fruit and Vegetable Consumption

Intervention-related search terms. Adapted from Brunner et al [1], Foster et al [2], Kremers et al [3]. Combined using the "or" operator.

explode "Communication"/ all subheadings

explode "Practice-Guidelines"/ all subheadings

explode "Counseling" tree: 3/ all subheadings

explode "Diet-Therapy"/ all subheadings

explode "Health-Education"/ all subheadings

explode "Life-Style"/ all subheadings

diet* adj (therap* or educat* or counsel* or intervention* or treatment*)

nutriti* adj (therap* or educat* or counsel* or intervention*)

health adj (therap* or counsel* or educat*)

group adj counsel*

brief adj intervention*

health adj behav* adj intervention*

advice

leaflet*

video*

guideline*

lifestyle* near chang*

diet* near chang*

intake* near (increas* or decreas* or reduc* or rais* or low* or chang* or restrict* or high*)

consumption near (increas* or decreas* or reduc* or rais* or low* or chang* or restrict* or high*)

(salt or sodium) near (decreas* or reduc* or low* or chang* or restrict*)

(fat* or cholesterol) near (decreas* or reduc* or low* or chang* or restrict*)

(fish or fruit* or vegetable*) near (increas* or rais* or chang* or high*)

Health education/

Primary prevention/

Health promotion/

Behaviour therapy

Cognitive therapy

Primary health care

Workplace/

promot$.tw.

educat$.tw.

program$.tw.

health promotion

intervention

Diet-related search terms. Adapted from Brunner et al [1] and Kremers et al [3]. Combined using the "or" operator.

explode "Fruit"/ all subheadings

explode "Vegetables"/ all subheadings

explode "Food-Habits"/ all subheadings

"Diet"/ all subheadings

"Diet-Therapy"/ all subheadings

diet* in ti,ab

food* in ti,ab

mediterranean* in ti,ab

vegetable* in ti,ab

fruit* in ti,ab

legum* in ti,ab

roughage in ti,ab

5-a-day

food habits

food

fruit and vegetable consumption

fruits

nutrition

Diet- and intervention-related search terms were combined using the "and" operator.

References

1. Brunner EJ, Rees K, Ward K, Burke M, Thorogood M. Dietary advice for reducing cardiovascular risk. Cochrane Database Syst Rev 2007(4):CD002128.
2. Foster C, Hillsdon M, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005(1):CD003180. pub2.
3. Kremers SP, de Bruijn GJ, Droomers M, van Lenthe F, Brug J. Moderators of environmental intervention effects on diet and activity in youth. Am J Prev Med 2007;32(2):163-72.