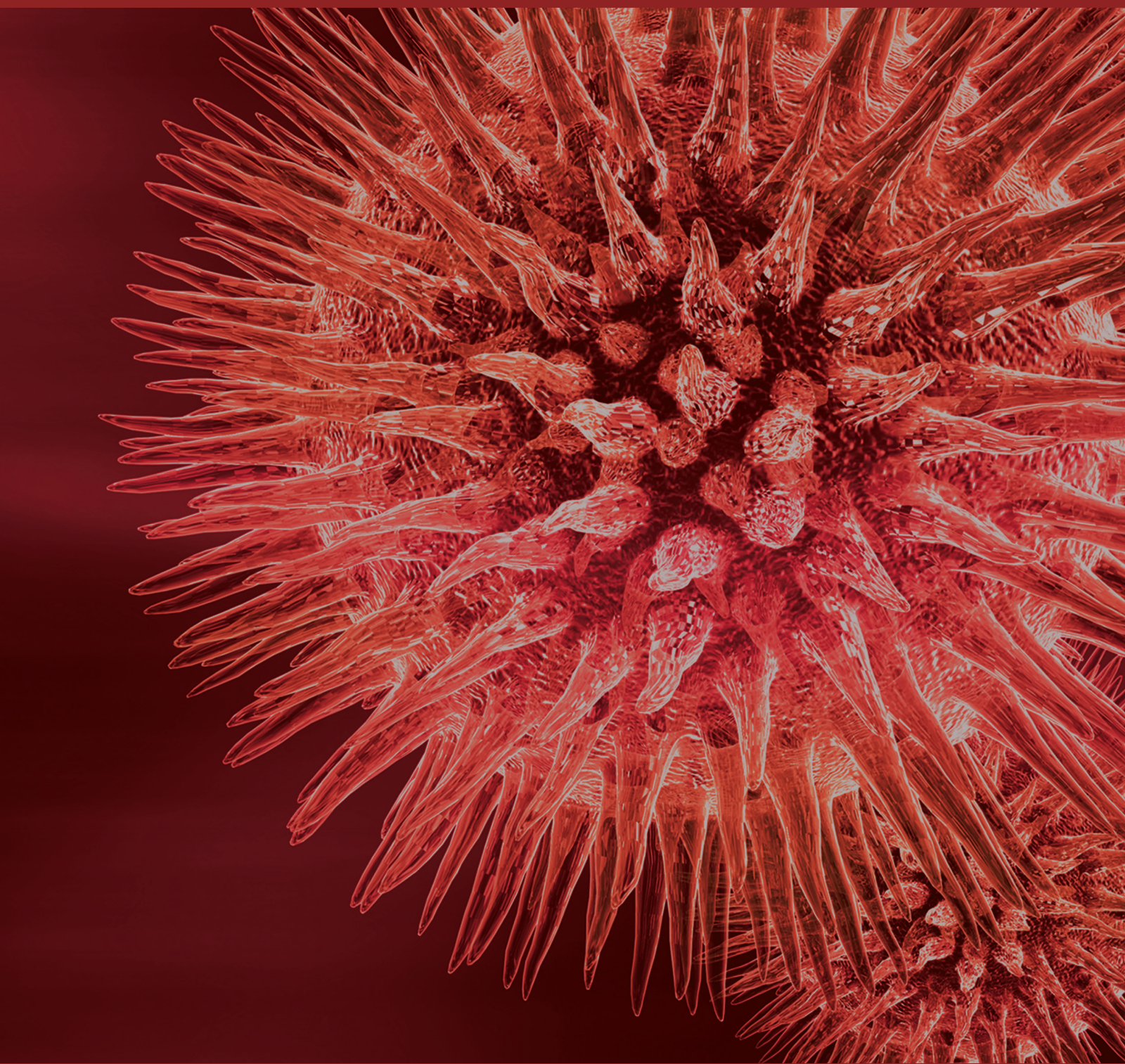
BioMed Research International

Evidence-Based Public Health

Guest Editors: Giedrius Vanagas, Malgorzata M. Bala, and Stefan K. Lhachimi



Evidence-Based Public Health

BioMed Research International

Evidence-Based Public Health

Guest Editors: Giedrius Vanagas, Malgorzata M. Bala, and Stefan K. Lhachimi

Contents

Hindawi Publishing Corporation
BioMed Research International
Volume 2016, Article ID 5681409, 2 pages
http://dx.doi.org/10.1155/2016/5681409

Editorial

Evidence-Based Public Health

Stefan K. Lhachimi,[1,2,3] Malgorzata M. Bala,[4] and Giedrius Vanagas[5]

[1]*Cooperative Research Group for Evidence-Based Public Health (ebph), Institute for Public Health and Nursing Research, University of Bremen and Leibniz Institute for Prevention Research and Epidemiology-BIPS, Achterstraße 30, 28359 Bremen, Germany*
[2]*Health Sciences Bremen, University of Bremen, Bibliothekstraße 1, 28359 Bremen, Germany*
[3]*Institute of Health Service Research and Health Economics, Faculty of Medicine, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany*
[4]*Department of Hygiene and Dietetics, Jagiellonian University Medical College, 31-034 Krakow, Poland*
[5]*Department of Preventive Medicine, Lithuanian University of Health Sciences, LT-49264 Kaunas, Lithuania*

Correspondence should be addressed to Giedrius Vanagas; giedrius.vanagas@lsmuni.lt

Received 20 December 2015; Accepted 20 December 2015

Public health decision making can be a complicated process because of the complex nature of inputs and the need for group decision making. Nevertheless, public health research and practice during the last century gained many notable achievements and contributed to the 30-year gain in life expectancy. Despite these accomplishments, a greater attention to evidence-based approaches may be helpful. In its most straightforward definition, evidence-based public health (EBPH) means applying the principles of evidence-based medicine (EBM) to the field of public health [1]. However, the randomized clinical trials—the gold standard in EBM—are not always applicable in investigating public health problems. Key components of a broader definition of EBPH include making decisions on the basis of the best available scientific evidence by using sound data collection and research methods while engaging with the affected community in decision making. An evidence-based approach to public health could potentially have numerous direct and indirect benefits, including access to more and higher-quality information on best practice, a higher likelihood of successful prevention programs and policies, greater workforce productivity, and more efficient use of public and private resources.

In this special issue we collect several contributions to the emerging field of evidence-based public health, covering different aspects of the EBPH policy cycle, that is, from evidence generation and evidence synthesis to evidence communication and policy recommendation [2]. We are particularly delighted about the broad geographical coverage and methodological range of the included paper. The utilized methodology ranges from randomized controlled trial (RCT) to simulation studies with observational studies still being the most popular research design. The study titled "Worksite Tobacco Prevention: A Randomized, Controlled Trial of Adoption, Dissemination Strategies, and Aggregated Health-Related Outcomes across Companies" by V. Friedrich et al. conducted an RCT to compare various approaches to worksite tobacco prevention. An instructive finding of this study is that special attention must be paid to the dissemination of findings if adoption of proven public health measure should increase. Similarly, the article "Strong Public Health Recommendations from Weak Evidence? Lessons Learned in Developing Guidance on the Public Health Management of Meningococcal Disease" by G. Hanquet et al. shows the necessity to be transparent about the quality of the underlying evidence when giving policy recommendations.

Systematic reviews and meta-analysis are a workhorse of evidence-based medicine and two studies in our special issue demonstrate clearly that this is also the case for evidence-based public health. G. A. Kelley et al. conducted a systematic review titled "Exercise and BMI in Overweight and Obese Children and Adolescents: A Systematic Review and Trial Sequential Meta-Analysis" utilizing a trial sequential

meta-analysis approach. This approach combines conventional meta-analysis methodology with meta-analytic sample size considerations; inferences derived from this method may potentially improve reliability of estimates. The systematic review "Sexual Risk Behaviors and HIV Infection among Men Who Have Sex with Men and Women in China: Evidence from a Systematic Review and Meta-Analysis" by H.-Y. Wang et al. sheds light on potential transmission mechanism of HIV/AIDS in China, indicating that more emphasis must be put on prevention measures for men who have sex with men and women.

A crucial, but at times overlooked, factor for a successful public health policy is the availability of trained medical staff. In our special issue two articles study this important topic. Interestingly enough, both are from Sub-Saharan Africa, a region that constantly faces the challenge of brain drain, that is, the outmigration of well-trained individuals. The study from Ethiopia titled "The Prevalence of Skilled Birth Attendant Utilization and Its Correlates in North West Ethiopia" by M. Alemayehu and W. Mekonnen looks at the prevalence of skilled birth attendants demonstrating that both social and technical factors must be addressed if this prevalence should be improved in the future. The study titled "Working Atmosphere and Job Satisfaction of Health Care Staff in Kenya: An Exploratory Study" by K. Goetz et al. investigates the factors that influence job satisfaction for health care works in Kenya. Clearly, improving job satisfaction may prove to be a very cost-effective tool in increasing retention rates among health care staff in Africa.

The contribution "Intervention Mapping to Adapt Evidence-Based Interventions for Use in Practice: Increasing Mammography among African American Women" by L. Highfield et al. uses the topic of mammography among African American women to demonstrate the use of intervention mapping as a tool for a systematic planning process. Their use of a simplified framework (IM Adapt) allowed adapting and implementing an evidence-based intervention to help underserved African American women to keep appointments for mammography screening. A separate paper titled "Evaluation of the Effectiveness and Implementation of an Adapted Evidence-Based Mammography Intervention for African American Women" by L. Highfield et al. evaluated the effectiveness of the identified intervention and added evidence to the finding that sequentially measuring efficacy and effectiveness of an evidence-based intervention, followed by implementation, may be missing important contextual information.

A paper from Croatia titled "Perinatal Health Statistics as the Basis for Perinatal Quality Assessment in Croatia" by U. Rodin et al. describes how the national perinatal health audit improved after introducing reporting criteria as recommended by WHO and PERISTAT. This database now allows comparison in perinatal outcome with other countries and targeting areas of improvement in a more evidence-based fashion. The contribution "Health Impacts of Increased Physical Activity from Changes in Transportation Infrastructure: Quantitative Estimates for Three Communities" by T. J. Mansfield and J. M. Gibson utilizes the approach of quantitative health impact assessment to assess the implications of increased physical activity from changes in transportation infrastructure. To this end they use a simulation tool (DYNAMO-HIA) that was specifically developed for such applications.

An Australian paper titled "Public Concern about the Sale of High-Caffeine Drinks to Children 12 Years or Younger: An Australian Regulatory Perspective" by C. M. Pollard et al. presents a study connected to the epidemiology of caffeine intake. Usually this is attributed mostly to coffee consumption. However, this paper demonstrates that growing concerns regarding the consumption of caffeinated beverages such as energy drinks by children and adolescents exist in communities in Western Australia, in particular by females and those living with children. These concerns increase with age. The final paper of this special issue titled "Smoke-Free Workplaces Are Associated with Protection from Second-Hand Smoke at Homes in Nigeria: Evidence for Population-Level Decisions" by D. Kaleta et al. aims to support decision making at the population level. These authors find evidence that smoke-free workplaces have the important additional effect of stimulating smoke-free homes in Nigeria and in turn reduce second hand smoking exposure of children.

In summary such collection of papers covering different issues relevant to EBPH will provide evidence for decision makers and will contribute to the development of this field.

Acknowledgments

The guest editors of this special issue thank the reviewers and all authors for providing important contributions in this issue.

Stefan K. Lhachimi
Malgorzata M. Bala
Giedrius Vanagas

References

[1] T. L. Pettman, R. Armstrong, J. Doyle et al., "Strengthening evaluation to capture the breadth of public health practice: ideal vs. real," *Journal of Public Health*, vol. 34, no. 1, pp. 151–155, 2012.

[2] A. Gerhardus, J. Breckenkamp, and O. Razum, "Evidence-based public health. Prevention and health promotion in the context of science, values and (vested) interests," *Medizinische Klinik*, vol. 103, no. 6, pp. 406–412, 2008.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 850132, 12 pages
http://dx.doi.org/10.1155/2015/850132

Research Article

Sexual Risk Behaviors and HIV Infection among Men Who Have Sex with Men and Women in China: Evidence from a Systematic Review and Meta-Analysis

Hong-Yi Wang,[1,2] Jun-Jie Xu,[1,2] Hua-Chun Zou,[3] Kathleen Heather Reilly,[4] Christiana Meng Zhang,[5] Ke Yun,[1,2] Yong-Ze Li,[1,2] Yong-Jun Jiang,[1,2] Wen-Qing Geng,[1,2] Hong Shang,[1,2] and Ning Wang[1,5]

[1]*Key Laboratory of AIDS Immunology of National Health and Family Planning Commission, Department of Laboratory Medicine, The First Affiliated Hospital, China Medical University, Shenyang 110001, China*
[2]*Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou 310000, China*
[3]*Kirby Institute, University of New South Wales, Sydney, NSW 2052, Australia*
[4]*National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China*
[5]*Georgetown University School of Medicine, 3900 Reservoir Road Northwest, Washington, DC 20007, USA*

Correspondence should be addressed to Hong Shang; hongshang100@hotmail.com and Ning Wang; wangnbj@163.com

Received 26 January 2015; Revised 7 June 2015; Accepted 16 June 2015

Academic Editor: Malgorzata Bala

Objectives. To understand the current risk of HIV infection and transmission and further elucidate the underlying risk factors among men who have sex with men and women (MSMW) in China. *Methods*. Following PRISMA guidelines, we conducted a systematic review and meta-analysis of searching through Chinese and English available literature databases between January 2000 and June 2014 to identify articles. *Results*. Thirty-six articles (including 19,730 MSMW and 53,536 MSMO) met the selection criteria and the aggregated results found that MSMW have significantly higher HIV prevalence than MSMO (6.6% versus 5.4%, OR = 1.27, 95% CI = 1.01–1.58). A higher proportion of MSMW had commercial male partners in the past 6 months (18.3% versus 12.2%, OR = 1.56, 95% CI = 1.01–2.42). Additionally, substance use in the past 6 months was significantly more frequent among MSMW than MSMO (alcohol use: 27.1% versus 13.1%, OR = 2.53, 95% CI = 2.14–2.99; illicit drug use: 5.3% versus 2.5%, OR = 2.09, 95% CI = 1.48–2.95). *Conclusion*. A higher proportion of commercial sex and substance use among MSMW may be a potentially indicative factor for significantly higher HIV prevalence compared to MSMO. Targeted interventions should aim at increasing the frequency of HIV/STIs screening and preventing high risk commercial sex and substance use among MSMW to decrease their HIV transmission to the general population.

1. Introduction

Men who have sex with men (MSM) are considered as a group of high risk for HIV infection in many countries. According to The Joint United Nations Programme on HIV and AIDS (UNAIDS), the prevalence of HIV infection among MSM in capital cities from nearly 80 countries is on average 13 times higher than that of general populations in these countries [1]. This disparity is also prominent in China; a recent meta-analysis reported that almost 6.5% of MSM were living with HIV [2], which was more than 100 times higher than that in the general population (0.058%) [3]. The latest national report revealed that the proportion of newly diagnosed HIV cases due to male homosexual contact has increased from 12.2% in 2007 to 23.4% in 2014 [4, 5].

Homosexuality is still highly discriminated against in China, which partly results in bisexual behavior becoming one of the biggest obstacles in implementing HIV/AIDS prevention and control interventions among MSM [6, 7]. Additionally, marriage is a family obligation of Chinese traditional culture, which prompts over 80% of MSM eventually getting married to female to hide their real sexual identity

and/or carry on family name [8, 9]. Only one in ten MSMW discloses homosexuality to their female partners [10, 11], and unprotected sexual behaviors with female sex partners were prevalent among MSMW [12], which may bridge HIV from the MSM population to the general population.

Previous studies have estimated peraction transmission probability of HIV during anal or vaginal sex among MSM [13, 14]. Clear understanding of the relative risk may help inform MSM of the potential risk which may in turn render them to avoid high risk sexual behaviors, especially among those who make condom use decisions based on penetrative sex type and role in anal sex (receptive or insertive anal sex) and self-reported HIV status of partner. However, until now little literature had vigorously compared the relative risk of HIV infection among MSMW, based on all available data in China.

Existing studies had contradictory findings on the relative risk of HIV infection comparing MSMW and MSMO in China. A recent published national cross-sectional survey of 47,231 MSM from 61 cities in China indicated that MSMW have a lower HIV prevalence than MSMO [15]. However, this conclusion is contradictory to a previous meta-analysis study, which found that MSMW in China have a significantly higher HIV prevalence than MSMO [16]. Until now, no new studies indicated whether or not Chinese MSMW really have higher risk of HIV infection than MSMO. An explicit assessment of the risk of HIV infection in MSMW populations is useful for understanding the dynamic of national HIV epidemic among MSM and developing targeted interventions. Additionally, many new researches on bisexual behavior of MSM have been published in recent years. As such, we also include the latest extensive literature in order to gain a more comprehensive and current understanding of national HIV epidemic among MSM in China. Besides, previous meta-analysis on bisexual behavior among MSM in China failed to explore the underlying differences in HIV-related behavioral risk factors between MSMW and MSMO, which are essential to contextualize behavioral interventions of high risk subpopulations among MSM in China. Therefore, we conducted this systematic review and meta-analysis to compare the disparities in HIV prevalence between MSMW and MSMO and examine behavioral factors underlying the disparities. Three research questions were addressed: (1) do MSMW in China have significantly higher HIV prevalence than MSMO; (2) what are the differences between risk behavioral factors of MSMW and MSMO in China; (3) do MSMW engage in more risk sexual behaviors in different proportions than MSMO that might help to explain the differences in the effect size for HIV prevalence?

2. Methods

2.1. Literature Search. This systematic review and meta-analysis adheres to the PRISMA guidelines [17, 18]; comprehensive literature search was conducted using the following databases for literature published between January 2000 and June 2014 to identify articles: PubMed, Web of Knowledge, Google Scholar, Chinese National Knowledge Infrastructure, VIP, and Wanfang Data. In addition, the reference lists of pertinent articles were examined for additional relevant studies. The combinations of search terms include "bisexual behavior," "bisexual men," "MSMW," "men who have sex with men and women," "BBM," "behaviorally bisexual men," "MSM," "men who have sex with men," "homosexual men," "gay," "MSMO," "men who have sex with men only," "HIV" and "risk behaviors," "risk factor," "unprotected sex," and "condom use."

2.2. Inclusion/Exclusion Criteria. Studies were included in the meta-analysis if they reported quantitative data on HIV prevalence and risky behaviors rate among both MSMW and MSMO. Studies were excluded if (1) they were duplicate reports; (2) they failed to report risky behaviors or HIV prevalence among both MSMW and MSMO; (3) HIV status was self-reported and not confirmed by test; (4) they did not mention the period of the recall window of risk behaviors; and (5) they were postintervention studies. For research data repeatedly be published in multiple articles, the most comprehensive article was included in the meta-analysis.

2.3. Quality Assessment. The quality of studies was assessed using the quality assessment checklist for observational studies (QATSO score), a validated quality assessment tool for HIV prevalence/risk behaviors among MSM [19]. Items were scored as 1, 0, and NA, which represents "yes," "no," and "not applicable," respectively. The total score of each eligible study must have been above 33% (0% and 33%, 33% and 66%, and 67% and 100%, corresponded to "bad," "satisfactory," and "good quality," resp.).

2.4. Data Extraction. Bisexuality was operationalized using a classification of MSMW or MSMO over any timeframe (behavior recall window) assessed by researchers. Data from eligible studies were extracted with the following information by two reviewers independently: (1) general information about each selected article was extracted (e.g., first author, publication year, study period, study location, method of recruitment, sampling method, and behavior recall window). (2) HIV infection: the prevalence of HIV infection. (3) Behavioral information: (1) condom use: unprotected anal intercourse (UAI) with males, unprotected receptive anal intercourse with males (URAI), and unprotected insertive anal intercourse (UIAI) with males and UAI with commercial male partner; (2) anal sex role: mainly insertive anal intercourse (IAI), mainly receptive anal intercourse (RAI), and both of the two roles; (3) male partner: multiple (≥2) male partners have casual male partners and have commercial male partner (include purchasing sex from male sex workers (money boys and MBs) or selling sex to males); (4) substance use: alcohol use and ever illicit drug use (including opiate, cocaine, amphetamine-type stimulants (ATS), cannabis, and hallucinogens [20]).

2.5. Statistical Analysis. To assess the differences in HIV infection and related risky behaviors between MSMW and MSMO, the random/fixed effect models were used to compute the pooled effect rates and odds ratios (OR) (i.e., prevalence and OR of HIV infection and proportion and OR

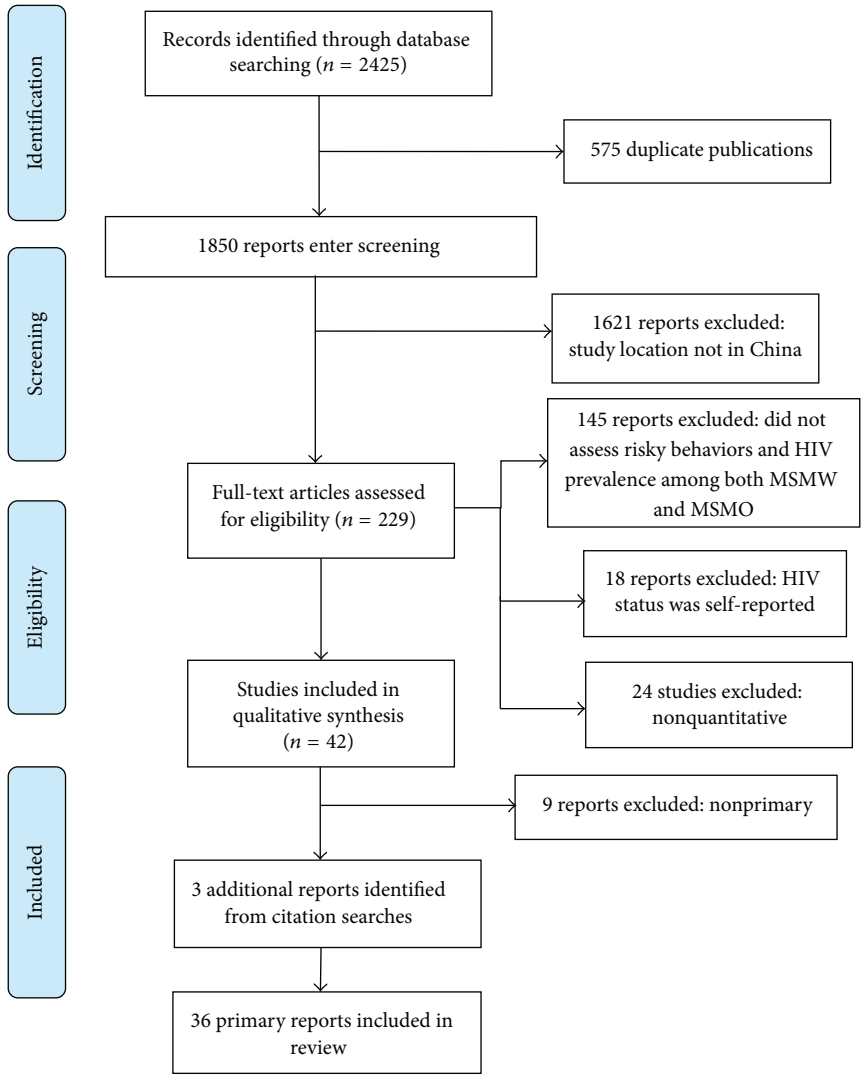


Figure 1: Flow diagram of studies included in analysis.

of condom use, anal sex role, sex partners, and substance use) and relevant 95% confidence intervals (CI). Statistical heterogeneity was qualitatively tested using Cochran's Q statistic ($p < 0.10$ indicates significant heterogeneity) and quantified by the I^2 index ($I^2 < 25\%$, low heterogeneity; $I^2 = 25$–50%, moderate heterogeneity; $I^2 > 75\%$, high heterogeneity). If significant homogeneity was detected ($I^2 < 75.0$, $p \geq 0.10$), fixed effect models were employed to calculate the pooled effect rates and ORs; otherwise random effects models were employed [21].

There were many factors that may have affected the homogeneity between studies, such as sample characteristics and methodological differences. Therefore, we performed subgroup analyses to explore the potential sources of between-study heterogeneity on (1) study region: Southwest China, East China, Northeast, North China, and multiple regions; (2) data collection period: prior to 2009, 2009 and later, and unidentified; (3) data collection method: interviewer-administered, self-administered, and unidentified.

A sensitivity analysis was conducted by omitting each study one at a time to assess the influence of each study on the overall estimate. To investigate publication bias, we utilized Egger's regression test and examined the symmetry of funnel plots for each comparative meta-analytic domain. All the statistical analyses were done using STATA V11.2.

3. Results

3.1. Study Selection. As shown in Figure 1, a total of 2425 relevant articles were identified, of which 42 articles entered further screening, and 36 articles (16 published in English and 20 in Chinese, including 19,730 MSMW and 53,536 MSMO) were finally included in our systematic review and meta-analysis (Figure 1) [15, 22–56]. The characteristics of these included studies are summarized in Table 1.

Table 1: Study characteristics of 36 samples included in meta-analysis.

Author	Publication year	Location	Study periods	Setting	Sampling method	Survey method	Recall window (months)	QACO score (%)
Tang [55]	2007	Harbin	N/A[#]	Multiple recruiting methods	Snowball sampling	Interviewer-administered	6	40
Chen [51]	2008	Fuzhou	2008	MSM venue	Multiple sampling methods	Interviewer-administered	6	60
Ouyang et al. [50]	2008	Chongqing	2006	Multiple recruiting methods	Snowball sampling	Interviewer-administered	6	33
Bai and Feng [43]	2010	Liuzhou	2008	Multiple recruiting methods	N/A	Interviewer-administered	6	75
Han et al. [35]	2012	China	2008	N/A	N/A	N/A	6	60
She et al. [34]	2012	4 cites	2008	N/A	Snowball sampling	Interviewer-administered	6	60
Li et al. [48]	2009	Chongqing	2006	MSM venue	Snowball sampling	Interviewer-administered	6	40
Zhang et al. [31]	2012	Wuhan	N/A	MSM venue	Stratified cluster sampling	Interviewer-administered	6	80
Ruan et al. [47]	2009	Jinan	2008	N/A	RDS[##]	Interviewer-administered	6	60
Chow et al. [29]	2013	Changsha,Tianjin	2011	MSM venue	Convenience sampling	Interviewer-administered	6	75
Cai et al. [24]	2014	Shenzhen	2010	Community-based	RDS	CASI[*]	6	80
Liao et al. [41]	2011	4 cities in Shandong Province	2008-2009	Multiple recruiting methods	Multiple sampling methods	Interviewer-administered	6	80
Wang et al. [52]	2007	Harbin	2006	MSM venue	N/A	Interviewer-administered	6	60
Ruan et al. [53]	2007	Beijing	2005	N/A	Multiple sampling methods	Interviewer-administered	6	80
Xu et al. [44]	2010	Liaoning	2006	Recruiting by NGO[**]	N/A	Interviewer-administered	12	60
Lau et al. [49]	2008	Kunming	2003–2005	Multiple recruiting methods	Convenience sampling	Interviewer-administered	6	60
Cao et al. [37]	2012	4 cities	2008	Recruiting by NGO	Snowball sampling	Interviewer-administered	6	80
Liao et al. [22]	2014	4 cities in Shandong Province	2011	Multiple recruiting methods	N/A	Interviewer-administered	6	70
Tao et al. [28]	2013	Beijing	2010-2011	Multiple recruiting methods	Multiple sampling methods	Interviewer-administered	12	60
Wu et al. [15]	2013	61 cities	2008-2009	N/A	RDS + Snowball sampling	Interviewer-administered	6	90
Guo et al. [36]	2012	Beijing	2009	Multiple recruiting methods	N/A	Self-administered	6	70

Table 1: Continued.

Author	Publication year	Location	Study periods	Setting	Sampling method	Survey method	Recall window (months)	QACO score (%)
Wang et al. [33]	2012	Harbin	2006–2010	Community-based	Snowball sampling	Self-administered	12	65
Cao et al. [30]	2013	Nanchang	2011	Multiple recruiting methods	Multiple sampling methods	Interviewer-administered	1 week	60
Li et al. [56]	2014	Beijing	2009–2011	N/A	RDS	CASI	12	80
Feng et al. [46]	2010	Chengdu	N/A	Multiple recruiting methods	Snowball sampling	Interviewer-administered	6	50
Wang et al. [27]	2013	Beijing	2008-2009	Multiple recruiting methods	Multiple sampling methods	Interviewer-administered	6	80
Zhang et al. [32]	2012	Chongqing	2009	N/A	RDS	CASI	6	80
Zhang et al. [25]	2013	Harbin	2011	Recruiting by NGO	N/A	Interviewer-administered	12	60
Cai et al. [23]	2014	Shenzhen	2009–2012	N/A	Snowball sampling + RDS	Interviewer-administered	6	75
Liu et al. [40]	2011	Beijing	N/A	MSM venue	Snowball sampling	Interviewer-administered	12	50
Wang et al. [26]	2013	Mianyang	2011-2012	N/A	Snowball sampling	Interviewer-administered	6	55
Qu et al. [39]	2011	Inner Mongolia	2010	Community-based	Snowball sampling	Interviewer-administered	6	60
Xu et al. [45]	2010	4 cities	N/A	N/A	Snowball sampling + RDS	Interviewer-administered	6	70
Xi [38]	2011	Hangzhou	2009-2010	MSM venue	Snowball sampling	Interviewer-administered	6	50
Chen et al. [42]	2011	Lanzhou	2006–2010	MSM venue	Snowball sampling	Interviewer-administered	6	50
Li [54]	2007	Beijing	2005	Multiple recruiting methods	Multiple sampling methods	Interviewer-administered	6	70

[#]N/A: not available, [##]RDS: respondent-driven sampling, [*]CASI: computer-assisted self-administered interview, and [**]NGO: Non-Governmental Organizations.

3.2. Quantitative Data Synthesis

3.2.1. HIV Prevalence Analysis and Comparison. As seen in Table 2, across 23 samples, a significantly higher prevalence of HIV was found among MSMW compared with MSMO (6.6% versus 5.4%, OR = 1.27, 95% CI = 1.01–1.58) (Figure 2). Subgroup analysis by region showed that in the Southwest China, the difference in HIV prevalence between MSMW and MSMO was the highest (14.8% versus 6.7%, OR = 2.43, 95% CI = 1.58–3.72), followed by Northeast China (6.5% versus 5.7%, OR = 1.27, 95% CI = 0.89–1.82), the East China (2.7% versus 5.7%, OR = 0.67, 95% CI = 0.32–1.42), North China (5.7% versus 4.1%, OR = 1.32, 95% CI = 0.89–1.94), and other regions (6.3% versus 6.0%, OR = 1.03, 95% CI = 0.78–1.36). For studies that collected data in 2009 and later, the gap in HIV prevalence between MSMW and MSMO was larger than in other periods (8.3% versus 5.6%, OR = 1.59, 95% CI = 1.08–2.33). The prevalence of HIV for studies recruiting participants through RDS and all other sampling methods was 6.9% versus 6.3%, OR = 1.16, 95% CI = 0.65–2.09 and was 6.7% versus 5.2%, OR = 1.30, 95% CI = 1.05–1.61, respectively.

3.2.2. Possible Explanatory Behavioral Factors for Significant Difference in HIV Infection Risk. As seen in Table 3, for anal sex role with male, MSMW were more likely to have insertive anal intercourse (IAI) and less likely to have receptive anal intercourse (RAI) than MSMO (IAI: 53.2% versus 41.1%, OR = 1.74, 95% CI = 1.26–2.42; RAI: 23.7% versus 38.7%, OR = 0.42, 95% CI = 0.28–0.64). No significant differences were found in condom use between MSMW and MSMO (UAI: 56.7% versus 57.8%, OR = 1.19, 95% CI = 0.99–1.26; UIAI: 63.9% versus 58.3%, OR = 1.19, 95% CI = 0.89–1.60; URAI: 45.9% versus 53.3%, OR = 0.73, 95% CI = 0.50–1.06; UAI with commercial male partner: 53.2% versus 41.1%, OR = 1.07, 95% CI = 0.76–1.50). In comparison with sex partners in the past 6 months, MSMW were significantly more likely to engage

Table 2: The odds and prevalence of HIV among MSMW and MSMO by study and design characteristics.

Category	Subgroup	Comparison group	Number of studies	Prevalence estimate (%) and 95% CI	Heterogeneity I^2 (%)	Heterogeneity p value	OR and 95% CI	p value	Heterogeneity I^2 (%)	Heterogeneity p value
Overall		MSMW	23	6.6 (5.3, 7.8)	89.2	<0.001	1.27 (1.01, 1.58)	0.018	74.4	<0.001
		MSMO	23	5.4 (4.7, 6.2)	83.9	<0.001				
Regions	Southwest China	MSMW	4	14.8 (9.3, 20.3)	79.5	0.002	2.43 (1.58, 3.72)	0.004	53.3	0.093
		MSMO	4	6.7 (4.0, 9.5)	83.2	<0.001				
	East China	MSMW	3	2.7 (0.4, 4.9)	87.4	<0.001	0.67 (0.32, 1.42)	0.236	61.2	0.076
		MSMO	3	5.7 (1.7, 9.6)	89.5	<0.001				
	Northeast China	MSMW	3	6.5 (2.3, 10.7)	76.1	0.015	1.27 (0.89, 1.82)	0.101	0.0	0.740
		MSMO	3	5.7 (2.3, 9.2)	73.4	0.023				
	North China	MSMW	7	5.2 (4.0, 6.5)	20.0	0.277	1.32 (0.89, 1.94)	0.088	43.0	0.104
		MSMO	7	4.1 (3.1, 5.1)	48.8	0.069				
	Several cities from different regions or those that cannot be classified into the above four categories	MSMW	6	6.3 (4.5, 8.2)	82.3	<0.001	1.03 (0.78, 1.36)	0.304	61.6	0.023
		MSMO	6	6.0 (4.8, 7.2)	72.0	0.003				
Sampling method	RDS	MSMW	5	6.9 (3.8, 10.0)	95.5	<0.001	1.16 (0.65, 2.09)	0.218	88.9	<0.001
		MSMO	5	6.3 (5.0, 7.6)	85.4	<0.001				
	All other nonprobability sampling	MSMW	18	6.7 (5.1, 8.3)	84.8	<0.001	1.30 (1.05, 1.61)	0.037	47.4	0.014
		MSMO	18	5.2 (4.2, 6.3)	83.3	<0.001				
Data collection period	Before 2009	MSMW	7	4.9 (2.8, 6.9)	94.4	<0.001	0.91 (0.63, 1.32)	0.103	75.1	<0.001
		MSMO	7	5.8 (4.2, 7.4)	92.6	<0.001				
	2009 and later	MSMW	10	8.3 (5.9, 10.7)	71.8	<0.001	1.59 (1.08, 2.33)	<0.001	65.3	0.002
		MSMO	10	5.6 (4.1, 7.1)	77.4	<0.001				
	Not reported or cannot be classified into the above two categories	MSMW	6	6.5 (4.6, 8.5)	75.8	0.001	1.33 (0.97, 1.84)	0.376	48.8	0.082
		MSMO	6	5.0 (3.8, 6.2)	61.9	0.022				
Data collection method	Interviewer-administered	MSMW	19	6.5 (5.1, 7.9)	90.5	<0.001	1.21 (0.94, 1.56)	0.245	77.1	<0.001
		MSMO	19	5.6 (4.8, 6.5)	84.3	<0.001				
	Self-administered	MSMW	2	5.4 (2.8, 8.1)	23.5	0.253	1.37 (0.91, 2.08)	0.211	0.0	0.567
		MSMO	2	3.9 (2.9, 4.8)	0.0	0.449				
	Not reported or cannot be classified into the above two categories	MSMW	2	8.5 (3.4, 13.6)	72.7	0.056	1.58 (1.01, 2.50)	0.030	0.0	0.354
		MSMO	2	5.7 (0.1, 11.2)	89.2	0.002				

Table 3: The odds and proportion of HIV-related risky behaviors among MSMW and MSMO.

Behaviors	Number of studies	Comparison group	Prevalence estimate (%) and 95% CI	Heterogeneity I^2 (%)	Heterogeneity p value	OR and 95% CI	p value	Heterogeneity I^2 (%)	Heterogeneity p value
Condom use in the past 6 months									
UAI with male	9	MSMW	56.7 (46.2, 67.2)	94.6	<0.001	1.19 (0.99, 1.26)	0.123	40.7	0.096
		MSMO	57.8 (50.4, 65.2)	95.0	<0.001				
UIAI with male	3	MSMW	63.9 (57.8, 70.0)	0	0.857	1.19 (0.89, 1.60)	0.101	23.2	0.272
		MSMO	58.3 (46.2, 70.4)	94.3	<0.001				
URAI with male	3	MSMW	45.9 (13.4, 78.4)	95.1	<0.001	0.73 (0.50, 1.06)	0.317	35.9	0.210
		MSMO	53.3 (32.9, 73.8)	98.0	<0.001				
UAI with commercial male partner	3	MSMW	53.2 (22.8, 83.5)	90.1	<0.001	1.07 (0.76, 1.50)	0.785	35.7	0.211
		MSMO	41.1 (10.2, 72.6)	94.7	<0.001				
Anal sexual role									
Mainly insertive	8	MSMW	53.2 (37.5, 68.9)	97.9	<0.001	1.74 (1.26, 2.42)	0.004	81.6	<0.001
		MSMO	41.1 (29.5, 52.8)	98.4	<0.001				
Mainly receptive	8	MSMW	23.7 (13.9, 33.4)	96.8	<0.001	0.42 (0.28, 0.64)	0.013	83.9	<0.001
		MSMO	38.7 (23.0, 54.5)	99.2	<0.001				
Both	5	MSMW	59.4 (53.9, 64.9)	96.8	<0.001	0.92 (0.69, 1.24)	0.256	62.2	0.032
		MSMO	35.8 (20.3, 51.3)	98.7	<0.001				
Sex partner in the past 6 months									
Multiple (⩾2) male partners	5	MSMW	57.0 (40.8, 73.2)	95.3	<0.001	1.19 (0.85, 1.67)	0.090	69.6	0.010
		MSMO	51.9 (42.8, 60.9)	94.7	<0.001				
Having casual male partners	4	MSMW	65.3 (56.8, 73.7)	78.6	0.003	1.16 (0.85, 1.58)	0.141	56.6	0.075
		MSMO	61.6 (53.1, 70.1)	94.5	<0.001				
Having commercial male partners	10	MSMW	18.3 (11.6, 25.1)	95.3	<0.001	1.56 (1.01, 2.42)	0.023	86.8	<0.001
		MSMO	12.2 (8.1, 16.4)	96.8	<0.001				
Substance use in the past 6 months									
Alcohol use	6	MSMW	27.1 (15.7, 38.2)	96.6	<0.001	2.53 (2.14, 2.99)	<0.001	33.5	0.185
		MSMO	13.1 (7.6, 18.6)	96.7	<0.001				
Illicit drug use	3	MSMW	5.3 (4.1, 6.5)	0	0.558	2.09 (1.48, 2.95)	<0.001	0	0.435
		MSMO	2.5 (1.9, 3.1)	0	0.536				

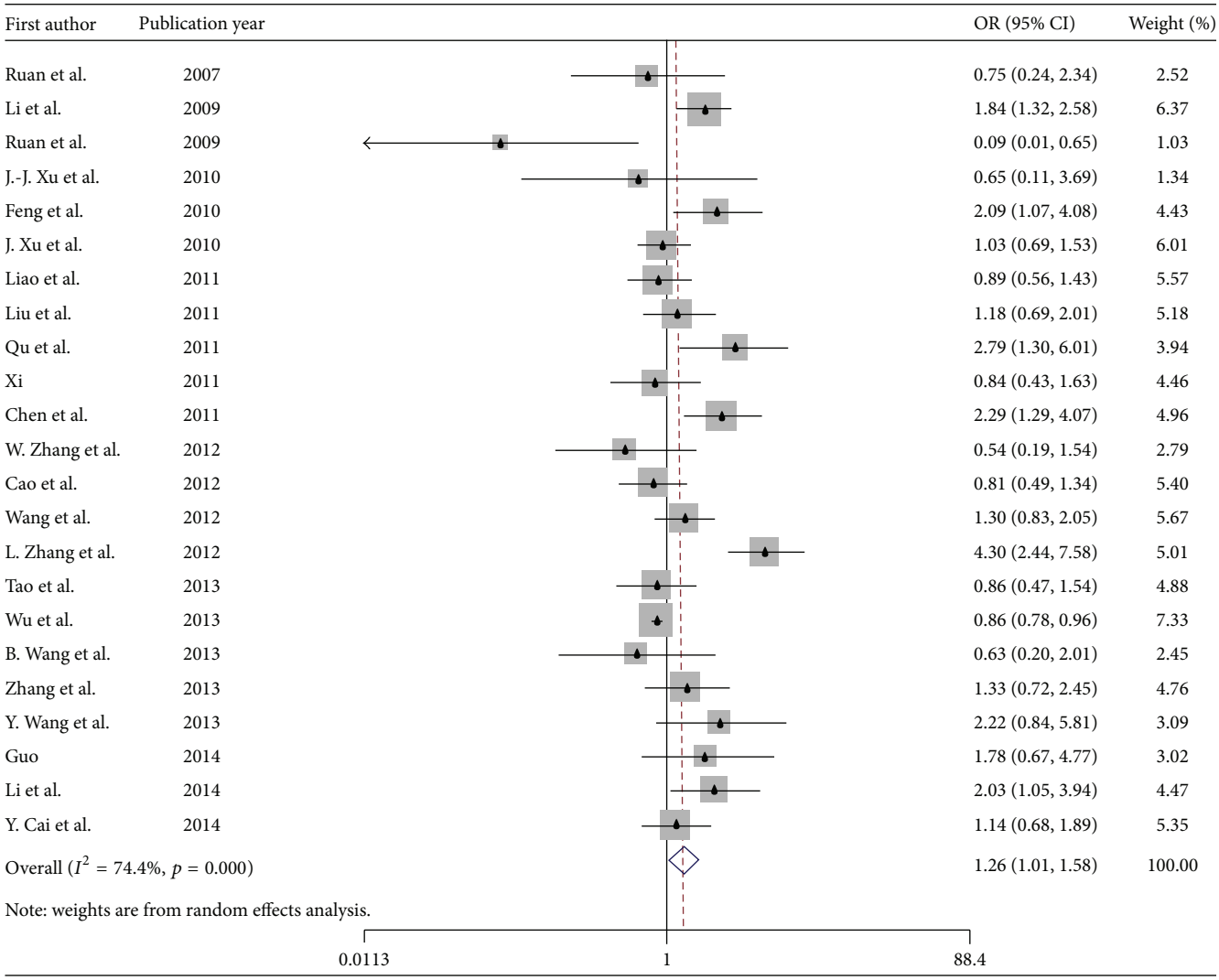


Figure 2: Forest plot of OR pool estimates of HIV infection among MSMW, compared to MSMO.

in commercial sex than MSMO (18.3% versus 12.2%, OR = 1.56, 95% CI = 1.01–2.42). Yet a marginally significant higher proportion of MSMW have multiple male partners (57% versus 51.9%, OR = 1.19, 95% CI = 0.85–1.67, $p = 0.090$) and casual male partners (65.3% versus 61.3%, OR = 1.16, 95% CI = 0.85–1.58) in the past 6 months. Additionally, substance use was significantly more frequent among MSMW than MSMO (alcohol use in the past 6 months: 27.1% versus 13.1%, OR = 2.53, 95% CI = 2.14–2.99; illicit drug use in the past 6 months: 5.3% versus 2.5%, OR = 2.09, 95% CI = 1.48–2.95).

3.3. Sensitivity Analysis and Publication Bias. The pooled rates and ORs were not significantly affected in the sensitivity analysis, suggesting that the results were robust. The funnel plot (dots nearly symmetrically distributed) and Egger's test (all p values for Egger's test > 0.05) reflected that there was no evidence of publication bias in any comparison model (data not provided).

4. Discussion

To our knowledge, this is the first systematic review to compare both HIV prevalence and various risk behaviors between Chinese MSMW and MSMO. It provides an insight into the latest HIV epidemic of MSMW and specific HIV transmission risk from MSM to general population and proposes a crucial suggestion for China's health department to make more effective prevention strategy and policy in the future. Building on our earlier research, we integrated the latest extensive 31 additional published pieces of literature and further validated that Chinese MSMW have a significantly higher HIV prevalence than MSMO, which is similar to our previous conducted meta-analysis results [16]. Remarkably, our present systematic review further clarified risk behavioral differences among MSMW and MSMO in China, which have not yet been disclosed before.

The aggregated ORs among MSMW were 1.3 times more likely to be HIV-positive than MSMO, which is quite different

from a meta-analysis of American MSM suggesting that HIV prevalence among MSMW is somewhat lower than MSMO (OR = 0.41, 95% CI: 0.31–0.54) [57]. These divergences may be attributed to relatively lower rates of risky behavior of MSMW than MSMO in the United States, such as fewer UAI and URAI exposures. On the contrary, higher rates of commercial sex and substance use were found in Chinese MSMW in present systematic review.

The HIV prevalence among MSMW varies according to geographic differences: the extraordinarily high HIV prevalence of MSMW in the Southwest compared to other regions in China (Southwest China: 12.9%; Northeast China: 6.5%; North China: 5.1%; East China: 2.7%). Due to the fact that Southwest China includes several high HIV prevalence areas in China such as Guangxi and Yunnan [3, 58], MSMW in this region engaged in similar sexual risk behaviors may have higher HIV prevalence than in other regions of China. So targeted measures should be taken, which consider the risk profile of MSMW in high HIV prevalence areas of China in order to curb the spread of HIV.

Moreover, our findings further showed that Chinese MSMW have a high rate of commercial sex in the last 6 months. As MSMW feared exposing their homosexual orientation to their female spouse or partner, some may choose to buy sex from MBs [59]. On the other hand, MBs in China often sell sex to both males and females to increase their income [24]. MBs have become one of the emerging high risk subgroups of MSM communities in recent years [60–62], so higher rate of commercial sex in MSMW was associated with HIV transmission. The current systematic review also estimated that MSMW were 2.1 and 3.0 times more likely to engage in alcohol use than MSMO. Moreover, MSMW more likely use illicit drugs in this study. Such social environments may lead MSMW to hide their sexual orientation by unwillingly engaging in heterosexual relationships, given that they may seek out substance use (alcohol/illicit drugs) in an attempt to modulate pressure from reality [22]. As substance use may relax safer sex norms and increase unprotected anal sex, a higher substance use rate in MSMW may increase the risk of HIV transmission [63–65]. Our findings are consistent with other studies that MSMW prefer insertive anal sex role [28, 30, 31, 34, 57, 66–68], but we failed to find significant differences in condom use with male partners between MSMW and MSMO; hence, insertive anal sex role reducing the risk of HIV transmission may be offset by significant higher rate of commercial sex and using alcohol/illicit drugs.

The findings yielded in this study have some important implications for national HIV prevention and care planning and intervention development. The results suggest that MSMW have composed a significant special subproportion of the population of MSM who are both vulnerable to HIV infection from other higher risk MSM and also in high risk of bridging HIV epidemic to their female partners. However, there are currently no HIV prevention targeted interventions policies for MSMW in China. Given this challenge, we should develop specific prevention strategies targeting MSMW from the following three aspects. First, strategies should focus on increasing HIV/STIs screening programs for MSMW subpopulations to help them to learn their HIV infection status. We could promote and implement VCT or STIs clinic-based interventions to increase the frequency of HIV/STIs screening among MSMW through more frequent calling or sending of text message to remind them with their clinic visits [69]. Second, we should take measure of MSMW on preventing high risk commercial sex and substance use; meanwhile, we should enhance their consciousness of protecting themselves as well as their partners (both male and female). Third, given high risk of HIV infection among MSMW, MSMW's family is also a potential source of HIV discordant couples in China [8, 70–72]. In recent years a number of international studies have shown that the early initiation of antiretroviral therapy (ART) could reduce rates of sexual transmission of HIV-1 and clinical events [73–75], in China; early ART initiation was also reported that may reduce HIV transmission in discordant couples [76]. Hence, early ART initiation for HIV-positive MSMW should also be a good candidate strategy to improve the quality of MSMW's lives and decrease their HIV transmission risk in discordant couples. Additionally, it is essential to actively advocate for the respect and social equality of people with bisexual and homosexual orientations and reduce the social discrimination and family pressures for MSM.

This systematic review and meta-analysis has several important limitations. First, the paucity of existing research did not allow for subgroup analyses of HIV prevalence by race/ethnicity and other demographic differences both within MSMW and compared to MSMO. Second, though we attempted to be as inclusive as possible, our available database searches may have excluded relevant studies from this systematic review and meta-analysis. Third, the majority of included articles of this study recruited participants in urban locations, so rural representation was limited. Fourth, for lack of data from Northwest and other areas in China, it is difficult to reflect the comprehensive national HIV epidemic. Fifth, the results of this study suggest that a higher proportion of commercial sex and substance use among MSMW may be a potentially indicative factor for significantly higher HIV prevalence compared to MSMO, but whether commercial sex and/or substance use significantly moderated HIV prevalence differences between MSMW and MSMO should also be verified in future research.

More emphasis needs to be placed on Chinese MSMW in order to develop more targeted prevention measures for these potentially hidden MSM. Strategies targeting MSMW focused on increasing the frequency of HIV/STIs screening, preventing high risk commercial sex and substance use, encouraging safe sex practices, and actively carrying out early antiretroviral therapy for HIV-positive MSMW; meanwhile, actively advocating for the respect and social equality of people with bisexual and homosexual orientations and reducing the social discrimination and family pressures for MSM could help to slow the spread of HIV/AIDS.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors' Contribution

The work presented here was carried out in collaboration between all authors. All authors read and approved the final paper. The first two authors contributed equally to this meta-analysis. Jun-Jie Xu defined the research theme. Hong-Yi Wang, Ke Yun, and Yong-Ze Li worked on data collection and data abstraction. Hong-Yi Wang and Jun-Jie Xu analysed the data, interpreted the results, and wrote the paper. Hua-Chun Zou, Christiana Meng Zhang, and Kathleen Heather Reilly revised the language. Yong-Jun Jiang, Wen-Qing Geng, Ning Wang, and Hong Shang supplied comments and helped with interpretation of the results.

Acknowledgments

The authors appreciate all the researchers whose articles were included in the study. In addition, the authors would like to thank Dr. Xu for reviewing the present paper and providing great guidance. This work was funded in part by Grants from the Mega-Projects of National Science Research for the 12th Five-Year Plan (2012ZX10001006), "Innovation Team Development Programme 2012" of The Ministry of Education, and Key Laboratory Program of Liaoning Education Department (LZ2014038).

References

[1] UNAIDS, *HIV/AIDS, Global Report: UNAIDS Report on the Global AIDS Epidemic: 2012*, UNAIDS, 2012.

[2] Y. Zhou, D. Li, D. Lu, Y. Ruan, X. Qi, and G. Gao, "Prevalence of HIV and syphilis infection among men who have sex with men in China: a meta-analysis," *BioMed Research International*, vol. 2014, Article ID 620431, 12 pages, 2014.

[3] Ministry of Health, People's Republic of China, Joint United Nations Programme on HIV/AIDS, and World Health Organization, *2011 Estimates for the HIV/AIDS Epidemic in China*, Joint United Nations Programme on HIV/AIDS, 2011, http://www.unaids.org.cn/pics/20130521161757.pdf.

[4] Ministry of Health, People's Republic of China, Joint United Nations Programme on HIV/AIDS, and World Health Organization, "Update on the AIDS/STD epidemic in China and main response in control and prevention in October, 2014," *Chinese Journal of AIDS & STD*, vol. 20, no. 12, p. 885, 2014.

[5] State Council AIDS Working Committee Office (SCAWCO), *China 2010 UNGASS Country Progress Report (2008-2009)*, Ministry of Health of the People's Republic of China, Beijing, China, 2010.

[6] Z. Cao, J. Xu, H. Zhang et al., "Risk factors for syphilis among married men who have sex with men in China," *Sexually Transmitted Diseases*, vol. 41, no. 2, pp. 98–102, 2014.

[7] H. Shang, J. Xu, X. Han, J. Spero Li, K. C. Arledge, and L. Zhang, "HIV prevention: bring safe sex to China," *Nature*, vol. 485, no. 7400, pp. 576–577, 2012.

[8] L. Feng, X. Ding, R. Lu et al., "High HIV prevalence detected in 2006 and 2007 among men who have sex with men in China's largest municipality: an alarming epidemic in Chongqing, China," *Journal of Acquired Immune Deficiency Syndromes*, vol. 52, no. 1, pp. 79–85, 2009.

[9] Y. Wang, L.-L. Li, G.-G. Zhang, J. Fan, X.-H. Zhao, and K. Li, "Analysis on the intention of marriage and the influence factors among unmarried men who have sex with men," *Zhonghua Liu Xing Bing Xue Za Zhi*, vol. 33, no. 10, pp. 1031–1035, 2012.

[10] X. Li, B. Zhang, Q. Chu et al., "Correlation between AIDS and homosexuals: a study of 2046 male homosexuals in nine major cities of China," *The Chinese Journal of Human Sexuality*, vol. 17, no. 8, pp. 6–10, 2008.

[11] Y. M. Cai, P. Pan, F. C. Hong et al., "Survey of KABP on STD/AIDS among male sexual workers in Shenzhen City," *China Tropical Medicine*, vol. 7, no. 11, pp. 2131–2133, 2007.

[12] E. P. F. Chow, D. P. Wilson, and L. Zhang, "What is the potential for bisexual men in China to act as a bridge of HIV transmission to the female population? Behavioural evidence from a systematic review and meta-analysis," *BMC Infectious Diseases*, vol. 11, article 242, 2011.

[13] B. Varghese, J. E. Maher, T. A. Peterman, B. M. Branson, and R. W. Steketee, "Reducing the risk of sexual HIV transmission: quantifying the per-act risk for HIV on the basis of choice of partner, sex act, and condom use," *Sexually Transmitted Diseases*, vol. 29, no. 1, pp. 38–43, 2002.

[14] F. Jin, J. Jansson, M. Law et al., "Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART," *AIDS*, vol. 24, no. 6, pp. 907–913, 2010.

[15] Z. Wu, J. Xu, E. Liu et al., "HIV and syphilis prevalence among men who have sex with men: a cross-sectional survey of 61 cities in China," *Clinical Infectious Diseases*, vol. 57, no. 2, pp. 298–309, 2013.

[16] K. Yun, J. J. Xu, K. H. Reilly et al., "Prevalence of bisexual behaviour among bridge population of men who have sex with men in China: a meta-analysis of observational studies," *Sexually Transmitted Infections*, vol. 87, no. 7, pp. 563–570, 2011.

[17] A. Liberati, D. G. Altman, J. Tetzlaff et al., "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *Journal of Clinical Epidemiology*, vol. 62, Article ID e1-34, 2009.

[18] E. Stovold, D. Beecher, R. Foxlee, and A. Noel-Storr, "Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram," *Systematic Reviews*, vol. 3, article 54, 2014.

[19] W. C. W. Wong, C. S. K. Cheung, and G. J. Hart, "Development of a quality assessment tool for systematic reviews of observational studies (QATSO) of HIV prevalence in men having sex with men and associated risk behaviours," *Emerging Themes in Epidemiology*, vol. 5, article 23, 2008.

[20] W. Hao, S. Xiao, T. Liu et al., "The second National Epidemiological Survey on illicit drug use at six high-prevalence areas in China: prevalence rates and use patterns," *Addiction*, vol. 97, no. 10, pp. 1305–1315, 2002.

[21] J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring inconsistency in meta-analyses," *British Medical Journal*, vol. 327, no. 7414, pp. 557–560, 2003.

[22] M. Liao, D. Kang, X. Tao et al., "Alcohol Use, stigmatizing/discriminatory attitudes, and HIV high-risk sexual behaviors among men who have sex with men in china," *BioMed Research International*, vol. 2014, Article ID 143738, 8 pages, 2014.

[23] Y. Cai, H. Liu, Y. Song, and F. Hong, "Factors associated with HIV and syphilis infection among men who have sex with men blood donors in Shenzhen," *Chinese Journal of Preventive Medicine*, vol. 48, no. 2, pp. 128–132, 2014.

[24] R. Cai, J. Zhao, W. Cai, L. Chen, J. H. Richardus, and S. J. de Vlas, "HIV risk and prevention behaviors in men who have sex

with men and women: a respondent-driven sampling study in Shenzhen, China," *AIDS and Behavior*, vol. 18, no. 8, pp. 1560–1568, 2014.

[25] L. Zhang, D. Zhang, B. Yu et al., "Prevalence of HIV infection and associated risk factors among men who have sex with men (MSM) in Harbin, P. R. China," *PLoS ONE*, vol. 8, no. 3, Article ID e58440, 2013.

[26] Y. Wang, L. Li, J. Fan et al., "Investigation on the sexual behavior and HIV infection among MSM in Mianyang City and analysis of the influencing factors," *Pratical Preventive Medicine*, vol. 20, pp. 1166–1169, 2013.

[27] B. Wang, X. Li, B. Stanton, Y. Liu, and S. Jiang, "Socio-demographic and behavioral correlates for HIV and syphilis infections among migrant men who have sex with men in Beijing, China," *AIDS Care*, vol. 25, no. 2, pp. 249–257, 2013.

[28] J. Tao, Y. Ruan, L. Yin et al., "Sex with women among men who have sex with men in China: prevalence and sexual practices," *AIDS Patient Care and STDs*, vol. 27, no. 9, pp. 524–528, 2013.

[29] E. P. F. Chow, J. Jing, Y. Feng et al., "Pattern of HIV testing and multiple sexual partnerships among men who have sex with men in China," *BMC Infectious Diseases*, vol. 13, no. 1, article 549, 2013.

[30] Y. Cao, S. Li, L. Lu et al., "Comparetive analysis on the characteristics of bisexually-and gay-identified men in Nanchang City," *Journal of Preventive Medicine & Public Health*, vol. 24, pp. 62–66, 2013.

[31] W. Zhang, P. Shi, W. Shi et al., "Characteristics of sexual orientation and sexual behavior among men who have sex both with men and women," *Acta Medicinae Universitatis Scientiae et Technologiae Huazhong*, vol. 41, pp. 375–378, 2012.

[32] L. Zhang, X. Ding, R. Lu et al., "Predictors of hiv and syphilis among men who have sex with men in a chinese metropolitan city: comparison of risks among students and non-students," *PLoS ONE*, vol. 7, no. 5, Article ID e37211, 2012.

[33] K. Wang, H. Yan, Y. Liu, Z. Leng, B. Wang, and J. Zhao, "Increasing prevalence of HIV and syphilis but decreasing rate of self-reported unprotected anal intercourse among men who had sex with men in Harbin, China: results of five consecutive surveys from 2006 to 2010," *International Journal of Epidemiology*, vol. 41, no. 2, pp. 423–432, 2012.

[34] M. She, H.-B. Zhang, J. Wang et al., "Investigation of HIV and syphilis infection status and risk sexual behavior among men who have sex with men in four cities of China," *Chinese journal of Preventive Medicine*, vol. 46, no. 4, pp. 324–328, 2012.

[35] L. Han, Y. Wen, X. Cheng et al., "Survey on risk factors of HEV infection among man who have sex with man," *Chinese Journal of Disease Control and Prevention*, vol. 16, pp. 586–589, 2012.

[36] Y. Guo, X. Li, Y. Song, and Y. Liu, "Bisexual behavior among Chinese young migrant men who have sex with men: implications for HIV prevention and intervention," *AIDS Care*, vol. 24, no. 4, pp. 451–458, 2012.

[37] Z. Cao, H.-B. Zhang, M. She et al., "Prevalence of HIV infection and sexual behaviors with both men and women among currently married men who have sex with men," *Chinese Journal of Epidemiology*, vol. 33, no. 5, pp. 488–491, 2012.

[38] S. Xi, *Study on AIDS prevention knowledge, behavioral characteristics, HIV infection and its inflnuencing factors among men who have sex with men (MSM) in Hangzhou City [M.S. thesis]*, Department of Social Medicine & Health Management, College of Public Health, Medical School, Zhejiang University, Zhejiang, China, 2011.

[39] L. Qu, Y. Gao, J. Liu et al., "Prevalence of HIV infection and its associated factors among 805 MSM in inner Mongolia," *Practical Preventive Medicine*, vol. 18, pp. 2081–2083, 2011.

[40] Y.-J. Liu, S.-L. Jiang, Y. Hu, L. Song, M. Yu, and S.-M. Li, "Characteristics of sexual behaviors and infection status of AIDS and other sexually transmitted diseases among men who have sex with men in 2009 in Beijing," *Chinese Journal of Preventive Medicine*, vol. 45, no. 11, pp. 971–974, 2011.

[41] M. Liao, D. Kang, B. Jiang et al., "Bisexual behavior and infection with HIV and syphilis among men who have sex with men along the east coast of China," *AIDS Patient Care and STDs*, vol. 25, no. 11, pp. 683–691, 2011.

[42] J. J. Chen, X. Y. Zhang, X. Y. Chen et al., "The analysis of influence factors of HIV in MSM in Lanzhou," *The Chinese Journal of Dermatovenereology*, vol. 25, no. 12, pp. 954–957, 2011.

[43] Y. Bai and W. D. Feng, "HIV related risk sexual behavior among men having sex with men withdifferent sex orientation in Liuzhou," *J Prev Med Inf*, vol. 26, 2010.

[44] J.-J. Xu, M. Zhang, K. Brown et al., "Syphilis and HIV seroconversion among a 12-month prospective cohort of men who have sex with men in Shenyang, China," *Sexually Transmitted Diseases*, vol. 37, no. 7, pp. 432–439, 2010.

[45] J. Xu, D.-L. Han, Z. Liu et al., "The prevalence of HIV infection and the risk factors among MSM in 4 cities, China," *Chinese Journal of Preventive Medicine*, vol. 44, no. 11, pp. 975–980, 2010.

[46] Y. Feng, Z. Wu, R. Detels et al., "HIV/STD prevalence among men who have sex with men in Chengdu, China and associated risk factors for HIV infection," *Journal of Acquired Immune Deficiency Syndromes*, vol. 53, no. 1, pp. S74–S80, 2010.

[47] S. Ruan, H. Yang, Y. Zhu et al., "Rising HIV prevalence among married and unmarried among men who have sex with men: Jinan, China," *AIDS and Behavior*, vol. 13, no. 4, pp. 671–676, 2009.

[48] C.-M. Li, Y.-J. Jia, X.-B. Ding, J.-B. Liu, and Y. Xiao, "HIV infections and heterosexual behaviors among men who have sex with men in Chongqing municipality, China," *Chinese Journal of Epidemiology*, vol. 30, no. 9, pp. 882–886, 2009.

[49] J. T. F. Lau, M. Wang, H. N. Wong et al., "Prevalence of bisexual behaviors among men who have sex with men (MSM) in China and associations between condom use in MSM and heterosexual behaviors," *Sexually Transmitted Diseases*, vol. 35, no. 4, pp. 406–413, 2008.

[50] L. Ouyang, X. B. Ding, C. Zhou et al., "HIV risk sexual behavior among MSM with different sex orientation in Chongqing," *South China Journal of Preventive Medicine*, vol. 34, no. 2, pp. 16–19, 2008.

[51] K. Chen, *Investigation on HIV/AIDS related risk behaviors and infection among men who have sex with men [M.S. thesis]*, Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, 2008.

[52] L. Wang, H. Tang, D. Zhang et al., "Behavioral features of men who have sex with men with different sexual orientations," *Chinese Journal of AIDS & STD*, vol. 13, pp. 123–126, 2007.

[53] Y. Ruan, D. Li, X. Li et al., "Relationship between syphilis and HIV infections among men who have sex with men in Beijing, China," *Sexually Transmitted Diseases*, vol. 34, no. 8, pp. 592–597, 2007.

[54] X. Li, *High risk behaviors and HIV/syphilis prevalence among men who have sex with men in Beijing, China [M.S. thesis]*, Chinese Center for Disease Control and Prevention, 2007.

[55] H. L. Tang, *Characteristics of sexual network and HIV/AIDS related risk behaviors among men who have sex with men [M.S. thesis]*, Chinese Center for Disease Control and Prevention, 2007.

[56] X. Li, H. Lu, C. Cox et al., "Changing the landscape of the HIV epidemic among MSM in China: results from three consecutive respondent-driven sampling surveys from 2009 to 2011," *BioMed Research International*, vol. 2014, Article ID 563517, 10 pages, 2014.

[57] M. R. Friedman, C. Wei, M. L. Klem, A. J. Silvestre, N. Markovic, and R. Stall, "HIV infection and sexual risk among men who have sex with men and women (MSMW): a systematic review and meta-analysis," *PLoS ONE*, vol. 9, Article ID e87139, 2014.

[58] Y. Yang, "AIDS epidemic trend in the second-round national comprehensive AIDS prevention and control demonstration zones in Guangxi," *Guangxi Medical Journal*, no. 8, pp. 1123–1125, 2014.

[59] S. S. Solomon, S. H. Mehta, A. Latimore, A. K. Srikrishnan, and D. D. Celentano, "The impact of HIV and high-risk behaviours on the wives of married men who have sex with men and injection drug users: implications for HIV prevention," *Journal of the International AIDS Society*, vol. 13, supplement 2, article S7, 2010.

[60] Y. Lai, Y. Cai, Y. Song et al., "HIV/syphilis infection and high risk behaviors among men who have sex with men previous to be money boys," *China Tropical Medicine*, vol. 13, no. 6, pp. 680–683, 2013.

[61] S. Liu, J. Zhao, K. Rou et al., "A Survey of condom use behaviors and HIV/STI prevalence among venue-based money boys in Shenzhen, China," *AIDS and Behavior*, vol. 16, no. 4, pp. 835–846, 2012.

[62] E. J. Nehl, N. He, L. Lin et al., "Drug use and sexual behaviors among MSM in China," *Substance Use & Misuse*, vol. 50, no. 1, pp. 123–136, 2015.

[63] Y. Ding, N. He, W. Zhu, and R. Detels, "Sexual risk behaviors among club drug users in Shanghai, China: prevalence and correlates," *AIDS and Behavior*, vol. 17, no. 7, pp. 2439–2449, 2013.

[64] Y. Guo, X. Li, and B. Stanton, "HIV-related behavioral studies of men who have sex with men in China: a systematic review and recommendations for future research," *AIDS and Behavior*, vol. 15, no. 3, pp. 521–534, 2011.

[65] Q. He, Y. Wang, P. Lin et al., "High prevalence of risk behaviour concurrent with links to other high-risk populations: a potentially explosive HIV epidemic among men who have sex with men in Guangzhou, China," *Sexually Transmitted Infections*, vol. 85, no. 5, pp. 383–390, 2009.

[66] C. Maulsby, F. Sifakis, D. German, C. P. Flynn, and D. Holtgrave, "HIV risk among men who have sex with men only (MSMO) and men who have sex with men and women (MSMW) in Baltimore," *Journal of Homosexuality*, vol. 60, no. 1, pp. 51–68, 2013.

[67] C. Beyrer, G. Trapence, F. Motimedi et al., "Bisexual concurrency, bisexual partnerships, and HIV among Southern African men who have sex with men," *Sexually Transmitted Infections*, vol. 86, no. 4, pp. 323–327, 2010.

[68] A. E. Phillips, C. M. Lowndes, M. C. Boily et al., "Men who have sex with men and women in Bangalore, South India, and potential impact on the HIV epidemic," *Sexually Transmitted Infections*, vol. 86, no. 3, pp. 187–192, 2010.

[69] H. Zou, C. K. Fairley, R. Guy et al., "Automated, computer generated reminders and increased detection of gonorrhoea, chlamydia and syphilis in men who have sex with men," *PLoS ONE*, vol. 8, no. 4, Article ID e61972, 2013.

[70] B. Zhang, X. Li, Q. Chu et al., "A survey of HIV/AIDS related behaviors among 2250 MSM in nine major cities of China," *Chinese Journal of AIDS & STD*, vol. 14, pp. 541–547, 2008.

[71] Q. He, Y. Wang, P. Lin et al., "Potential bridges for HIV infection to men who have sex with men in Guangzhou, China," *AIDS and Behavior*, vol. 10, no. 7, pp. S17–S23, 2006.

[72] K.-H. Choi, D. R. Gibson, L. Han, and Y. Guo, "High levels of unprotected sex with men and women among men who have sex with men: a potential bridge of HIV transmission in Beijing, China," *AIDS Education and Prevention*, vol. 16, no. 1, pp. 19–30, 2004.

[73] M. S. Cohen, Y. Q. Chen, M. McCauley et al., "Prevention of HIV-1 infection with early antiretroviral therapy," *The New England Journal of Medicine*, vol. 365, no. 6, pp. 493–505, 2011.

[74] D. Pao, D. Pillay, and M. Fisher, "Potential impact of early antiretroviral therapy on transmission," *Current Opinion in HIV & AIDS*, vol. 4, no. 3, pp. 215–221, 2009.

[75] A. Sáez-Cirión, C. Bacchus, L. Hocqueloux et al., "Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study," *PLoS Pathogens*, vol. 9, no. 3, Article ID e1003211, 2013.

[76] N. He, S. Duan, Y. Ding et al., "Antiretroviral therapy reduces HIV transmission in discordant couples in rural Yunnan, China," *PLoS ONE*, vol. 8, no. 11, Article ID e77981, 2013.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 537318, 9 pages
http://dx.doi.org/10.1155/2015/537318

Review Article

Perinatal Health Statistics as the Basis for Perinatal Quality Assessment in Croatia

Urelija Rodin,[1] Boris Filipović-Grčić,[2] Josip Đelmiš,[3] Tatjana Glivetić,[4] Josip Juras,[3] Željka Mustapić,[2] and Ruža Grizelj[2]

[1]*School of Medicine, University of Zagreb, School of Public Health "A. Štampar", Croatian Institute of Public Health, Rockefellerova 4, 10000 Zagreb, Croatia*
[2]*School of Medicine, University of Zagreb, Clinical Hospital Center Zagreb, Department of Pediatrics, Kišpatićeva 12, 10000 Zagreb, Croatia*
[3]*School of Medicine, University of Zagreb, Clinical Hospital Center Zagreb, Department of Obstetrics and Gynecology, Petrova 13, 10000 Zagreb, Croatia*
[4]*General Hospital Zabok, Trg Dragutina Domjanića 6, 49210 Zabok, Croatia*

Correspondence should be addressed to Urelija Rodin; urelija.rodin@hzjz.hr

Received 5 June 2015; Revised 7 October 2015; Accepted 9 November 2015

Academic Editor: Malgorzata Bala

Context. Perinatal mortality indicators are considered the most important measures of perinatal outcome. The indicators reliability depends on births and deaths reporting and recording. Many publications focus on perinatal deaths underreporting and misclassification, disabling proper international comparisons. *Objective*. Description of perinatal health care quality assessment key indicators in Croatia. *Methods*. Retrospective review of reports from all maternities from 2001 to 2014. *Results*. According to reporting criteria for birth weight ≥500 g, perinatal mortality (PNM) was reduced by 31%, fetal mortality (FM) by 32%, and early neonatal mortality (ENM) by 29%. According to reporting criteria for ≥1000 g, PNM was reduced by 43%, FM by 36%, and ENM by 54%. PNM in ≥22 weeks' (wks) gestational age (GA) was reduced by 28%, FM by 30%, and ENM by 26%. The proportion of FM at 32–36 wks GA and at term was the highest between all GA subgroups, as opposed to ENM with the highest proportion in 22–27 wks GA. Through the period, the maternal mortality ratio varied from 2.4 to 14.3/100,000 live births. The process indicators have been increased in number by more than half since 2001, the caesarean deliveries from 11.9% in 2001 to 19.6% in 2014. *Conclusions*. The comprehensive perinatal health monitoring represents the basis for the perinatal quality assessment.

Dedicated to the memory of Professor Ante Dražančić, our teacher and founder of Croatian Perinatology

1. Introduction

Perinatal health care, as well as the other areas of health care, requires the usage of useful indicators for quality assessment and evaluation, which will enable sustainable planning in accordance with limited resources. A weighted sum of all essential indicators, including fetal and maternal, short-term, and long-term outcomes, as well as maternal satisfaction and the impact on future pregnancies and deliveries, would represent the ideal measure of quality [1]. However, all recommended perinatal health indicators cannot be produced and gathered through routine national health statistics system. Some of the indicators are already available in the international databases but not presented by subgroups, which would make them more specific and sensitive perinatal health measurements. According to the World Health Organization (WHO) recommendations for international comparisons, the countries should calculate the perinatal indicators for

total births, fetal and early neonatal deaths ≥1000 g birth weight (BW), or ≥28 weeks' (wks) gestational age (GA). Moreover, the inclusion of fetuses and infants weighing 500–999 g or 22–27 completed wks GA in national statistics is recommended by WHO because it improves the coverage of reporting according to criteria for international comparisons and enables better evaluation outcomes [2]. Also, on the European community's research agenda, there was a need for defining measures of maternal and child health care and outcomes for use in evaluating health care and public health programmes. As a part of the EU's Health Monitoring Programme, PERISTAT (Perinatal Statistics) project has been launched in 1999 [3]. The objective of the PERISTAT project is to establish a high quality, innovative, internationally recognized, and sustainable European perinatal health information system. This system's goal is to produce data and analysis on a regular basis for use by national, European, and international stakeholders who make decisions about the health and health care of pregnant women and newborns. PERISTAT scientific advisory committee defined the core perinatal indicators list in order to monitor the perinatal health more precisely. These indicators are sufficient for international comparisons, measuring fetal and infant health outcomes and key interventions implemented to prevent death and morbidity [4, 5].

The aim of this study was to analyze the key indicators for perinatal health care quality assessment in Croatia for the period 2001–2014.

The feasible perinatal indicators for Croatia's perinatal health care assessment were as follows:

(i) Perinatal outcome indicators: perinatal mortality (PNM), fetal mortality (FM), early neonatal mortality (ENM) by BW subgroups (≥500 g and ≥1000 g) and GA subgroups (≥22 wks and ≥28 wks), and maternal mortality.

(ii) Process indicators: antenatal visits, ultrasound (US) examinations, and caesarean deliveries (CS).

2. Material and Methods

2.1. Croatian Population Characteristics. According to the 2011 Census, the Croatian Bureau of Statistics (CBS) data, the Croatian population amounted to approximately 4.3 million, with approximately 40,000 deliveries and 50,000 deaths per year with negative natural trend [6]. Sociodemographic characteristics of population by WHO's estimates indicated low birth rate (9.3/1,000), low total fertility rate (1.5/per woman, 15–49 years old), moderate death rate (12.0/1,000), and moderate life expectancy at birth for both sexes (76 years) [7]. Numerous sociodemographic characteristics of mothers remain unknown since CBS collects a limited data set of these data like permanent residence, marital status, parity, and professional birth attendance [6]. According to CBS data in the period 2000–2013, more than 80% of births were from marriages and more than 99% in health institutions [8]. Almost half of all deliveries were first deliveries, 35% second, and 15% third or higher birth order. Concerning the mother age, the deliveries were most common at the age 25–29 (91.8 deliveries per 1,000 females of the same age), followed by the age 30–34 (89.7 deliveries per 1,000 females of this age), and at the age 20–24 (47.5/1,000). The overall average birth age was 30, while the average age at first birth among women was 27 years [6, 9]. About 3% were births from multiple pregnancies [9]. CBS vital statistics data about births are limited to sex, live birth (LB) or stillbirth, and residence. It was a basic reason for routine health statistics system improvement and introduction of new medical birth notification for birth monitoring with a broader set of data in 2001.

Despite efforts for better data entry in health institutions, the Croatian Institute of Public Health (CIPH) as the main producer of routine health statistics disposes of some basic data of newborns like sex, GA, and BW. According to CIPH data, 20,283 males and 19,505 females were born in the year 2014 and there was 1,04 : 1 male : female newborn ratio. In this period, there were 5.24% newborns with BW less than 2500 g: 0.51% newborns with extremely low birth weight (less than 1000 g (ELBW) newborns), 0.47% newborns of BW 1000–1499 g, 1.08% newborns of BW 1500–1999 g, and 3.18% newborns of BW 2000–2499 g [9].

During the war and the postwar period in Croatia, the share of preterm births amounted to 7-8% [10]; in the period 2001–2010, this share decreased to 5.8% [11], while in the period 2010–2014, the increase was present up to 6.2% [12].

2.2. Study Design. The reports were retrieved from CIPH for the period 2001–2014, after collecting and processing birth, fetal deaths, and early neonatal deaths data from maternities and neonatal intensive care units (NICUs). These reports derived from hospital data were obtained as a part of the CIPH and Croatian Society of Perinatal Medicine (CSPM) Programme for perinatal health surveillance and reporting. According to these reports, there were 587,356 total births ≥22 wks and 4,633 perinatal deaths ≥22 wks (Table 1).

Perinatal data collecting according to the WHO recommendations by the BW and GA groups has been introduced in Croatian maternities since 2001 by the CIPH, in cooperation with CSPM, through two ways: reports with aggregated birth and perinatal death data by the BW and GA subgroups and individual birth and perinatal death notifications based on hospital discharges [10].

WHO encourages countries to build perinatal monitoring capacity and collect data for key perinatal mortality indicators, relying on the same definitions in order to allow for the comparison of these statistics. The events related to birth, death, and the perinatal period, as well as the reporting requirements for the data from which internationally comparable statistics are drawn, are defined in detail in the International Classification of Diseases, 10th Edition (ICD-10), Instruction Manual, [2].

WHO recommended definition for national purposes [2]:

$$\text{PNM} = \frac{\text{Fetal deaths and early neonatal deaths} \geq 500\,\text{g or} \geq 22\,\text{wks}}{\text{Total births} \geq 500\,\text{g or} \geq 22\,\text{wks}} \times 1000,$$

Table 1: Total births, fetal, early neonatal and perinatal deaths in Croatian maternities in the period 2001–2014.

Year	Total births ≥22 wks numbers	Fetal deaths ≥22 wks numbers	Early neonatal deaths ≥22 wks numbers	Perinatal deaths ≥22 wks numbers
2001	41,487	235	173	408
2002	40,493	234	152	386
2003	40,013	227	156	383
2004	40,759	231	139	370
2005	43,030	237	146	383
2006	41,964	218	138	356
2007	42,456	195	139	334
2008	44,315	202	118	320
2009	45,071	191	149	340
2010	43,842	215	108	323
2011	41,556	158	102	260
2012	42,074	161	78	239
2013	40,319	156	92	248
2014	39,977	158	125	283

$$\text{FM} = \frac{\text{Fetal deaths} \geq 500\,\text{g or} \geq 22\,\text{wks}}{\text{Total births} \geq 500\,\text{g or} \geq 22\,\text{wks}} \times 1000,$$

$$\text{ENM} = \frac{\text{Early neonatal deaths or} \geq 500\,\text{g or} \geq 22\,\text{wks}}{\text{LB} \geq 500\,\text{g or} \geq 22\,\text{wks}} \times 1000. \quad (1)$$

WHO recommended definition for international comparison:

$$\text{PNM} = \frac{\text{Fetal deaths and early neonatal deaths} \geq 1000\,\text{g or} \geq 28\,\text{wks}}{\text{Total births} \geq 1000\,\text{g or} \geq 28\,\text{wks}} \times 1000,$$

$$\text{FM} = \frac{\text{Fetal deaths} \geq 1000\,\text{g or} \geq 28\,\text{wks}}{\text{Total births} \geq 1000\,\text{g or} \geq 28\,\text{wks}} \times 1000, \quad (2)$$

$$\text{ENM} = \frac{\text{Early neonatal deaths or} \geq 1000\,\text{g or} \geq 28\,\text{wks}}{\text{LB} \geq 1000\,\text{g or} \geq 28\,\text{wks}} \times 1000.$$

Following the year 2001, according to the new methodology, the calculations for PNM and FM based on reports from the maternities have been obtained according to the national and international reporting criteria [13]. However, for the purpose of obtaining the vital statistics, all dead LB, independently of BW or GA, have been included in infant mortality.

Maternal mortality data from medical death records in CIPH were matched with CBS vital statistics data. The record of each death during pregnancy, childbirth, and puerperium was verified together by experts from CIPH and CSPM in order to get complete and reliable data. The WHO defines maternal death as the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, of any cause related to or aggravated by the pregnancy or its management but not of accidental or incidental causes. This definition allows identification of maternal deaths based on their causes as either direct or indirect. The direct obstetric deaths are those resulting from obstetric complications of the pregnant state (pregnancy, delivery, and postpartum period), from interventions, omissions, or incorrect treatment, or from a chain of events resulting from any of the above. The indirect obstetric deaths are those resulting from previous existing disease or diseases that develop during pregnancy and which were not due to direct obstetric causes but were aggravated by physiological effects of pregnancy [14]. The WHO Conference agreed that since the number of LB was more universally available than the number of total births, it should be used as the denominator in the ratios related to maternal mortality [2, 14]. Therefore, the maternal mortality ratio (MMR) in Croatia has been calculated including direct and indirect causes of woman death in pregnancy, delivery, or puerperium on 100,000 LB [13].

Data for process indicators (antenatal visits, US, and CS) have been obtained from maternities and CS proportions on 100 LB compared with WHO Health For All (WHO-HFA) database indicators.

3. Results

3.1. Perinatal Outcome Indicators

3.1.1. Perinatal, Fetal, and Early Neonatal Mortality by Birth Weight Subgroups (≥500 g and ≥1000 g). In the period 2001–2014, PNM for all ≥500 g BW total births was reduced by 30.6%, from 9.8‰ to 6.8‰, FM for all ≥500 g by 32.1%, from 5.6‰ to 3.8‰, and ENM by 28.6%, from 4.2‰ to 3.0‰. In the same period of time, PNM for all ≥1000 g BW total births was reduced by 42.5% (from 7.3‰ in 2001 to 4.2‰ in 2014).

Table 2: Perinatal, fetal, and early neonatal mortality rates (‰) in total births' birth weight ≥500 g and birth weight ≥1000 g in Croatia in the period 2001–2014.

Year	PNM ≥ 500 g (‰)	PNM ≥ 1000 g (‰)	FM ≥ 500 g (‰)	FM ≥ 1000 g (‰)	ENM ≥ 500 g (‰)	ENM ≥ 1000 g (‰)
2001	9.8	7.3	5.6	4.5	4.2	2.8
2002	9.3	6.9	5.6	4.3	3.7	2.6
2003	9.5	6.3	5.7	4.1	3.8	2.2
2004	8.7	5.8	5.3	3.9	3.4	1.9
2005	8.8	6.4	5.4	4.2	3.4	2.2
2006	8.3	5.3	5.1	3.4	3.2	1.9
2007	7.8	4.9	4.5	3.1	3.2	1.8
2008	7.0	4.6	4.4	3.2	2.6	1.4
2009	7.2	4.4	4.2	3.0	3.0	1.5
2010	7.2	4.7	4.8	3.5	2.4	1.2
2011	5.9	3.5	3.6	2.5	2.3	1.0
2012	5.4	3.6	3.6	2.7	1.7	0.9
2013	5.8	3.5	3.6	2.5	2.2	1.0
2014	6.8	4.2	3.8	2.9	3.0	1.3

Table 3: Perinatal, fetal, and early neonatal mortality rates (‰) in gestational age subgroups ≥22 weeks and ≥28 weeks in Croatia in the period 2001–2014.

Year	PNM ≥ 22 wks (‰)	PNM ≥ 28 wks (‰)	FM ≥ 22 wks (‰)	FM ≥ 28 wks (‰)	ENM ≥ 22 wks (‰)	ENM ≥ 28 wks (‰)
2001	9.8	7.4	5.7	4.7	4.2	2.7
2002	9.5	7.4	5.8	4.8	3.8	2.6
2003	9.6	6.6	5.7	4.4	3.9	2.1
2004	9.1	6.3	5.7	4.4	3.4	1.9
2005	8.9	6.7	5.5	4.5	3.4	2.2
2006	8.5	5.8	5.2	4.0	3.3	1.8
2007	7.9	5.5	4.6	3.6	3.3	1.9
2008	7.2	4.8	4.6	3.5	2.7	1.3
2009	7.5	4.7	4.2	3.2	3.3	1.5
2010	7.4	5.0	4.9	3.8	2.5	1.2
2011	6.3	3.8	3.8	2.9	2.5	0.9
2012	5.7	3.9	3.8	3.0	1.9	0.9
2013	6.2	3.9	3.9	2.9	2.3	1.0
2014	7.1	4.4	4.0	3.0	3.1	1.5

The FM was reduced by 35.6% (from 4.5‰ to 2.9‰) and ENM by more than half (53.6%, from 2.8‰ to 1.3‰), with slight variations in the rates over the years (Table 2).

3.1.2. Perinatal, Fetal, and Early Neonatal Mortality by Gestational Age Subgroups (≥22 wks and ≥28 wks). In the period 2001–2014, PNM in all ≥22 wks GA group was reduced by about one-quarter (27.6%, from 9.8‰ in 2001 to 7.1‰ in 2014). FM was reduced by 29.8% (from 5.7‰ to 4.0‰) and ENM by 26.2% (from 4.2‰ to 3.1‰). In the same period of time, PNM in all ≥28 wks GA was reduced by 40.6% (from 7.4‰ to 4.4‰), FM in all ≥28 wks GA by 36.1% (from 4.7‰ to 3.0‰), and ENM by 44.6% (from 2.7‰ to 1.5‰) (Table 3).

Figure 1 reports FM attributable to each of the four GA groups: 22–27; 28–31; 32–36; and ≥37 wks. This figure illustrates the impact of FM subgroup differences on overall rates. FM ranged from 1.2‰ to 1.8‰ in 32–36 wks GA subgroup and from 0.9‰ to 2.2‰ in 37–41 wks GA subgroup, for both subgroups more than a half of overall FM throughout the period 2001–2014.

ENM was the highest in 22–27 wks GA subgroup, representing more than one-third to more than a half of the total of all early neonatal deaths in the period 2001–2014. ENM ranged from 1.0‰ to 1.8‰ in this GA (Figure 2).

3.2. Maternal Mortality. Table 4 reports MMR related to the direct and indirect obstetric causes for the period 2001–2014. Total MMR varied from 2.4 to 14.3/100,000 LB. In the period 2010–2014, the decreasing trend in direct obstetrical causes could be observed. In the period 2001–2014, the direct obstetric deaths due to pregnancy, labor, or puerperium caused 63.8% of all maternal deaths. The indirect obstetric deaths due to maternal chronic diseases, malignant diseases, and other

Table 4: Maternal deaths and maternal mortality ratios related to the direct and indirect obstetric causes/100,000 live births.

Year	MD: all causes		MD: direct obstetric causes		MD: indirect obstetric causes	
	Numbers	MMR	Numbers	MMR	Numbers	MMR
2001	1	2.4	1	2.4	0	0
2002	4	10.0	4	10.0	0	0
2003	3	7.6	3	7.6	0	0
2004	3	7.4	2	5.0	1	2.5
2005	3	7.1	1	2.4	2	4.7
2006	4	9.7	2	4.8	2	4.8
2007	6	14.3	3	7.2	3	7.2
2008	3	6.9	1	2.3	2	4.6
2009	6	13.5	6	13.5	0	0
2010	4	9.2	1	2.3	3	6.9
2011	4	9.7	3	7.3	1	2.4
2012	3	7.2	1	2.4	2	4.8
2013	2	5.0	1	2.5	1	2.5
2014	1	2.5	1	2.5	0	0
2001–2014	47	8.1	30	5.2	17	2.9

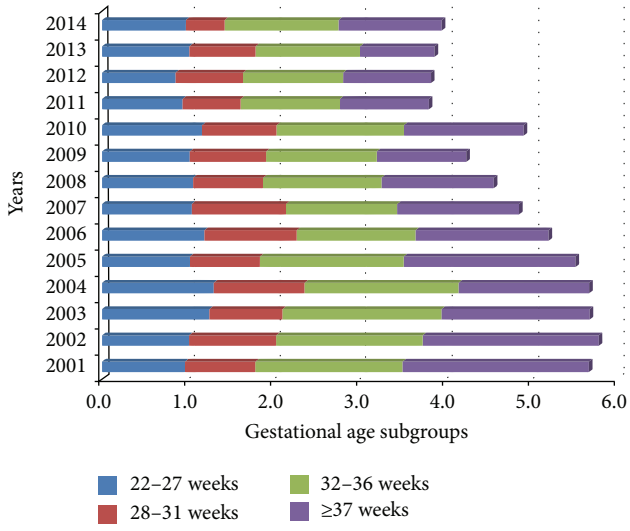


Figure 1: Fetal mortality attributed to gestational age subgroup (per 1,000 total births).

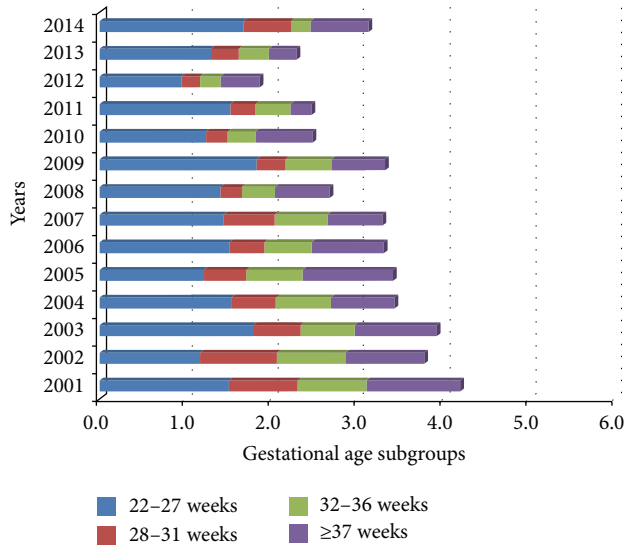


Figure 2: Early neonatal mortality attributed to gestational age subgroup (per 1,000 live births).

causes unrelated to pregnancy, labor, or puerperium caused the remaining 36.2% of all maternal deaths.

3.3. Perinatal Process Indicators

3.3.1. Antenatal Visits and Ultrasound Examinations. Figure 3 reports the proportion of pregnant women with 0–2, 3–5, 6–8, and ≥9 antenatal visits. The percentage of pregnant women with ≥9 visits increased from 43.0% in 2001 to 72.3% in 2014, followed by other subgroups' proportion decrease.

Figure 4 reports the proportion of pregnant women with 0, 1, 2, 3, and ≥4 US in pregnancy. The percentage of pregnant women with ≥4 visits increased from 63.1% in 2001 to 93.4% in 2014, followed by other subgroups' proportion decrease.

3.3.2. Caesarean Section. The frequency of CS is continuously rising. Figure 5 reports the comparison with EU average.

4. Discussion

The key indicators for perinatal health care quality assessment in Croatia were analyzed after introducing new reporting criteria in routine health statistics on national level for monitoring PNM, FM, and ENM in 2001. The changes in

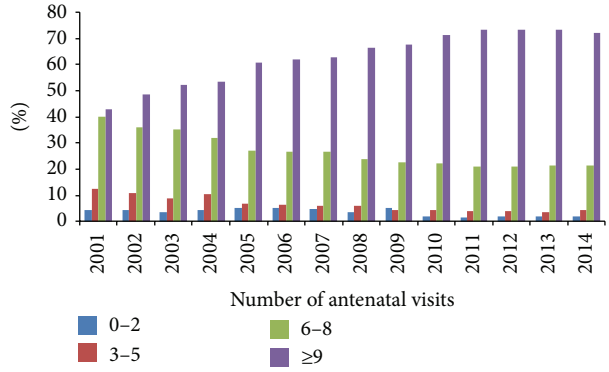


Figure 3: Distribution of pregnant women in relation to antenatal visits in the period 2001–2014.

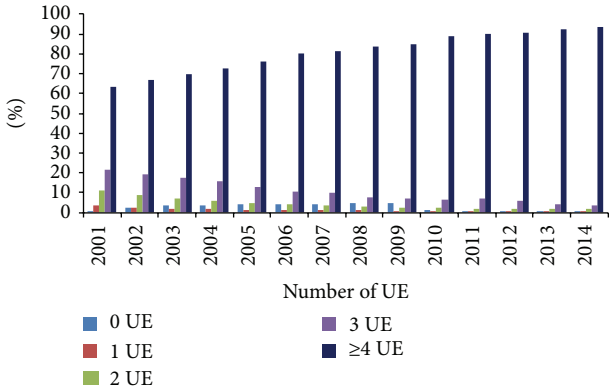


Figure 4: Distribution of pregnant women by ultrasound examination frequency in the period 2001–2014.

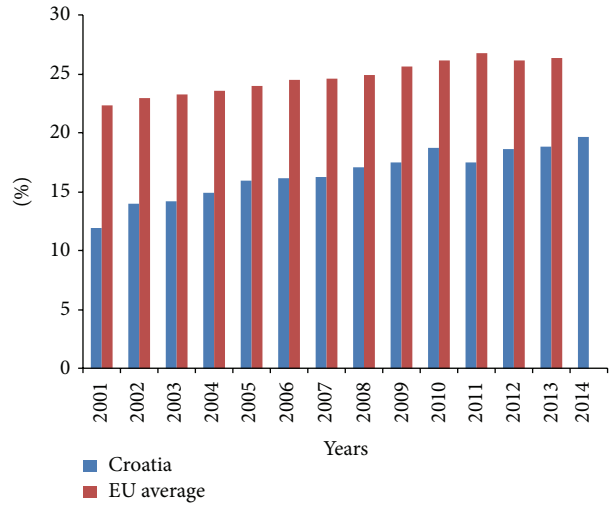


Figure 5: The proportion of caesarean sections in Croatia and EU average per 100 live births in the period 2001–2014.

reporting criteria have enabled us to have a deeper insight into PNM, FM, and ENM trends in births <1000 g BW and <28 completed wks GA. According to reporting criteria for international comparisons and WHO-HFA indicators, PNM and FM in Croatia were below EU average [15]. However, this study showed that the inclusion of perinatal deaths <1000 g BW and <28 completed wks GA considerably changes picture about perinatal outcomes in Croatia. Furthermore, the study showed that the highest ENM was in 22–27 wks GA, by more than third to more than a half of the total ENM in the period 2001–2014, which is reflected in the increase in total PNM, in particular in 2013-2014.

Another important finding of this study was that, according to antenatal visits and US examinations which are prenatal care measures, the highest number of mothers was included in the optimal number of examinations in accordance with national recommendations [16]. This implies that some other factors like perinatal health care organization could have influenced ENM and PNM trends in 22–27 wks GA in Croatia and should be additionally investigated.

The mortality indicators are considered the most important measures of perinatal outcome, encouraging the health care professional efforts to prevent avoidable deaths. The data reliability depends on credible reporting and births and deaths recording process. The numerous researches have been emphasizing the problem of stillbirths and early neonatal deaths underreporting and misclassification [17–27]. The civil registration systems from many countries provide only basic information related to numbers of births and deaths and registration is required by law. However, the majority of civil registers do not collect birth or perinatal death data according to BW or GA. As in majority of countries, the main source of perinatal mortality indicators in Croatia was civil registration system, which was based on the birth and death certificates from CBS. According to the CBS methodology, FM is calculated as a number of stillbirths after 28th completed wks of pregnancy on 1,000 total births, irrespective of GA, up to the year 2001. ENM is calculated as a number of newborns who died in the first 168 hours (7 days) of life on 1,000 LB, irrespective of BW or GA. Whereas CBS does not collect data according to WHO and PERISTAT recommendations, it was impossible to carry out precise perinatal monitoring and international comparisons. Therefore, CSPM and CIPH introduced perinatal monitoring according to reports from maternities and developed new medical birth and perinatal death certificates by BW and GA, implementing their usage in the national routine health statistics system which covers more than 99% of births and perinatal deaths. However, the new monitoring system based on individual records needed a few years for developing and improving [13, 17]. Since 2001, Croatia has been providing the data about all perinatal deaths ≥1000 g to WHO-HFA for PNM calculations according to criteria for international comparison. For the purpose of the national analyses and evaluation of perinatal health care, both rates (according to BW ≥ 500 g and ≥ 1000 g) are used.

A similar reporting problem was detected in many European countries, preventing the perinatal, neonatal, postneonatal, and infant mortality comparisons by the BW or GA subgroups per country [22]. The health statistics systems differ in data collection methodology and area of coverage. Many countries use some form of linkage procedure to merge

data from different sources. In international databases such as WHO-HFA and the Statistical Office of the European Communities (Eurostat) database, these indicators can be found to be related to the different methodologies: for WHO-HFA database countries, the perinatal mortality indicator for BW group ≥1000 g or group ≥28 wks of GA is provided, while for Eurostat, it is according to different national statistical offices data for vital statistics registration. These data are not sufficient for perinatal outcome measures as opposed to the PERISTAT subgroup and GA division due to the fact that more than 70% of perinatal deaths in developed countries are connected with preterm birth and low BW [17–19, 21, 22, 28]. The lack of BW and GA data for late neonatal and postneonatal deaths hinders the analyses of the long-term consequences caused by ELBW and/or GA: physical, neurological, and cognitive impairments.

PNM, FM, and ENM by BW and GA have been regularly analyzed and discussed at annual national perinatal mortality conferences as a form of perinatal surveillance with the basic aim of preventing unfavorable perinatal outcomes [10]. Croatian PNM, calculated by WHO-HFA methodology criteria of BW ≥ 1000 g, seems to be lower than European Union (EU) average, amounting to below 5‰ for the period of 2007 onwards [15]. FM rates were 5.7‰–4.0‰ for ≥22 wks in the period 2001–2014, mildly decreasing from 2001 to 2014. In comparison with the PERISTAT survey data for 2010, the European countries range from 2.6‰ to 8.9‰ [29]. The first Croatian data originated in the year 1950, indicating FM rate of 17.8‰, which gradually decreased to today's value [10]. ENM rates were 4.4–3.1‰ for ≥22 wks in the period 2001–2014, decreasing in 2012 and increasing in the years 2013 and 2014. In comparison with the PERISTAT survey data for 2010, the European countries range from 1.0‰ to 4.0‰, but mostly in the scope from 1.5‰ to 2.0‰ [29]. The first officially published ENM rate in Croatia was 27.7‰ in the year 1950 with substantial decrease, especially after 1996 when FM rate was surpassed [10]. The perinatal health in Croatia, measured by PNM, FM, and ENM, has improved considerably in recent decades with the evident increase in ENM during the last two years, especially in 22–27 GA subgroup, which represents the cause for concern and requires detailed new analyses. According to other national studies, the increased number of ELBW newborns, mostly from multiple pregnancies, led to the rise of ENM resulting with the consequent rise of PNM over the last two years [12, 30–32].

The increase in the number of <1500 g BW newborns, LB, and deaths directs the perinatal health care endeavor towards the prevention, early diagnosis, and appropriate treatment of threatened early preterm labors, harmless delivery of those children, and thereafter appropriate treatment of those newborns in the NICUs [10, 23]. This may be achieved by implementing a regional organization of perinatal health care according to evidence-based studies and observations [33–43]. The routine perinatal health monitoring system is an important tool which enables the health care planning process in accordance with the requirements for appropriate level of health care, including human resources and adequate equipment. In order to improve the structure of Croatian perinatal health care system, all maternities and neonatal units are organized in a network, regionalized according to the professional guidelines [44]. However, the network is not officially confirmed by the Ministry of Health of the Republic of Croatia. The pregnant women are referred, as well as postpartum sick newborns, to the facilities of appropriate level, according to maternal/infant health condition. The referrals are mainly towards the maternities with NICUs, level III, situated in own perinatal region. The most complicated pregnancies and newborns can be referred to the National Center of Perinatal Medicine or to the National Center of Neonatal Intensive Medicine (level IV). The transfer of the sick newborns is organized as "one-way transport" [44].

The pregnancy and childbirth still involve risk for pregnant women and their babies and health in the perinatal period, while remaining an important public health priority. Although poor outcomes are increasingly rare, mothers in Europe still die in childbirth (5–15 women per 100,000 LB) [45]. Not only does MMR represent a key perinatal health outcome, but also it indicates the quality of obstetrical care, since many direct maternal deaths are associated with substandard care. The analysis of maternal deaths revealed that one-third of them are avoidable. The indicators of maternal mortality are extremely sensitive to underreporting, both in developing and in developed countries [46, 47]. The ascertainment of maternal deaths requires an effort by governments to ensure that deaths during or within one year after pregnancy are identified on death certificates or using other measures. Their precise registration depends on cause of death coding rules.

MMR for Croatia displays the substantial variation over time and the average is 8.1/100,000 LB during the period of 14 years, slightly higher than EU average considering the last few years. In 1954 (first known data), maternal mortality was 168/100,000 LB, rapidly decreased in the period 1960–1980, afterwards showing the values below 10/100,000 LB [48]. Following the introduction of the new reporting criteria and registration in the year 2001, the death causes of women in pregnancy, labor, and 42 days after labor are double-checked by CIPH and CSPM. The increase of the overall maternal deaths might have been caused by the fact that CIPH and CSPM have been including the indirect causes of maternal death as a part of the overall maternal deaths count since 2001. Up to the year 2001, Croatia reported only direct obstetrical maternal deaths.

The number of clinical visits and of US examinations of pregnant women is continuously increasing. The recommendations of CSPM, officially adopted and completely financed by Croatian Institute of Health Insurance, are 10 clinical visits per healthy pregnant woman and 3 US examinations [44]. The low proportion of pregnant women without adequate antenatal care and ultrasound examinations represents the indicators of prenatal care availability quality.

In view of the ongoing debates about the safest path to delivery, there is not yet clear consensus achieved. The spontaneous deliveries represented the majority of births in all countries, but the proportion of CS has been increased in the majority of European countries [15]. The rise of CS in Croatia, in comparison to the majority of European and other countries, is relatively mild. In relation to USA

and some European countries, there is a certain lag period present [49, 50]. There is an endeavor to stop further CS rise. The measurements of differences in mode of delivery under different circumstances including breech presentation, previous CS, parity, and multiple gestation pregnancies would offer better insight into necessity, risks, and benefits in specific circumstances.

The strength of this study is the perinatal mortality audit based on national routine health statistics which enables calculation of PNM, FM, and ENM rates adjusted for GA and BW for the whole population. The results can be considered as fairly reliable and representative for the entire country. Our review of perinatal mortality outcomes related to BW and GA specific mortality rates over the period of 14 years and comparisons with PERISTAT 2010 report can be considered a way to improve the health care process for all pregnant women and their newborns. It provides an opportunity to learn from adverse events, identifying and analyzing them and providing the future preventive measures.

The limitations of this study include the lack of the other perinatal health care indicators required for detailed insight into the provided perinatal health care, as well as perinatal morbidity outcomes with long-term physical, neurological, and cognitive impairment.

5. Conclusions

The perinatal health audit in Croatia has been improved after introducing recommended reporting criteria by WHO and PERISTAT which enable comparison in perinatal outcome with other countries. The outcome of this research provides an opportunity to identify problems and to prepare the plan for perinatal health care improvement. The perinatal monitoring system should be further improved while analyzing other perinatal indicators, except for the presented few core outcomes and process indicators, in order to achieve more complete image of perinatal care effectiveness and availability. The comprehensive perinatal health monitoring represents the basis for the perinatal quality assessment.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank all perinatologists and neonatologists in Croatia for their participation in data delivery.

References

[1] G. Lindmark and J. Langhoff-Ros, "Perinatal quality indicators and perinatal audit," in *Textbook of Perinatal Medicine*, A. Kurjak and F. A. Chervenak, Eds., vol. 1, pp. 264–268, 2nd edition, 2006.

[2] World Health Organization, "Standards and reporting requirements related to fetal, perinatal, neonatal and infant mortality," in *International Statistical Classification of Diseases and Related Health Problems. 10th Revision*, vol. 2 of *Instruction Manuel*, chapter 5, pp. 151–156, World Health Organization, Valletta, Malta, 2011, http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf.

[3] J. Zeitlin, K. Wildman, G. Bréart et al., "PERISTAT. Indicators for monitoring and evaluating perinatal health in Europe," *European Journal of Public Health*, vol. 13, supplement 3, pp. 29–37, 2003.

[4] J. Zeitlin, K. Wildman, and G. Bréart, "Perinatal health indicators for Europe: an introduction to the PERISTAT project," *European Journal of Obstetrics Gynecology and Reproductive Biology*, vol. 111, supplement 1, pp. S1–S4, 2003.

[5] J. Zeitlin, K. Wildman, G. Bréart et al., "Selecting an indicator set for monitoring and evaluating perinatal health in Europe: criteria, methods and results from the PERISTAT project," *European Journal of Obstetrics Gynecology & Reproductive Biology*, vol. 111, supplement 1, pp. S5–S14, 2003.

[6] L. Ostroški, "Statistical Yearbook of the Republic of Croatia 2014," Croatian Bureau of Statistics, Zagreb, Croatia, 2014, http://www.dzs.hr/Hrv_Eng/ljetopis/2014/sljh2014.pdf.

[7] World Health Organization, *Croatia: WHO Statistical Profile*, World Health Organization, Geneva, Switzerland, 2015, http://www.who.int/gho/countries/hrv.pdf?ua=1.

[8] M. Pekeč and V. Petrić, "Is there a change occurring in Croatia in the 21 st century?" in *Eurostat. Statistics Explained. Marriages and Births in Croatia*, 2015, http://ec.europa.eu/eurostat/statistics-explained/index.php/Marriages_and_births_in_Croatia/2015.

[9] T. Poljičanin and T. Benjak, "Croatian Health Service Yearbook 2014," Croatian Institute of Public Health, 2015, http://www.hzjz.hr/wp-content/uploads/2015/10/ljetopis_20142.pdf.

[10] A. Dražančić, U. Rodin, and B. Filipović-Grčić, "Perinatal care in Croatia yesterday, today, tomorrow," *Liječnički Vjesnik*, vol. 129, no. 3-4, pp. 87–99, 2007.

[11] U. Rodin, B. Filipović-Grčić, and A. Dražančić, "Births and perinatal mortality of very low birth weight newborns in Croatia in the 2001–2010 period," *Gynaecologia et Perinatologia*, vol. 20, supplement 2, pp. 46–50, 2011.

[12] J. Đelmiš, J. Juras, and U. Rodin, "Perinatal mortality in Republic of Croatia in the year 2014," *Gynaecologia et Perinatologia*, vol. 24, no. 1, pp. 3–18, 2015.

[13] U. Rodin, "Quality of data from new birth notification in Croatian maternities," *Gynaecologia et Perinatologia*, vol. 11, supplement 1, pp. 25–29, 2002.

[14] World Health Organization, *Reviewing Maternal Deaths and Complication to Make Pregnancy Safer; Beyond the Numbers*, Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland, 2004.

[15] World Health Organization, *Health for All Database*, World Health Organization, 2014, http://data.euro.who.int/hfadb.

[16] Ministry of Health, "Plan and Programme of Health Care Measures in Compulsory Health Insurance," *Official Gazette*, 126/2006.

[17] J. B. Gould, "Vital records for quality improvement," *Pediatrics*, vol. 103, no. 1, supplement E, pp. 278–290, 1999.

[18] T. A. Slagle, "Perinatal information systems for quality improvement: visions for today," *Pediatrics*, vol. 103, no. 1, supplement E, pp. 266–277, 1999.

[19] M. S. Kramer, S. Liu, Z. Luo, H. Yuan, R. W. Platt, and K. S. Joseph, "Analysis of perinatal mortality and its components: time for a change?" *American Journal of Epidemiology*, vol. 156, no. 6, pp. 493–497, 2002.

[20] P. H. T. Cartlidge and J. H. Stewart, "Effect of changing the stillbirth definition on evaluation of perinatal mortality rates," *The Lancet*, vol. 346, no. 8973, pp. 486–488, 1995.

[21] P. W. Setel, S. B. Macfarlane, S. Szreter et al., "A scandal of invisibility: making everyone count by counting everyone," *The Lancet*, vol. 370, no. 9598, pp. 1569–1577, 2007.

[22] A. Macfarlane, M. Gissler, F. Bloomer, and S. Rasmussen, "The availability of perinatal health indicators in Europe," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 111, supplement 1, pp. 15–32, 2003.

[23] J. E. Lawn, S. Cousens, and J. Zupan, "4 million neonatal deaths: when? Where? Why?" *The Lancet*, vol. 365, no. 9462, pp. 891–900, 2005.

[24] N. Lack, J. Zeitlin, L. Krebs, W. Künzel, and S. Alexander, "Methodological difficulties in the comparison of indicators of perinatal health across Europe," *European Journal of Obstetrics Gynecology and Reproductive Biology*, vol. 111, supplement 1, pp. S33–S44, 2003.

[25] G. L. Darmstadt, Z. A. Bhutta, S. Cousens, T. Adam, N. Walker, and L. De Bernis, "Evidence-based, cost-effective interventions: how many newborn babies can we save?" *The Lancet*, vol. 365, no. 9463, pp. 977–988, 2005.

[26] K. Shibuya, S. Scheele, and T. Boerma, "Health statistics: time to get serious," *Bulletin of the World Health Organization*, vol. 83, no. 10, p. 722, 2005.

[27] J. E. Lawn, D. Osrin, A. Adler, and S. Cousens, "Four million neonatal deaths: counting and attribution of cause of death," *Paediatric and Perinatal Epidemiology*, vol. 22, no. 5, pp. 410–416, 2008.

[28] S. Buitendijk, J. Zeitlin, M. Cuttini, J. Langhoff-Roos, and J. Bottu, "Indicators of fetal and infant health outcomes," *European Journal of Obstetrics Gynecology and Reproductive Biology*, vol. 111, supplement 1, pp. S66–S77, 2003.

[29] J. Zeitlin, A. Mohangoo, and M. Delnord, "The European Perinatal Health Report," 2010, http://www.europeristat.com/reports/european-perinatal-health-report-2010.html.

[30] J. Đelmiš, J. Juras, and U. Rodin, "Perinatal mortality in Republic of Croatia in the year 2013," *Gynaecologia et Perinatologia*, vol. 23, no. 1, pp. S3–S18, 2014.

[31] U. Rodin, B. Filipović-Grčić, T. Ćorić, and J. Juras, "Perinatal deaths causes in Croatia in the year 2014," *Gynaecologia et Perinatologia*, vol. 24, no. 1, pp. 19–25, 2015.

[32] B. Filipović-Grčić, Ž. Mustapić, U. Rodin, D. Bartoniček, and R. Grizelj, "Mortality of newbornsto discharge from hospital in Croatia in the year 2014," *Gynaecologia et Perinatologia*, vol. 24, no. 1, pp. 26–32, 2015.

[33] E. Papiernik and L. G. Keith, "The regionalization of perinatal care in France—description of a missing policy," *European Journal of Obstetrics and Gynecology*, vol. 61, no. 2, pp. 99–103, 1995.

[34] R. Vieux, J. Fresson, J.-M. Hascoet et al., "Improving perinatal regionalization by predicting neonatal intensive care requirements of preterm infants: an EPIPAGE-based cohort study," *Pediatrics*, vol. 118, no. 1, pp. 84–90, 2006.

[35] K. Beeckman, F. Louckx, S. Downe, and K. Putman, "The relationship between antenatal care and preterm birth: the importance of content of care," *European Journal of Public Health*, vol. 23, no. 3, pp. 366–371, 2013.

[36] Committee on Perinatal Health, *Toward Improving the Outcome of Pregnancy. Recommendations for the Regional Development of Maternal and Perinatal Health Services*, The National Foundation—March of Dimes, White Plains, NY, USA, 1976.

[37] P. R. Swyer, "The regional organisation of special care for the neonate," *Pediatric Clinics of North America*, vol. 17, no. 4, pp. 761–776, 1970.

[38] K. P. Russell, S. H. Gardiner, and E. E. Nichols, "A conceptual model for regionalization and consolidation of obstetric-gynecologic services," *American Journal of Obstetrics and Gynecology*, vol. 121, no. 6, pp. 756–764, 1975.

[39] R. Usher, "Changing mortality rates with perinatal intensive care and regionalization," *Seminars in Perinatology*, vol. 1, no. 3, pp. 309–319, 1977.

[40] J. Zeitlin, E. Papiernik, and G. Bréart, "Regionalization of perinatal care in Europe," *Seminars in Neonatology*, vol. 9, no. 2, pp. 99–110, 2004.

[41] V. Y. H. Yu and P. M. Dunn, "Development of regionalized perinatal care," *Seminars in Neonatology*, vol. 9, no. 2, pp. 89–97, 2004.

[42] D. L. Levy, K. Noelke, and J. P. Goldsmith, "Maternal and infant transport program in Louisiana," *Obstetrics & Gynecology*, vol. 57, no. 4, pp. 500–504, 1981.

[43] American Academy of Pediatrics and Committee on Fetus and Newborn, "Policy statement, levels of neonatal care," *Pediatrics*, vol. 114, no. 5, pp. 1341–1347, 2004.

[44] Croatian Society for Perinatal Medicine and Working Group for Perinatal Health Care, "Proposal of organization of perinatal care in the Republic of Croatia," *Gynaecologia et Perinatologia*, vol. 12, pp. 87–99, 2003.

[45] G. Lewis, "Reviewing maternal deaths to make pregnancy safer," *Best Practice and Research: Clinical Obstetrics and Gynaecology*, vol. 22, no. 3, pp. 447–463, 2008.

[46] D. Karimian-Teherani, G. Haidinger, T. Waldhoer, A. Beck, and C. Vutuc, "Under-reporting of direct and indirect obstetrical deaths in Austria, 1980–98," *Acta Obstetricia et Gynecologica Scandinavica*, vol. 81, no. 4, pp. 323–327, 2002.

[47] S. Alexander, K. Wildman, W. Zhang, M. Langer, C. Vutuc, and G. Lindmark, "Maternal health outcomes in Europe," *European Journal of Obstetrics Gynecology and Reproductive Biology*, vol. 111, supplement 1, pp. S78–S87, 2003.

[48] A. Dražančić, "Maternal mortality," *Gynaecologia et Perinatologia*, vol. 14, no. 1, pp. 7–17, 2005.

[49] C. McCourt, J. Weaver, H. Statham, S. Beake, J. Gamble, and D. K. Creedy, "Elective cesarean section and decision making: a critical review of the literature," *Birth*, vol. 34, no. 1, pp. 65–79, 2007.

[50] American College of Obstetricians and Gynecologists and Society for Maternal-Fetal Medicine, "Safe prevention of primary cesarean delivery," Obstetric-Care-Consensus 1, 2014, http://www.acog.org/Resources-And-Publications/Obstetric-Care-Consensus-Series/Safe-Prevention-of-the-Primary-Cesarean-Delivery.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 569235, 10 pages
http://dx.doi.org/10.1155/2015/569235

Review Article

Strong Public Health Recommendations from Weak Evidence? Lessons Learned in Developing Guidance on the Public Health Management of Meningococcal Disease

Germaine Hanquet,[1,2] Pawel Stefanoff,[3,4] Wiebke Hellenbrand,[5] Sigrid Heuberger,[6] Pierluigi Lopalco,[7] and James M. Stuart[8,9]

[1]*Consultant Epidemiologist (Independent), 1081 Brussels, Belgium*
[2]*Department of Vaccinology, University of Antwerp, 2000 Antwerp, Belgium*
[3]*Division of Epidemiology, Norwegian Institute of Public Health, 0403 Oslo, Norway*
[4]*National Institute of Public Health-National Institute of Hygiene, 400-791 Warsaw, Poland*
[5]*Department of Infectious Disease Epidemiology, Immunization Unit, Robert Koch Institute, 13086 Berlin, Germany*
[6]*Meningococcal Reference Laboratory, Austrian Agency for Food and Health Safety, 8010 Graz, Austria*
[7]*Vaccine-Preventable Diseases Programme, European Centre for Disease Prevention and Control, 171 83 Stockholm, Sweden*
[8]*School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK*
[9]*Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK*

Correspondence should be addressed to Pawel Stefanoff; pawel.stefanoff@gmail.com

Received 5 June 2015; Revised 22 September 2015; Accepted 26 October 2015

Academic Editor: Malgorzata Bala

The evidence underpinning public health policy is often of low quality, leading to inconsistencies in recommended interventions. One example is the divergence in national policies across Europe for managing contacts of invasive meningococcal disease. Aiming to develop consistent guidance at the European level, a group of experts reviewed the literature and formulated recommendations. The group defined eight priority research questions, searched the literature, and formulated recommendations using GRADE methodology. Five of the research questions are discussed in this paper. After taking into account quality of evidence, benefit, harm, value, preference, burden on patient of the intervention, and resource implications, we made four strong recommendations and five weak recommendations for intervention. Strong recommendations related not only to one question with very low quality of evidence as well as to two questions with moderate to high quality of evidence. The weak recommendations related to two questions with low and very low quality of evidence but also to one question with moderate quality of evidence. GRADE methodology ensures a transparent process and explicit recognition of additional factors that should be considered when making recommendations for policy. This approach can be usefully applied to many areas of public health policy where evidence quality is often low.

1. Introduction

The incidence of invasive meningococcal disease (IMD), caused by *Neisseria meningitidis*, is low in Europe, but case fatality is high (0.8 cases/100,000 inhabitants and 8.7%, resp., in 2011) [1]. Outbreaks of IMD may generate significant anxiety in the population, and even a single case may have important public health implications [2, 3]. *N. meningitidis* is transmitted from person to person and the risk of disease is highest in contacts from the same household as a case [4, 5]. In 2007, a European survey showed that recommendations for chemoprophylaxis to eliminate nasopharyngeal carriage in close contacts of IMD cases varied across Europe, in particular regarding the type of antibiotic and the groups that should be targeted [6, 7]. Discrepancies were partly due to differences in policy, medical practices, and health systems

but could also be explained by uncertainty surrounding the effectiveness of preventive measures or the differences in methods used to develop recommendations [6].

Divergences in national policies are particularly problematic in cross-border settings, as they lead to differences in disease management among population groups with the same exposure. For instance, passengers sharing an aeroplane flight with a case of IMD might or might not receive chemoprophylaxis depending on their country of residence [8]. The European Centre for Disease Prevention and Control (ECDC) therefore commissioned a group of experts to develop guidance for European countries on the management of contacts of IMD.

It was clear from the outset that high quality evidence in this area would be limited as in many areas of public health [9]. This is because randomized clinical trials on public health interventions are often difficult to organize (particularly when incidence of the outcome is low) and the use of placebo is no longer considered ethical when the intervention studied is already a recommended standard of care. In addition, evidence may be indirect as when only surrogate (proxy) endpoints are available. Therefore, information often comes from observational studies that are more prone to bias and are considered to provide a lower quality of evidence [10]. However, such studies can still be used in developing recommendations if systematically researched and graded appropriately [11]. In addition, evidence obtained for a public health intervention in one country may not be fully applicable to another setting, as public health interventions are strongly dependent not only on the epidemiological context but also on cultural and economic context of countries in which they are implemented.

Here we share our experience and lessons learned in using different types and quality of evidence to develop guidance on the public health management of IMD for European countries within a short time period using GRADE methodology. The aim of this guidance, available on the ECDC website, was to assist countries across Europe in making decisions about appropriate measures to control and prevent IMD in contacts of cases at national and subnational levels [12].

2. Description of the Process

We adapted existing methods for producing evidence-based guidelines to deal with the short time frame and the scarcity of direct evidence (see Section 2.12) [10, 13–17]. The main steps are described below.

2.1. Setting Up Expert Groups. We set up a consortium of national experts: four in the area of epidemiology and public health surveillance and one in the area of microbiology of meningococcal disease. The consortium members represented five EU countries, previously involved in the assessment of national practices for IMD management across the European Union [6]. The consortium identified research questions, developed protocols, identified, assessed, and graded evidence, and formulated and graded recommendations. This work was contracted for completion within 6 months, had a budget of 20,000 Euros, and was conducted through two face-to-face meetings, three teleconferences, and close to 500 e-mail exchanges. Each member of the expert group completed a declaration of potential conflict of interests.

The consortium identified other national epidemiologists and microbiologists working with meningococcal disease from all EU countries through two established European networks (the European Meningococcal Disease Society and the ECDC European Invasive Bacterial Diseases Surveillance Network). These EU experts were asked to provide any related grey literature and technical advice during the process. Additionally, the consortium consulted two patient group networks, both based in the United Kingdom, on patient-related values and preferences.

2.2. Defining the Area of Guidance and Formulating the Research Questions. The consortium defined research questions for guidance focused on the prevention of subsequent cases following sporadic cases of IMD and based on the needs identified through two previous surveys among public health representatives of EU countries [6, 18]. Five research questions for evidence assessment are discussed in this paper. A summary of the evidence and recommendations for all research questions can be found in the ECDC guidance [12].

Research Questions for Evidence Assessment

(A) What is the effectiveness of chemoprophylaxis to a case of IMD before discharge from hospital in preventing further cases of IMD?

(B) What is the effectiveness of chemoprophylaxis to household contacts of an IMD case in preventing further cases?

(C) What is the effectiveness of chemoprophylaxis to contacts of an IMD case in pre-school and school settings in preventing further cases?

(D) What is the effectiveness of chemoprophylaxis to those sharing the same transport vehicle as an IMD case in preventing further cases?

(E) Which antibiotic regimes are most effective in eradicating carriage among adults, children and pregnant women?

2.3. Defining the Methodology. We opted for GRADE (Grading of Recommendations Assessment, Development, and Evaluation) methodology to assess evidence and produce guidance; we used the GRADE guidance available at the time of this study (2008-09) [14, 17, 19]. GRADE not only considers the balance between the benefits and harm and the quality of evidence but also includes additional factors on which to base recommendations, such as burden of the intervention on the patient, patients' values and preferences, and resource implications, which were not addressed by a number of other grading systems. GRADE also provides clear criteria to qualify the strength of a recommendation. Although GRADE has been considered by some as being too resource intensive

and difficult to apply in public health guidance, especially under time constraint and when evidence is limited [20], it is recognized by others to provide a systematic approach, promote dialogue, and ensure documentation of the process that leads to a given recommendation [21, 22]. This makes decision-making more transparent. We referred to previous experience reported in two World Health Organization (WHO) publications on rapid advice guidelines that made use of GRADE methodology and to Cochrane guideline for public health interventions [14, 15, 23].

The consortium developed protocols, templates, and checklists for screening abstracts/papers retrieved in the literature searches to ensure a homogenous process across the research questions and across reviewers. The process was also reviewed against the criteria for guideline development as defined by the AGREE collaboration [16].

2.4. Search Strategy and Selection Criteria for Systematic Reviews. When defining the most suitable terms for the population, intervention, comparison, and outcome (PICO) to define our research questions for the search strategy (see examples in Table 1) [15], we took into account prior knowledge and a preassessment of the literature. Our preassessment suggested that measurements of direct outcomes would not be available for several research questions due to the low incidence of IMD. For these questions, we defined and included proxy outcomes in our search strategy (see examples below Section 2.5).

We defined inclusion and exclusion criteria for selecting studies, applied to each research question. All European languages were included to avoid publication biases. As most of our studied interventions either were standard clinical practice or involved rare outcomes, clinical trials had not been conducted for ethical or logistical reasons. We thus did not limit inclusion to experimental studies but also included observational studies that involved comparison groups and case series with at least 10 cases.

We searched MEDLINE, Embase, Global Health, the Cochrane database of systematic reviews, and the Cochrane central register of controlled trials. The search terms for each question were agreed by at least two members of the consortium. Due to the short time frame, we applied some of the strategies previously used in rapid reviews by limiting the search to the period from January 1990 to the date of the literature search (December 2008) and giving priority to systematic reviews [24]. If a relevant systematic review was identified, we only considered abstracts published from the date of search for the last review up to the end of 2008. If no relevant review was identified, we screened all abstracts published from 1990 to 2008. We reviewed full papers of abstracts identified as relevant. One reviewer only was involved in reviewing abstracts and full papers of each research question due to time constraints. In case of doubts, ad hoc opinion of a second reviewer was requested.

We examined reference lists of the selected papers for other relevant publications and searched Google Scholar for citations of identified key papers. For instance, for the research question on effectiveness of antibiotic regimens (Question E), we found a nonindexed but peer reviewed trial on antibiotic prophylaxis by using Google Scholar, though this study had not been retrieved by a previous Cochrane systematic review [25].

2.5. Use of Indirect Evidence. As studies measuring direct evidence on outcomes could not be found in four of the five research questions discussed in this paper, we defined and searched for indirect evidence on outcomes. For example, the relevant direct outcome for the question on the effectiveness of chemoprophylaxis before hospital discharge in preventing further cases among contacts (Question A) would be the incidence of subsequent cases in household contacts of the IMD patients who received antibiotics prior to discharge from hospital. A prior systematic review did not identify relevant studies measuring this outcome but showed that eradicating nasopharyngeal carriage in household contacts reduced the risk of further cases [7]. We thus searched for data on the proxy outcome, that is, the prevalence of meningococcal carriage in discharged patients (Table 1).

We also did not find direct evidence on the research question regarding whether chemoprophylaxis of contacts in school settings would prevent further cases (Question C). However, we obtained indirect evidence by comparing the risk of subsequent cases in school contacts (not receiving chemoprophylaxis) with the background incidence rates of IMD in the relevant population [26].

Even when the literature search provided direct evidence on the benefits of an intervention (e.g., effectiveness in preventing secondary cases), the evidence was often insufficient on its harm. In particular, direct evidence on the adverse events of antibiotics administered as chemoprophylaxis (Question E) was insufficient, but we found and reviewed indirect evidence on adverse events of these antibiotics when administered for indications other than chemoprophylaxis (e.g., ciprofloxacin used in cystic fibrosis).

2.6. Analysis of Extracted Data. We extracted and summarized the evidence on benefits and harm and prepared evidence profiles. When possible, we pooled estimates retrieved from selected studies. For instance, for Question B on chemoprophylaxis for household contacts, we extracted results from a recent study published after a systematic review and analysed these together with the three former studies from the review [7, 27–29]. As the results of the four studies were statistically homogeneous, we calculated a common pooled estimate (Figure 1) [12]. In case of heterogeneity between studies, we performed stratified analysis when possible. In particular, analyses on the effectiveness of chemoprophylaxis to contacts of an IMD case in preschool and school settings (Question C) were stratified by each educational setting [26].

If the retrieved systematic reviews did not provide the level of detail needed to calculate pooled estimates of effectiveness or to fully answer the research questions, we extracted the necessary data from primary studies when appropriate. For instance, in the research question on antibiotic regimes for different subgroups (Question E), we identified Cochrane systematic review on antibiotics for preventing

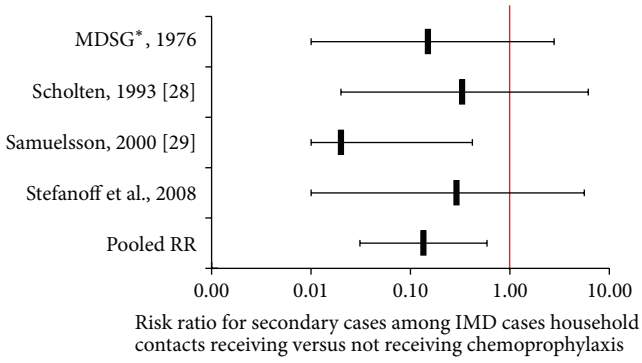
Table 1: Summary of findings and recommendations for three research questions.

Assessment of evidence quality						Assessment for recommendations				
Number of studies selected/number of studies reviewed	Design and quality	Inconsistency	Indirectness	Other modifying factors*	Grade of evidence	Benefits	Harm	Costs and burdens	Values and preferences	Grade of recommendation
Research question: What is the effectiveness of chemoprophylaxis to a case of IMD before discharge from hospital in preventing further cases of IMD?										
Population: cases of IMD and their household contacts										
Intervention: administration of chemoprophylaxis (including therapeutic treatment if effective in carriage eradication) to case prior to discharge from hospital										
Comparison: no chemoprophylaxis administered to case prior to discharge from hospital										
Outcome: carriage in IMD cases; incidence of IMD in household contacts										
4/349	All observational. No studies addressed the intervention	Different antibiotic regimens, times of swabbing, sampling, and study populations	Proxy outcome: prevalence of carriage following discharge from hospital	Small sample sizes. Study results are statistically homogenous	Very low	Potential reduction of the disease burden among close contacts of discharged cases	Low risk of treatment side effects	Very low cost and low burden for the patient (oral single dose)	Treatment widely accepted	Strong
Recommendation: chemoprophylaxis is recommended for patients with IMD on discharge from hospital unless an antibiotic regimen effective in eradicating carriage was used during hospital treatment										
Implication for practice: easy to implement in hospital procedures; treatment is widely accepted										
Research question: What is the effectiveness of chemoprophylaxis to contacts of a case of IMD who have shared the same transport vehicle as a case of IMD in preventing further cases among those contacts?										
Population: contacts of diagnosed IMD cases sharing the same transport vehicle, for example, plane, boat, bus, and car										
Intervention: administration of chemoprophylaxis to contacts sharing transport vehicle, following IMD diagnosis in a case										
Comparison: no chemoprophylaxis administered to contacts sharing transport vehicle, following IMD diagnosis in a case										
Outcome: incidence rate of IMD in contacts sharing a transport vehicle with IMD cases (up to 30 days)										
7/103	Only reports on sporadic cases and 3 clusters linked to travel. No studies addressed the intervention	No consistency across case reports	Proxy outcome: risk of subsequent cases among fellow passengers whether prophylaxis was given or not	No studies clearly established evidence of transmission in transport vehicles	Very low	No evidence of reduction of subsequent cases among contacts sharing the same transport and taking prophylaxis	Low risk of treatment side effects but potential anxiety among those not receiving prophylaxis if targeted	Low cost of the intervention. However, contact tracing can lead to considerable cost and may not be feasible	Treatment likely to be accepted even if objective risk is low. Possible public pressure to give prophylaxis	Weak

Table 1: Continued.

Number of studies selected/number of studies reviewed	Assessment of evidence quality: Design and quality	Inconsistency	Indirectness	Other modifying factors*	Grade of evidence	Assessment for recommendations: Benefits	Harm	Costs and burdens	Values and preferences	Grade of recommendation
Recommendation: sharing the same transport vehicle as a case of IMD is not, in itself, an indication for chemoprophylaxis										
Implication for practice: a consistent European policy is highly desirable because high potential for confusion related to divergent cross-border policies. However, achieving consensus may not be easy										
Research question: Which antibiotic regimes are most effective in eradicating carriage among adults? *Population*: adult carriers of *N. meningitidis* *Intervention*: administration of antibiotic (type, dose, duration, and route) *Comparison*: no antibiotics, alternative type of antibiotic, alternative dose, alternative duration, or alternative route *Outcome*: carriage of *N. meningitidis* at ≥7 days of follow-up. Occurrence of resistant strains of *N. meningitidis* after treatment										
28**	17 RCTs and 3 observational studies; no serious limitations	High consistency of results across studies	Proxy outcome: eradication of carriage. Only assessed in students/army for azithromycin and cefixime	High associations	High: rifampicin, ciprofloxacin, and ceftriaxone Moderate: azithromycin and cefixime	The 5 antibiotics are highly effective (eradication in 79–100%)	Limited harm of antibiotics and mild side effects. Emergence of resistance with rifampicin and interactions with other drugs	Low cost. Lower burden for ciprofloxacin and azithromycin as single oral dose. Ceftriaxone is intramuscular	High acceptability of intervention. A single oral dose is likely to be preferred	Strong for the 5 antibiotics. Weak for "ciprofloxacin, azithromycin and ceftriaxone are preferred"
Recommendation: rifampicin, ciprofloxacin, ceftriaxone, azithromycin, and cefixime can be advised for chemoprophylaxis in adults. Ciprofloxacin, azithromycin, and ceftriaxone are preferred										
Implication for practice: using ciprofloxacin, azithromycin, or ceftriaxone would require a change of practice in several countries but has a high feasibility at similar or lower cost. Surveillance of resistance is essential										

*Strength of association and imprecision [17].
**Total reviewed not relevant because we identified a systematic review and selected 16 of the included studies as well as 9 studies through references of selected papers. Of the update search conducted in the period after the systematic review, 3/67 studies were included.
RCT: randomized clinical trial; IMD: invasive meningococcal disease.



* Meningococcal Disease Surveillance Group, CDC, Atlanta, GA, USA

Figure 1: Estimate of effect of chemoprophylaxis to household contacts following a sporadic IMD case.

meningococcal infections, but this review did not present detailed analyses by antibiotic dosage and duration of therapy and did not stratify estimates by subgroups such as children and pregnant and lactating women [25]. We thus retrieved the nine primary studies that involved different dosages, treatment durations, and subgroups, extracted the required data, and appraised the studies based on the full papers.

2.7. Reviewing and Grading Evidence. Evidence was classified for all questions as either direct or indirect. We graded bodies of evidence according to GRADE guidelines and classified them as high, moderate, low, or very low, based on study design and quality, inconsistency, indirectness, imprecision, and strength of the association [17, 19]. In particular, we made a judgment of whether the evidence was sufficiently indirect to warrant downgrading. For example, we downgraded the quality of evidence for the four studies retrieved for the research question on chemoprophylaxis for cases before hospital discharge (Question A), as they measured a proxy outcome, had very small study populations ranging from 14 to 51, and used different therapeutic antibiotic regimes. As observational studies start with "low quality" evidence rating according to GRADE [13], this led to classification of the evidence quality to "very low."

We found high or moderate quality evidence for only two of the five research questions, that is, regarding chemoprophylaxis to household contacts (Question B) and antibiotic regimes (Question E). Only low or very low quality of evidence was found for the remaining three research questions (Questions A, C, and D). Table 1 describes evidence review and grading for three research questions (Questions A, D, and E) for which the quality of evidence was very low (A and D) or moderate to high (E).

2.8. Burden on Patient, Values, Preferences, and Resource Implications. Because these factors varied across settings [21, 22], we first outlined, for each question, what should be considered from the patients perspective as burden of the intervention and their values and preferences and resource implications in an EU setting. For instance, for the burden of prophylactic antibiotics (Question E), we considered for each antibiotic the potential side effects, inconvenience (e.g., number of days of treatment) for contacts, ease of administration, and the number of contacts needed to be treated (where possible to calculate) to prevent one IMD case among contacts according to each setting. We also considered the implications of contact tracing, as this can lead to considerable costs when, for instance, tracing close contacts on the same aeroplane as an IMD case is required (Question D) and may even not be feasible, for example, in case of free seating.

We found little information in the literature on burden of intervention perceived by patients and on their values and preferences. We searched for alternative data sources: for instance, information on perceived burden and values was requested from EU experts and national IMD representatives as well as from two meningitis patient associations. This confirmed that IMD is perceived as a severe disease that generates a high level of anxiety, and thus prevention measures are widely accepted, even if associated with some level of discomfort.

2.9. Developing and Grading the Strength of Recommendations. The consortium met face-to-face to develop recommendations according to GRADE, based on the quality of evidence and the balance between the benefits and harm, taking into account burden, values, preferences, and costs (see examples in Table 1).

Recommendations were classified as strong or weak as recommended by GRADE [10, 14, 17]. The GRADE guidance available at the time of developing this guidance did not provide an objective method for assessing the balance between benefit, harm, burden, values, and costs [13, 17]. We decided that the entire consortium should participate and agree on this appraisal process and we included advice of two patients' groups regarding values and preferences related to the recommendations. Based on these criteria, four strong and five weak recommendations for intervention were made for the five research questions. Strong recommendations were made not only in relation to two research questions with moderate to high quality of evidence (Questions B and E) but also in relation to one research question with very low quality of evidence (Question A). The weak recommendations for or against intervention referred not only to two research questions with low or very low quality of evidence (Questions C and D) but also to two aspects of one research question with moderate quality of evidence (Question E).

The strong recommendation for which the quality of evidence on the benefit was very low was related to the research question on chemoprophylaxis of IMD cases before hospital discharge (Question A). Despite the very low quality evidence on the benefits, the consortium considered that harm, cost, burden, and values were strongly in favour of the intervention: the low cost of the intervention, the low number of patients not treated with an eradicating antibiotic regimen prior to discharge, and the potential benefit in reducing risk from a life-threatening disease were balanced against

limited harm from antibiotics (Table 1). In general, higher quality evidence is more likely to be associated with strong recommendations than lower quality evidence [11]. However, the GRADE methodology indicates that a particular quality of evidence does not imply a particular strength of recommendation. A number of public health guidances and GRADE clinical guidelines issued strong recommendations in the face of a very low quality of evidence [11, 13, 14, 23, 30]. WHO also considers that strong recommendations can be made despite low or very low quality evidence in specific circumstances, as it is the net result of all relevant factors that are important [21]. For instance, WHO rapid advice guidelines for management of sporadic human infection with avian influenza A (H5N1) virus made a strong recommendation to treat H5N1 patients with oseltamivir, although the quality of the underlying evidence was rated as very low in part because of the severity of the disease [23]. A recent review highlighted that over than half (55%) of strong recommendations in WHO guidelines were based on low or very low confidence in effect estimates [30]. A GRADE guidance published after our review describes five situations in which a strong recommendation is warranted despite low or very low confidence in effect estimates [31]. The most relevant to our review was "when low quality evidence suggests benefit in a life-threatening situation," as all of our recommendations aimed at preventing a life-threatening invasive infection.

Consensus on the recommendations and grading of their strength was difficult to reach regarding chemoprophylaxis in day care settings (Question C). Here, the quality of evidence was low, and divergent recommendations were in place in the consortium members' native countries. Thus each expert was probably influenced by his/her existing national policy. This highlighted that recommendations are built not only on rigorous scientific reviews but also on expert interpretation and judgment of the evidence. An advantage of the GRADE approach is to promote useful dialogue and ensure transparency by making these value judgments explicit [17, 21].

We involved stakeholders and potential users of the guidance in the final steps. As our aim was to produce guidance that could be adapted to the needs of different EU countries, the draft document was circulated through EU experts and patient groups and reviewed by representatives of EU countries in the ECDC Advisory Forum. The feedback from ECDC and EU experts on the draft report allowed useful additions to the guidance [12]. For instance, we added recommendations on use of antibiotics by lactating women on request from representatives from a patient association.

2.10. Implications for Practice. We described how the guidance would potentially change current practice in EU countries. For instance, for the research question (E) on effectiveness of antibiotic regimens in IMD prophylaxis, we described what policy changes would be required and potential obstacles to the implementation of this guidance in a EU setting, based on whether the intervention or the specific drug is available and whether the recommended regimen differs from those currently recommended. In particular, some effective dosages did not correspond to recommendations and formulations available in EU countries and would require a change in current guidance. For instance, high quality evidence was available for the effectiveness of a single dose of 750 mg ciprofloxacin for the eradication of meningococcal carriage. However, in many countries, ciprofloxacin is recommended as a 500 mg single dose, although the effectiveness of this lower dosage has not been assessed in a controlled trial.

2.11. Strengths of the Process. The guidance was successfully completed within six months and was approved and endorsed by ECDC in 2010 [12]. The GRADE approach allowed transparent judgments on the quality of evidence and the formulation of recommendations. Our process met most of the criteria for guideline development as defined by the AGREE collaboration (2003 version) [16]. We complied with the following criteria: definition of the scope and purpose, stakeholder involvement, rigor of development, clarity and presentation, application, and editorial independence. On the other hand, our review process did not fulfil criteria pertaining to tools for application and audit.

An advantage of GRADE process in developing public health recommendations is the integrated appraisal of related values, preferences, burden to the patient, and resource implications in addition to quality of evidence and the balance between benefits and harm. Based on GRADE 2004–08 guidance, we made strong recommendations for some areas in which the quality of evidence was low or very low. The long deliberations often required to arrive at final agreement of recommendations were facilitated by frequent communication, mainly by e-mail. It should be noted that GRADE work published later provides a systematic approach by describing circumstances in which a strong recommendation is warranted despite low or very low confidence in effect estimates, but these were not available at the time of developing our guidance [30, 31].

The influence of national policies on the judgment of each consortium expert to formulate recommendations (described above) was dealt with by explicitly discussing each recommendation in the entire group. One advantage of having experts from five EU countries in the consortium was also that they had knowledge of current practices and health systems when considering implications for practice of the guidance.

The development of this guidance led to the identification of areas of uncertainty and research gaps, and we identified priorities for further research in each area. It was also a unique opportunity to progress towards common European health policy. Divergent health policies may cause confusion among the public and the media. The most objective argument for common health policies consists of a systematic and transparent search for and evaluation of available evidence. In 2013, we evaluated the impact of this guidance on the recommendations for public health management of IMD in European countries and found out that 90% of the 31 EU countries or regions found it useful at the national level and that 50% used it to update their national guidelines within the three years following the publication of the guidance [32].

WHO adopted a very similar process for developing evidence-based immunization recommendations, published after we initiated this work [21]. Immunization is an area of public health prevention in which the evidence may also be indirect (e.g., immunological surrogate for clinical efficacy) as in our example. GRADE was selected by WHO because it improves transparency in decision-making, promotes dialogue, and provides opportunities to reassess the evidence as required [21].

2.12. Limitations of the Process. The consortium included mostly experts in epidemiology and microbiology. It could have benefited from including clinical experts and members of patient organization groups, but the short time frame was already challenging for finalizing the project.

The limited time (6 months) and available resources imply that our literature reviews could not meet the standards of a full systematic review. In addition, we could not cover all aspects of IMD public health management.

Indeed, the comprehensive application of the GRADE methodology including exhaustive systematic reviews may require substantial resources and more time is often required for rigorous development of guidelines [24]. Thus, in this project, we applied some strategies of rapid reviews, such as focusing on existing systematic reviews, having only one expert for evidence reviews, and limiting the search period [24]. These may have introduced biases in the selection and appraisal of studies [24]. However, we enhanced our searches through inclusion of older studies, searching manually the references of retrieved studies, not restricting literature search by language or database, and asking experts for unpublished data and potentially missed studies. Additionally, though our initial search focused on updating systematic reviews, we nonetheless retrieved relevant primary studies to extract all relevant data, if they were not provided in the systematic reviews. Some authors have suggested that when the timeframe is limited, combination of electronic searching, hand searching of relevant reference lists, and consultation with experts on potentially missed articles may provide the most comprehensive results [24]. In this regard, WHO publication on rapid advice guideline—that met similar time constraints—was a particularly useful reference in helping us to ensure transparency of the process [14]. It is likely that the specialist expertise of those performing the review as well as input from other EU experts minimized the risk that relevant studies would be missed [24].

We relied mostly on systematic reviews (including one Cochrane review) for the quality appraisal of individual studies for Questions B and E [7, 25], but these described the risk of biases and not the other GRADE criteria for assigning grades of evidence [11]. In particular we did not fully appraise bodies of evidence for each outcome for imprecision, also due to limited instructions in the GRADE guidance available in 2008 [17]. New guidance published after our review describes each criterion for appraising evidence more explicitly, including imprecision [33], allowing further downgrading for indirectness, imprecision, and reporting bias [11]. We also did not explicitly define which outcomes were critical to a decision and which ones were important for grading overall quality of evidence [24, 34]. It is likely that the strict application of the newer GRADE guidance could have led to further downgrading of the quality of evidence for some of our outcomes, although this may not have changed our recommendations.

As explained above, one of the challenges was that we only found a low quality of evidence (according to GRADE) in most areas, as evidence from RCTs was only available on the effectiveness of antibiotic regimes in eradicating carriage.

The GRADE guidance required defining the burden of the intervention to the patient as well as patients' values, preferences, and resource implications to aid in the development of recommendations. However, the GRADE guidance available at the time of developing these recommendations did not provide a methodology to collect and appraise the evidence in these areas. We found scarce information in the literature on the burden, values, and preferences surrounding interventions and limited data on cost in a few countries, and these may differ across countries. Although we questioned EU representatives and two UK-based meningococcal patient organizations, a representative survey of patients across Europe would be required for obtaining sound and representative evidence. However, we did not have the resources to initiate a multinational public survey on these issues. Furthermore, the GRADE guidance did not standardize how the data on burden and values should affect the recommendation; this is left to deliberation on the part of the decision-making group and has been criticized as a weakness of the GRADE process [20]. The updated GRADE guidance also provides a more structured way to incorporate values and preferences in the development of recommendations [31].

3. Conclusions

We developed evidence-based guidance on the public health management of meningococcal disease for EU countries in a short time frame and with limited resources. A number of recommendations in this guidance were based on a low quality of sometimes indirect evidence due to the impracticability of conducting clinical trials on interventions for outcomes that are rare or that have become standard practice. However, the recommendations were generated systematically and transparently, following GRADE and AGREE standards. This approach, that explicitly integrates additional criteria with the quality of evidence, can be usefully applied to the many areas of public health policy in which quality of evidence is often low or indirect. A recent survey of European countries showed that the majority found the guidance based on this process useful, about half had used the guidance to update their national recommendations, and a higher proportion of countries since 2013 compared to that in 2007 recommended evidence-based measures for IMD public health management [32].

Disclosure

The funder had no role in the design, collection, analysis, and interpretation of data, in the preparation of the paper, and in the decision to submit it for publication.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Authors' Contribution

James M. Stuart, Germaine Hanquet, Pawel Stefanoff, Wiebke Hellenbrand, and Sigrid Heuberger designed the study, performed the literature review, and developed the guidance. Germaine Hanquet, Pawel Stefanoff, and James M. Stuart wrote the paper with substantial contributions from Wiebke Hellenbrand, Sigrid Heuberger, and Pierluigi Lopalco. All authors approved the final version. Germaine Hanquet and Pawel Stefanoff contributed equally to this paper.

Acknowledgments

The authors wish to thank the European meningococcal experts who commented on the ECDC report and provided information. The European Centre for Disease Prevention and Control (ECDC) commissioned and partly funded the development of the guidance and covered the paper-processing charges.

References

[1] European Centre for Disease Prevention and Control, *Surveillance of Invasive Bacterial Diseases in Europe, 2011*, ECDC, Stockholm, Sweden, 2013.

[2] K. Perrett, W. al-Wali, C. Read, P. Redgrave, and U. Trend, "Outbreak of meningococcal disease in Rotherham illustrates the value of coordination, communication, and collaboration in management," *Communicable Disease and Public Health*, vol. 3, no. 3, pp. 168–171, 2000.

[3] I. Zuschneid, A. Witschi, L. Quaback et al., "Invasive meningococcal disease with fatal outcome in a Swiss student visiting Berlin," *Euro Surveillance*, vol. 13, no. 45, 2008.

[4] P. De Wals, L. Hertoghe, I. Borlée-Grimée et al., "Meningococcal disease in Belgium. Secondary attack rate among household, day-care nursery and pre-elementary school contacts," *Journal of Infection*, vol. 3, no. 1, supplement, pp. 53–61, 1981.

[5] L. Hastings, J. Stuart, N. Andrews, and N. Begg, "A retrospective survey of clusters of meningococcal disease in England and Wales, 1993 to 1995: estimated risks of further cases in household and educational settings," *Communicable Disease Report. CDR Review*, vol. 7, no. 13, pp. R195–R200, 1997.

[6] M. Hoek, G. Hanquet, S. Heuberger et al., "A European survey on public health policies for managing cases of meningococcal disease and their contacts," *Eurosurveillance*, vol. 13, no. 10, 2008.

[7] B. Purcell, S. Samuelsson, S. J. M. Hahné et al., "Effectiveness of antibiotics in preventing meningococcal disease after a case: systematic review," *British Medical Journal*, vol. 328, no. 7452, pp. 1339–1342, 2004.

[8] T. Rachael, K. Schubert, W. Hellenbrand, G. Krause, and J. M. Stuart, "Risk of transmitting meningococcal infection by transient contact on aircraft and other transport," *Epidemiology and Infection*, vol. 137, no. 8, pp. 1057–1061, 2009.

[9] J. Latham, L. Murajda, F. Forland, and A. Jansen, "Capacities, practices and perceptions of evidence-based public health in Europe," *Eurosurveillance*, vol. 18, no. 10, Article ID 20421, 2013.

[10] D. Atkins, M. Eccles, S. Flottorp et al., "Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. The GRADE Working Group," *BMC Health Services Research*, vol. 4, no. 1, supplement 38, 2004.

[11] H. Balshem, M. Helfand, H. J. Schünemann et al., "GRADE guidelines: 3. Rating the quality of evidence," *Journal of Clinical Epidemiology*, vol. 64, no. 4, pp. 401–406, 2011.

[12] European Centre for Disease Prevention and Control, *Public Health Management of Sporadic Cases of Invasive Meningococcal Disease and Their Contacts*, ECDC, Stockholm, Sweden, 2010.

[13] G. H. Guyatt, A. D. Oxman, G. E. Vist et al., "GRADE: an emerging consensus on rating quality of evidence and strength of recommendations," *British Medical Journal*, vol. 336, no. 7650, pp. 924–926, 2008.

[14] H. J. Schünemann, S. R. Hill, M. Kakad et al., "Transparent development of the WHO rapid advice guidelines," *PLoS Medicine*, vol. 4, no. 5, article e119, 2007.

[15] The Cochrane Collaboration, *Systematic Reviews of Health Promotion and Public Health Interventions*, The Cochrane Collaboration, Melbourne, Fla, USA, 2008.

[16] AGREE Collaboration, *Appraisal of Guidelines for Research and Evaluation. Instrument Training Manual*, AGREE Collaboration, 2003.

[17] D. Atkins, D. Best, P. A. Briss et al., "Grading quality of evidence and strength of recommendations," *British Medical Journal*, vol. 328, no. 7454, pp. 1490–1494, 2004.

[18] D. Boccia, N. Andrews, S. Samuelsson, S. Heuberger, A. Perrocheau, and J. M. Stuart, "Effectiveness of different policies in preventing meningococcal disease clusters following a single case in day-care and pre-school settings in Europe," *Epidemiology and Infection*, vol. 134, no. 4, pp. 872–877, 2006.

[19] G. H. Guyatt, G. Vist, Y. Falck-Ytter, R. Kunz, N. Magrini, and H. Schunemann, "An emerging consensus on grading recommendations?" *Evidence-Based Medicine*, vol. 11, no. 1, pp. 2–4, 2006.

[20] European Centre for Disease Prevention and Control, *Evidence-Based Methodologies for Public Health—How to Assess the Best Available Evidence When Time is Limited and There is Lack of Sound Evidence*, ECDC, Stockholm, Sweden, 2011.

[21] P. Duclos, D. N. Durrheim, A. L. Reingold, Z. A. Bhutta, K. Vannice, and H. Rees, "Developing evidence-based immunization recommendations and GRADE," *Vaccine*, vol. 31, no. 1, pp. 12–19, 2012.

[22] D. Matysiak-Klose, F. Ahmed, P. Duclos et al., "Report on the 1st international workshop on procedures for the development of evidence-based vaccination recommendations, Berlin, Germany, 22-23 November 2010," *Vaccine*, vol. 30, no. 14, pp. 2399–2404, 2012.

[23] H. J. Schünemann, S. R. Hill, M. Kakad et al., "WHO Rapid Advice Guidelines for pharmacological management of sporadic human infection with avian influenza A (H5N1) virus," *Lancet Infectious Diseases*, vol. 7, no. 1, pp. 21–31, 2007.

[24] R. Ganann, D. Ciliska, and H. Thomas, "Expediting systematic reviews: methods and implications of rapid reviews," *Implementation Science*, vol. 5, no. 1, article 56, 2010.

[25] A. Fraser, A. Gafter-Gvili, M. Paul, and L. Leibovici, "Antibiotics for preventing meningococcal infections," *Cochrane Database of Systematic Reviews*, no. 4, Article ID CD004785, 2006.

[26] W. Hellenbrand, G. Hanquet, S. Heuberger, S. Nielsen, P. Stefanoff, and J. M. Stuart, "What is the evidence for giving chemoprophylaxis to children or students attending the same preschool, school or college as a case of meningococcal disease?" *Epidemiology & Infection*, vol. 139, no. 11, pp. 1645–1655, 2011.

[27] P. Stefanoff, M. Rosinska, G. Karczewski, and A. Zielinski, "The detection of meningococcal household clusters and their prophylaxis in the changing epidemiological situation of invasive meningococcal disease in Poland, 2003–2006," *Eurosurveillance*, vol. 13, no. 10, 2008.

[28] R. J. P. M. Scholten, H. A. Bijlmer, J. Dankert, and H. A. Valkenburg, "Secondary cases of meningococcal disease in the Netherlands, 1989-1990, a reappraisal of chemoprophylaxis," *Nederlands Tijdschrift voor Geneeskunde*, vol. 137, no. 30, pp. 1505–1508, 1993.

[29] S. Samuelsson, E. T. Hansen, M. Osler, and B. Jeune, "Prevention of secondary cases of meningococcal disease in Denmark," *Epidemiology and Infection*, vol. 124, no. 3, pp. 433–440, 2000.

[30] P. E. Alexander, L. Bero, V. M. Montori et al., "World Health Organization recommendations are often strong based on low confidence in effect estimates," *Journal of Clinical Epidemiology*, vol. 67, no. 6, pp. 629–634, 2014.

[31] J. C. Andrews, H. J. Schünemann, A. D. Oxman et al., "GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength," *Journal of Clinical Epidemiology*, vol. 66, no. 7, pp. 726–735, 2013.

[32] S. H. W. Vygen, W. Hellenbrand, P. Stefanoff, G. Hanquet, S. Heuberger, and J. M. Stuart, "European public health policies for managing contacts of invasive meningococcal disease cases better harmonized in 2013 than 2007," *Eurosurveillance*, In press.

[33] G. H. Guyatt, A. D. Oxman, R. Kunz et al., "GRADE guidelines 6. Rating the quality of evidence—imprecision," *Journal of Clinical Epidemiology*, vol. 64, no. 12, pp. 1283–1293, 2011.

[34] A. R. Jadad, D. J. Cook, A. Jones et al., "Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals," *Journal of the American Medical Association*, vol. 280, no. 3, pp. 278–280, 1998.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 160103, 11 pages
http://dx.doi.org/10.1155/2015/160103

Research Article

Intervention Mapping to Adapt Evidence-Based Interventions for Use in Practice: Increasing Mammography among African American Women

Linda Highfield,[1] Marieke A. Hartman,[2] Patricia Dolan Mullen,[2] Serena A. Rodriguez,[2] Maria E. Fernandez,[2] and L. Kay Bartholomew[2]

[1]*Department of Management, Policy and Community Health, School of Public Health, University of Texas, 1200 Pressler Street, Houston, TX 77030, USA*
[2]*Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas, 7000 Fannin Street, Houston, TX 77030, USA*

Correspondence should be addressed to Linda Highfield; linda.d.highfield@uth.tmc.edu

Received 23 June 2015; Revised 15 September 2015; Accepted 20 September 2015

Academic Editor: Giedrius Vanagas

This paper describes and demonstrates the use of the systematic planning process, Intervention Mapping, to adapt an evidence-based public health intervention (EBI). We used a simplified version of Intervention Mapping (IM Adapt) to increase an intervention's fit with a new setting and population. IM Adapt guides researchers and practitioners in selecting an EBI, making decisions about whether and what to adapt, and executing the adaptation while guarding the EBI's essential elements (those responsible for effectiveness). We present a case study of a project in which we used IM Adapt to find, adapt, implement, and evaluate an EBI to improve mammography adherence for African American women in a new practice setting in Houston, Texas. IM Adapt includes the following (1) assess needs and organizational capacity; (2) find EBIs; (3) plan adaptations based on fit assessments; (4) make adaptations; (5) plan for implementation; and (6) plan for evaluation of the adapted EBI. The case study shows an example of how public health researchers and practitioners can use the tool to make it easier to find and use EBIs, thus encouraging greater uptake. IM Adapt adds to existing dissemination and adaptation models by providing detailed guidance on how to decide on effective adaptation, while maintaining the essential elements of the EBI.

1. Background

Using evidence-based interventions (EBIs) to improve the health of the public improves the likelihood of program effectiveness and saves resources used in "reinventing the wheel" to address a particular health problem [1]. An evidence-based intervention (EBI) (including programs, policies, or practices) is one that has been shown to be effective through the application of sound scientific testing. Population impact on health is determined not only by the effectiveness of specific interventions, but also by how widely they are used. Governments have worked to improve opportunities for scale-up of EBIs. Nevertheless, uptake of EBIs is less than optimal and barriers to use are significant [2–4].

Challenges to using EBIs in practice include finding EBIs and their materials and deciding whether and how to adapt for a new setting. Planners need to assure a good match between the EBI and the new setting's capacity, health problem, context, and the at risk population [5, 6]. Furthermore, practitioners need to carefully consider whether to make changes in an EBI. Even small adaptations in the EBI are not trivial since adaptations may harm essential elements (also known as core elements or active ingredients) that made the EBI effective. Therefore, when adaptation is necessary to improve program fit, planners must determine not only whether or not a program works, but also which essential elements make the program successful [7, 8]. Unfortunately, program evaluations rarely report on which features of a program constitute these "essential elements." Because separate intervention elements are not usually tested independently, new users may not be able to identify and thereafter protect essential elements [9].

If planners find an EBI with reasonable fit, but decide adaptation is needed, a systematic approach can help them retain the balance between fidelity to original program design and adaptation to improve fit [7, 10–13]. For example, Lee and colleagues assess differences between the new population and the original population, execute content adaptation and pretesting, and plan an evaluation of an adapted EBI. Van Deale and colleagues [13] present a framework for a high level of community involvement to implement the essential elements of a program with fidelity while still allowing for adaptation to fit the needs of the new population or setting. The authors also advise planners to use Intervention Mapping as a way of identifying and articulating the essential elements of a program that should be maintained and implemented with fidelity. Intervention Mapping provides a systematic approach that adds detailed "how tos" to existing frameworks.

In this report, we describe Intervention Mapping to adapt EBIs for use in practice (IM Adapt) and present a case study that used Intervention Mapping to find, adapt, implement, and evaluate an EBI. We used this systematic approach in a community-based project to improve mammography rates for African American women in a mobile mammography practice setting in Houston, Texas [14]. Epidemiologic research has shown that African American women are less likely to use mammography screening [15–17] and more likely to miss scheduled mammography appointments [18] and to be diagnosed at a later stage of breast cancer [18] than their Caucasian counterparts.

2. Methods

2.1. Case Study. In the project we found, adapted, implemented, and evaluated an EBI to help underserved African American women in Houston, Texas, keep appointments for mammography screening. The needs assessment and basic program search are described elsewhere [19], as are the evaluation results [20]. The largest mobile mammography provider in the Houston area served as the new implementation setting.

2.2. IM Adapt: Intervention Mapping for Adaptation. For this community project we used a modified version of Intervention Mapping to guide the steps and tasks for adapting and implementing an evidence-based program [21–23]. Intervention Mapping is a systematic approach for developing theory- and evidence-based health promotion interventions that consists of six steps. Figure 1 presents a simplification of Intervention Mapping to help planners compare candidate EBIs to their community program needs and adapt when necessary [23, 24]. From an Intervention Mapping perspective, systematic adaptation requires that planners make adaptation decisions by comparing the logic of change in the EBI with the needs of the new community. Planners should only make changes that correspond with mismatches between the EBI and community needs.

The case study applied the steps of IM Adapt: (1) conduct a needs assessment and assess organizational capacity; (2) search for EBIs; (3) assess fit and plan adaptations; (4) make adaptations; (5) plan for implementation; and (6) plan for evaluation with a focus on changes to the EBI (see Figure 1).

Step 1 (conduct a needs assessment and assess organizational capacity). The first step of adapting an EBI following IM Adapt is to fully understand the health problem in the new site. The planning group completes four tasks: (1A) assess organizational capacity; (1B) conduct a needs assessment and develop a logic model of the problem; (1C) develop a logic model of change; and (1D) write program goals for expected outcomes from implementing the EBI at the new site. The logic model of the problem illustrates how risk behaviors and environmental factors are causally related to the health problem. Following the creation of the logic model of the problem, planners transit to a logic model of change to describe desired change. Planners show how theory-based change methods are proposed to influence first the determinants of behavior and environment, then the behavior and environmental factors, and finally the health problem and quality of life.

Step 2 (search for EBIs). The second step involves two tasks: (2A) search for an EBI and (2B) judge basic fit to identify interventions to review in more detail in Step 3. Basic fit is an initial assessment of how well an intervention tested in one setting might fit the needs and resources in another setting.

In general, planners would first be looking for an EBI described in a web-based database of interventions [25, 26]. Planners may want to review websites with evidence-based strategies derived from systematic reviews to gain a broad view of effective interventions. Since these usually provide descriptions of general approaches (e.g., one on one education), these may require more effort to obtain specific interventions and materials [27–29].

To judge basic fit of the EBIs identified, the planner considers whether the focus of the EBI matches the health problem, behaviors, environmental conditions, organizational resources, and characteristics of the population in the new setting or community. Planners may not need to reject an EBI immediately if the population for the original EBI is not the same as that in the new site. Different populations or subpopulations might have sufficient characteristics in common that carrying the EBI forward to the next step would still be worthwhile. Furthermore, it is strongly recommended to develop organizational capacity rather than cutting an EBI component.

Step 3 (assess fit and plan adaptations). With the materials for each candidate EBI in hand, the tasks for the third step are to (3A) judge how well the candidate EBI fits the desired behavioral and environmental conditions from the community's logic model of change; (3B) judge whether the determinants of behavior and environmental conditions and the change methods used to influence them in the original EBI are adequate in the new setting; (3C) judge how well original delivery, design features, and cultural elements fit the new setting and population; (3D) judge the fit of implementation

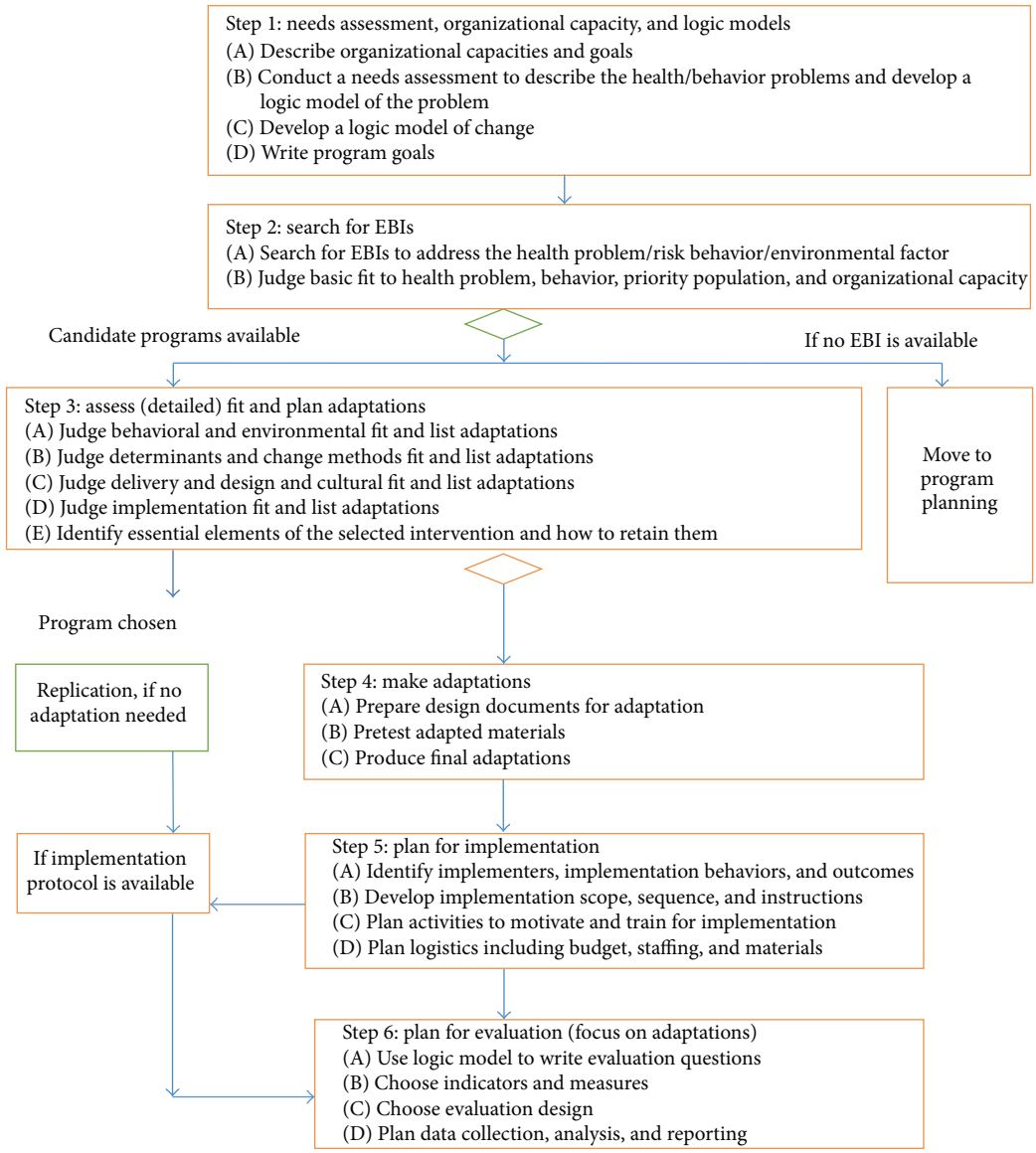


Figure 1: The IM Adapt Framework to adapt evidence-based interventions for use in practice.

strategies to the new setting; and (3E) consider which EBI elements are essential and decide how to retain them. If the original logic model of change and theoretical grounding are not published, planners have to work backwards from available intervention materials to try to figure out what change methods were used, which determinants were addressed, and which behaviors and environmental conditions were promoted. In addition, planners can contact the original EBI developers to obtain more information about the logic of the original EBI. Based on these assessments, planners can make a grounded selection of one EBI, and they will have a to do list of what to adapt and ideas about how to adapt. Planners should remember that less adaptation saves resources and protects an evidence-based intervention from changes that may make it less effective.

Step 4 (modify materials and activities). The tasks of the fourth step are the following: (4A) prepare design documents for the adaptations and drafting changes; (4B) pretest the adapted materials; and (4C) produce final adapted materials. Design documents provide detailed descriptions of planned changes, link the change to their location in the EBI materials, and provide an outline of messages. If proposed adaptations include the *addition* of behaviors, environmental conditions, determinants, and change methods, then planners can create a separate matrix of change objectives with the behavior or environmental condition being targeted and its determinants. This matrix arrangement is described in literature about Intervention Mapping [21, 24]. IM Adapt recommends planners make sure not to make unintended changes to the logic model of the EBI when making adaptations or revising adaptations based on pretest outcomes.

Step 5 (plan for implementation). The fifth step consists of the following tasks: (5A) identify implementers, implementation behaviors, and outcomes; (5B) develop implementation and

maintenance scope, sequence, and instructions; (5C) plan activities to motivate and train implementers; and (5D) plan logistics including budget, staffing, and materials. Planners compare the implementation protocol of the EBI, if available, to implementation considerations and constraints for the new site to create a revised protocol. Adaptations that resulted in modifications to the program components or delivery may require a different way to implement the modified EBI. The new implementation plan should include expected implementation outcomes—delivered to whom? when? how much?—and list the persons who will implement the program and how (required implementation steps or behaviors). Next, planners specify how much of the EBI will be implemented in what sequence over what period of time and write explicit instructions for new program implementers to bring that into action. Planners explore determinants of implementation and the change methods and practical applications that would influence them. Usually, for implementation, these change methods are woven into trainings, consultation, and technical support activities [30].

Step 6 (plan for evaluation). In the sixth step, planners (6A) write evaluation questions; (6B) choose indicators and measures; (6C) choose the evaluation design; and (6D) plan data collection, analysis, and reporting. The purpose of evaluating an adapted EBI is to determine whether the intervention achieves outcomes in the new setting comparable with outcomes in the original evaluation ("effect evaluation") and whether the new setting can successfully implement the adapted EBI (e.g., by measuring reach and fidelity). Evaluation questions for adapted EBIs can be borrowed from the original EBI evaluation. If the target behavior or environmental condition has been adapted, the indicators and measures must match the new logic model.

3. Results: Case Study

Step 1: Organizational Capacity, Needs Assessment, and Logic Models

(1A) Organizational Capacity. A local hospital-based charity organization (the Charities) initiated the planning for the project. The Charities proposed to find an evidence-based program to reduce the no-show rate for appointments made at mobile mammography sites with primarily African American women. The planning team included representatives from the lead agency research arm, breast cancer provider organizations, the local Breast Health Collaborative (an organization to establish linkages between organizations with breast health missions), and the local school of public health. All partners had missions that encompassed improving breast health in the Houston area. The researchers from the school of public health and the Charities also had a commitment to using and evaluating evidence-based programs.

(1B) Needs Assessment and Logic Model of the Problem. The planning group conducted a needs assessment to examine barriers to mammography screening and appointment keeping among African American women in Houston. We present a brief summary of the assessment outcomes here, with a detailed description of methods and results previously published [14].

Research suggests that recent reductions in breast cancer mortality are related to early detection (mammography) and enhanced cancer treatment [31]. However, African American women are less likely to schedule and attend mammography screening, with appointment no-show rates in some sites of 30–50% [15–18, 31–33]. Local women described the following barriers to appointment keeping: (1) fear of the outcome (it will be cancer); (2) competing demands (taking care of everyone but myself); (3) logistical barriers, such as insurance, cost, and transportation; (4) fear of partner abandonment if mastectomy results (loss of womanhood); (5) lack of education (nobody talks about mammography/breast cancer); (6) fear the mammogram would hurt; and (7) no need for a mammogram because their faith would protect them from cancer. Next, the planning group organized the data from the needs assessment in a logic model of the problem. The planning group focused on failure to keep mammography appointments. They then included all of the information from the community data collection as determinants of the lack of mammograms.

(1C-D) Logic Model of Change and Intervention Goal. Next, the planning group converted the logic model of the problem to a logic model of change to create the foundation for comparing EBIs to the intervention needs in the community (see Figure 2). The model focused on the behaviors of African American women because the environmental factors (i.e., access to treatment options and access to primary care) were not changeable in the scope of this project. The group worked from the list of local barriers to create categories of counter arguments to barriers by theoretical constructs. The barriers and counter arguments could be summarized with the following Social Cognitive Theory [34] constructs: knowledge, outcome expectations, modeling/vicarious reinforcement, skills, and self-efficacy. For example, *negative outcome expectations* such as "a diagnosis of breast cancer leads to death" could be countered by "early detection can lead to treatment and cure" and *low self-efficacy* such as "logistical problems of caring for others make mammography impossible for me to do" could be countered by "I can use the problem solving skills I use for other problems." These theoretical category labels did not replace the natural language used by the women to describe barriers. The group completed the logic model of change by adding theory- and evidence-based change methods that are suited to influencing outcome expectations, self-efficacy, and the other determinants with methods such as persuasion, role model stories, culturally congruent role models, and guided practice for problem solving. Based on Step 1, the group set the goal to decrease missed appointments of low-income African American women by 20% in the first year of program implementation.

Step 2: Finding an Evidence-Based Intervention

(2A) Searching for an EBI. To increase the number of possible interventions found, the planning group searched

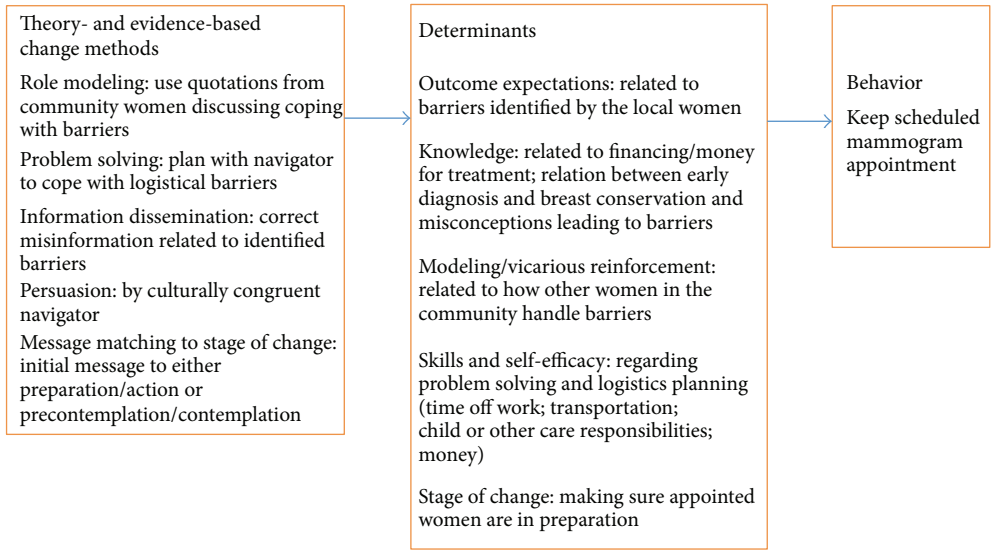


Figure 2: Case study logic model of change.

for EBIs focused on improving adherence to mammograms in African American women rather than more narrowly on appointment keeping. Prior to searching, they reviewed the Community Guide [28] to understand the types of strategies recommended to improve mammography screening. The team then searched for a full intervention (one with both description and available materials) using RTIPs [14, 26]. The group found four potential candidate programs [14]. The team then performed a second search of databases of peer reviewed studies to find reports of the original evaluations conducted on the EBIs they had located.

(2B) Assessing Basic Fit. Using the peer reviewed articles that described the programs and their evaluations, the planning group assessed basic fit: Was the health promoting behavior the same as for the new community? Could the organizational capacity support the program? Was the program acceptable for the risk group in the new community? All of the programs fit with the goal of encouraging mammography, but no programs focused explicitly on the behavior of appointment keeping. Therefore, the group recognized that the behavioral focus had to be specified no matter which of the four candidate programs they would work with [35–38]. All of the programs targeted either multiple ethnicities or African American women and any of them might be a basic fit to the priority population. No matter the final program chosen, the planning team would need to adapt it by integrating the information from African American women about local barriers, the way they feel about and talk about mammography, and their screening intentions. The planners selected a telephone counseling program after considering program fit with the implementation capacity of the clinical partner [35]. Other programs including community, church, and home visiting programs were outside of the Charities' mission, scope, and resources.

Step 3: Assess [Detailed] Fit and Plan Adaptations. To judge detailed fit and plan adaptations, the planning group obtained the program manual from the original developers (available on RTIPs). The manual contained instructions and scripts for the telephone counselors to address barriers. The team reviewed each type of fit and noted planned changes on an adaptation "to do list" (see Table 1).

(3A) Behavioral Fit. Referencing their logic model of change, the group described adaptations required to change the behavioral focus from getting a mammogram in general to appointment keeping for women who already had appointments scheduled.

(3B) Determinants and Change Methods. While judging the EBI's change methods and determinants, the planning group thought that several additions should be made to the scripts. The group members could discern how the original program guided the telephone counselor to ascertain a woman's stage of change, but they were unclear about how the counselors matched change methods to stages. Therefore, the planning group decided to assess only two stage categories (precontemplation/contemplation and preparation/action) as measured by women's certainty that they would keep their appointments and then match dialogue to the stage. For example, if a woman seemed unsure (precontemplation/contemplation), the telephone counselor would explore intensively for barriers. Additionally, the planners were unclear from the manual about which determinants were targeted and recommended that Social Cognitive Theory constructs of self-efficacy, skills, and outcome expectations be added with matching change methods of persuasion, cultural congruence, role modeling, and problem solving. Finally, the group found in the manual a comprehensive list of barriers, but barriers were addressed as if most beliefs could be remedied by provision of information.

Table 1: Adaptation "to do list" for telephone counseling program.

Breast Cancer Screening Among Nonadherent Women	
Fit category	Adaptation ideas
Behaviors from logic model of change Adherence to mammography	Change behavior to "appointment keeping" rather than general mammogram
Environmental conditions from logic model of change	No change
Change methods (with determinants) for at risk group: Information dissemination (barriers) Staging (but does not seem closely related to change methods beyond information)	Role modeling: quotations from women in the community regarding barriers Problem solving: regarding logistical barriers Information dissemination: correcting misinformation Persuasion: by culturally congruent navigator Message matching to stage of change: initial message to either preparation/action or precontemplation/contemplation
Change methods (with determinants) for environmental agents	Not applicable
Delivery for components, *at risk group*: Telephone counseling call. Note: does not have conversational structure comfortable for navigators to implement	Change staging question and scripts to be less research oriented and more "real-world" navigator approach Develop an active listening framework for barrier scripts Retain staging but with only two classifications (precontemplation/contemplation; preparation/action) Editing for local idiom (ways of speaking of breast cancer and barriers)
List delivery for components, environmental agents	Not applicable
List design features and cultural relevance Barriers are general and information-based	Add barriers described by local women: perceived likelihood of no cancer; expectation that God will protect against cancer; no money for treatment; becoming less than a woman with the loss of a breast; fear of losing partner; cancer being a death sentence; time only for caring for others Logistics: no time off work; no transportation; responsibilities caring for a child or others; money/lack of awareness of programs that can pay for breast cancer treatment
Describe implementation plan Research-based; elaborate staging; unclear how script changes based on staging; unclear transition between one barrier and the next	Add script with conversational transitions Add local barriers Simplify stages Match script to stage

The group noted that most belief change requires change methods beyond information (i.e., role models, persuasion, and guided practice).

(3C) Delivery Fit, Design Features, and Cultural Relevance. In this task, the planning group considered the acceptability of the EBI to the new population: whether the original delivery will reach the new population, how the design of program materials will resonate with the new participants, and how culturally congruent the entire program will feel to users. The planning group understood from the clinical partner that telephone reminders were an effective way to reach the priority population, so telephone delivery was acceptable. The group also judged that the scripts for barriers were up-to-date, accurate, and understandable but were not targeted to the exact concerns of local African American women or expressed in the ways that local women talked about their concerns. Furthermore, the original program did not outline an underlying communication approach; therefore the group agreed on an active listening framework for the scripts to maximize the connection of the navigator with the women through listening to their concerns and validating them [39].

(3D) Implementation Fit. Implementation fit is closely related to delivery, and the team made several adaptations to delivery as mentioned above. For instance, they decided to develop a conversational script to inquire about barriers in ways that fit with each stage category, which dealt with transitions in the conversation, and to enable the counselor to develop rapport with the women. However, implementation also has to do with how the implementing agency will manage the logistics of getting the EBI in place and maintaining it. The planning team worked with the clinical partner to plan staff placement and training for the telephone counseling. They also made sure that contact information was available for women with scheduled appointments and that data from each call and from appointment records could be recorded and accessed for the evaluation.

(3E) Considering Essential Elements. The planning group considered essential elements of the counseling program. The program was developed over a decade ago, and the developers were not available to answer questions. Therefore, the group independently considered the program to decide the program features that might have been essential to its effectiveness.

They listed the following characteristics: (1) barrier-focused counseling (change method), (2) telephone call delivered by a person (rather than a computer) (delivery), and (3) assessment of stage of change (prerequisite for matching change methods). Looking back at their adaptation "to do list," they made sure that their suggestions for change did not eliminate these essential program elements and sought only to enhance their intensity.

Step 4: Make Adaptations

(4A) Preparing Design Documents and Adaptation Drafts. The planning team noted the planned changes, described the program materials and activities in which the change should be made, and then wrote or edited messages that supported the change. For example, the team members proposed a foundation conversational structure based on active listening [39]. They then wrote the script to support the change. Another important adaptation was the addition of specific barriers as described by local African American women in the assessment and acknowledgement that local women had both described each barrier and also had described strategies for overcoming it (role model change method) (based on the assessment and feedback from the community advisors).

(4B) Pretesting Adapted Materials. Once the team adapted the program manual, it pretested the scripts with local African American women. Fourteen pretesters worked in pairs to role play the scripts with one woman as counselor and the other, as patient (caller). The pretesters noted needed changes in the scripts to make them as relevant as possible to local women. For example, they recommended taking care not to talk about a cancer diagnosis (and engender fear) in women simply being prepared to undergo screening. The pretesters then thoroughly debriefed the role plays with the entire project team. Pretesters also strongly recommended that the calls be made by a culturally congruent counselor.

(4C) Producing Final Adaptations. The team produced the revised manual of barriers and foundational conversation scripts organized in hard copy form. The hard copy format was used for the initial implementation and evaluation of the adapted EBI [40]. Following the initial evaluation study, the manual was converted to computer-assisted scripts for use by a live counselor [19].

Step 5: Plan for Implementation

(5A) Identifying Implementers, Behaviors, and Outcomes. The mammography team found the implementation protocol for the original program (prepared for research staff rather than patient counselors) to be focused on assessment of stage of change and barriers with little guidance to the implementer about how to transition from assessment to barriers and how to transition between barriers. The original manual was helpful, but it was not sufficient for this new site. Therefore the team identified implementers as patient navigators or community health workers familiar with making reminder calls and with identified implementation tasks or behaviors including (1) making standard reminder calls; (2) making protocol-driven, barrier-focused counseling calls for African American women already appointed for a mammogram; and (3) documenting the content of each call. Other implementers were the clinical partner managers who would provide space and access to appointment records for the navigator. The partner would also provide data on patient appointment attendance for the evaluation study. The desired implementation outcome was the completion of at least 100 EBI calls and 100 standard reminder calls in twelve months.

(5B) Developing Scope, Sequence, and Instructions. The scope of the adapted EBI was one call per woman. The sequence was seen as the sequence of the call to include assessment of stage, query regarding barriers, and solutions to barriers based on barrier scripts. The technique for moving the conversations forward was based on active listening [39].

(5C) Planning Activities to Motivate and Train Implementers. The mammography team developed training to encourage self-efficacy, outcome expectations, and skills of the implementers. The skills included opening, moving, and closing conversations, establishing rapport, conducting active listening, and addressing barriers. In the training sessions, we explained the theory behind the program but spent the majority of the time in the sessions conducting role play practice with feedback.

Step 6: Plan for Evaluation with a Focus on Adaptations. The evaluation plan sought to accomplish two aims: (1) determine the effectiveness of the adapted EBI in improving appointment keeping for mammography in African American women and (2) describe processes of implementation of an EBI in a practice setting. For evaluation results, see Highfield et al. in this issue [20].

(6A) Writing Evaluation Questions. We wrote the following evaluation questions for the effectiveness evaluation: (1) What was the effectiveness in decreasing appointment "no-show" rates in the new setting? (2) How did the effectiveness of the adapted EBI compare to the effectiveness of the original EBI? The questions for implementation evaluation (process) included the following: (1) Was the adapted EBI delivered to the intended population (i.e., low-income African American women with mobile mammography appointments)? (2) Did the implementers follow the protocol (i.e., implemented with fidelity?)? (3) What barriers were discussed in the phone calls? (4) Did the women who received the adapted EBI find it helpful and acceptable? (5) What problems occurred during implementation of the adapted EBI?

(6B) Choosing Indicators and Measures. To measure effectiveness, we obtained kept and missed appointments from the electronic database of the clinical partner. We also collected site of mammography, time between phone call and appointment, age, date, and time of appointment, counselor information, and contact information including phone number. We compared intervention phone calls to the protocol (whether the navigator asked the staging question, used the

barrier scripts, conducted logistical planning, and used active listening). In addition, we interviewed randomly selected intervention patients regarding their perceptions of the EBI calls and systems barriers encountered.

(6C) Choosing the Evaluation Design. We used the type-1 hybrid design to test the intervention's effectiveness and to gather information on the implementation [40, 41]. We used a quasi-experimental, sequential recruitment design in which we assigned contacted women to usual care or adapted intervention in groups of 50 patients.

(6D) Planning Data Collection, Analysis, and Reporting. We enrolled African American females who were aged between 35 and 64, uninsured, and had income of ≤200% of the federal poverty level (FPL) and with an upcoming appointment for a mobile screening mammogram. We tracked all data for the pilot either in an access database or in paper data collection forms. We calculated descriptive statistics and then conducted logistic regression analysis to report attendance in the intervention group as compared to the comparison group while controlling for potential confounders. Following the basic analysis, we further evaluated the effectiveness of the EBI using intent to treat analysis [41–46].

3.1. Project Outcomes and Current Status. The evaluation for this project is completed and the results have been used to acquire funding for a larger implementation of the adapted EBI [19, 20]. For the evaluation results, see Highfield et al. [20]. The effectiveness results were in the range of the results from the original intervention evaluation and indicated improved EBI effectiveness as a result of the systematic adaptation process [20, 35]. The implementation evaluation allowed us to discover problems in the initial implementation and correct them with a change in evaluation design and eventually in personnel.

4. Discussion

IM Adapt provides a systematic "how to" process to guide intervention adaptation, implementation, and evaluation. While we illustrate the approach using a mammography appointment attendance EBI, the processes used and described in this case study are widely applicable for the adaptation of any EBI. Intervention Mapping has been used worldwide to help planners develop and implement EBIs for a variety of public health problems including cancer prevention (cervical and breast), nutrition, parent education to reduce violence, and asthma, to name a few [47–52]. Other researchers and planners have used models for guiding EBI adaptation, such as Planned Adaptation [10]. For example, in their study of Planned Adaptation, Lee et al. noted that future studies should explicitly consider the roles of practitioners in participating in EBI translation. It has also been noted that, for cancer disparities in particular, EBIs should be adopted and tailored at the community level by partnerships that include both researchers and practitioners, adding an additional level of complexity to the translation of these EBIs [53]. IM Adapt specifically incorporates practitioners throughout the adaptation process and gives them a structured role through the planning team.

Additionally, Lee et al. [10] noted the need for models of adaptation to guide users through the process of developing appropriate evaluation and measurement plans. IM Adapt addresses this gap through having users consider evaluation questions that need to be addressed. Additionally, the IM Adapt process guides planners to incorporate both effectiveness and process evaluation. Curran et al. [40] have suggested the need for blended EBI effectiveness and implementation trials that consider both effectiveness outcomes and implementation process [40]. Blended effectiveness/implementation trials are a key area of study which are currently underreported [2, 40, 54]. As noted by Wandersman et al., 2008 [55], publication of these kinds of studies is necessary to guide future efforts to disseminate EBIs into practice. Practitioners lack sufficient insight into this process and as a result are forced to make decisions based on limited information [24].

4.1. Limitations of the IM Adapt Process and of the Case Study Project. IM Adapt helps planners compare their local health problem and needs to available EBIs to judge fit of the intervention in terms of the health problem and its causes and then to continue the comparison to the logic model of change. This is a fairly complex process that is made more difficult by inadequate reports of EBIs in the scientific literature. Often original investigators neither publish their logic model of the problem (evidence and assumptions about the problem at the developing site of the EBI) nor explicitly address the logic model (theory) of change. This situation leaves the adopter of an EBI sometimes peering into a black box and making guesses about what change methods are contained in a program, what determinants they were meant to influence, and whether original investigators consider them to be essential program elements that should not be adapted. Improving intervention reporting would significantly improve the ability of those who want to use EBIs to choose one, decide whether adaptations are advisable, and use a systematic process to carry them out if they are needed. We applied the IM Adapt process to a case study of an adapted mammography appointment adherence EBI. Future studies are necessarily applying IM Adapt to a variety of EBIs and public health issues to allow for continued evaluation and refinement of the approach.

5. Conclusion

In this report, we have presented a case study of a community project for which we used the IM Adapt framework to find, adapt, implement, and evaluate an EBI to help underserved African American women in Houston, Texas, keep appointments for mammography screening. IM Adapt should be useful for planners who are considering EBIs to avoid developing an intervention from the beginning. Not wanting to develop an intervention de novo can be from awareness of insufficient resources for developing and evaluating a theory- and evidence-based intervention or because a funder has

required the use of an existing EBI. If a planning group is able to find an EBI that addresses its priority health problem, it will face a core question, does this program fit with our community and with the characteristics of the health problem in the new setting and can it be adapted so that it better fits and still works? IM Adapt is a guide for answering this question and performing a systematic adaptation.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to acknowledge the contributions of our community partners The Rose, The Breast Health Collaborative of Texas, and St. Luke's Episcopal Health Charities throughout the steps of intervention adaptation. The authors would also like to acknowledge the Avon Foundation for funding this study (Project nos. 05-2010-004 and 05-2011-008). Marieke A. Hartman was supported by the postdoctoral fellowship, University of Texas School of Public Health Cancer Education and Career Development Program, National Cancer Institute (NIH Grant R25 CA57712).

References

[1] R. C. Brownson, J. E. Fielding, and C. M. Maylahn, "Evidence-based public health: a fundamental concept for public health practice," *Annual Review of Public Health*, vol. 30, pp. 175–201, 2009.

[2] R. E. Glasgow, E. Lichtenstein, and A. C. Marcus, "Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition," *American Journal of Public Health*, vol. 93, no. 8, pp. 1261–1267, 2003.

[3] L. N. Krivitsky, S. J. Parker, A. Pal, L. Meckler, R. Shengelia, and M. C. Reid, "A systematic review of health promotion and disease prevention program adaptations: how are programs adapted?" in *Research for the Public Good: Applying the Methods of Translational Research to Improve Health and Well-Being*, E. Wethington and R. E. Dunifon, Eds., pp. 73–99, American Psychological Association, Washington, DC, USA, 2012.

[4] V. S. McKleroy, J. S. Galbraith, B. Cummings et al., "Adapting evidence-based behavioral interventions for new settings and target populations," *AIDS Education & Prevention*, vol. 18, no. supplement, pp. 59–73, 2006.

[5] S. F. Mihalic, A. A. Fagan, and S. Argamaso, "Implementing the LifeSkills Training drug prevention program: factors related to implementation fidelity," *Implementation Science*, vol. 3, article 5, 16 pages, 2008.

[6] Selecting the program that's right for you: a feasibility assessment tool, http://hhd.org/resources/assessmenttools/selecting-program-s-right-you-feasibility-assessment-tool.

[7] D. S. Elliott and S. Mihalic, "Issues in disseminating and replicating effective prevention programs," *Prevention Science*, vol. 5, no. 1, pp. 47–53, 2004.

[8] T. E. Backer, *Finding the Balance: Program Fidelity and Adaptation in Substance Abuse Prevention: A State-of-the-Art Review*, Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Prevention, Rockville, Md, USA, 2001.

[9] G. J. Botvin, "Advancing prevention science and practice: challenges, critical issues, and future directions," *Prevention Science*, vol. 5, no. 1, pp. 69–72, 2004.

[10] S. J. Lee, I. Altschul, and C. T. Mowbray, "Using planned adaptation to implement evidence-based programs with new populations," *American Journal of Community Psychology*, vol. 41, no. 3-4, pp. 290–303, 2008.

[11] J. J. Card, J. Solomon, and S. D. Cunningham, "How to adapt effective programs for use in new contexts," *Health Promotion Practice*, vol. 12, no. 1, pp. 25–35, 2011.

[12] J. Shen, H. Yang, H. Cao, and C. Warfield, "The fidelity-adaptation relationship in non-evidence-based programs and its implication for program evaluation," *Evaluation*, vol. 14, no. 4, pp. 467–481, 2008.

[13] T. van Daele, C. van Audenhove, D. Hermans, O. van den Bergh, and S. van den Broucke, "Empowerment implementation: enhancing fidelity and adaptation in a psycho-educational intervention," *Health Promotion International*, vol. 29, no. 2, pp. 212–222, 2014.

[14] L. Highfield, L. K. Bartholomew, M. A. Hartman, M. M. Ford, and P. Balihe, "Grounding evidence-based approaches to cancer prevention in the community: a case study of mammography barriers in underserved African American women," *Health Promotion Practice*, vol. 15, no. 6, pp. 904–914, 2014.

[15] J. Legler, H. I. Meissner, C. Coyne, N. Breen, V. Chollette, and B. K. Rimer, "The effectiveness of interventions to promote mammography among women with historically lower rates of screening," *Cancer Epidemiology Biomarkers and Prevention*, vol. 11, no. 1, pp. 59–71, 2002.

[16] S. R. Crump, R. M. Mayberry, B. D. Taylor, K. P. Barefield, and P. E. Thomas, "Factors related to noncompliance with screening mammogram appointments among low-income African-American women," *Journal of the National Medical Association*, vol. 92, no. 5, pp. 237–246, 2000.

[17] K. M. Schueler, P. W. Chu, and R. Smith-Bindman, "Factors associated with mammography utilization: a systematic quantitative review of the literature," *Journal of Women's Health*, vol. 17, no. 9, pp. 1477–1498, 2008.

[18] R. Smith-Bindman, D. L. Miglioretti, N. Lurie et al., "Does utilization of screening mammography explain racial and ethnic differences in breast cancer?" *Annals of Internal Medicine*, vol. 144, no. 8, pp. 541–553, 2006.

[19] L. Highfield, *Evidence-Based Dissemination for Mammography Adherence in Safety Net Communities*, 2014.

[20] L. Highfield, M. A. Hartman, L. K. Bartholomew, P. Balihe, and V. M. Ausborn, "Evaluation of the effectiveness and implementation of an adapted evidence-based mammography intervention for African American women," *BioMed Research International*, In press.

[21] L. K. Bartholomew and P. D. Mullen, "Five roles for using theory and evidence in the design and testing of behavior change interventions," *Journal of Public Health Dentistry*, vol. 71, supplement 1, pp. S20–S33, 2011.

[22] J. N. Leerlooijer, R. A. C. Ruiter, J. Reinders, W. Darwisyah, G. Kok, and L. K. Bartholomew, "The world starts with me: using intervention mapping for the systematic adaptation and transfer of school-based sexuality education from Uganda to Indonesia," *Translational Behavioral Medicine*, vol. 1, no. 2, pp. 331–340, 2011.

[23] S. R. Tortolero, C. M. Markham, G. S. Parcel et al., "Using intervention mapping to adapt an effective HIV, sexually transmitted disease, and pregnancy prevention program for high-risk minority youth," *Health Promotion Practice*, vol. 6, no. 3, pp. 286–298, 2005.

[24] L. K. Bartholomew, G. S. Parcel, G. Kok, N. H. Gottlieb, and M. E. Fernàndez, *Planning Health Promotion Programs: An Intervention Mapping Approach*, Jossey-Bass, San Francisco, Calif, USA, 3rd edition, 2011.

[25] Center for Training and Research Translation (Center TRT), http://www.centertrt.org/.

[26] Research-tested intervention programs (RTIPs): moving science into programs for people, http://rtips.cancer.gov/rtips/index.do.

[27] The Guide to Community Preventive Services, http://www.thecommunityguide.org/index.html.

[28] Model Practice Database, http://www.naccho.org/topics/model-practices/search.cfm.

[29] Trip database, http://tripdatabase.com/.

[30] A. Wandersman, V. H. Chien, and J. Katz, "Toward an evidence-based system for innovation support for implementing innovations with quality: tools, training, technical assistance, and quality assurance/quality improvement," *American Journal of Community Psychology*, vol. 50, no. 3-4, pp. 445–459, 2012.

[31] K. P. Williams, V. B. Sheppard, D. Todem, A. Mabiso, J. T. Wulu Jr., and R. D. Hines, "Family matters in mammography screening among African-American women age >40," *Journal of the National Medical Association*, vol. 100, no. 5, pp. 508–515, 2008.

[32] I. Menashe, W. F. Anderson, I. Jatoi, and P. S. Rosenberg, "Underlying causes of the black-white racial disparity in breast cancer mortality: a population-based analysis," *Journal of the National Cancer Institute*, vol. 101, no. 14, pp. 993–1000, 2009.

[33] K. L. Margolis, N. Lurie, P. G. McGovern, and J. S. Slater, "Predictors of failure to attend scheduled mammography appointments at a public teaching hospital," *Journal of General Internal Medicine*, vol. 8, no. 11, pp. 602–605, 1993.

[34] A. Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1986.

[35] I. M. Lipkus, B. K. Rimer, S. Halabi, and T. S. Strigo, "Can tailored interventions increase mammography use among HMO women?" *American Journal of Preventive Medicine*, vol. 18, no. 1, pp. 1–10, 2000.

[36] E. D. Paskett, C. M. Tatum, R. D'Agostino Jr. et al., "Community-based interventions to improve breast and cervical cancer screening: results of the Forsyth County Cancer Screening (FoCaS) project," *Cancer Epidemiology Biomarkers and Prevention*, vol. 8, no. 5, pp. 453–459, 1999.

[37] J. S. Slater, C. Nim Ha, M. E. Malone et al., "A randomized community trial to increase mammography utilization among low-income women living in public housing," *Preventive Medicine*, vol. 27, no. 6, pp. 862–870, 1998.

[38] D. O. Erwin, T. S. Spatz, R. C. Stotts, and J. A. Hollenberg, "Increasing mammography practice by African American women," *Cancer Practice*, vol. 7, no. 2, pp. 78–85, 1999.

[39] M. B. Wanzer, M. Booth-Butterfield, and K. Gruber, "Perceptions of health care providers' communication: relationships between patient-centered communication and satisfaction," *Health Communication*, vol. 16, no. 3, pp. 363–383, 2004.

[40] G. M. Curran, M. Bauer, B. Mittman, J. M. Pyne, and C. Stetler, "Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact," *Medical Care*, vol. 50, no. 3, pp. 217–226, 2012.

[41] A. C. Bernet, D. E. Willens, and M. S. Bauer, "Effectiveness-implementation hybrid designs: implications for quality improvement science," *Implementation Science*, vol. 8, supplement 1, article S2, 2013.

[42] S. K. Gupta, "Intention-to-treat concept: a review," *Perspectives in Clinical Research*, vol. 2, no. 3, pp. 109–112, 2011.

[43] D. J. Newell, "Intention-to-treat analysis: implications for quantitative and qualitative research," *International Journal of Epidemiology*, vol. 21, no. 5, pp. 837–841, 1992.

[44] R. T. Wertz, "Intention to treat: once randomized, always analyzed," *Clinical Aphasiology*, vol. 23, pp. 57–64, 1995.

[45] S. R. Heritier, V. J. Gebski, and A. C. Keech, "Inclusion of patients in clinical trial analysis: the intention-to-treat principle," *Medical Journal of Australia*, vol. 179, no. 8, pp. 438–440, 2003.

[46] M. P. LaValley, "Intent-to-treat analysis of randomized clinical trials," in *Proceedings of the ACR/ARHP Annual Scientific Meeting*, Orlando, Fla, USA, October 2003.

[47] T. L. Byrd, K. M. Wilson, J. L. Smith et al., "Using intervention mapping as a participatory strategy: development of a cervical cancer screening intervention for Hispanic women," *Health Education and Behavior*, vol. 39, no. 5, pp. 603–611, 2012.

[48] S.-I. Hou, M. E. Fernandez, E. Baumler, and G. S. Parcel, "Effectiveness of an intervention to increase pap test screening among Chinese women in Taiwan," *Journal of Community Health*, vol. 27, no. 4, pp. 277–290, 2002.

[49] I. C. Scarinci, L. Bandura, B. Hidalgo, and A. Cherrington, "Development of a theory-based (pen-3 and health belief model), culturally relevant intervention on cervical cancer prevention among latina immigrants using intervention mapping," *Health Promotion Practice*, vol. 13, no. 1, pp. 29–40, 2012.

[50] D. M. Hoelscher, A. Evans, G. S. Parcel, and S. H. Kelder, "Designing effective nutrition interventions for adolescents," *Journal of the American Dietetic Association*, vol. 102, no. 3, pp. S52–S63, 2002.

[51] C. Markham, S. Tyrrell, R. Shegog, M. E. Fernandez, and L. K. Bartholomew, "Partners in school asthma program," in *Intervention Mapping: Designing Theory- and Evidence-Based Health Promotion Programs*, L. K. Bartholomew, G. S. Parcel, G. Kok, and N. H. Gottlieb, Eds., pp. 387–424, Mayfield, Thousand Oaks, Calif, USA, 2001.

[52] N. Murray, S. Kelder, D. Hoelscher, J. Conroy, P. Cribb, and G. Parcel, "Diffusion of the child and adolescent trial for cardiovascular health (CATCH) program in Texas," in *Proceedings of the 14th National Conference on Chronic Disease Prevention and Control*, Dallas, Tex, USA, 1999.

[53] S. M. Ahmed and A.-G. S. Palermo, "Community engagement in research: frameworks for education and peer review," *American Journal of Public Health*, vol. 100, no. 8, pp. 1380–1387, 2010.

[54] S. R. Tunis, D. B. Stryer, and C. M. Clancy, "Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy," *Journal of the American Medical Association*, vol. 290, no. 12, pp. 1624–1632, 2003.

[55] A. Wandersman, J. Duffy, P. Flaspohler et al., "Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation," *American Journal of Community Psychology*, vol. 41, no. 3-4, pp. 171–181, 2008.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 704539, 17 pages
http://dx.doi.org/10.1155/2015/704539

Review Article

Exercise and BMI in Overweight and Obese Children and Adolescents: A Systematic Review and Trial Sequential Meta-Analysis

George A. Kelley,[1] Kristi S. Kelley,[1] and Russell R. Pate[2]

[1]*Department of Biostatistics, School of Public Health, West Virginia University, Morgantown, WV 26506, USA*
[2]*Department of Exercise Science, University of South Carolina, Columbia, SC 29208, USA*

Correspondence should be addressed to George A. Kelley; gkelley@hsc.wvu.edu

Received 13 June 2015; Revised 11 September 2015; Accepted 13 September 2015

Academic Editor: Giedrius Vanagas

Objective. Determine the effects of exercise on body mass index (BMI in kg·m^{-2}) among overweight and obese children and adolescents. *Methods*. Trial sequential meta-analysis of randomized controlled exercise intervention trials ≥ 4 weeks and published up to November 11, 2014. *Results*. Of the 5,436 citations screened, 20 studies representing 971 boys and girls were included. Average length, frequency, and duration of training were 13 weeks, 3 times per week, for 46 minutes per session. Overall, random-effects models showed that exercise decreased BMI by 3.6% (mean: −1.08; 95% CI: −0.52 to −1.64; $Q = 231.4$; $p < 0.001$; $I^2 = 90.9\%$; 95% CI: 87.6% to 93.4%; $D^2 = 91.5\%$). Trial sequential meta-analysis showed that changes in BMI crossed the monitoring boundary for a type 1 error in 2010, remaining stable thereafter. The number needed to treat was 5 while the percentile improvement was 26.9. It was estimated that approximately 2.5 million overweight and obese children in the US and 22.0 million overweight and obese children worldwide could reduce their BMI by participating in a regular exercise program. Overall quality of evidence was rated as moderate. *Conclusions*. Exercise is associated with improvements in BMI among overweight and obese children and adolescents. This trial is registered with PROSPERO Trial Registration #CRD42015017586.

1. Introduction

The prevalence of overweight and obesity in children and adolescents is a pandemic problem both in the United States (US) and worldwide. Recently, Ogden et al. reported that the prevalence of overweight and obesity in the US, defined as a body mass index (BMI) in kg·m^{-2} ≥ 85th percentile based on Centers for Disease Control Growth Charts, was 31.8% among children and adolescents 2 to 19 years of age, while the prevalence of obesity, defined as a BMI in kg·m^{-2} ≥ 95th percentile, was 16.9% [1]. When compared to 30 years ago, this represents an obesity prevalence that is more than two times higher in US children and more than four times higher in adolescents [1, 2]. From a worldwide perspective, the prevalence of overweight and obesity in 2013 has been reported to be approximately 23% among children and adolescents in developed countries and 13% among children and adolescents from developing countries [3]. Collectively, this represents an approximate 47% increase in the worldwide prevalence of overweight and obesity among children and adolescents between 1980 and 2013 [3].

The economic costs associated with overweight and obesity among children and adolescents are also substantial. For example, Finkelstein et al. estimated that the incremental lifetime medical cost of an obese 10-year-old child in the US, in relation to a normal weight child who maintained normal weight throughout adulthood, was $19,000 [4]. Based on the current number of obese 10-year-olds in the US, the total direct medical costs associated with obesity were estimated at $14 billion for this age only [4].

The negative health consequences of obesity in children and adolescents are both immediate and long-term. For example, in a population-based sample of US children and adolescents 5 to 17 years of age from the Bogalusa Heart Study,

approximately 70% of obese youth had at least one cardiovascular disease risk factor [5]. In addition, obese children and adolescents, in relation to their normal weight peers, suffer from a greater prevalence and/or incidence of other conditions that include, but are not necessarily limited to, musculoskeletal pain, injuries and fractures [6], obstructive sleep apnea [7], and poorer self-esteem and quality of life [8]. From a long-term perspective, overweight and obesity during childhood and adolescence have been shown to track into adulthood [9], thereby placing this population group at an increased risk for premature all-cause mortality [10]. This is a major problem since overweight and obesity have been reported to be the third leading cause of preventable death in the US, responsible for 216,000 deaths in 2005 [10]. Globally, the World Health Organization has estimated that approximately 3.4 million adults die each year as a result of being overweight or obese [11]. The issue of obesity has become so problematic that it is now recognized by the American Medical Association as a disease [12].

Exercise has been recommended for the prevention and treatment of overweight and obesity in children and adolescents [13–18]. In a recent systematic review with meta-analysis of studies published until the year 2012, the investigative team reported a statistically significant decrease of approximately 3% in BMI z-score in overweight and obese children and adolescents [19]. However, body mass index BMI in $kg \cdot m^{-2}$ continues to be the most commonly assessed and reported metric and is easily recognized and interpreted by practitioners. Unfortunately, the effects of exercise on BMI in $kg \cdot m^{-2}$ have been underwhelming. For example, with the exception of one previous systematic review with meta-analysis that focused on exercise [20], others reported a nonsignificant decrease in BMI in $kg \cdot m^{-2}$ among children and adolescents [17, 21–23]. However, all five suffer from potential limitations. These include (1) the pooling of a small number of exercise-only studies [17, 21, 22], (2) the inclusion of nonrandomized trials [20, 22], (3) inclusion of children and adolescents who were not overweight or obese [20, 22, 23], and (4) overall quality scores ranging from only 45% to 82% when the Assessment of Multiple Systematic Reviews (AMSTAR) instrument was applied to the studies [24]. In addition, none of the studies used trial sequential analysis, an approach that can provide data regarding (1) adequate information size, (2) a threshold for a statistically significant effect, and (3) a threshold for futility [25]. Given the former, the purpose of the current study was to conduct a systematic review and trial sequential meta-analysis of randomized controlled trials addressing the *overall* effects of exercise (aerobic training, strength training, or both) on BMI in $kg \cdot m^{-2}$ among overweight and obese children and adolescents.

2. Methods

2.1. Registration and General Procedure. This systematic review with trial sequential meta-analysis is registered in PROSPERO (#CRD42015017586), an international prospective registry of systematic reviews. The conduct and reporting of this study was accomplished according to the general guidelines recommended by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [26].

2.2. Study Eligibility. The *a priori* inclusion criteria for this study were as follows: (1) randomized controlled trials (assignment at participant level only), (2) control group (nonintervention, usual care, wait-list control, and attention control), (3) exercise (aerobic training, strength training, or both) $\geq$ 4 weeks as an independent intervention, (4) overweight and obese children and adolescents, as defined by the authors, (5) boys and/or girls 2 to 18 years of age, (6) studies published in full in any language between January 1, 1990, and November, 11, 2014, and (7) data available for calculating changes in BMI in $kg \cdot m^{-2}$. Studies were excluded based on an inappropriate population, intervention, comparison, outcome, study type, or lack of requisite data for BMI in $kg \cdot m^{-2}$.

2.3. Data Sources. The following databases were searched from January 1, 1990, to December 31, 2012: (1) Academic Search Complete, (2) CINAHL, (3) Cochrane Central Register of Controlled Trials (CENTRAL), (4) Education Research Complete, (5) ERIC, (6) LILACS, (7) Medline, (8) Proquest, (9) Scopus, (10) Sport Discus, and (11) Web of Science. In addition, an updated PubMed search was conducted for potentially eligible studies published between August 1, 2012, and November 11, 2014. A brief description of each database is shown in Supplementary File 1 (see Supplementary Material available online at http://dx.doi.org/10.1155/2015/704539) while the updated search strategy for PubMed can be found in Supplementary File 2. Database searches were supplemented by cross-referencing for potentially eligible studies, including reviews, as well as expert review by the third author. All studies were stored in Reference Manager, version 12.0 [27]. Overall precision of the searches was computed by dividing the number of studies included by the total number of studies screened while the number needed to read (NNR) was calculated as the inverse of the precision [28].

2.4. Study Selection. Independent, dual-selection of eligible studies was conducted by the first two authors who then met and reviewed their choice for inclusion. Disagreements were resolved by consensus and, if necessary, consultation with the third author.

2.5. Data Abstraction. Codebooks were developed in an electronic spreadsheet program [29] that included items that fell within the following four major categories: (1) study characteristics, (2) physical characteristics of participants, (3) training program characteristics, and (4) outcomes and outcome characteristics. Independent, dual-selection of eligible studies was conducted by the first two authors who then met and reviewed their choice for inclusion. Disagreements were resolved by consensus and, if necessary, consultation with the third author. Using Cohen's kappa statistics (κ) [30], the

overall agreement rate prior to correcting discrepancies was 0.94.

2.6. Risk of Bias Assessment. The Cochrane Risk of Bias Assessment Instrument was used to assess potential risk of bias [31]. Items were rated as being at a high, low, or unclear risk for bias [31]. In addition to the six basic items, an additional item about whether the participants were exercising regularly prior to study participation, as defined by the original study authors, was included [31]. Assessment for risk of bias was limited to the primary outcome, BMI in $kg{\cdot}m^{-2}$. Dual and independent assessment for risk of bias was conducted by the first two authors who subsequently met and reviewed every item for agreement. Disagreements were resolved by consensus and, if necessary, consultation with the third author. Using Cohen's kappa statistics (κ) [30], the overall agreement rate prior to correcting discrepancies was 0.70.

2.7. Data Synthesis

2.7.1. Calculation of Effect Sizes for BMI in $kg{\cdot}m^{-2}$. The primary outcome for this study was changes in BMI in $kg{\cdot}m^{-2}$. Secondary outcomes included body weight, percent body fat, fat mass, fat-free mass, changes in maximum oxygen consumption in $mL{\cdot}kg^{-1}\,min^{-1}$ (VO_{2max}), and upper and lower body strength and kilocalorie intake. Effect sizes (ES) using the original metrics were calculated by subtracting the change score difference in the exercise group from the change score difference in the control group. Variances were calculated from the pooled standard deviations of change scores in the exercise and control groups. If change score standard deviations were not reported, they were calculated from pre- and poststandard deviations according to procedures developed by Follmann et al. [32]. Each ES was then weighted by the inverse of its variance.

2.7.2. Pooled Estimates for Changes in BMI in $kg{\cdot}m^{-2}$. Changes in BMI in $kg{\cdot}m^{-2}$ and all secondary outcomes were pooled using random-effects, method-of-moments models that incorporate between-study heterogeneity into the final estimate [33]. Ninety-five percent confidence intervals (CI) were calculated while z-based two-tailed alpha values ≤ 0.05 were considered statistically significant. Heterogeneity was examined using the Q statistics [34], with an alpha value ≤ 0.10 representative of statistically significant heterogeneity. Inconsistency was examined using I^2 [35] and diversity using D^2 [36]. For both I^2 and D^2 values < 25%, 25% to <50%, 50% to <75%, and 75% or greater were considered to represent very low, low, moderate, and large amounts of inconsistency and diversity [37]. Statistically significant outliers were considered to be those with standardized residual alpha values ≤ 0.05. Multiple exercise groups in the same study were analyzed independently as well as collapsing multiple groups so that only one ES represented each study while the sample size for the control group was divided by the number of exercise groups [38]. In addition to 95% CI, 95% prediction intervals (PI) were also calculated [39, 40] for any result that was statistically significant. Based on recent recommendations [41], small-study effects (publication bias, etc.) were examined both qualitatively and quantitatively using funnel plots and Egger's regression intercept test [42]. A one-tailed probability value < 0.05 was considered to be indicative of statistically significant small-study effects. The influence of each result on the overall findings was examined by deleting each result from the model once. *Post hoc*, the fail-safe N test was used to estimate the number of studies that would be needed to reverse our finding of a statistically significant, that is, $p < 0.05$, improvement in BMI in $kg{\cdot}m^{-2}$ [43]. This test was used because four studies that met all of our inclusion criteria except for the provision of sufficient postintervention data were excluded from the meta-analysis.

To enhance practical application, the number needed to treat (NNT) was calculated for changes in BMI in $kg{\cdot}m^{-2}$ assuming a conservative control group risk of 10% and only if changes in BMI in $kg{\cdot}m^{-2}$ were statistically significant. If the NNT was calculated, gross estimates were determined for the number of obese children and adolescents in the US as well as worldwide that could potentially benefit from exercise. These estimates were based on 12.5 [1] and 110 million [44, 45] overweight and obese children in the US and worldwide, respectively. In addition to NNT, Cohen's U_3 index, an index used to determine the percentile gain in an intervention group, was calculated for any statistically significant results with respect to BMI in $kg{\cdot}m^{-2}$ and secondary outcomes [46]. Finally, The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) instrument was used to assess the overall quality of evidence and was limited to the primary outcome, BMI in $kg{\cdot}m^{-2}$ [47]. Overall quality was categorized as very low, low, moderate, or high [47].

Based on empirical evidence that consideration of information size and adjusted significance thresholds may avoid false statistical inferences due to imprecision and repeated significance testing in meta-analysis [25, 48–50], information size estimates and trial sequential analysis were performed [51] for BMI in $kg{\cdot}m^{-2}$. Trial sequential analysis is an approach that combines conventional meta-analysis methodology with meta-analytic sample size considerations as well as previously established methods for repeated significance testing on accruing data in randomized trials [51]. Inferences derived from using trial sequential analysis may be more reliable than using conventional meta-analysis procedures [51]. More specifically, previous research suggests that information size considerations as well as adjusted significance thresholds may eliminate early false positive findings due to a lack of precision and repeated significance testing in meta-analyses [25, 48–51].

The *a priori* plan was to estimate the required information size based on previous research suggesting that a $0.1\,kg/m^2$ change in BMI in $kg{\cdot}m^{-2}$ can be clinically important [52]. However, because of the inability to obtain variance statistics, a *post hoc* decision was made to estimate the required information size using the pooled mean difference and variance, adjusted for between-study heterogeneity, from the current study. A two-tailed type 1 error rate of 5% and power of 80% were employed. To control for multiple tests, trial sequential

monitoring boundaries for both type 1 (5%) and type 2 (20%) error rates were established using O'Brien-Fleming adjustments [53, 54].

2.7.3. Metaregression Analysis. Simple, random-effects metaregression (method of moments) models were used to examine associations between changes in BMI in kg·m^{-2} and potential predictors [33]. An *a priori* decision was made to not conduct any type of multiple metaregression analyses because of missing data for different variables from different studies. Metaregression analysis was limited to those studies in which there were at least four results for continuous variables or four results per group for categorical variables. Continuous variables, determined *a priori*, included year of publication, percent dropout, age, baseline BMI in kg·m^{-2}, and exercise intervention (length, frequency, duration, compliance, minutes per week, unadjusted and adjusted for compliance, and total minutes for the intervention, unadjusted and adjusted for compliance). Categorical variables examined included country, type of control group, funding, *a priori* sample size estimates, adverse events, risk of bias (sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and whether subjects were inactive prior to enrollment), gender, race/ethnicity, changes in exercise and/or physical activity outside the exercise intervention, pubertal stage, type of exercise (aerobic, strength, and both), exercise supervision, setting that exercise took place, type of participation, type of analysis, and exercise intensity (low, moderate, and high) [55].

2.7.4. Reporting and Software Utilization. Changes in primary and secondary outcomes are reported in their natural direction of benefit, that is, negative values for changes in BMI in kg·m^{-2} and positive values for increases in fat-free mass. All statistical analyses were conducted using Comprehensive Meta-Analysis (version 3.3) [56], Microsoft Excel 2010 [57], Trial Sequential Analysis (version 0.9) [51], GradePro (version 3.6) [58], and two add-ins for Microsoft Excel, SSC-stat (version 2.18) [59] and EZ-Analyze (version 3.0) [60].

3. Results

3.1. Characteristics of Included Studies. After removing duplicates, a total of 5,436 articles were screened. Of these, 20 studies representing 42 groups (22 exercise, 20 control) and final assessment of BMI in kg·m^{-2} in 971 participants (575 exercise, 396 control) met all eligibility criteria [61–80]. Overall precision of the searches was 0.004 while the NNR was 272. The major reasons for exclusion were inappropriate study design (51.6%), intervention (31.2%), population (14.0%), comparison (2.3%), and outcome(s) (0.9%). Another four studies comprising less than 1% of the reasons for exclusion were omitted because data necessary for conducting trial sequential meta-analysis were not available [81–84]. This included (1) lack of both post and change outcome values as well as standard deviations, or data for conversion to standard deviations (e.g., standard error of the mean), for BMI in kg·m^{-2} [81–83] and (2) lack of separate sample sizes for exercise and control groups [84]. A flow diagram of the search process is shown in Figure 1 while a list of excluded studies, including the specific reason(s) for exclusion, can be found in Supplementary File 3.

Table 1 describes the characteristics of each included study. Studies were conducted in 12 different countries and published between 2004 and 2014 [61–80], with all but one [62] published in English-language journals. Assessment of primary and secondary outcomes took place after six [68, 69], eight [63, 67, 77–79], 10 [76], 12 [61, 64, 66, 71, 72, 75, 80], 13 [65], 16 [73], 22 [74], and 24 [62, 70] weeks of exercise.

Two studies used some type of matching procedure, one according to age, gender, and BMI in kg·m^{-2} [67] and another according to sex and degree of overweight [74]. For those studies in which it could be determined, six used the per-protocol approach to analyze their data [61, 65, 69, 70, 73, 75], two used intention-to-treat [76, 77], and three used both [62, 64, 74]. Only five of the studies included sample size estimates [62, 64, 74, 75, 79] while the majority (80%) received some type of funding for their work [61, 63–65, 67–71, 73–76, 78–80]. Nine studies received singular support from either university [69, 70, 75, 80], government [63, 65, 76], or private [78, 79] entities while seven other studies reported multiple sources of support from government and private [67, 71], government and university [64, 68], government, university, and private [73, 74], or private and university [61] entities.

For those studies in which data were available, the dropout rate for studies in which data were available ranged from 0% to 34% in the exercise groups ($\overline{X} \pm$ SD, 16.9% ± 14.0, median = 23) and 0% to 26% in the control groups ($\overline{X} \pm$ SD, 12.6%±12.0, median = 14). Reasons for dropping out included time, lack of interest, unhappiness with group assignment, moving, and medical condition. Four studies reported no serious adverse events during the intervention period [74, 77–79].

Participant characteristics are shown in Tables 1 and 2. Twelve of the 20 studies (60%) included both boys and girls [61, 62, 64, 65, 67, 70, 71, 74, 76–79], seven (35%) were limited to boys [63, 66, 68, 69, 73, 75, 80], and one (5%) was limited to girls [72]. Participants included Whites, Blacks, Asians, and Hispanics. With respect to maturational development, the studies represented boys and girls at the prepubertal, pubertal, and postpubertal stages of development; two studies reported boys and girls at the prepubertal stage [61, 64], one at the postpubertal stage [74], two at the prepubertal and pubertal stage [67, 76], and one each at either the pubertal and postpubertal [73] or prepubertal, pubertal, and postpubertal [63] stages.

Characteristics of the exercise programs from each study are also shown in Table 1. Thirteen of the 22 groups participated in aerobic exercise, two in strength training, and seven in both. For those studies and groups in which data were available, length of training ranged from six to 24 weeks ($\overline{X} \pm$ SD, 13.4 ± 5.7, median = 12) and frequency from one to five times per week ($\overline{X} \pm$ SD, 3.3 ± 1.1, median = 3). Intensity of training was classified as low for one group, moderate for

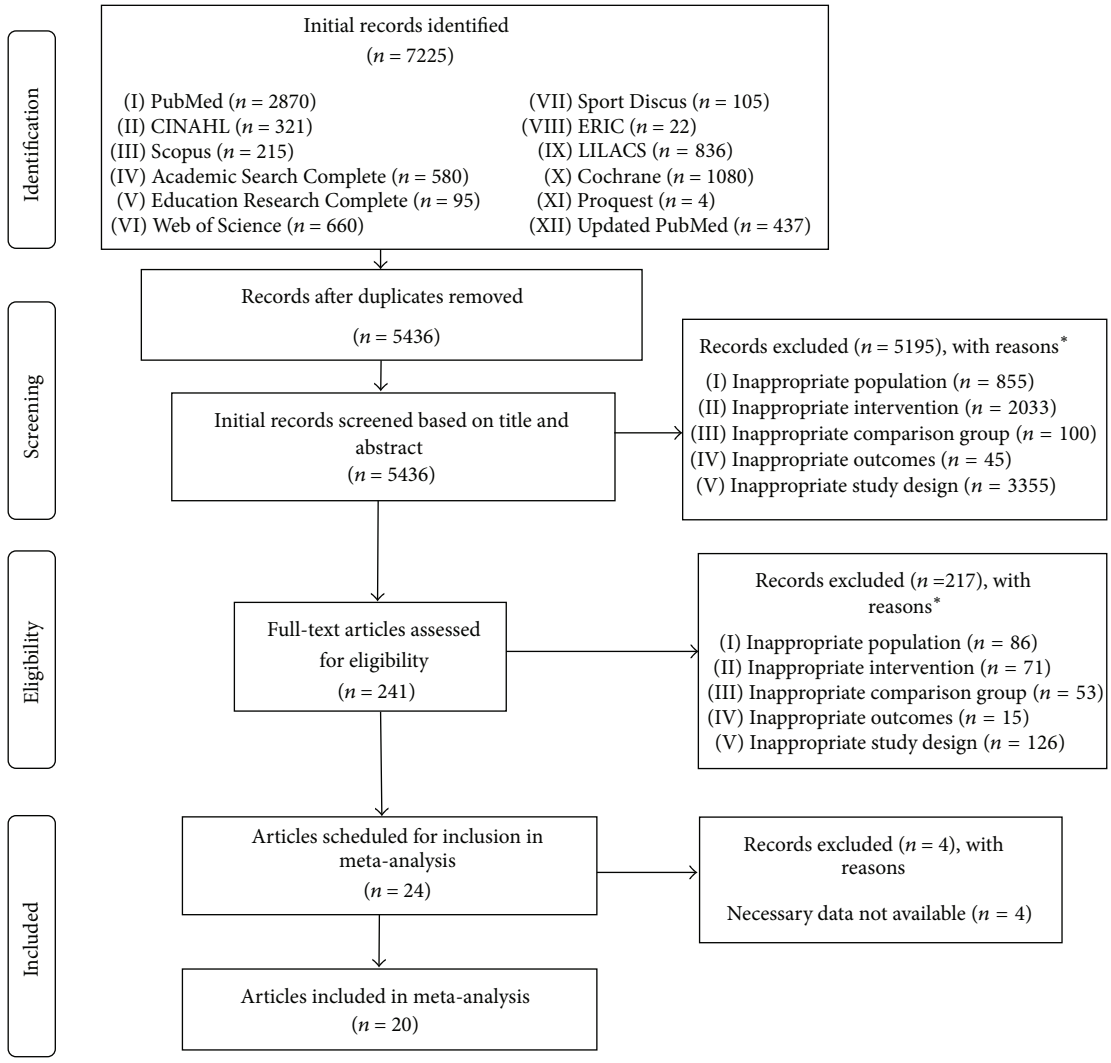


Figure 1: Flow diagram describing the search for relevant literature. *Number of reasons exceeds the number of records excluded.

four groups, moderate to high for two groups, and high for seven groups. When limited to aerobic exercise, duration of training for the 18 groups in which data were available ranged from 20 to 75 minutes per session ($\overline{X}$ ± SD, 45.9 ± 15.4, median = 45). Total minutes of training, per week, ranged from 40 to 224 minutes ($\overline{X}$ ± SD, 148.1 ± 55.0, median = 155) while total minutes of training for the entire intervention period ranged from 480 to 5,040 minutes ($\overline{X}$ ± SD, 1979 ± 1302, median = 1540). When adjusted for compliance to the exercise protocol, total minutes per week for the three groups in which data could be calculated ranged from 39 to 75 minutes ($\overline{X}$ ± SD, 67.3 ± 25.2, median = 75) while total minutes over the entire intervention period ranged from 470 to 896 minutes ($\overline{X}$ ± SD, 749 ± 241, median = 880). Aerobic exercises included walking, jogging, cycling, swimming, jumping rope, stair climbing, aerobic dance, and games (soccer, handball, basketball, volleyball, etc.) as well as other various activities.

For strength training groups, the within-study number of sets for the five groups in which data were provided ranged from one to three while the number of repetitions per set for the six groups in which data were available ranged from three to 25. Two strength training groups reported within-study rest periods between sets that ranged between 60 and 180 seconds. For the five groups that reported data, the number of strength training exercises ranged from seven to 13 ($\overline{X}$ ± SD, 9.2 ± 2.5, median = 9). Types of resistance training equipment used included free weights, machine weights, elastic bands, medicine balls, and the participants' own body weight. Across all exercise groups, 18 participated in supervised exercise, one in unsupervised exercise, and three in both. Compliance to the exercise interventions for the four groups in which data could be calculated ranged from 55% to 98% ($\overline{X}$ ± SD, 83.4 ± 20.2, median = 90).

3.2. Risk of Bias Assessment. The results for pooled risk of bias assessment are shown in Figure 2 while study level results are shown in Supplementary File 4. As can be seen, 95% of the included studies adequately described the process for random sequence generation while none of the studies suffered from incomplete outcome reporting. In contrast, more than half of

Table 1: Study characteristics.

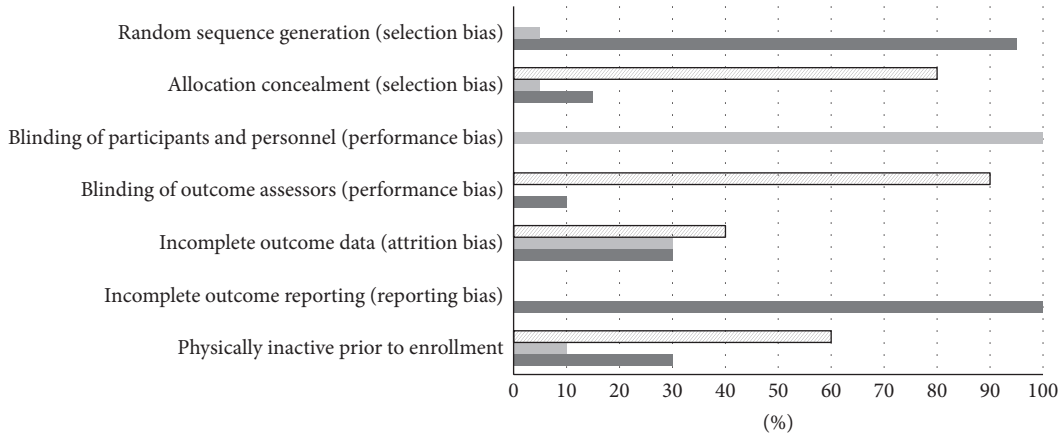
Study	Year	Sex M/F	*N* Ex/Con	Age Ex/Con	Exercise intervention									
					Length (Weeks)	Frequency (Days)	Intensity (%)				Duration (min)		Mode A/S	Compliance (%)
							MHR	HRR	VO_{2max}	1RM	A	S		
Alberga et al. [61]	2013	M/F	12/7	10.0 ± 1.0/10.0 ± 2.0	12	2	65–70	—	—	65–86	20	45	A + S	98%
Alves et al. [62]	2008	M/F	39/39	8.0 ± 1.8/7.9 ± 1.5	24	3	—	—	—	—	50	—	A	—
Elloumi et al. [63]	2011	M	7/8	13.1 ± 1/13.2 ± 0.2	8	4	—	—	—	—	—	—	A	—
Farpour-Lambert et al. [64]	2009	M/F	22/22	9.1 ± 1.4/8.8 ± 1.6	12	3	—	—	55–65	—	30	20	A + S	83%
Hagströmer et al. [65]	2009	M/F	16/15	13.7 ± 2.0/13.6 ± 2.2	13	1	—	—	—	50–70	60		A + S	—
Karacabey [66]	2009	M	20/20	11.8 ± 0.5/11.2 ± 0.80	12	3	—	60–65	—	—	20–45	—	A	—
Kelly et al. [67]	2004	M/F	10/10	11.0 ± 0.6/11.0 ± 0.7	8	4	—	—	50–80	—	30–50	—	A	≥80%
Kim et al. [68]	2007	M	14/12	17.0 ± 0.4/17.0 ± 0.4	6	5	—	—	—	—	30	—	A	—
Kim et al. [69]	2008	M	8/9	11.0/11.0	12	2	—	55–75	—	70	30–35	50	A + S	—
Meyer et al. [70]	2006	M/F	33/34	13.7 ± 2.1/14.1 ± 2.4	24	3	—	—	—	—	60–90	—	A	—
Murphy et al. [71]	2009	M/F	23/12	7–12	12	5	—	—	—	—	10–30	—	A	75% ≥ 5x weeks; 15% ≥ 3x weeks
Saygin and Öztürk [72]	2011	F	20/19	10–12	12	3	—	43	—	—	75	—	A	—
Shaibi et al. [73]	2006	M	11/11	15.1 ± 0.5/15.6 ± 0.5	16	2	—	—	—	62–97	—	—	S	96%
Sigal et al. [74][a]	2014	M/F	75 (A)	15.5 ± 1.4 (A)	22	4	65–85	—	—	—	20–50	—	A	62% (A)
			78 (S)	15.9 ± 1.5 (S)			—	—	—	—	—	—	S	56% (S)
			75 (A + S)	15.5 ± 1.3 (A ± S)			65–85	—	—	—	20–50		A + S	64% (A + S)
			76 (Con)	15.6 ± 1.3 (Con)										
Song et al. [75]	2012	M	12/10	12.7 ± 0.7/12.6 ± 0.6	12	3	60–70	—	—	—	30	—	A	≥80%
Sun et al. [76]	2011	M/F	25/17	13.6 ± 0.7	10	4	—	—	40–60	—	40	—	A	55%
Tan et al. [77][b]	2010	M/F	30/30	9.4 ± 0.5/9.5 ± 0.5	8	5	—	—	—	—	40	—	A	—
Watts et al. [78][c]	2004	M/F	19/19	14.3 ± 1.5	8	3	65–85	—	—	55–70	60	—	A + S	≥90%
Watts et al. [79]	2004	M/F	14/14	8.9 ± 1.6	8	3	66–85	—	—	—	60	—	A	≥90%
Wong et al. [80][c]	2008	M	12/12	13.8 ± 1.1/14.3 ± 1.5	12	2	50–85	—	—	—	55		A + S	—

Notes. M/F: males/females; Ex/Con: exercise/control; days: days per week, minutes: minutes per session; A/S: aerobic/strength; compliance: percentage of exercise sessions attended; data reported as mean ± standard deviation; VO_{2max}: maximum oxygen consumption in $mL \cdot kg^{-1} \cdot min^{-1}$; 1RM: one-repetition maximum; MHRR: maximal heart rate reserve; MHR: maximum heart rate; [a]compliance based on medians versus means; [b]participants trained at intensity equivalent to 78% of lactate threshold; [c]study consisted of circuit training.

Table 2: Initial physical characteristics of participants.

Variable	Exercise					Control				
	Groups/participants	Missing data (%)*	$\overline{X} \pm$ SD	Mdn	Range	Groups/participants	Missing data (%)*	$\overline{X} \pm$ SD	Mdn	Range
Age (years)	21/489	4.5	12.4 ± 2.5	13	8–17	18/301	10.0	12.0 ± 2.6	12	8–17
Height (cm)	19/443	13.6	155.7 ± 12.5	163	130–174	17/263	15.0	154.7 ± 13.4	154	127–175
Body weight (kg)	20/370	9.1	74.5 ± 18.7	75	35–104	18/312	10.0	73.3 ± 18.5	74	34–99
BMI ($kg{\cdot}m^{-2}$)	22/433	0.0	29.7 ± 4.0	30	21–36	20/376	0	29.6 ± 3.9	30	21–36
Fat mass (kg)	14/228	36.4	33.4 ± 9.5	32	22–50	12/181	40.0	31.8 ± 7.1	31	22–47
Body fat (%)	17/321	22.7	39.0 ± 6.2	36	31–50	15/265	25.0	38.5 ± 5.0	38	31–48
Fat-free mass (kg)	15/261	31.8	45.2 ± 7.9	47	28–54	13/208	35.0	44.9 ± 9.4	46	26–62
VO_{2max} ($mL{\cdot}kg^{-1}{\cdot}min^{-1}$)[a]	12/229	40.0	30.2 ± 6.1	31	20–39	10/175	50.0	30.0 ± 6.8	29	20–40
Muscular strength (kg)[b]										
Upper	3	66.7	86.8 ± 84.1	39	37–18	2	75.0	124.4 ± 36.6	124	37–212
Lower	3	66.7	84.3 ± 35.7	100	43–110	2	75.0	80.0 ± 34.2	80	56–104
Energy intake (kcals)	6/108	72.7	2465 ± 577	2319	1813–3278	4/44	81.8	2511 ± 771	2586	1614–3259

Notes. Groups/participants: number of groups and participants in which data were available; $\overline{X} \pm$ SD: mean ± standard deviation; Mdn: median; BMI: body mass index; VO_{2max}: maximum oxygen consumption; kcals: kilocalories; [a]data limited to those groups in which aerobic exercise was an intervention; [b]data limited to those groups in which strength training was an intervention; *percentage of missing data calculated based on the premise that (1) all studies should have assessed and reported data for age, height, body weight, BMI, fat mass, percent body fat, fat-free mass, and energy intake, (2) all studies that included an aerobic exercise component should have assessed and reported data on VO_{2max} in $mL{\cdot}kg^{-1}{\cdot}min^{-1}$, and (3) all studies that included a strength training component should have assessed and reported data on upper and lower body strength.



	Physically inactive prior to enrollment	Incomplete outcome reporting (reporting bias)	Incomplete outcome data (attrition bias)	Blinding of outcome assessors (performance bias)	Blinding of participants and personnel (performance bias)	Allocation concealment (selection bias)	Random sequence generation (selection bias)
Unclear risk	60%	0%	40%	90%	0%	80%	0%
High risk	10%	0%	30%	0%	100%	5%	5%
Low risk	30%	100%	30%	10%	0%	15%	95%

Figure 2: Cochrane risk of bias results.

Table 3: Changes in primary and secondary outcomes.

Variable	ES (#)	$\overline{X}$ (95% CI)	Z (p)	Q (p)	I^2% (95% CI)	D^2%	95% PI
Primary							
BMI (kg·m^{-2})	22	**−1.08 (−0.52, −1.64)***	−3.81 (<0.001)	**231.4 (<0.001)***	90.9 (87.6, 93.4)	91.5	−3.74, 1.58
Secondary							
Body weight (kg)	20	**−1.66 (−0.87, −2.45)***	−4.11 (<0.001)	**57.3 (<0.001)***	66.8 (47.0, 79.2)	78.0	−4.48, 1.17
Fat mass (kg)	14	**−1.07 (−0.36, −1.79)***	−2.93 (0.003)	**29.5 (0.006)***	55.9 (19.8, 75.8)	62.6	−3.33, 1.19
Body fat (%)	17	**−1.13 (−0.58, 1.67)***	−4.05 (<0.001)	**56.2 (<0.001)***	71.5 (53.6, 82.6)	77.1	−3.16, 0.91
Fat-free mass (kg)	15	−0.006 (−0.24, 0.22)	−0.05 (0.96)	21.8 (0.08)	35.8 (0, 65.4)	39.2	—
VO_{2max} (mL·kg^{-1}·min^{-1})[a]	10	**3.1 (1.1, 5.2)***	2.95 (<0.001)	**197.8 (<0.001)***	95.4 (93.3, 96.9)	96.8	−4.53, 10.76
Muscular strength (kg)[b]							
Upper	3	**7.7 (4.4, 10.9)***	4.58 (<0.001)	0.9 (0.63)	0 (0, 92.7)	0	**6.3, 9.0***
Lower	3	**44.2 (29.5, 59.0)***	5.88 (<0.001)	0.3 (0.86)	0 (0, 77.3)	0	**42.9, 45.6***
Energy intake (kcals)	6	−141 (−294, 13)	−1.80 (0.07)	8.9 (0.11)	43.6 (0, 77.7)	59.4	—

Notes.[#]Number; ES: effect size; $\overline{X}$ (95% CI): mean and 95% confidence interval; $Z(p)$: Z value and alpha value for Z; $Q(p)$: Cochrane's Q statistic and alpha value for Q; I^2 (%): I-squared; 95% PI: 95% prediction intervals; D^2: D-squared; BMI: body mass index; VO_{2max}: maximum oxygen consumption; kcals: kilocalories; [a]data limited to those groups in which aerobic exercise was an intervention; [b]data limited to those groups in which strength training was an intervention; [*]statistically significant; —: not calculated; boldface items indicate statistical significance.

the studies were at a high or unclear risk of bias with respect to allocation concealment (85%), blinding of participants and personnel (100%), blinding of outcome assessors (90%), incomplete outcome data, that is, attrition bias (70%), and boys and girls being physically inactive prior to enrollment (70%).

3.3. Data Synthesis

3.3.1. Primary Outcome. Pooled results for changes in BMI in kg·m^{-2} are shown in Table 3 and Figure 3. Across all categories, a statistically significant reduction equivalent to 3.6% was found for BMI in kg·m^{-2} along with statistically significant heterogeneity, a large amount of inconsistency and diversity, and overlapping prediction intervals. Changes in BMI in kg·m^{-2} ranged from 0.59 to −7.30 kg/m^2. With one outlier deleted from the model [66], reductions were not as large (26% difference) but remained statistically significant along with statistically significant heterogeneity as well as a large amount of inconsistency and diversity ($\overline{X}$: −0.80; 95% CI: −0.40 to −1.20; $z = -3.94$; $p < 0.001$; $Q = 108.1$; $p < 0.001$; $I^2 = 81.5\%$; 95% CI = 72.7 to 87.5; $D^2 = 83.2\%$). Reductions in BMI in kg·m^{-2} also remained statistically significant along with statistically significant heterogeneity as well as a large amount of inconsistency and diversity when results were collapsed so that only one result represented each

Study	Subgroup	Analysis	Statistics for each study: Difference in means	Lower limit	Upper limit
Alberga et al., 2013	None	abp	0.59	−0.13	1.31
Alves et al., 2008	None	itt	−0.53	−1.05	−0.01
Elloumi et al., 2011	None	Unknown	−1.30	−3.02	0.42
Farpour-Lambert et al., 2009	None	itt	−0.40	−0.91	0.11
Hagströmer et al., 2009	None	abp	−0.45	−1.12	0.22
Karacabey, 2009	None	Unknown	−7.20	−8.31	−6.09
Kelly et al., 2004	None	Unknown	0.11	−0.52	0.74
Kim et al., 2007	None	Unknown	−0.70	−1.50	0.10
Kim et al., 2008	None	abp	−1.89	−2.52	−1.26
Meyer et al., 2006	None	abp	−2.90	−4.01	−1.79
Murphy et al., 2009	None	Unknown	−0.40	−0.98	0.18
Saygin and Öztürk, 2011	None	Unknown	−2.11	−2.65	−1.57
Shaibi et al., 2006	None	abp	0.00	−0.98	0.98
Sigal et al., 2014	Aerobic	Combined	−0.65	−1.76	0.46
Sigal et al., 2014	Aerobic + resistance	Combined	−1.35	−2.44	−0.26
Sigal et al., 2014	Resistance	Combined	−0.60	−1.74	0.54
Song et al., 2012	None	Unknown	−0.40	−1.41	0.61
Sun et al., 2011	None	itt	−0.40	−1.29	0.49
Tan et al., 2010	None	itt	−2.10	−2.62	−1.58
Watts et al., 2004a	None	Unknown	−0.30	−1.28	0.68
Watts et al., 2004b	None	Unknown	0.00	−1.61	1.61
Wong et al., 2008	None	Unknown	−1.10	−2.43	0.23
			−1.08	−1.64	−0.52

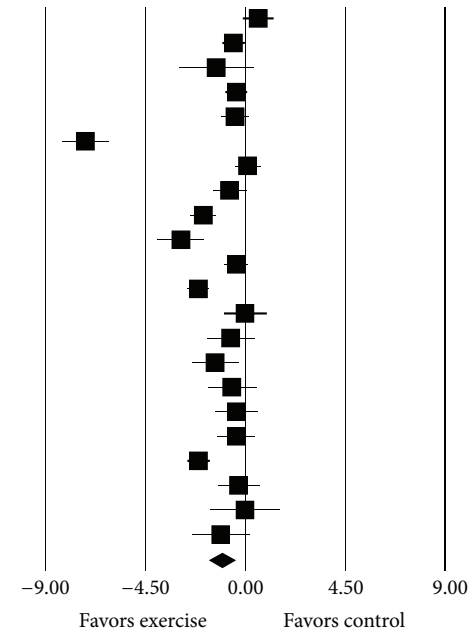


Figure 3: Forest plot for point estimate changes in BMI in kg·m^{-2}. The black squares represent the mean difference while the left and right extremes of the squares represent the corresponding 95% confidence intervals. The middle of the black diamond represents the overall mean difference while the left and right extremes of the diamond represent the corresponding 95% confidence intervals.

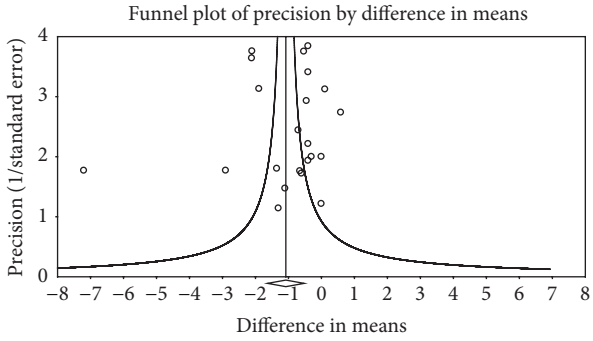


Figure 4: Funnel plot for changes in BMI in kg·m^{-2}.

study ($\overline{X}$: −1.10; 95% CI: −0.52 to −1.68; $z = -3.71$; $p < 0.001$; $Q = 230.3$; $p < 0.001$; $I^2 = 91.8\%$; 95% CI = 88.7 to 94.0; $D^2 = 92.2\%$). No small-study effects were observed as indicated by a lack of funnel plot asymmetry (Figure 4) and Egger's regression-intercept test (β_0: −0.92, $p = 0.34$). With each result deleted from the model once, reductions in BMI in kg·m^{-2} remained statistically significant across all deletions, with changes ranging from −0.80 to −1.16, a difference of 31% (Figure 5). The NNT was 5 (95% CI = 3 to 12) while the percentile improvement was 26.9 (95% CI = 15.3 to 36.0). It was estimated that approximately 2.5 million overweight and obese children in the US (95% CI, 1.0 to 4.2) and 22.0 million overweight and obese children worldwide (95% CI, 9.2 to 36.7) could reduce their BMI in kg·m^{-2} by participating in a regular exercise program.

For the four studies excluded because of insufficient data for BMI in kg·m^{-2} [81–84], one reported a statistically significant exercise minus control group reduction ($p = 0.02$) in BMI in kg·m^{-2} [84]. Another study that did not report results for BMI in kg·m^{-2} did report a statistically significant exercise minus control group reduction in BMI z-score ($p = 0.02$) for the high-dose group as well as a trend for improvement ($p = 0.06$) in the low-dose group [81]. The remaining two studies did not report any BMI-related results, although both reported statistically significant reductions of $p < 0.01$ [82] and $p < 0.001$ [83] for percent body fat. In addition, fail-safe N results indicated that a total of 774 studies with null findings would be needed to reverse our findings of a statistical significant reduction in BMI in kg·m^{-2}.

The results for trial sequential meta-analysis are shown in Figure 6. As can be seen, these findings confirm that the maximum information size has been reached and the stability of findings has been achieved with respect to exercise-induced reductions in BMI in kg·m^{-2} among overweight and obese children and adolescents. More specifically, changes in BMI in kg·m^{-2} crossed the monitoring boundary for a type 1 error in 2010 and have remained stable thereafter. This confirms the statistical significance of exercise-induced

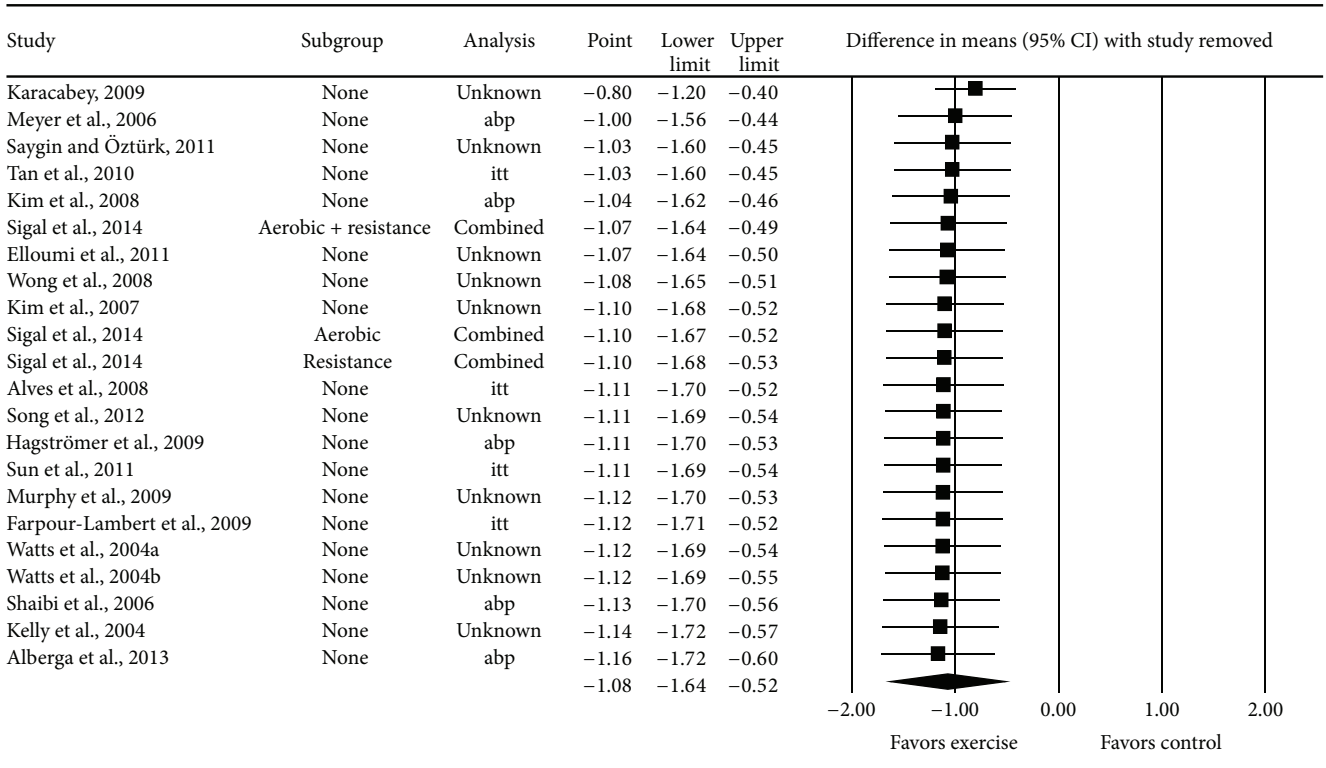


Figure 5: Influence analysis results for point estimate changes in BMI in $kg \cdot m^{-2}$ with each result deleted from the model once. The black squares represent the mean difference while the left and right extremes of the squares represent the corresponding 95% confidence intervals. The middle of the black diamond represents the overall mean difference while the left and right extremes of the diamond represent the corresponding 95% confidence intervals. Results are ordered from smallest to largest reductions.

reductions in BMI in $kg \cdot m^{-2}$ since 2010 among overweight and obese children and adolescents and suggests that the accumulation of additional studies in future years will not change these findings to one of nonsignificance. Simple metaregression results are shown in Supplementary File 5. No statistically significant associations were observed for those variables in which metaregression analysis was possible. Findings were similar when the one outlier for changes in BMI in $kg \cdot m^{-2}$ was deleted from each of the analyses (results not shown) [66].

The results for GRADE with respect to changes in BMI in $kg \cdot m^{-2}$ are shown in Supplementary File 6. Despite potential biases as well as heterogeneity, inconsistency, diversity, and overlapping prediction intervals, the overall quality of evidence was upgraded from low to moderate based on the magnitude of effect observed, trial sequential analysis results, and lack of adverse events.

3.3.2. Secondary Outcomes. Secondary outcomes are shown in Table 3. Statistically significant improvements were found for body weight, fat mass, percent body fat, VO_{2max} in $mL \cdot kg^{-1} \cdot min^{-1}$, and upper and lower body strength. No statistically significant differences were observed for fat-free mass or energy intake. Changes were equivalent to relative improvements of 2.2% (body weight), 3.2% (fat mass), 2.9% (percent body fat), 10.3% (VO_{2max} in $mL \cdot kg^{-1} \cdot min^{-1}$), 8.9% (upper body strength), and 52.4% (lower body strength). With the exception of changes in upper and lower body strength, statistically significant heterogeneity as well as moderate to large inconsistency and diversity was observed for body weight, fat mass, percent body fat, and VO_{2max} in $mL \cdot kg^{-1} \cdot min^{-1}$. Prediction intervals were overlapping for all outcomes except for upper and lower body strength.

Statistically significant outliers ($p < 0.05$) were identified for changes in body weight [61, 66], fat mass [69], percent body fat [69], and VO_{2max} in $mL \cdot kg^{-1} \cdot min^{-1}$ [72]. With two outliers deleted from the model for body weight [61, 66], reductions remained statistically significant along with statistically significant heterogeneity, low inconsistency, and moderate diversity ($\overline{X}$: −1.62; 95% CI: −0.93 to −2.31; $z = -4.61$; $p < 0.001$; $Q = 33.3$; $p < 0.001$; $I^2 = 49.0\%$; 95% CI = 12.0 to 70.4; $D^2 = 68.1\%$). For fat mass, decreases remained statistically significant along with no statistically significant heterogeneity as well as very low inconsistency and diversity when one outlier was deleted from the model [69] ($\overline{X}$: −0.69; 95% CI: −0.17 to −1.22; $z = -2.58$; $p = 0.01$; $Q = 13.9$; $p = 0.31$; $I^2 = 13.9\%$; 95% CI = 0 to 62.5; $D^2 = 21.1\%$). With the same study deleted [69], reductions in percent body fat remained statistically significant along with statistically significant heterogeneity, moderate inconsistency, and large diversity ($\overline{X}$: −1.01; 95% CI: −0.47 to −1.54; $z = -3.68$; $p < 0.001$; $Q = 49.6$; $p < 0.001$; $I^2 = 69.7\%$; 95% CI = 49.5 to

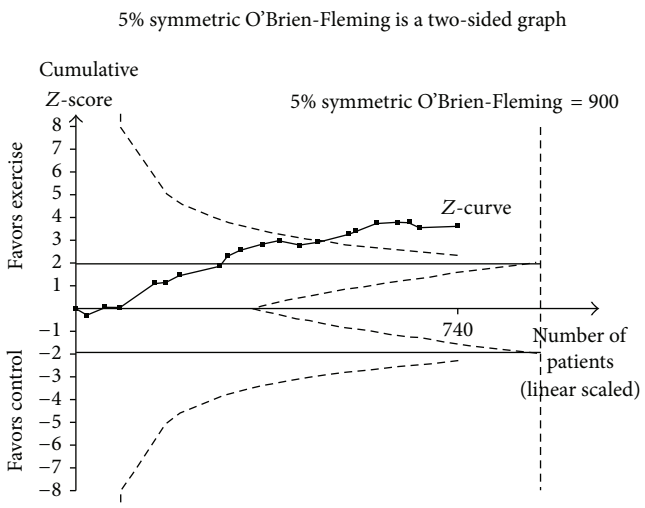


Figure 6: Trial sequential analysis results. Trial sequential meta-analysis of exercise versus control for changes in BMI in kg·m^{-2}. The dashed inward sloping lines to the left represent trial sequential monitoring boundaries while the outward dashed sloping lines to the right represent futility boundaries. The solid black line represents the Z-curve and the black squares represent the cumulative results with each accumulating study from earliest (2004) to most recent (2014) year. The cumulative Z-curve, that is, black solid line with filled squares, crossed the monitoring boundaries in 2010, confirming that exercise reduces BMI in kg·m^{-2} in overweight and obese children and adolescents and is unlikely to be reversed with additional studies in future years.

81.9; D^2 = 75.7%). Increases in VO_{2max} in mL·kg^{-1}·min^{-1} also remained statistically significant along with statistically significant heterogeneity, moderate inconsistency, and large diversity when one outlier [72] was deleted from the model ($\overline{X}$: 2.35; 95% CI: 1.38 to 3.31; z = 4.76; $p < 0.001$; Q = 22.3; p = 0.004; I^2 = 64.1%; 95% CI = 26.4 to 82.4; D^2 = 81.3%). No outliers were identified for changes in lower and upper body muscular strength.

For those secondary outcomes in which statistically significant improvements were found, statistically significant small-study effects were observed for changes in percent body fat (β_0: 1.92, p = 0.03). No statistically significant small-study effects were observed for body weight (β_0: −0.54, p = 0.22), fat mass (β_0: −0.76, p = 0.26), VO_{2max} in mL·kg^{-1}·min^{-1} (β_0: −0.01, p = 0.50), or upper (β_0, 1.99, p = 0.15) and lower (β_0, 4.77, p = 0.27) body strength.

With each result deleted from the model once, changes remained statistically significant for all secondary outcomes in which the original findings were statistically significant. Changes ranged from −1.38 to −1.86 kg for body weight (25.3% difference), −0.69 to −1.25 kg for fat mass (44.8% difference), −1.00 to −1.23 for percent body fat (23.0% difference), 2.34 to 3.42 mL·kg·min^{-1} for VO_{2max} (31.6% difference), 6.9 to 9.4 kg for upper body strength (26.6% difference), and 41.5 to 46.0 kg for lower body strength (9.8% difference).

4. Discussion

4.1. Findings. The overall findings of the current meta-analysis suggest that exercise is associated with reductions in BMI in kg·m^{-2} among overweight and obese children and adolescents. Support for this interpretation is derived from (1) the overall magnitude of effect, (2) nonoverlapping 95% CI, (3) continued significance when each study was deleted from the model once, including the one outlier [66], (4) apparent absence of small-study effects, (5) trial sequential analysis results demonstrating that the maximum information size had been reached and been stable since 2010, (6) the low NNT, and (7) the number of overweight and obese children and adolescents in the US and worldwide who might potentially improve their BMI in kg·m^{-2} from the uptake of regular exercise. In addition, the magnitude of change in BMI in kg·m^{-2} observed in this study (−1.08 kg/m^2 or 3.6%) may be clinically relevant as previous research has found significant improvements in selected health outcomes with a decrease in BMI in kg·m^{-2} of approximately 4.8% [85]. While the results of the current meta-analysis were 1.2% smaller, they may still be clinically important. Regardless, the observed reductions in BMI in kg·m^{-2} are most likely important at the population level. For example, a recent meta-analysis that reported a reduction of only 0.17 kg·m^{-2} in BMI [20] as a result of school-based interventions suggested that their findings may result in important health benefits at the population level. This suggestion was based on the work of Rose [86] who contended that a small shift in population distribution can be an effective primary preventative strategy because more events occur among the large number of individuals at moderate risk than the small number at high risk. Importantly, the results of the current meta-analysis were more than six times larger than those of Lavelle et al. [20]. However, it is important to realize that whether an intervention should be recommended at the population level depends not only on the size of the effect but also on the costs associated with achieving such an effect as well as society's willingness to pay for this. While the willingness of a society to pay for this most likely varies between countries and there is limited evidence regarding the cost-effectiveness of exercise interventions for the treatment of overweight and obesity in children and adolescents, one cost-effectiveness study found that the number of disability-adjusted life years was greater for a multifaceted school-based intervention that included physical education (8000) versus one without physical education (500) [87]. Clearly, further research in this area is needed.

Finally, considering that the results for GRADE were increased from low to moderate provides justification for recommending exercise for improving BMI in kg·m^{-2} in overweight and obese children and adolescents. This is especially relevant given that a low rating is based on the belief that additional evidence in the future would most likely change the direction of effect, something that the investigative team does not believe will happen, especially given the trial sequential analysis results.

In contrast to the investigative team's findings that support the effects of exercise for reducing BMI in kg·m^{-2} as

well as the fact that a random-effects model that incorporates heterogeneity into the analysis was used, no potential sources of heterogeneity were identified as a result of metaregression analyses. Thus, the current results could be compromised. This may be especially important given the large amount of inconsistency and diversity observed for BMI in kg·m^{-2} in the current meta-analysis. However, while such analyses are important, covariate analyses in meta-analysis are considered observational given that studies are not randomly assigned to covariates [88]. As a result, such analyses do not support causal inferences [88]. Thus, while such analyses may generate important findings about potential sources of heterogeneity, they would still need to be tested in adequately powered randomized controlled trials [88]. A second finding that may weaken the BMI in kg·m^{-2} results is the overlapping PI observed for changes in BMI in kg·m^{-2}. However, it is important to understand that PI are different compared to CI as the former are based on random-mean effects [35].

While no variables that accounted for heterogeneity with respect to changes in BMI in kg·m^{-2} were found, it may be that factors that were unable to be assessed could account for some or all of the observed heterogeneity between the included studies. These include such things as (1) differences or changes in diet during the exercise intervention [89], (2) physical activity compensation [65, 90], and (3) genetic factors [91].

The results of the current meta-analysis are in agreement with one previous systematic review with meta-analysis that focused on exercise [20] but disagreement with four others that reported a nonsignificant decrease in BMI in kg·m^{-2} among children and adolescents [17, 21–23]. Possible reasons for these discrepancies include (1) the small number of exercise-only studies that were included and pooled in these meta-analyses [17, 21, 22], (2) the inclusion of nonrandomized trials [20, 22], and (3) the inclusion of children and adolescents who were not overweight or obese [20, 22, 23].

The reductions in BMI in kg·m^{-2} observed in the current meta-analysis also compare favorably to orlistat, the only weight-loss medication currently approved by the US Food and Drug Administration for the treatment of obese adolescents. In a recent meta-analysis, changes in BMI in kg·m^{-2} that included two studies representing 579 participants resulted in a statistically significant decrease of −0.76 kg·m^{-2} (95% CI, −1.07, −0.44) as a result of the use of orlistat [92]. These findings are approximately 30% less than those found for BMI in kg·m^{-2} and exercise in the current meta-analysis.

The reductions in BMI in kg·m^{-2} found in the current meta-analysis are also similar to the results reported in a recent systematic review of diet-only interventions in which decreases ranged from 0.8 to 2.7 kg/m^2 [89]. This suggests that either exercise or diet can reduce BMI in kg·m^{-2} in a similar fashion. In contrast, the results of this previous systematic review when diet and exercise were combined were equivocal, with changes in BMI in kg·m^{-2} ranging from −4.4 to 0.27 kg/m^2 for aerobic exercise (4 studies), −0.2 to 1.1 kg/m^2 for resistance training (3 studies), and −0.5 to −2.02 kg/m^2 for combined aerobic and resistance exercise (3 studies) [89]. However, whether these changes differ significantly according to type of exercise, type of diet, or some other factor(s) is not known.

In addition to changes in BMI in kg·m^{-2}, statistically significant and clinically important improvements in body weight, fat mass, percent body fat, relative VO_{2max}, and upper and lower body strength were observed. The changes in fat mass as well as percent body fat are particularly noteworthy since both are more relevant than BMI in kg·m^{-2} with respect to improvements in body composition. However, because they are not as practical to assess, BMI in kg·m^{-2} continues to be the preferred method of assessing and classifying overweight and obesity. In addition, the significant changes observed for the six secondary outcomes support the multiple benefits that can be derived from regular participation in exercise. The multiple benefits observed are in contrast to treatments such as pharmacological interventions, approaches that are usually intended to treat one outcome. In addition, orlistat, the only pharmacological intervention currently approved in the United States for the treatment of obesity in children and adolescents [93], has been shown to be less cost-effective than several nonpharmacologic interventions, including exercise [87, 94], and has also been accompanied by side-effects such as gastrointestinal distress [95]. With respect to exercise in the current meta-analysis, four studies that did include information on side-effects reported no serious adverse events [74, 77–79], defined as any intervention that results in death, a life threatening condition, hospitalization (initial or prolonged), disability, or permanent damage [96]. For these same four studies, adverse events, defined as any undesirable experience associated with an intervention, included primarily acute musculoskeletal injury or discomfort in 7.9% of exercise participants in one study [74] and none in the other three [77–79]. However, it is important to realize that 16 [61–73, 75, 76, 80] of the 20 studies in the current meta-analysis did not report adequate information with respect to adverse events.

4.2. Implications for Research. The results of this meta-analysis have several implications for both the reporting and conduct of future research. First, it is suggested that future studies report complete information regarding (1) allocation concealment, (2) blinding of outcome assessment, (3) dropouts according to each group, including reasons for dropping out, (4) adverse events, (5) the physical activity levels of participants prior to and during the intervention, (6) intensity of the exercise intervention, and (7) compliance to the exercise intervention. It is also suggested that investigators analyze and report results using both the per-protocol and intention-to-treat results. This will allow one to understand both the efficacy and effectiveness of exercise for improving BMI in kg·m^{-2} as well as other outcomes in overweight and obese children and adolescents. In addition, since both energy intake and energy expenditure are critical in determining weight loss, future studies should collect and report data on energy intake and total daily energy expenditure. Finally, future studies should report complete

information on all outcomes assessed, partitioned by group. At a minimum, these data should include pre- and postsample sizes, means, and standard deviations as well as change outcome results along with their standard deviations.

It appears that a need exists for a four-arm randomized controlled trial in overweight and obese children and adolescents that includes an aerobic, strength, and combined aerobic and strength training group as well as a control group. Furthermore, to aid practitioners, a need exists for dose-response studies to determine the optimal exercise program(s) for overweight and obese children and adolescents. This may be especially important since it is currently recommended that children and adolescents participate in 60 or more minutes of moderate to vigorous physical activity per day (420 minutes per week) [97] but the current meta-analysis found statistically significant reductions in BMI in $kg{\cdot}m^{-2}$ as well as several other outcome variables (body weight, fat mass, percent body fat, VO_{2max}, and upper and lower body strength) when the average total minutes per week was less than currently recommended. Finally, since cost is an important factor when deciding what intervention to recommend over another, a need exists for cost-effectiveness studies in overweight and obese children and adolescents.

4.3. Implications for Practice. The results of the current meta-analysis suggest that exercise results in important improvements in BMI in $kg{\cdot}m^{-2}$ as well as body weight, fat mass, percent body fat, VO_{2max} in $mL{\cdot}kg^{-1}{\cdot}min^{-1}$, and muscular strength in both upper and lower body. Lending further support for this contention is the low NNT, percentile improvement, and the estimated number of overweight and obese children in the United States and worldwide who could potentially benefit. Furthermore, no serious adverse events were reported for the four groups in which sufficient information was available. Unfortunately, the dose-response effects of exercise on BMI in $kg{\cdot}m^{-2}$ and other outcomes in overweight and obese children and adolescents remain elusive. Thus, in order to not withhold a potentially beneficial and safe intervention and until more definitive evidence is available, it would appear prudent to recommend that practitioners follow the guidelines specific to children and adolescents as denoted in the 2008 Physical Activity Guidelines for Americans [98]. This includes at least 60 minutes per day of moderate to vigorous physical activity, primarily aerobic activity (running, hopping, skipping, jumping rope, swimming, dancing, and bicycling) as well as muscle strengthening activities and bone strengthening activities (running, jumping rope, basketball, tennis, hopscotch, etc.) [98].

While the focus of the current meta-analysis was on the effects of exercise on BMI in $kg{\cdot}m^{-2}$ in overweight and obese children and adolescents, it would appear plausible to suggest that the addition of reduced caloric intake combined with exercise may result in even greater reductions in BMI in $kg{\cdot}m^{-2}$ However, a recent meta-analysis of randomized controlled trials by Ho et al. found no statistically significant differences in BMI in $kg{\cdot}m^{-2}$ between exercise and diet versus diet-only groups [89]. Importantly, the authors concluded that further randomized controlled trials with a rigorous design are needed to confirm their findings. Until that time, it would appear plausible to suggest that practitioners follow the recent recommendations that, in addition to exercise, include (1) the avoidance of sugar-sweetened beverages, (2) less food with high caloric density, and (3) increased intake of fruits and vegetables [99].

4.4. Strengths and Potential Limitations of Current Study. In the investigative team's opinion, there are at least four *strengths* of the current meta-analysis. First, to the best of the authors' knowledge, this is the first trial sequential meta-analysis that has examined the effects of exercise on BMI in $kg{\cdot}m^{-2}$ in overweight and obese children and adolescents, something that was not done in previous work by the investigative team [19]. This is important because it suggests that changes in BMI in $kg{\cdot}m^{-2}$ are stable and not subject to a type 1 or type 2 error. Second, the inclusion of data regarding NNT, percentile improvement, relative improvement, and absolute number of overweight and obese children who might benefit from participation in a regular exercise program provides practical information to decision-makers with respect to what treatment, or combination of treatments, to recommend over others for overweight and obese children and adolescents. Third, the calculation of PI provides future researchers with an estimate of what effect they might expect to find for BMI in $kg{\cdot}m^{-2}$ and several secondary outcomes (body weight, fat mass, percent body fat, VO_{2max} in $mL{\cdot}kg^{-1}{\cdot}min^{-1}$, and upper and lower muscular strength) if they were to conduct a randomized controlled exercise intervention trial in overweight and obese children and adolescents. Fourth, this supports previous work by the investigative team in which an exercise minus control group improvement of approximately 3% was found for BMI z-score [19]. From the investigative team's perspective, the similar improvements observed for both BMI z-score and BMI in $kg{\cdot}m^{-2}$ are important given the continued controversy regarding which metric is the most valid and reliable for assessing changes in adiposity among children and adolescents. For example, while one study reported that BMI z-score is the best BMI measure for assessing adiposity in children and/or adolescents [100], another [101], as well as more recent research [102], suggests that both absolute and relative changes in BMI in $kg{\cdot}m^{-2}$ are better proxies for changes in adiposity. Thus, regardless of which BMI measure is superior for measuring changes in adiposity, something that is unlikely to be resolved in the near future, the investigative team's previous [19] as well as current findings support similar exercise-induced improvements for both.

As opposed to the strengths of the current meta-analysis, there are at least five *potential limitations*. First, given the statistically significant heterogeneity as well as high inconsistency and diversity of the current findings as well as overlapping PI and GRADE findings, one might conclude that insufficient evidence currently exists to conclude that exercise is associated with statistically significant improvements in BMI in $kg{\cdot}m^{-2}$ and selected secondary outcomes (body weight, fat mass, percent body fat, VO_{2max} in $mL{\cdot}kg^{-1}{\cdot}min^{-1}$, and

upper and lower muscular strength). Second, the statistically significant findings for increases in upper and lower body strength may need to be viewed with caution given that these findings were limited to three results. Consequently, the generalizability of these findings may be limited. Third, the generalizability of the current findings beyond the populations and intervention protocols included may be limited. Fourth, the results of the current meta-analysis, like any meta-analysis, may suffer from ecological fallacy, phenomena in which incorrect inferences about individual findings are made based upon aggregate statistics [103]. Fifth, since the search for studies focused on BMI in $kg{\cdot}m^{-2}$ as the primary outcome, the results for all eight secondary outcomes may represent a biased sample.

5. Conclusions

The results of the current systematic review of previous meta-analyses suggest that exercise is associated with reductions in BMI in $kg{\cdot}m^{-2}$ among overweight and obese children and adolescents. A need exists for randomized controlled trials to identify sources of heterogeneity, including dose-response studies.

Disclosure

Kristi S. Kelley and Dr. Russell R. Pate are the coauthors of this paper.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research reported in this paper was supported by the American Heart Association, Great Rivers Affiliate, Grant-in-Aid 12GRNT11670019 (GAK, Principal Investigator) and the National Institute of General Medical Sciences of the National Institutes of Health under Award no. U54GM104942. The content is solely the responsibility of the authors and does not necessarily represent the views of the American Heart Association or National Institutes of Health.

References

[1] C. L. Ogden, M. D. Carroll, B. K. Kit, and K. M. Flegal, "Prevalence of childhood and adult obesity in the United States, 2011-2012," *The Journal of the American Medical Association*, vol. 311, no. 8, pp. 806–814, 2014.

[2] National Center for Health Statistics, *Health, United States, 2011: With Special Feature on Socioeconomic Status and Health*, National Center for Health Statistics, Hyattsville, Md, USA, 2012.

[3] M. Ng, T. Fleming, M. Robinson et al., "Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, vol. 384, no. 9945, pp. 766–781, 2014.

[4] E. A. Finkelstein, W. C. K. Graham, and R. Malhotra, "Lifetime direct medical costs of childhood obesity," *Pediatrics*, vol. 133, no. 5, pp. 854–862, 2014.

[5] D. S. Freedman, Z. Mei, S. R. Srinivasan, G. S. Berenson, and W. H. Dietz, "Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study," *Journal of Pediatrics*, vol. 150, no. 1, pp. 12–17, 2007.

[6] W. D. Paulis, S. Silva, B. W. Koes, and M. van Middelkoop, "Overweight and obesity are associated with musculoskeletal complaints as early as childhood: a systematic review," *Obesity Reviews*, vol. 15, no. 1, pp. 52–67, 2014.

[7] R. Arens and H. Muzumdar, "Childhood obesity and obstructive sleep apnea syndrome," *Journal of Applied Physiology*, vol. 108, no. 2, pp. 436–444, 2010.

[8] L. J. Griffiths, T. J. Parsons, and A. J. Hill, "Self-esteem and quality of life in obese children and adolescents: a systematic review," *International Journal of Pediatric Obesity*, vol. 5, no. 4, pp. 282–304, 2010.

[9] A. S. Singh, C. Mulder, J. W. R. Twisk, W. van Mechelen, and M. J. M. Chinapaw, "Tracking of childhood overweight into adulthood: a systematic review of the literature," *Obesity Reviews*, vol. 9, no. 5, pp. 474–488, 2008.

[10] G. Danaei, E. L. Ding, D. Mozaffarian et al., "The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors," *PLoS Medicine*, vol. 6, no. 4, Article ID e1000058, 2009.

[11] World Health Organization, *Obesity and Overweight*, World Health Organization, Geneva, Switzerland, 2014.

[12] American Medical Association, "Obesity as a disease," *Policy Statement*, vol. 420, no. 13, pp. 6–18, 2013.

[13] S. Lipnowski and C. M. A. Leblanc, "Healthy active living: physical activity guidelines for children and adolescents," *Paediatrics & Child Health*, vol. 17, pp. 209–210, 2012.

[14] European Commission, *EU Action Plan on Childhood Obesity 2014–2020*, 2014.

[15] D. M. Hoelscher, S. Kirk, L. Ritchie, and L. Cunningham-Sabo, "Position of the Academy of Nutrition and Dietetics: interventions for the prevention and treatment of pediatric overweight and obesity," *Journal of the Academy of Nutrition and Dietetics*, vol. 113, no. 10, pp. 1375–1394, 2013.

[16] World Health Organization, *Global Recommendations on Physical Activity for Health: 5–17 Year Olds*, World Health Organization, Geneva, Switzerland, 2011.

[17] L. McGovern, J. N. Johnson, R. Paulo et al., "Treatment of pediatric obesity: a systematic review and meta-analysis of randomized trials," *Journal of Clinical Endocrinology and Metabolism*, vol. 93, no. 12, pp. 4600–4605, 2008.

[18] C. C. Kamath, K. S. Vickers, A. Ehrlich et al., "Behavioral interventions to prevent childhood obesity: a systematic review and metaanalyses of randomized trials," *Journal of Clinical Endocrinology and Metabolism*, vol. 93, no. 12, pp. 4606–4615, 2008.

[19] G. A. Kelley, K. S. Kelley, and R. R. Pate, "Effects of exercise on BMI z-score in overweight and obese children and adolescents: a systematic review with meta-analysis," *BMC Pediatrics*, vol. 14, article 225, 2014.

[20] H. V. Lavelle, D. F. MacKay, and J. P. Pell, "Systematic review and meta-analysis of school-based interventions to reduce body mass index," *Journal of Public Health*, vol. 34, no. 3, pp. 360–369, 2012.

[21] E. Atlantis, E. H. Barnes, and M. A. F. Singh, "Efficacy of exercise for treating overweight in children and adolescents: a systematic review," *International Journal of Obesity*, vol. 30, no. 7, pp. 1027–1040, 2006.

[22] K. C. Harris, L. K. Kuramoto, M. Schulzer, and J. E. Retallack, "Effect of school-based physical activity interventions on body mass index in children: a meta-analysis," *Canadian Medical Association Journal*, vol. 180, no. 7, pp. 719–726, 2009.

[23] P. H. Guerra, M. R. C. Nobre, J. A. C. da Silveira, and J. A. de Aguiar Carrazedo Taddei, "The effect of school-based physical activity interventions on body mass index: a meta-analysis of randomized trials," *Clinics*, vol. 68, no. 9, pp. 1263–1273, 2013.

[24] B. J. Shea, C. Hamel, G. A. Wells et al., "AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews," *Journal of Clinical Epidemiology*, vol. 62, no. 10, pp. 1013–1020, 2009.

[25] J. Wetterslev, K. Thorlund, J. Brok, and C. Gluud, "Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis," *Journal of Clinical Epidemiology*, vol. 61, no. 1, pp. 64–75, 2008.

[26] A. Liberati, D. G. Altman, J. Tetzlaff et al., "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *Annals of Internal Medicine*, vol. 151, no. 4, pp. W-65–W-94, 2009.

[27] Reference Manager, (12.0.3), Thompson ResearchSoft, Philadelphia, Pa, USA, 2009.

[28] E. Lee, M. Dobbins, K. Decorby, L. McRae, D. Tirilis, and H. Husson, "An optimal search filter for retrieving systematic reviews and meta-analyses," *BMC Medical Research Methodology*, vol. 12, article 51, 2012.

[29] Microsoft Corporation, *Microsoft Excel. (2007)*, Microsoft Corporation, Redmond, Wash, USA, 2010.

[30] J. Cohen, "Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit," *Psychological Bulletin*, vol. 70, no. 4, pp. 213–220, 1968.

[31] J. P. T. Higgins, D. G. Altman, P. C. Gøtzsche et al., "The Cochrane Collaboration's tool for assessing risk of bias in randomised trials," *British Medical Journal*, vol. 343, Article ID d5928, 2011.

[32] D. Follmann, P. Elliott, I. Suh, and J. Cutler, "Variance imputation for overviews of clinical trials with continuous response," *Journal of Clinical Epidemiology*, vol. 45, no. 7, pp. 769–773, 1992.

[33] R. Dersimonian and N. Laird, "Meta-analysis in clinical trials," *Controlled Clinical Trials*, vol. 7, no. 3, pp. 177–188, 1986.

[34] W. G. Cochran, "The combination of estimates from different experiments," *Biometrics*, vol. 10, no. 1, pp. 101–129, 1954.

[35] J. P. T. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring inconsistency in meta-analyses," *British Medical Journal*, vol. 327, no. 7414, pp. 557–560, 2003.

[36] J. Wetterslev, K. Thorlund, J. Brok, and C. Gluud, "Estimating required information size by quantifying diversity in random-effects model meta-analyses," *BMC Medical Research Methodology*, vol. 9, article 86, 2009.

[37] J. P. T. Higgins and S. G. Thompson, "Quantifying heterogeneity in a meta-analysis," *Statistics in Medicine*, vol. 21, no. 11, pp. 1539–1558, 2002.

[38] J. P. T. Higgins and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0, The Cochrane Collaboration, 2011.

[39] J. P. Higgins, S. G. Thompson, and D. J. Spiegelhalter, "A re-evaluation of random-effects meta-analysis," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 172, no. 1, pp. 137–159, 2009.

[40] G. A. Kelley and K. S. Kelley, "Impact of progressive resistance training on lipids and lipoproteins in adults: another look at a meta-analysis using prediction intervals," *Preventive Medicine*, vol. 49, no. 6, pp. 473–475, 2009.

[41] J. A. C. Sterne, A. J. Sutton, J. P. A. Ioannidis et al., "Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials," *British Medical Journal*, vol. 343, article d4002, 2011.

[42] M. Egger, G. D. Smith, M. Schneider, and C. Minder, "Bias in meta-analysis detected by a simple, graphical test," *British Medical Journal*, vol. 315, no. 7109, pp. 629–634, 1997.

[43] R. Rosenthal, "The 'file drawer' problem and tolerance for null results," *Psychological Bulletin*, vol. 86, no. 3, pp. 638–641, 1979.

[44] A. M. G. Cali and S. Caprio, "Obesity in children and adolescents," *Journal of Clinical Endocrinology and Metabolism*, vol. 93, no. 11, pp. s31–s36, 2008.

[45] D. W. Haslam and W. P. T. James, "Obesity," *The Lancet*, vol. 366, no. 9492, pp. 1197–1209, 2005.

[46] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Academic Press, New York, NY, USA, 1988.

[47] G. H. Guyatt, A. D. Oxman, G. E. Vist et al., "GRADE: an emerging consensus on rating quality of evidence and strength of recommendations," *British Medical Journal*, vol. 336, no. 7650, pp. 924–926, 2008.

[48] K. Thorlund, P. J. Devereaux, J. Wetterslev et al., "Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses?" *International Journal of Epidemiology*, vol. 38, no. 1, pp. 276–286, 2009.

[49] J. Brok, K. Thorlund, C. Gluud, and J. Wetterslev, "Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses," *Journal of Clinical Epidemiology*, vol. 61, no. 8, pp. 763–769, 2008.

[50] J. Brok, K. Thorlund, J. Wetterslev, and C. Gluud, "Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses," *International Journal of Epidemiology*, vol. 38, no. 1, pp. 287–298, 2009.

[51] K. Thorlund, J. Engstrom, J. Brok, G. Imberger, and C. Gluud, *User Manual for Trial Sequential Analysis (TSA)*, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen, Denmark, 2011.

[52] R. Lazarus, M. Wake, K. Hesketh, and E. Waters, "Change in body mass index in Australian primary school children, 1985–1997," *International Journal of Obesity*, vol. 24, no. 6, pp. 679–684, 2000.

[53] J. P. T. Higgins, A. Whitehead, and M. Simmonds, "Sequential methods for random-effects meta-analysis," *Statistics in Medicine*, vol. 30, no. 9, pp. 903–921, 2011.

[54] I. van der Tweela and C. Bollenb, "Sequential meta-analysis: an efficient decision-making tool," *Clinical Trials*, vol. 7, no. 2, pp. 136–146, 2010.

[55] American College of Sports Medicine, *ACSM's Guidelines for Exercise Testing and Prescription*, Lippincott Williams & Wilkins, Baltimore, Md, USA, 2006.

[56] BioStat, *Comprehensive Meta-Analysis. (3.3)*, BioStat, Englewood Cliffs, NJ, USA, 2015.

[57] Microsoft Corporation, *Microsoft Excel. (2010)*, Microsoft Corporation, Redmond, Wash, USA, 2010.

[58] H. Schunemann, J. Brozek, and A. E. Oxman, *GRADE Handbook for Grading Quality of Evidence and Strength of Recommendation*, version 3.2, GRADE Working Group, 2009.

[59] Statistical Services Center, *SSC-Stat. (2.18)*, University of Reading, Statistical Services Center, Reading, UK, 2007.

[60] T. A. Poynton, *EZ Analyze (3.0)*, 2007.

[61] A. S. Alberga, B.-C. Farnesi, A. Lafleche, L. Legault, and J. Komorowski, "The effects of resistance exercise training on body composition and strength in obese prepubertal children," *The Physician and Sportsmedicine*, vol. 41, no. 3, pp. 103–109, 2013.

[62] J. G. B. Alves, C. R. Galé, E. Souza, and G. D. Batty, "Effect of physical exercise on bodyweight in overweight children: a randomized controlled trial in a Brazilian slum," *Cadernos de Saude Publica*, vol. 24, pp. S353–S359, 2008.

[63] M. Elloumi, E. Makni, O. B. Ounis et al., "Six-minute walking test and the assessment of cardiorespiratory responses during weight-loss programmes in obese children," *Physiotherapy Research International*, vol. 16, no. 1, pp. 32–42, 2011.

[64] N. J. Farpour-Lambert, Y. Aggoun, L. M. Marchand, X. E. Martin, F. R. Herrmann, and M. Beghetti, "Physical activity reduces systemic blood pressure and improves early markers of atherosclerosis in pre-pubertal obese children," *Journal of the American College of Cardiology*, vol. 54, no. 25, pp. 2396–2406, 2009.

[65] M. Hagströmer, K. Elmberg, S. Mårild, and M. Sjöström, "Participation in organized weekly physical exercise in obese adolescents reduced daily physical activity," *Acta Paediatrica*, vol. 98, no. 2, pp. 352–354, 2009.

[66] K. Karacabey, "The effect of exercise on leptin, insulin, cortisol and lipid profiles in obese children," *Journal of International Medical Research*, vol. 37, no. 5, pp. 1472–1478, 2009.

[67] A. S. Kelly, R. J. Wetzsteon, D. R. Kaiser, J. Steinberger, A. J. Bank, and D. R. Dengel, "Inflammation, insulin, and endothelial function in overweight children and adolescents: the role of exercise," *The Journal of Pediatrics*, vol. 145, no. 6, pp. 731–736, 2004.

[68] E. S. Kim, J.-A. Im, K. C. Kim et al., "Improved insulin sensitivity and adiponectin level after exercise training in obese Korean youth," *Obesity*, vol. 15, no. 12, pp. 3023–3030, 2007.

[69] H. J. Kim, S. Lee, T. W. Kim et al., "Effects of exercise-induced weight loss on acylated and unacylated ghrelin in overweight children," *Clinical Endocrinology*, vol. 68, no. 3, pp. 416–422, 2008.

[70] A. A. Meyer, G. Kundt, U. Lenschow, P. Schuff-Werner, and W. Kienast, "Improvement of early vascular changes and cardiovascular risk factors in obese children after a six-month exercise program," *Journal of the American College of Cardiology*, vol. 48, no. 9, pp. 1865–1870, 2006.

[71] E. C.-S. Murphy, L. Carson, W. Neal, C. Baylis, D. Donley, and R. Yeater, "Effects of an exercise intervention using Dance Dance Revolution on endothelial function and other risk factors in overweight children," *International Journal of Pediatric Obesity*, vol. 4, no. 4, pp. 205–214, 2009.

[72] Ö. Saygin and M. A. Öztürk, "The effect of twelve week aerobic exercise programme on health related physical fitness components and blood lipids in obese girls," *African Journal of Pharmacy and Pharmacology*, vol. 5, no. 12, pp. 1441–1445, 2011.

[73] G. Q. Shaibi, M. L. Cruz, G. D. C. Ball et al., "Effects of resistance training on insulin sensitivity in overweight Latino adolescent males," *Medicine & Science in Sports & Exercise*, vol. 38, no. 7, pp. 1208–1215, 2006.

[74] R. J. Sigal, A. S. Alberga, G. S. Goldfield et al., "Effects of aerobic training, resistance training, or both on percentage body fat and cardiometabolic risk markers in obese adolescents: the healthy eating aerobic and resistance training in youth randomized clinical trial," *JAMA Pediatrics*, vol. 168, no. 11, pp. 1006–1014, 2014.

[75] J.-K. Song, C. L. Stebbins, T.-K. Kim, H.-B. Kim, H.-J. Kang, and J.-H. Chai, "Effects of 12 weeks of aerobic exercise on body composition and vascular compliance in obese boys," *Journal of Sports Medicine and Physical Fitness*, vol. 52, no. 5, pp. 522–529, 2012.

[76] M.-X. Sun, X.-Q. Huang, Y. Yan et al., "One-hour after-school exercise ameliorates central adiposity and lipids in overweight Chinese adolescents: a randomized controlled trial," *Chinese Medical Journal*, vol. 124, no. 3, pp. 323–329, 2011.

[77] S. Tan, C. Yang, and J. Wang, "Physical training of 9- to 10-year-old children with obesity to lactate threshold intensity," *Pediatric Exercise Science*, vol. 22, no. 3, pp. 477–485, 2010.

[78] K. Watts, P. Beye, A. Siafarikas et al., "Exercise training normalizes vascular dysfunction and improves central adiposity in obese adolescents," *Journal of the American College of Cardiology*, vol. 43, no. 10, pp. 1823–1827, 2004.

[79] K. Watts, P. Beye, A. Siafarikas et al., "Effects of exercise training on vascular function in obese children," *Journal of Pediatrics*, vol. 144, no. 5, pp. 620–625, 2004.

[80] P. C. H. Wong, M. Y. H. Chia, I. Y. Y. Tsou et al., "Effects of a 12-week exercise training programme on aerobic fitness, body composition, blood lipids and C-reactive protein in adolescents with obesity," *Annals of the Academy of Medicine, Singapore*, vol. 37, no. 4, pp. 286–293, 2008.

[81] C. L. Davis, N. K. Pollock, J. L. Waller et al., "Exercise dose and diabetes risk in overweight and obese children: a randomized controlled trial," *The Journal of the American Medical Association*, vol. 308, no. 11, pp. 1103–1112, 2012.

[82] B. Gutin, S. Owens, G. Slavens, S. Riggs, and F. Treiber, "Effect of physical training on heart-period variability in obese children," *Journal of Pediatrics*, vol. 130, no. 6, pp. 938–943, 1997.

[83] B. Gutin, S. Owens, T. Okuyama, S. Riggs, M. Ferguson, and M. Litaker, "Effect of physical training and its cessation on percent fat and bone density of children with obesity," *Obesity Research*, vol. 7, no. 2, pp. 208–214, 1999.

[84] R. Maddison, L. Foley, C. Ni Mhurchu et al., "Effects of active video games on body composition: a randomized controlled trial," *American Journal of Clinical Nutrition*, vol. 94, no. 1, pp. 156–163, 2011.

[85] S. Kirk, M. Zeller, R. Claytor, M. Santangelo, P. R. Khoury, and S. R. Daniels, "The relationship of health outcomes to improvement in BMI in children and adolescents," *Obesity Research*, vol. 13, no. 5, pp. 876–882, 2005.

[86] G. Rose, "Sick individuals and sick populations," *International Journal of Epidemiology*, vol. 14, no. 1, pp. 32–38, 1985.

[87] M. M. Haby, T. Vos, R. Carter et al., "A new approach to assessing the health benefit from obesity interventions in children and adolescents: the assessing cost-effectiveness in obesity project," *International Journal of Obesity*, vol. 30, no. 10, pp. 1463–1475, 2006.

[88] J. H. Littell, J. Corcoran, and V. Pillai, *Systematic Reviews and Meta-Analysis*, Oxford University Press, New York, NY, USA, 2008.

[89] M. Ho, S. P. Garnett, L. A. Baur et al., "Impact of dietary and exercise interventions onweight change andmetabolic outcomes in obese children and adolescents: a systematic review and meta-analysis of randomized trials," *JAMA Pediatrics*, vol. 167, no. 8, pp. 759–768, 2013.

[90] M. Hagströmer, P. Oja, and M. Sjöström, "Physical activity and inactivity in an adult population assessed by accelerometry," *Medicine and Science in Sports and Exercise*, vol. 39, no. 9, pp. 1502–1508, 2007.

[91] M. Teran-Garcia, T. Rankinen, and C. Bouchard, "Genes, exercise, growth, and the sedentary, obese child," *Journal of Applied Physiology*, vol. 105, no. 3, pp. 988–1001, 2008.

[92] H. Oude Luttikhuis, L. Baur, H. Jansen et al., "Interventions for treating obesity in children," *Cochrane Database of Systematic Reviews*, vol. 21, no. 1, 2009.

[93] K. L. Matson and R. M. Fallon, "Treatment of obesity in children and adolescents," *Journal of Pediatric Pharmacology and Therapeutics*, vol. 17, pp. 45–57, 2012.

[94] A. Zohrabian, "Clinical and economic considerations of antiobesity treatment: a review of orlistat," *ClinicoEconomics and Outcomes Research*, vol. 2, no. 1, pp. 63–74, 2010.

[95] R. M. Viner, Y. Hsia, T. Tomsic, and I. C. K. Wong, "Efficacy and safety of anti-obesity drugs in children and adolescents: systematic review and meta-analysis," *Obesity Reviews*, vol. 11, no. 8, pp. 593–602, 2010.

[96] Food and Drug Administration, *What is a Serious Adverse Event?* United States Food and Drug Administration, Silver Spring, Md, USA, 2014.

[97] Centers for Disease Control and Prevention, *How Much Physical Activity Do Children Need?* Centers for Disease Control and Prevention, Atlanta, Ga, USA, 2013.

[98] Physical Activity Guidelines Advisory Committee, *Physical Activity Guidelines Advisory Report*, US Department of Health and Human Services, Washington, DC, USA, 2008.

[99] S. R. Daniels, S. G. Hassink, and Committee on Nutrition, "The role of the pediatrician in primary prevention of obesity," *Pediatrics*, vol. 136, no. 1, pp. e275–e292, 2015.

[100] M. Inokuchi, N. Matsuo, J. I. Takayama, and T. Hasegawa, "BMI z-score is the optimal measure of annual adiposity change in elementary school children," *Annals of Human Biology*, vol. 38, no. 6, pp. 747–751, 2011.

[101] T. J. Cole, M. S. Faith, A. Pietrobelli, and M. Heo, "What is the best measure of adiposity change in growing children: BMI, BMI %, BMI z-score or BMI centile?" *European Journal of Clinical Nutrition*, vol. 59, no. 3, pp. 419–425, 2005.

[102] L. Kakinami, M. Henderson, A. Chiolero, T. J. Cole, and G. Paradis, "Identifying the best body mass index metric to assess adiposity change in children," *Archives of Disease in Childhood*, vol. 99, no. 11, pp. 1020–1024, 2014.

[103] M. C. Reade, A. Delaney, M. J. Bailey, and D. C. Angus, "Bench-to-bedside review: Avoiding pitfalls in critical care meta-analysis—funnel plots, risk estimates, types of heterogeneity, baseline risk and the ecologic fallacy," *Critical Care*, vol. 12, no. 4, article 220, 2008.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 436938, 8 pages
http://dx.doi.org/10.1155/2015/436938

Research Article

The Prevalence of Skilled Birth Attendant Utilization and Its Correlates in North West Ethiopia

Mulunesh Alemayehu[1] and Wubegzier Mekonnen[2]

[1]*Department of Public Health, College of Medicine and Health Sciences, Debre Markos University, P.O. Box 269, Debre Markos, Ethiopia*
[2]*School of Public Health, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia*

Correspondence should be addressed to Mulunesh Alemayehu; mulu2019@gmail.com

Received 26 February 2015; Accepted 7 July 2015

Academic Editor: Stefan K. Lhachimi

The low utilization of skilled birth attendants sustained high maternal mortality. The aim of this study was to assess its magnitude and correlates in Northwest Ethiopia. A study was conducted on 373 randomly selected women who gave birth in the 12 months preceding the survey. Correlates were identified using binary logistic regression. Skilled birth attendance was 18.8%. Inability to perform cultural practices in health facilities (65.5%), expecting smooth delivery (63.4%), and far distance (62%) were the main barriers. Women with urban residence (AOR = 5.46: 95% CI [2.21–13.49]), primary (AOR = 2.10: 95% CI [0.71–6.16]) and secondary-plus (AOR = 6.12: [1.39–26.92]) educational level, four-plus ANC visits (AOR = 17.33: 95% CI [4.22–71.29]), and proximity to health centers (AOR = 5.67: 95% CI [1.47–25.67]) had higher odds of using skilled birth attendants though women with no labor complications had lower odds (AOR = 0.02: 95% CI [0.01–0.05]). Skilled birth attendance use was low. Urban residence, primary-plus level of education, frequent ANC visits, living nearby the health centers, and a problem during labor were positively correlated with skilled birth attendance utilization. Stakeholders should enhance girls' education beyond primary level and ANC services and shorten distances to health facilities.

1. Background

Though pregnancy and childbirth are a natural phenomenon and is often an eventful process which most women aspire to have at some point in their lives, this normal life affirming process carries its own risk and complications to both the mother and the newborn [1, 2]. Globally over half a million reproductive age women (15–49 years) die every year from pregnancy and childbirth complications and 300 million women suffer from debilitating injuries [3, 4].

Many researchers documented that more than three-fourths of maternal deaths are related to direct obstetric causes, such as haemorrhage, sepsis, abortion, ruptured uterus, and hypertensive diseases of pregnancy which are easily preventable and treatable, and 77% of deaths occur during or soon after childbirth (within 24 hours) [3, 5]. Eighty-eight to ninety-eight percent of these problems are estimated to be avoidable, although over 99% of maternal deaths in Sub-Saharan Africa (SSA) could not be prevented [1, 2]. Maternal mortality in Ethiopia is the highest in the World with an estimated maternal mortality ratio of 676 deaths per 100,000 live births in 2011 which slightly increased from the 2005 maternal mortality ratio (MMR) level of 673 deaths per 100,000 live births [6, 7].

World leaders started to join their efforts together at the dawn of the millennium by the MDG through the safe motherhood initiative to decrease maternal morbidity and mortality globally [8]. Targets were set at the International Conference on Population and Development +5 (ICPD+5) to have more than 80% of deliveries assisted by skilled birth attendants globally by 2005, 85% by 2010, and 90% by 2015 [5]. In spite of all national and global efforts, the maternal and newborn morbidity or mortality indices have shown no change or only marginal reductions in the last five years indicating that MDGs targets by 2015 are unachievable using only current strategies [9]. One of the reasons for poor health

outcomes among women and newborn is not using modern health care services by sizable proportion of women [10–12].

Studies from both industrialized and developing countries indicated that maternal mortality has been generally low when a higher proportion of deliveries are attended by skilled birth attendants [13–16]. However, study result from Ethiopia, Malawi, and Tanzania showed that socioeconomic status, availability of facility, short labor duration, staff attitudes, lack of privacy, reproductive behavior, cultural traditions, and the patterns of decision making power within household are mainly responsible factors for low utilization of skilled birth attendants [17–19]. According to WHO, skilled birth attendants are health professionals who have basic midwifery and obstetric skills including nurses, midwives, and physicians (WHO, 2004).

In Ethiopia, only 6% of births were delivered in health facilities at 2000 and there is no significant increase in the proportions of health facility delivery service utilization as evidenced from the series of Ethiopian Demographic and Health Surveys (EDHS) conducted in 2000, 2005, and 2011 which was found to be 6%, 8%, and 10% respectively [6, 7, 20].

Institutional delivery service utilization in Amhara Region was only 10.1% according to the EDHS 2011 which is equivalent to the national level of 10% [7]. However, according to the service statistics report by Akansha Guagusa Woreda Health Department, skilled delivery coverage was 21.7% in 2013 [21].

Therefore, this study was very crucial to measuring the prevalence of skilled birth attendance utilization and its correlates in remote areas of the second populous country in Africa to enhance maternal health to thereby reduce maternal morbidity and mortality in Ethiopia. The result will also be very important for policy makers as documented in the reproductive health road map to improve maternal health in general and increase level of skilled birth attendance by promoting quality antenatal care, make mothers have skilled assistance during their delivery, strengthen capacity of community health workers and community development army to perform emergency obstetric care (EmOC).

2. Methods

This study was conducted in Akansha Guagusa *Woreda* (district) in Awi zone, Amhara Regional State of Ethiopia. The main town of the study district was located at a distance of 450 kilometers northwest of Addis Ababa. The *Woreda* was administratively divided into seven clusters, 29 rural and 4 urban *kebeles* (the lowest administrative units). Clusters were established according to availability of health center. The area coverage of the study is 79881.75 square kilometers and was inhabited by about 221,796 persons [22].

According to the *Woreda* Health Department report there were seven health centres and five private clinics. With regard to health professionals employed in the *woreda*, there were nine health officers, five B. Sc. nurses, 46 clinical nurses, 16 midwifery nurses, one pharmacy technologist, nine pharmacy technicians, two laboratory technologists, 11 laboratory technicians, and 80 health service extension workers (primary health workers who had 10 months training on mainly disease prevention and promotion but do not have obstetric skills) making a total of 179 health professionals. The same source indicated that 28.4% mothers had four or more ANC visits and 74% of children had taken Penta three immunization. The average number of expected deliveries per annum in the *Woreda* was 6520, that is, 2.9% of the total population [21].

A community based cross-sectional study was conducted in February 2014 among 373 women who had delivered in 12 months prior to the study.

The sample size was calculated using single population proportion formula considering the Amhara Region skilled birth attendance utilization level of 10.1% [5], a 95% confidence level, 4% margin of error, 1.5 design effects, and 10% nonresponse rate. First, Akansha Guagusa Woreda (district) was purposely selected from eight Woredas from Awi zone. Three clusters in the Woreda, from available seven, were selected randomly. Two kebeles from each of the selected three clusters were also drawn randomly which made it a total of six kebeles. In each cluster, there are five health posts and one health center which cater approximately 5000 and 25,000 households and people, respectively. The number of women who delivered in the past one year (about 6520 in size) was obtained from registries of Health Services Extension Workers (HSEWs) in selected kebeles. To select the estimated sample size of 373, we used a multistage sampling scheme. The sample size was then proportionally allocated (using probability proportionate to size technique) according to the size of women who delivered in the past one year in each selected kebele. Then, in each kebele, study mothers were randomly selected from the rosters of HSEWs. Finally, study participants were identified using HSEWs as field guides.

An interviewer-administered pretested and close ended questionnaire developed in the local language was used. Mothers were briefed about the purpose of the study and data were collected after a verbal informed consent. The data collection was done by 12 female nurses supervised by two senior supervisors with public health background. Data quality was assured through careful design of the questionnaire and training of the field staffs. Moreover, data were checked for completeness and consistency.

Data were entered in Epi Info and exported to SPSS for analysis. Bivariate and multivariate models using binary logistic regression were run to assess any relationship between independent variables and skilled birth attendant utilization. Crude and adjusted odds ratios were used to ascertain effect sizes for any association between the dependent and predictor variables while significance was determined using 95% confidence intervals. Independent variables found to be significant with p value less than 0.05 at the bivariate level were included in a multivariate logistic regression model to control for potential confounding variables.

Ethical approval was obtained from the Research Ethics Committee of the School of Public Health, Addis Ababa University, and approval letter was obtained from the *Woreda* Administration Council. The purpose of the study was explained to mothers and the survey was commenced after obtaining verbal consent. Confidentiality of information was

maintained by omitting any personal identifier from the questionnaires.

3. Results

Altogether, 373 mothers participated in the survey. Of these, more than three-fourths, 282 (75.6%), of them were from rural areas and almost all, 367 (98.39%), of them were Orthodox Christians. The main ethnic group was Agew, 290 (77.75%), and nearly one-third, 115 (30.83%), of them were in the age range of 25–29 years with a mean age of 30.41 (±5.56) years. With regard to educational status of respondents', nearly two-thirds, 233 (62.47%), of them could not read and write whereas only 54 (14.48%) of them had attained secondary educational level and above. Majority, 347 (93.03%), of them were currently married. Among married, 19 (5.5%) of respondents' husband were polygamous. Three quarters, 278 (74.53%), of the mothers were housewives and nearly half, 186 (49.87%), of them had five to seven household members. Concerning respondents' monthly income, majority, 124 (33.24%), of mothers were found in 701–1249 Ethiopian birrs income group which is less than two dollars per day (a dollar is approximately 20 ETB during our study period) (Table 1).

Analysis of the obstetric characteristics of study women showed that 165 (44.24%) of respondents were married before celebrating their 16th birthday and the mean age at first marriage was 16.38 (±3.77) years. More than a third, 138 (37.00%), got their first pregnancy before the age of 19 years with the mean age of 19.73 (±3.00). About 101 (27.08%) gave their first birth before celebrating their 19th year of age with the mean age at first childbirth of 20.52 (±2.99) years. Moreover, 145 (38.87%) of them had more than four children and 285 (76.41%) of them who recently gave birth to a child had received antenatal care visits. Among 285 mothers who had ANC visit, about 89 (31%) of them had only one visit while 45 (16%) had four and more antenatal visits (Figure 1). Out of the 285 pregnant mothers, only 150 (52.63%) of them started ANC follow-up during the fourth month of their recent pregnancy.

In addition, this study revealed that only 78 (20.91%) of the mothers gave birth to their last babies at health facilities including health posts where health extension workers (who do not have education and training about basic obstetric skills) are working, whereas only 70 (18.77%) of them were attended by skilled birth attendants with basic obstetric skills (with nursing and above level of training). Among institutional deliveries assisted by skilled birth attendants, 43 (61.43%) of them delivered by spontaneous vaginal deliveries while 21 (30.00%) and 6 (8.57%) of them were assisted to give birth using instruments and through caesarean section, respectively. On the other hand, 71 (19.0%) of women who gave birth 12 months preceding the survey had history of labor complications, mainly prolonged labor, vaginal bleeding, and severe headache.

Nevertheless, from mothers who were not attended by skilled birth attendants, 219 (58.71%) were assisted by family members and relatives, 60 (16.08%) by health extension workers, and 12 (3.22%) by traditional birth attendants, and 12 (3.22%) had to do it by themselves without any assistance.

Table 1: Sociodemographic characteristics of respondents in Ankasha Guagusa *Woreda*, Awi zone, Amhara Region, Northwest Ethiopia, 2014 (n = 373).

Characteristics of respondents	Numbers	Percent
Residence		
Rural	282	75.60
Urban	91	24.40
Religion		
Orthodox	367	98.39
Muslim	6	1.61
Ethnicity		
Agew	290	77.75
Amhara	83	22.25
Woman's education		
No education	233	62.47
Primary	86	23.06
Secondary and above	54	14.48
Age group		
≤24	53	14.21
25–29	115	30.83
30–34	110	29.49
≥35	95	25.47
Marital status		
In marital union	347	93.03
Not in union[†]	26	6.97
Polygamous		
No	328	94.5
Yes	19	5.5
Husband's education		
No education	192	55.33
Primary	88	25.36
Secondary and above	67	19.31
Woman's occupation		
Housewife	278	74.53
Others[††]	95	25.47
Husband's occupation		
Farmer	264	76.08
Others[††]	83	23.92
Household size		
≤4	130	34.85
5–7	186	49.87
≥8	57	15.28
Monthly income (Eth. Birr)[++]		
≤700	61	16.35
701–1249	124	33.24
1250–1799	121	32.44
≥1800	67	17.96
Source of income		
Women themselves	54	14.5
Family and relatives	52	13.9
Both women and husbands	178	47.7
Husbands	89	23.9

[†]Not in union: single, divorced, and widowed and [††]others: employee (government and private), petty traders, daily laborer, student, and handcrafter. [++]Monthly income is calculated by converting annual crop yields and other agricultural products they had in that year into monthly cash income using the then market price of each merchandise and ultimately dividing it by 12.

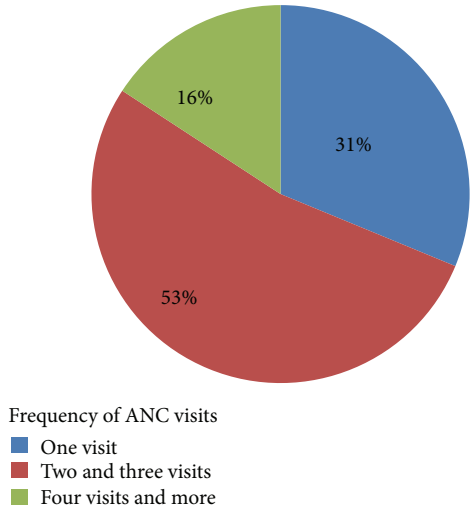


Figure 1: The distribution of respondents by their category of frequency of antenatal care visits in Akansha Guagusa *Woreda*, Awi zone, Amhara Region, Northwest Ethiopia, 2014.

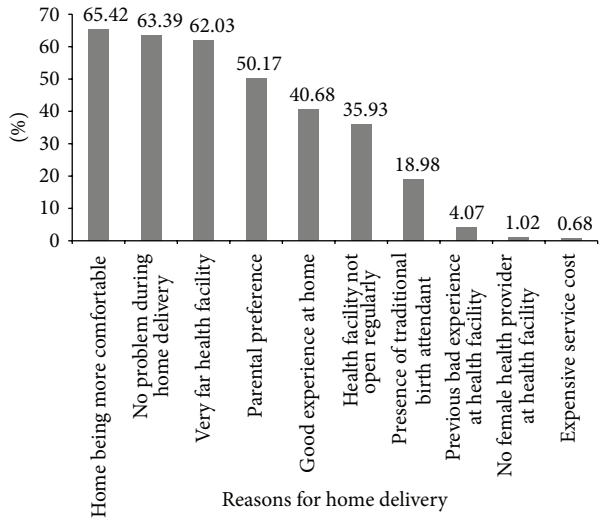


Figure 2: Reasons for home delivery among mothers in Akansha Guagusa *Woreda*, Awi zone, Amhara Region, Northwest Ethiopia, 2014 (n = 295).

About 295 (79.09%) mothers gave a variety of reasons for delivering at home; of these, 193 (65.42%) felt more comfortable to deliver at home since they will perform variety of cultural practices, 187 (63.39%) assumed no problem during home delivery as it is natural, 183 (62.03%) considered that health facilities are located far from their places of residences, and for 148 (50.17%) of the mothers their parents decided where they should have a baby and preferred home as their place of delivery (Figure 2).

Binary logistic regression models were fitted to identify correlates of skilled birth attendant utilization. The first model which attempted to calibrate sociodemographic correlates revealed that residence type has been found to be strongly and significantly associated with skilled birth attendance at both bivariate and multivariate levels.

Skilled birth attendance was COR = 7.15: 95% CI (4.07–12.28) times more likely to be utilized among urbanites compared with rural residents whereas when other confounding sociodemographic variables were controlled, the odds of skilled birth attendance utilization were AOR = 5.46: 95% CI (2.21–13.49) times higher among urban residents compared with their rural counterparts. On the other hand, skilled birth attendance was COR = 2.59: 95% CI (1.33–5.05) and COR = 9.83: 95% CI (4.95–19.52) times more likely to be utilized among mothers who completed primary and secondary and above level of education, respectively, compared with those who had never been into formal schooling. However, the significance of association vanishes at multivariate level for primary level of education. When all the other sociodemographic variables are controlled, the odds of skilled birth attendance were AOR = 6.12: 95% CI (1.39, 26.92) times higher among those mothers whose educational level was secondary and above compared with those with no education (Table 2).

The focus of the second model was to identify the reproductive health related correlates of skilled birth attendance utilization. Skilled birth attendant utilization among women who had four or more antenatal care visits was AOR = 17.33: 95% CI (4.22–71.29) times higher compared to those mothers with three and lower antenatal care visits (Table 3). On the other hand, women who had been living in an area nearby the health center were more likely to be assisted by skilled birth attendants in their parturition time than those mothers living closer to health post (AOR = 5.67: 95% CI [1.47–25.67]) (Table 3). Moreover, the odds of skilled birth attendant utilization among mothers who did not encounter labour complication during current delivery was 98% times less likely to deliver with the assistance of skilled professionals than women who encountered complication (AOR = 0.02: 95% CI [0.01–0.05]) (Table 3).

4. Discussion

This study showed that about one-third of the women in Akansha Guagusa Woreda had completed three visits of antenatal care services. However, institutional delivery and skilled birth attendant utilization were low, especially among women from rural areas, who had three and less antenatal care visits and were uneducated.

In this study less than one-fifth of the mothers (18.77%) were assisted by skilled birth attendants during their recent deliveries. Moreover, the study documented women who were not attended by anyone during delivery. The prevalence of skilled birth attendant utilization for the study district was somewhat similar to the findings of other studies done in Ethiopia, that is, Ephrtanagidim district, Kembata-Tembaro zone, and Raya Alamata district [23–25]. However, it was higher than the 2011 Ethiopian Demographic and Health Survey finding, a study done in Sekela district, Metekel zone of Ethiopia, and Afghanistan [7, 14, 23, 26]. This might be due to improvements in accessing and utilizing the service and community mobilization through health developmental

Table 2: Sociodemographic characteristics associated with skilled birth attendance in Akansha Guagusa *Woreda*, Awi zone, Amhara Region, North West Ethiopia, 2014.

Variables	SBA use		Crude odds ratio (95% CI)	Adjusted odds ratio (95% CI)
	No	Yes		
Residence				
Rural	253	29	1.00	1.00
Urban	50	41	**7.15 (4.07–12.58)*****	**5.46 (2.21–13.49)*****
Women's education				
No education	210	23	1.00	1.00
Primary	67	19	**2.59 (1.33–5.05)****	2.10 (0.71–6.16)
Secondary and above	26	28	**9.83 (4.95–19.52)*****	**6.12 (1.39–26.92)***
Polygamy				
No	271	57	**0.36 (0.14–0.96)***	0.36 (0.12–1.16)
Yes	12	7	1.00	1.00
Husbands' education				
No education	172	20	1.00	1.00
Primary	77	11	1.23 (0.56–2.69)	0.68 (0.23–1.98)
Secondary and above	34	33	**8.35 (4.29–16.25)*****	2.04 (0.45–9.33)
Women's occupation				
Housewife	245	33	1.00	1.00
Others^{+}	58	37	**4.74 (2.73–8.21)*****	0.07 (0.00–2.00)
Husbands' occupation				
Farmers	234	30	1.00	1.00
Others†	49	34	**5.41 (3.03–9.66)*****	7.72 (0.17–83.54)
Household size				
2–4	94	36	**5.07 (1.29–6.61)*****	0.83 (0.52–6.37)
5–7	156	30	1.06 (0.41–2.76)	0.76 (0.55–5.62)
8–12	53	4	1.00	1.00
Source of income				
Women themselves	36	18	**2.92 (1.92–6.61)****	1.43 (0.49–4.16)
Relatives and family	44	8	1.06 (0.41–2.76)	1.79 (0.58–5.49)
Both women and husbands	147	31	1.23 (0.61–2.49)	0.82 (0.34–1.99)
Husbands	76	13	1.00	1.00

Significance at $^{*}p$ value < 0.05, $^{**}p$ value < 0.01, and $^{***}p$ value < 0.001.
$^{+}$Others: governmental workers, housewives, merchants, and daily laborers for respondents.
†Others: governmental workers, merchants, and daily laborers for husbands.

army which has been implemented in the country in recent years. On the other hand, it was lower than those studies conducted outside of Ethiopia (Southern Tanzania, Nigeria, Namibia, and Nepal) and in Ethiopia (Bahir Dar, Woldia, and Kilte Awlaelo) [8, 12, 26–30]. The differences could be explained by the fact that women in those countries had better economic status, educational status, and antenatal care service coverage. However, Ethiopian studies were done in urban contexts for which ANC coverage is higher.

This study also showed the association of place of residence with skilled birth attendant utilization. In particular, mothers who resided in urban areas were more likely to get skilled birth attendance compared to their counterparts which is consistent with the findings of similar studies done in Ethiopia and other countries [4, 8, 13, 14, 24, 25, 29–35]. This might be due to the increased availability and access to health services and other infrastructures such as shorter distance to health facilities, better roads and transportation, and more information and education in urban than rural areas. The media promoting good health have been widely available in urban areas and rural residents might be influenced by traditional practices.

Educational status of mothers was also significantly associated with the utilization of skilled birth attendants. Mothers who had attained primary and secondary and above educational level were more likely to utilize skilled birth attendant than those mothers who were unable to read and write. This finding was in line with other studies that were done in Ethiopia and other developing countries [8, 12–14, 25, 31–34, 36, 37]. This could be due to the fact that educated women might have more access to written information and could adapt to modern cultural perspectives.

Table 3: Obstetric characteristics associated with skilled birth attendance in Akansha Guagusa *Woreda*, Awi zone, North West Ethiopia, 2014.

Variables	SBA use		Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)
	No	Yes		
Age at first marriage				
8–15 years	148	17	**0.21 (0.11–0.40)*****	1.64 (0.35–7.64)
16–18 years	84	14	**0.30 (0.15–0.60)****	0.76 (0.25–2.28)
19–27 years	71	39	**1.00**	1.00
Age at first pregnancy				
14–18 years	120	18	**0.53 (0.30–0.95)***	0.44 (0.11–1.67)
19–31 years	183	52	**1.00**	1.00
Parity				
1–4 children	173	55	**2.76 (1.49–5.09)****	1.82 (0.68–4.88)
5–12 children	130	15	1.00	1.00
Frequency of ANC visit				
One visit	78	11	1.00	1.00
Two and three visits	121	30	1.76 (0.83–3.71)	2.52 (0.81–7.83)
Four visits and more	22	23	**7.41 (3.14–17.52)*****	**17.33 (4.22–71.29)*****
Type of nearby HF				
Health post	253	29	1.00	1.00
Health center	50	41	**7.15 (4.07–12.58)*****	**5.67 (1.47–25.67)***
Travelling hour				
<1 hour	81	46	**9.94 (2.29–43.23)****	3.75 (0.54–25.67)
1 hour	187	22	2.06 (0.46–9.15)	1.50 (0.25–9.08)
>1 hour	35	2	1.00	1.00
Ever given birth at HF				
No	240	36	1.00	1.00
Yes	63	34	**3.60 (2.09–6.20)*****	0.77 (0.29–2.04)
Problem during current labour				
No	283	22	**0.04 (0.02–0.07)*****	**0.02 (0.01–0.05)*****
Yes	23	48	1.00	**1.00**

Significance at $^{*}p$ value < 0.05, $^{**}p$ value < 0.01, and $^{***}p$ value < 0.001.

Moreover, education empowers women so as to increase autonomy and self-confidence to make them decide for their better reproductive health needs.

Controlling other variables, frequency of antenatal care visits during their last pregnancy was also found to be a strong predictor of utilization of skilled birth attendants. Those mothers who had four antenatal care visits and above were more likely to deliver through the assistance of skilled birth attendants than those mothers who had three and less antenatal care visits. This finding was consistent with studies done in developing countries including Arsi, Alamata, Tigray, Dabat, Woldia, Mekelle, Fogera in Ethiopia, Nepal, Tanzania, and Cambodia [4, 8, 13, 14, 24, 25, 29–35]. This might be due to the fact that as the number of antenatal care visits increases, women will be acquainted with basic information on pregnancy and delivery related risks that require skilled providers' assistance.

Another important finding of this study was that mothers who did not face any labour complications during recent labour and delivery were 98% less likely to get skilled attendant during their delivery compared to those mothers who faced maternal complications. This implies that labouring mothers decide to give birth in health facilities and require the assistance of skilled birth attendants when they faced difficulties and when repeated trials at home failed. Similar findings were documented by other studies done in Sekela, North Shewa, Agemssa, Wollega, and Metekel zone [13, 14, 23, 38, 39].

5. Conclusions

Utilization of skilled birth attendant is still very low with a high number of deliveries being attended by unqualified persons at home. There are still women who deliver by themselves. Women's higher education, urban residence type, higher frequency of antenatal care visits, proximity of living to a health center, and encountering complications during current labour were found to be positively correlated with skilled birth attendant utilization. Therefore, the government should enhance secondary and above level of educational attainment among females, promote universal antenatal care follow-up service, and encourage mothers to utilize skilled birth attendants during pregnancy and delivery. Good referral linkages should be established, health facilities should be

staffed by professionals with basic obstetric skills and with necessary supplies and medicines, and different means of behavioral changing communication should be designed to improve the demand.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors' Contribution

Mulunesh Alemayehu designed the study, coordinated the data collection, entered data in preprepared EPI INFO template, and performed first draft data analysis and interpretation. Wubegzier Mekonnen designed the study and selected the study area. Moreover, Wubegzier Mekonnen checked the data analysis and interpretation, supervised the whole process, and made critical comments at each step in the research process and approved the final draft of the paper.

Acknowledgments

The authors would like to thank the Gates Institute at Johns Hopkins University for its financial support. Their deep gratitude also goes to their study participants who volunteered and took their time to give them all the relevant information for the study.

References

[1] WHO, "Maternal health and safe motherhood progress report update," 1991.

[2] WHO, *Reviewing Maternal Deaths and Complication to Make Pregnancy Safer; Beyond the Numbers*, Department of Reproductive Health and Research, Geneva, Switzerland, 2004.

[3] "Testing approaches for increasing skilled care durin g childbirth," Tech. Rep. The Skilled Care Initiative, Family Care International, New York, NY, USA, 2007.

[4] S. Rashid, Hashima-E-Nasreen, and M. A. Sarker, *Factors Influencing Utilization of Manoshi Delivery Centres in Urban Slums of Dhaka*, BRAC, Dhaka, Bangladesh, 2009.

[5] United Nations Population Fund (UNFPA), "Trends in maternal health in Ethiopia; challenges in achieving the MDG for maternal mortality," in *In-Depth Analysis of the EDHS 2000–2011*, United Nations Population Fund (UNFPA), Addis Ababa, Ethiopia, 2012.

[6] *Ethiopia Demographic and Health Survey 2005*, Central Statistical Agency, Addis Ababa, Ethiopia; ORC Macro, Calverton, Md, USA, 2006.

[7] Central Statistical Agency and ORC Macro, *Ethiopia Demographic and Health Survey 2011*, Central Statistical Agency, Addis Ababa, Ethiopia; ORC Macro, Calverton, Md, USA, 2012.

[8] E. Zere, D. Oluwole, J. M. Kirigia, C. N. Mwikisa, and T. Mbeeli, "Inequities in skilled attendance at birth in Namibia: a decomposition analysis," *BMC Pregnancy and Childbirth*, vol. 11, article 34, 2011.

[9] Nigerian Federal Ministry of Health and National Primary Health Care Development Agency [NPHCDA], *Accelerating Reduction in Maternal, Newborn and Child Mortality and Morbidity through Improved Access to Skilled Attendant at Birth*, Midwives Service Scheme (MSS), Abuja, Nigeria, 2009.

[10] S. Motherhood, "A challenge to midwifery practices," *World Health Forum*, vol. 12, no. 1, 1991.

[11] A. Binyam, *What Factors Determine Delivery Practices of Pregnant Women? Comparative Analysis of Findings from Behavioral Modeling and Follow Up oF Actual Practice*, 2005.

[12] B. Choulagai, S. Onta, N. Subedi et al., "Barriers to using skilled birth attendants' services in mid- and far-western Nepal: a cross-sectional study," *BMC International Health and Human Rights*, vol. 13, article 49, 2013.

[13] M. Abera, A. Gebremariam, and T. Belachew, "Predictors of safe delivery service utilization in Arsi Zone, South-East Ethiopia," *Ethiopian Journal of Health Sciences*, vol. 21, supplement 1, pp. 95–106, 2011.

[14] A. S. Teferra, F. M. Alemu, and S. M. Woldeyohannes, "Institutional delivery service utilization and associated factors among mothers who gave birth in the last 12 months in Sekela District, North West of Ethiopia: a community—based cross sectional study," *BMC Pregnancy and Childbirth*, vol. 12, article 74, 2012.

[15] L. N. Elizabeth and E. Joanne, "Achieving the millennium development goal of improving maternal health: determinants, interventions and challenges," HNP Discussion Paper, The International Bank for Reconstruction and Development, Washington, DC, USA, 2005.

[16] A. Jayaraman and S. Chandrasekhar, "Factors affecting maternal health care seeking behavior in Rwanda," DHS Working Paper 59, Demographic and Health Research, 2008.

[17] A. Tsehaynesh, *Assessment of utilization of skill birth attendant at delivery in Mekelle town, Northern Ethiopia [MPH thesis]*, 2007.

[18] R. Mikey, "'Women's groups' perceptions of maternal health issues in rural Malawi," *The Lancet*, vol. 368, pp. 1180–1188, 2006.

[19] D. Mushi, *Promoting Access to Obstetric Care through Community Volunteers in Mtwara Rural District, Tanzania*, Heidelberg University, Heidelberg, Germany, 2007.

[20] Central Statistical Authority and ORC Macro, *Ethiopia Demographic and Health Survey 2000*, Central Statistical Authority, Addis Ababa, Ethiopia; ORC Macro, Calverton, Md, USA, 2001.

[21] Department of Ankasha Guagusa Woreda, *Annual Health Service Delivery Report 2013*, Department of Ankasha Guagusa Woreda, 2013.

[22] Ethiopia Central Statistical Agency, *Population and Housing Census*, Ethiopia Central Statistical Agency, Addis Ababa, Ethiopia, 2007.

[23] G. Tura and A. Gebremariam, "Safe delivery utilization in Metkel zone, North West Ethiopia," *Ethiopian Science Journal of Public Health*, vol. 1, no. 1, pp. 18–23, 2013.

[24] F. G/Hiwot, *Assessment of factors for safe delivery service utilization among women of child bearing age in Ephratagidim District, North Shewa, Ethiopia [MPH thesis]*, School of Public Health, Addis Ababa University, Addis Ababa, Ethiopia, 2009.

[25] F. Tadese and A. Ali, "Determinants of use of skilled birth attendance among mothers who gave birth in the past 12 months in Raya Alamata District, North East Ethiopia," *Clinics in Mother and Child Health*, vol. 11, article 2, 2014.

[26] R. N. M. Mpembeni, J. Z. Killewo, M. T. Leshabari et al., "Use pattern of maternal health services and determinants of skilled care during delivery in Southern Tanzania: implications for achievement of MDG-5 targets," *BMC Pregnancy and Childbirth*, vol. 7, article 29, 2007.

[27] E. A. Envuladu, H. A. Agbo, S. Lassa, J. H. Kigbu, and A. I. Zoakah, "Factors determining the choice of a place of delivery among pregnant women in Russia village of Jos North, Nigeria: achieving the MDGs 4 and 5," *International Journal of Medicine and Biomedical Research*, vol. 2, no. 1, pp. 23–27, 2013.

[28] G. Abeje, M. Azage, and T. Setegn, "Factors associated with institutional delivery service utilization among mothers in Bahir Dar City administration, Amhara region: a community based cross sectional study," *Reproductive Health*, vol. 11, article 22, 2014.

[29] W. Awoke, J. Muhammed, and G. Abeje, "Institutional delivery service utilization in Woldia, Ethiopia," *Science Journal of Public Health*, vol. 1, no. 1, pp. 18–23, 2013.

[30] Y. A. Melaku, B. Weldearegawi, F. H. Tesfay et al., "Poor linkages in maternal health care services—evidence on antenatal care and institutional delivery from a community-based longitudinal study in Tigray region, Ethiopia," *BMC Pregnancy and Childbirth*, vol. 14, article 418, 2014.

[31] Z. B. Mengesha, G. A. Biks, T. A. Ayele, G. A. Tessema, and D. N. Koye, "Determinants of skilled attendance for delivery in Northwest Ethiopia: a community based nested case control study," *BMC Public Health*, vol. 13, no. 1, article 130, 2013.

[32] S. M. Tarekegn, L. S. Lieberman, and V. Giedraitis, "Determinants of maternal health service utilization in Ethiopia: analysis of the 2011 Ethiopian Demographic and Health Survey," *BMC Pregnancy and Childbirth*, vol. 14, no. 1, article 161, 2014.

[33] A. M. Mehari, "Levels and determinants of use of institutional delivery care services among women of childbearing age in Ethiopia: analysis of EDHS 2000 and 2005 data," DHS Working Papers 83, ICF International, Calverton, Md, USA, 2013.

[34] E. Desalegn, "Place of delivery after antenatal care: the case of Fogera district, Amhara region, North West, Ethiopia; 2013," *Journal of Gynecology and Obstetrics*, vol. 2, no. 1, pp. 1–6, 2014.

[35] S. Shiferaw, M. Spigt, M. Godefrooij, Y. Melkamu, and M. Tekie, "Why do women prefer home births in Ethiopia?" *BMC Pregnancy and Childbirth*, vol. 13, article 5, 2013.

[36] S. Yanagisawa, S. Oum, and S. Wakai, "Determinants of skilled birth attendance in rural Cambodia," *Tropical Medicine and International Health*, vol. 11, no. 2, pp. 238–251, 2006.

[37] M. Mayhew, P. M. Hansen, D. H. Peters et al., "Determinants of skilled birth attendant utilization in Afghanistan: a cross-sectional study," *The American Journal of Public Health*, vol. 98, no. 10, pp. 1849–1856, 2008.

[38] M. Øxnevad, *Perceptions and practices related to home based and facility based birth. A qualitative study from Agemssa, Ethiopia [M.S. thesis]*, Centre for International Health Faculty of Medicine and Dentistry University of Bergen, Bergen, Norway, 2011.

[39] A. Binyam, *What Factors Determine Delivery Practices of Pregnant Women? Comparative Analysis of Findings from Behavioral Modeling and Follow up of Actual Practice*, 2005.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 256205, 7 pages
http://dx.doi.org/10.1155/2015/256205

Research Article

Working Atmosphere and Job Satisfaction of Health Care Staff in Kenya: An Exploratory Study

Katja Goetz,[1] Michael Marx,[2] Irmgard Marx,[2] Marc Brodowski,[3] Maureen Nafula,[4] Helen Prytherch,[2] Irene K. E. Omogi Awour,[5] and Joachim Szecsenyi[1,3]

[1]*Department of General Practice and Health Services Research, University of Heidelberg, Vossstrasse 2, 69115 Heidelberg, Germany*
[2]*Evaplan at the University Hospital of Heidelberg, Ringstrasse 19b, 69115 Heidelberg, Germany*
[3]*AQUA-Institute for Applied Quality Improvement and Research in Health Care, Maschmühlenweg 8-10, 37073 Goettingen, Germany*
[4]*Institute of Health Policy, Management and Research, NHIF Building, 2nd Floor Ragati Road, Upperhill, P.O. Box 307-00202, Nairobi, Kenya*
[5]*Gesellschaft für Internationale Zusammenarbeit (GIZ), P.O. Box 41607, Nairobi 00100, Kenya*

Correspondence should be addressed to Katja Goetz; katja.goetz@med.uni-heidelberg.de

Received 4 May 2015; Accepted 9 September 2015

Academic Editor: Malgorzata Bala

Background. Job satisfaction and working atmosphere are important for optimal health care delivery. The study aimed to document working atmosphere and job satisfaction of health care professionals in Kenya and to explore associations between job satisfaction, staff characteristics, and working atmosphere. *Methods*. Data from the integrated quality management system (IQMS) for the health sector in Kenya were used. Job satisfaction was measured with 10 items and with additional 5 items adapted to job situation in Kenya. Working atmosphere was measured with 13 item questionnaire. A stepwise linear regression analysis was performed with overall job satisfaction and working atmosphere, aspects of job satisfaction, and individual characteristics. *Results*. Out of 832 questionnaires handed out, 435 questionnaires were completed (response rate: 52.3%). Health care staff indicated high commitment to provide quality services and low levels regarding the adequacy and functionality of equipment at their work station. The aspect "support of the ministry of health" ($\beta = 0.577$) showed the highest score of explained variance (32.9%) regarding overall job satisfaction. *Conclusions*. IQMS which also evaluates job satisfaction and working atmosphere of health care staff provides a good opportunity for strengthening the recruitment and retention of health care staff as well as improving the provision of good quality of care.

1. Introduction

"At the heart of each and every health system, the workforce is central to advancing health" [1]. Health care staff are crucial for health service delivery and the provision of quality care to patients. However, constraints such as limited career opportunities, insufficient workforce, and low remuneration are known to increase the risk that health care staff migrate from their countries but also within countries such as from faith-based to public hospitals [2–5]. Unattractive working conditions of health care staff in combination with increased risk of occupational exposure or political violence have been identified as critical push factors that cause health care workers based on low- and middle-income countries to try and migrate abroad, including OECD countries [6].

The shortages of health care staff in low- and middle-income countries are dramatic. In America 24.8 health care workers per 1000 population are available, whereas, in Africa, where the burden of disease is higher, there are only 2.3 health care workers per 1,000 population [7]. These shortages have important social and economic costs to the countries concerned. It has been demonstrated that the financial loss to a country caused by the emigration of a single nurse is US$ 338,868 [6]. This is without any attempt to capture the financial value of the social costs. The out-migration of health care staff results in a loss of institutional memory and absolute shortages of much needed skills and experience. For a low-income country like Kenya it is particularly cost-intensive to continually invest in the training of health care staff and policy makers are keen to find ways to strengthen

the so-called pull factors like career development, improvement of working conditions, and greater financial rewards to retain and motivate their health workforce [8]. A detailed description about the health system in Kenya is presented by the Global Health Observatory and WHO and within a report of the National Coordinating Agency for Population and Development, Kenya [9, 10].

Maternal and reproductive health outcomes are important markers of the functionality of health systems. Whilst progress has been made, achievement of the Millennium Development Goals (MDGs) for maternal and reproductive health by the year 2015 will remain elusive [11]. Weak health systems impede the performance of health care staff and prevent the delivery of quality care. Low levels of training, insufficient supervision, support, and recognition all serve to erode the motivation of health care staff [12], whilst the overall lack of staff and difficult working conditions leave health workers particularly those that provide maternal health care at high risk of burnout [13]. In the frame of the Millennium Development Goals maternal services have been prioritized and in many settings they are provided free of charge at the point of delivery which can cause demand for services to be high [14]. Moreover, maternal health has long been recognized to be an area where health workers and communities including providers of traditional health care have to work together, which makes relationships between the formal health system and the community particularly important.

Despite the important attention human resource issues have received in recent years, health workers in many low-income settings report their superiors taking little interest in their job satisfaction and work environment, although relatively simple and cost-effective steps can be taken to improve them. It was shown that recognition, responsibility, and training are the main motivational factors for retention of health workers [15]. These motivational factors are closely linked with the perception of job satisfaction. Therefore, the aim of this study was to evaluate the job satisfaction of health care staff working in maternal and reproductive health care in Kenya and to explore associations between job satisfaction, staff characteristics, and working atmosphere.

2. Methods

2.1. Design and Participants. In the frame of the collaboration between the Gesellschaft für Internationale Zusammenarbeit (GIZ) and the Kenyan Ministry of Health, a consortium including evaplan GmbH at the University of Heidelberg and the AQUA Institute in Germany and the Institute of Health Policy, Management and Research (IHPMR) in Nairobi was contracted to develop and implement an integrated quality management system (IQMS) that was initially focused on facilities providing reproductive and maternal health services. The development of the IQMS is described by Herrler et al. [16] and was inspired by the European Practice Assessment (EPA) methodology [17]. EPA represents a quality management program including validated instruments based on quality indicators for assessing practice management aiming at continuous improvement process [18].

Once the quality assessment tool IQMS had been developed it was field-tested at two facilities between January and February 2013. Public health authorities from different districts supported by GIZ (Kisumu East, Vihiga, Bondo, Butere, and Gucha) were asked to recruit interested health facilities. In total 36 health facilities responded to this call with a letter of motivation. Finally, 10 health facilities were selected to participate in this study. Inclusion criteria included the facilities provided services for the prevention of mother to child transmission and for survivors of gender-based violence.

One part of the IQMS focused upon evaluating the job satisfaction and working atmosphere of health care staff. After a pilot study, data were collected from health care staff working in 10 health facilities providing maternal and reproductive health care (district hospitals and health centres) across the aforementioned districts in Kenya.

2.2. Procedure and Measurement. To measure job satisfaction, aspects of working atmosphere, and other individual characteristics all participants completed a written questionnaire. Staff were encouraged to fill out the survey whilst the project coordinator was still at the facility. A collection box was left to also gather the responses from staff that were absent that day. All responses were anonymous. The questionnaire included the following items: structural questions about gender and age, how many hours a week the health care staff were contracted to work, and how many years they had worked in the facility. Job satisfaction was measured with 15 items that included a modified version of the Warr-Cook-Wall (WCW) job satisfaction scale developed by Warr et al. and additional 5 job satisfaction items for the specific job situation in Kenya [19]. The WCW-instrument measures overall job satisfaction and satisfaction with nine aspects of work which can be aligned to the theoretical background of the Two-Factor Theory found in Herzberg and colleagues [20]. These items were "amount of variety in job," "opportunity to use abilities," "amount of responsibility," "recognition for work," "freedom of working method," "physical working condition," "hours of work," "income," and "relations with colleagues." The additional 5 job satisfaction items for the specific job situation in Kenya were "quality of materials and equipment," "time needed to receive the results of laboratory tests," "opportunities to attend training," "opportunities for career advancement," and "support from ministry of health." Each item was rated on a 7-point Likert scale (1 = extreme dissatisfaction to 7 = extreme satisfaction). A higher overall mean score indicates higher satisfaction with job. The aspects of working atmosphere were measured by 13 items such as "the responsibilities within the team are clear," "offering suggestions to improve the facility," "I can speak my mind without fear of reprisal," "the work atmosphere in the team is good," and "communication with management is enough and frequent." Rating options ranged from "1" (strongly disagree) to "5" (strongly agree). A higher overall mean score indicates better perception of working atmosphere.

2.3. Data Analysis. The analyses were performed using SPSS version 20.0 (SPSS Inc., IBM, USA). Continuous data was

Table 1: Description of the study population.

Characteristics	Health care staff (n = 435)
Gender[a]	
Male	174 (40.0%)
Female	250 (57.5%)
Age, years; mean (SD)	35.2 (10.6)
Contracted working hours per week; mean (SD)	46.7 (21.3)
Time period of employment; mean (SD)	5.9 (5.6)
Type of health care staff[a]	
Nurse	120 (27.6%)
Cleaner	50 (11.5%)
Clinical officer	32 (7.4%)
Secretary/administration	23 (5.3%)
Laboratory technician	18 (4.1%)
Security	13 (3.0%)
Doctor	12 (2.8%)
Pharmacist	9 (2.1%)
Midwife	7 (1.6%)
Public health officer	3 (0.7%)
Other	94 (21.6)

[a]n various due to missing data; SD standard deviation.

summarized using means and standard deviations. Categorical data was presented as frequency counts and percentages. Furthermore, a stepwise linear regression analysis was performed with overall job satisfaction as the dependent outcome variable and the different aspects of satisfaction with work, working atmosphere, and individual characteristics as the potential predictors. An alpha level of $P < 0.05$ was used for tests of statistical significance.

2.4. Ethical Approval. Ethical clearance was obtained from the Institutional Research Ethics Committee at Moi University, Kenya. All participants gave their informed consent. All participants received a briefing of the findings and have been involved in subsequent steps to improve the working conditions at these facilities.

3. Results

3.1. Description of the Study Sample. The 10 participating facilities can be described as follows: 3/10 were located in urban areas, 2 were located in rural areas (information from 5 facilities was missing), 4/10 were assigned as health center, 5 were assigned as district hospital (information from 1 facility was missing), and 7/10 were government-owned and financed.

Out of 832 questionnaires handed out in 10 facilities, 435 questionnaires of health care staff were returned (response rate: 52.3%). Table 1 shows the characteristics of the participating health care staff and their professional group affiliation. The participants who returned their questionnaire had a mean age of 35.2 years (SD = 10.6). The mean of contracted working hours per week was 46.7 (SD = 21.3). The mean time period of employment at the facility was nearly 6 years (SD = 5.6). A high proportion of health care staff was nurses (27.6%).

Table 2 shows the results from the job satisfaction scale and the additional 5 job satisfaction items for the special job situation in Kenya. Health care staff in Kenya were highly satisfied with "colleagues" (mean = 5.21) and "recognition for work" (mean = 4.92) and less satisfied with "remuneration" (mean = 3.43) and "the needed materials and equipment" (mean = 3.71). Furthermore, the health care staff had a satisfactory feeling regarding their overall job satisfaction (mean = 4.87).

The different aspects of working atmosphere concerning health care staff in Kenya are presented in Table 3. They highly agree with the factors "my colleagues are committed to doing quality work" (mean = 3.99), "offering suggestions to improve the facility" (mean = 3.93), and "a good collaboration between my facility and community health workers" (mean = 3.91). The health care staff less agree with "the equipment in my work station is adequate and in good working condition" (mean = 2.57) and "a good collaboration between my facility and traditional birth attendants" (mean = 2.81).

Table 4 shows the stepwise regression analysis of the individual characteristics, working atmosphere, and satisfaction with aspects of work on overall satisfaction for health care staff in Kenya in this study. A model with 5 steps was carried out and explained more than 47% (R^2 = 0.476) of the variance on the dependent variable "overall job satisfaction." All regression coefficients of the items in the stepwise regression analysis were statistically significant. These were four items of job satisfaction: "the support of the ministry of health," "remuneration," "needed materials and equipment," and "physical working condition" and one item of working atmosphere "responsibilities within the team are clear." In the first step of the stepwise regression analysis the item "the support of ministry of health" showed the highest score (R^2 = 0.329) of explained variance. The last step was reported in Table 4.

4. Discussion

In the last years there has been a growing interest in the working situation of health care staff in low- and middle-income countries [21, 22]. As there is limited published research on the relationship between job satisfaction and working atmosphere, this study contributes additional knowledge to a field of increasing importance.

Our study population presented a broad range of health care staff. The largest group was the nurses with a proportion of 27.6%. However, a high proportion of participants did not mention their professional group. There is scope important to examine the working situation of staff other than nurses in further studies. The current study gives a first impression of job satisfaction of health care staff in Kenya. In our study population we found young staff combined with a high number of contracted working hours per week and with a short time period of employment of health care staff in

Table 2: Descriptive statistics of job satisfaction of health care staff in Kenya (n = 435).

Rate your satisfaction with following statements...[1]	Mean (SD)	CI (95%)
The physical working condition	4.37 (2.02)	4.12–4.56
The freedom to choose your own method of working	3.81 (2.20)	3.62–4.09
Your colleagues and fellow workers	5.21 (1.88)	5.00–5.40
The recognition of work	4.92 (2.08)	4.71–5.15
The amount of responsibility you are given	4.67 (2.20)	4.50–4.97
Your remuneration	3.43 (2.18)	3.32–3.79
The opportunity to use your abilities	4.73 (2.15)	4.54–5.00
Your hours of work	4.63 (2.31)	4.33–4.82
The amount of variety in your job	4.26 (2.13)	4.13–4.59
The materials and the equipment you need	3.71 (2.17)	3.46–3.92
The time needed to receive the results of laboratory tests	4.40 (2.10)	4.14–4.59
The opportunities to attend training	4.05 (2.37)	3.85–4.36
The opportunities for career advancement	3.88 (2.28)	3.65–4.14
The support from ministry of health	4.07 (2.17)	3.85–4.32
Taking everything into consideration, how do you feel about your job	4.87 (2.01)	4.68–5.11

[1]Range from 1 "very dissatisfied" to 7 "very satisfied."

Table 3: Descriptive statistics of working atmosphere of health care staff in Kenya (n = 435).

Rate your agreement with following statements...[1]	Mean (SD)	CI (95%)
Supervision is provided in a supportive manner at this facility	3.71 (1.40)	3.59–3.95
Responsibilities within the team are clear	3.81 (1.37)	3.61–3.98
I feel encouraged to offer suggestions to improve the facility	3.93 (1.42)	3.81–4.17
My suggestions for improvement are taken seriously	3.17 (1.50)	3.07–3.47
The working atmosphere in the team is good	3.64 (1.36)	3.37–3.75
My colleagues are committed to doing quality work	3.99 (1.27)	3.75–4.10
At the workstations I can speak my mind without fear of reprisal	3.51 (1.54)	3.23–3.65
The ministry headquarters keep employees informed about official matters	3.31 (1.53)	3.09–3.51
The communication with management is enough and frequent	3.29 (1.48)	3.07–3.46
The equipment in my work station is adequate and in good working condition	2.57 (1.47)	2.26–2.64
There is a good collaboration between my facility and traditional birth attendants	2.81 (1.53)	2.57–2.97
There is a good collaboration between my facility and community health workers	3.91 (1.28)	3.72–4.06
There is a good collaboration between my facility and community midwives	3.22 (1.49)	3.02–3.41

[1]Range from 1 "strongly disagree" to 5 "strongly agree."

the facility. A possible explanation could be that positions may be vacant and only few staff are working with a high number of contracted working hours. This would reinforce the impression of the high workload being shouldered by very few staff in Kenya. In general, a high turnover rate of health care staff in Kenya can be observed [23]. Between 1999 and 2007 over 22% of the nurses in Kenya applied to out-migrate mainly to the United States or United Kingdom [6, 23]. Unattractive working condition, limited career opportunities, and weak health care systems produce dissatisfaction and demotivation with work which are reasons for leaving the country [24–26]. The results of our study showed that health care staff were least satisfied with their remuneration, material and equipment, freedom of working methods, and possibilities for career development. Different studies reported similar results about job satisfaction of health care staff in low- and middle-income countries [12, 21, 26, 27]. An improvement of the working situation of health care staff should be an important aim to strengthen a health care system [28]. Investments in training, retention, and sustenance of skilled health care workers in combination with recognition of their performance are a promising approach [29]. Furthermore, it was shown that recognition for work seems to be a strong predictor for job satisfaction of health care staff [30]. Moreover, as assumed by our data the collaboration between the facility and traditional birth attendants should be improved. It has been strongly recommended that qualified and motivated staff are available to build relationships and facilitate good cooperation with community members like in our study and traditional birth attendants [28]. Reasons for the poor cooperation between facility and traditional birth attendants should be examined in further studies.

Table 4: Associations of individual characteristics, working atmosphere, and satisfaction of aspects of work of health care staff on overall job satisfaction (results of stepwise linear regression analysis, under specification of standardized beta coefficient, $\alpha = 5\%$).

	Step 1	Step 2	Step 3	Step 4	Step 5
The support of the ministry of health	0.577	0.440	0.382	0.326	0.281
Your remuneration		0.313	0.284	0.218	0.198
Responsibilities within the team are clear			0.221	0.193	0.176
The materials and the equipment you need				0.193	0.170
The physical working condition					0.137
Pseudo R^2	**0.329**	**0.405**	**0.446**	**0.467**	**0.476**

Our data showed that relationships with colleagues have a high impact on satisfaction with job. This result does not surprise and was commonly observed internationally before [31–33]. However, collegial relationships at the workplace have been shown to mitigate an excessive workload. It was demonstrated that a high satisfaction with colleagues could reduce stress leading to more positive work environment [34]. In addition, this underlines the importance of organizing teamwork as in shifts to allow staff time-off, to build relationships with colleagues and especially with communities. It was shown that working in teams could motivate health care staff on the one hand and on the other hand it could increase effectiveness, responsiveness, and job satisfaction. Moreover, it is a cost-effective motivational factor for personnel retention [28].

Effective human resource strategies which support working condition of health care staff in low- and middle-income countries are necessary and should focuse on different levels: health system (macro level), health facility (micro level), and health workers (individual level) [28]. As a result of our study, we found a strong association between overall job satisfaction and the support by the ministry of health. It could be assumed that supporting the facilities by the ministry influences the perception of job satisfaction by staff. With the implementation of the Emergency Hiring Program supported by the government more staff could be hired, trained, and deployed [35, 36]. On the level of health facility concerning our results investing in equipment and material could be recommended. An important factor related to out-migration pertains to workforce concerns regarding occupational risks associated with the availability of safety equipment [6]. Paying attention to protecting staff from occupational risk is also known to make staff feel more appreciated [37].

The study benefited from the usage of internationally validated measures for the evaluation of job satisfaction [17]. The results of this study have to be seen in relation to maternal and reproductive health service and no general conclusions of the results regarding health care staff could be drawn. Moreover, our sample may not be representative for all facilities that provide maternal and reproductive health care in Kenya because we only involved facilities that took part in the IQMS. Participation depended upon the facility management's interest to improve the quality of care as articulated in a letter of motivation. The recruitment of health facilities was conducted by public health authorities. Therefore, a potential selection bias is indicated. Moreover, we cannot analyze the job satisfaction of each professional group separately because there was no balanced proportion between the different health care professional groups. Furthermore, we cannot analyze the data regarding facility description because there is too much information which is missing. In addition, this was an exploratory study; P values should be interpreted carefully. Due to the exploratory character of this study a sample size calculation could not be done. Significant results might be due to chance and will need to be confirmed in further targeted studies.

5. Conclusions

Job satisfaction and working atmosphere are important indicators for recruitment and retention of health care staff but also for the provision of good quality of care. Financial and nonfinancial incentives serve as motivational factors like increasing the remuneration, to invest in continuously career development or to improve the work equipment. However, these incentives will only have the desired effect, if they are introduced in supportive work environments. Therefore, an implementation of IQMS for maternal and reproductive health care in Kenya which also evaluates job satisfaction and working atmosphere of health care staff in these facilities provides a good opportunity to develop improvement strategies. Such improvement strategies should be developed with the involvement of policy makers, health facility managers, and health workers themselves.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the participating facilities and their health care staff for filling in the questionnaire and taking an active part in the introduction of the IQMS in Kenya.

References

[1] World Health Organization, *The World Health Report 2006: Working Together for Health*, World Health Organization, Geneva, Switzerland, 2006, http://www.who.int/whr/2006/whr06_en.pdf.

[2] J. M. Gross, P. L. Riley, R. Kiriinya et al., "The impact of an emergency hiring plan on the shortage and distribution of

nurses in Kenya: the importance of information systems," *Bulletin of the World Health Organization*, vol. 88, no. 11, pp. 824–830, 2010.

[3] D. Dovlo, "Migration of nurses from Sub-Saharan Africa: a review of issues and challenges," *Health Services Research*, vol. 42, no. 3, pp. 1373–1388, 2007.

[4] G. Dussault and M. C. Franceschini, "Not enough there, too many here: understanding geographical imbalances in the distribution of the health workforce," *Human Resources for Health*, vol. 4, article 12, 2006.

[5] P. Tabatabai, H. Prytherch, I. Baumgarten, O. M. E. Kisanga, B. Schmidt-Ehry, and M. Marx, "The internal migration between public and faith-based health providers: a cross-sectional, retrospective and multicentre study from southern Tanzania," *Tropical Medicine and International Health*, vol. 18, no. 7, pp. 887–897, 2013.

[6] J. M. Kirigia, A. R. Gbary, L. K. Muthuri, J. Nyoni, and A. Seddoh, "The cost of health professionals' brain drain in Kenya," *BMC Health Services Research*, vol. 6, article 89, 2006.

[7] S. Naicker, J. Plange-Rhule, R. C. Tutt, and J. B. Eastwood, "Shortage of healthcare workers in developing countries—Africa," *Ethnicity & Disease*, vol. 19, no. 1, supplement1, pp. S1-60–S1-64, 2009.

[8] J. B. Eastwood, R. E. Conroy, S. Naicker, P. A. West, R. C. Tutt, and J. Plange-Rhule, "Loss of health professionals from sub-Saharan Africa: the pivotal role of the UK," *The Lancet*, vol. 365, no. 9474, pp. 1893–1900, 2005.

[9] Word Health Organization, "Global Health Observatory (GHO) data, Kenya: country profiles," September 2015, http://www.who.int/gho/countries/ken/country_profiles/en/.

[10] National Coordinating Agency for Population and Development (NCAPD), Ministry of Health (MOH), Central Bureau of Statistics (CBS), and ORC Macro, *Kenya Service Provision Assessment Survey 2004*, National Coordinating Agency for Population and Development, Ministry of Health, Central Bureau of Statistics, ORC Macro, Nairobi, Kenya, 2005, http://dhsprogram.com/pubs/pdf/SPA8/SPA8.pdf.

[11] World Health Organization, *World Health Statistics 2015*, World Health Organization, Geneva, Switzerland, 2015.

[12] M. Willis-Shattuck, P. Bidwell, S. Thomas, L. Wyness, D. Blaauw, and P. Ditlopo, "Motivation and retention of health workers in developing countries: a systematic review," *BMC Health Services Research*, vol. 8, article 247, 2008.

[13] V. C. Thorsen, A. L. T. Tharp, and T. Meguid, "High rates of burnout among maternal health staff at a referral hospital in Malawi: a cross-sectional study," *BMC Nursing*, vol. 10, article 9, 2011.

[14] World Health Organization, *Towards Reaching the Health-Related Millennium Development Goals: Progress Report and the Way Forward*, WHO Regional Office for Africa, 2010, http://www.afro.who.int/index.php?option=com_content&view=article&id=2871:towards-reaching-the-health-related-mdgs&catid=1893&Itemid=2673.

[15] M. Dieleman, J. Toonen, H. Touré, and T. Martineau, "The match between motivation and performance management of health sector workers in Mali," *Human Resources for Health*, vol. 4, article 2, 2006.

[16] C. Herrler, A. Bramesfeld, M. Brodowski et al., "Integrated Quality Management System (IQMS): ein Modell zur Förderung der Qualität der reproduktiven Gesundheitsversorgung im ländlichen Kenia," *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 2015.

[17] Y. Engels, M. Dautzenberg, S. Campbell et al., "Testing a European set of indicators for the evaluation of the management of primary care practices," *Family Practice*, vol. 23, no. 6, pp. 137–147, 2006.

[18] J. Szecsenyi, S. Campbell, B. Broge et al., "Effectiveness of a quality-improvement program in improving management of primary care practices," *CMAJ*, vol. 183, no. 18, pp. E1326–E1333, 2011.

[19] P. Warr, J. Cook, and T. Wall, "Scales for the measurement of some work attitudes and aspects of psychological well-being," *Journal of Occupational Psychology*, vol. 52, no. 2, pp. 129–148, 1979.

[20] F. Herzberg, B. Mausner, and B. B. Snyderman, *The Motivation to Work*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1967.

[21] L. Fogarty, Y. M. Kim, H.-S. Juon et al., "Job satisfaction and retention of health-care providers in Afghanistan and Malawi," *Human Resources for Health*, vol. 12, no. 1, article 11, 2014.

[22] W. Chimwaza, E. Chipeta, A. Ngwira et al., "What makes staff consider leaving the health service in Malawi?" *Human Resources for Health*, vol. 12, article 17, 2014.

[23] J. M. Gross, M. F. Rogers, I. Teplinskiy et al., "The impact of out-migration on the nursing workforce in Kenya," *Health Services Research*, vol. 46, no. 4, pp. 1300–1318, 2011.

[24] P. Mbindyo, L. Gilson, D. Blaauw, and M. English, "Contextual influences on health worker motivation in district hospitals in Kenya," *Implementation Science*, vol. 4, article 43, 2009.

[25] I. Mathauer and I. Imhoff, "Health worker motivation in Africa: the role of non-financial incentives and human resource management tools," *Human Resources for Health*, vol. 4, article 24, 2006.

[26] M. Van der Doef, F. B. Mbazzi, and C. Verhoeven, "Job conditions, job satisfaction, somatic complaints and burnout among East African nurses," *Journal of Clinical Nursing*, vol. 21, no. 11-12, pp. 1763–1775, 2012.

[27] H. Prytherch, D. C. V. Kakoko, M. T. Leshabari, R. Sauerborn, and M. Marx, "Maternal and newborn healthcare providers in rural Tanzania: In-depth interviews exploring influences on motivation, performance and job satisfaction," *Rural and Remote Health*, vol. 12, no. 3, article 2072, 2012.

[28] M. Dieleman and J. W. Harnmeijer, *Improving Health Worker Performance: In Search of Promising Practices*, World Health Organization, Geneva, Switzerland, 2006.

[29] S. C. E. Anyangwe and C. Mtonga, "Inequities in the global health workforce: the greatest impediment to health in Sub-Saharan Africa," *International Journal of Environmental Research and Public Health*, vol. 4, no. 2, pp. 93–100, 2007.

[30] K. Goetz, S. M. Campbell, J. Steinhaeuser, B. Broge, S. Willms, and J. Szecsenyi, "Evaluation of job satisfaction of practice staff and general practitioners: an exploratory study," *BMC Family Practice*, vol. 12, article 137, 2011.

[31] K. Goetz, S. Campbell, B. Broge et al., "Job satisfaction of practice assistants in general practice in Germany: an observational study," *Family Practice*, vol. 30, no. 4, pp. 411–417, 2013.

[32] A. Gavartina, S. Zaroti, J. Szecsenyi et al., "Practice assistants in primary care in Germany—associations with organizational attributes on job satisfaction," *BMC Family Practice*, vol. 14, article 110, 2013.

[33] M. F. Harris, J. G. Proudfoot, U. W. Jayasinghe et al., "Job satisfaction of staff and the team environment in Australian general practice," *Medical Journal of Australia*, vol. 186, no. 11, pp. 570–573, 2007.

[34] G. A. Zangaro and K. L. Soeken, "A meta-analysis of studies of nurses' job satisfaction," *Research in Nursing and Health*, vol. 30, no. 4, pp. 445–458, 2007.

[35] S. M. Vindigni, P. L. Riley, F. Kimani et al., "Kenya's emergency-hire nursing programme: a pilot evaluation of health service delivery in two districts," *Human Resources for Health*, vol. 12, article 16, 2014.

[36] U. Adano, "The health worker recruitment and deployment process in Kenya: an emergency hiring program," *Human Resources for Health*, vol. 6, article 19, 2008.

[37] International Council of Nurses, *Positive Practice Environments: Quality Workplaces = Quality Patient Care. Information and Action Tool Kit*, International Council of Nurses, Geneva, Switzerland, 2007.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 707149, 8 pages
http://dx.doi.org/10.1155/2015/707149

Research Article

Public Concern about the Sale of High-Caffeine Drinks to Children 12 Years or Younger: An Australian Regulatory Perspective

Christina Mary Pollard,[1,2] Catrina Lisa McStay,[1,2] and Xingqiong Meng[3]

[1]*School of Public Health, Curtin University, Kent Street, GPO Box U1987, Perth, WA 6845, Australia*
[2]*Department of Health in Western Australia, 189 Royal Street, East Perth, WA 6004, Western Australia, Australia*
[3]*Flinders Centre for Innovation in Cancer, School of Medicine, Flinders University, Adelaide, SA, Australia*

Correspondence should be addressed to Christina Mary Pollard; c.pollard@curtin.edu.au

Received 18 May 2015; Revised 27 August 2015; Accepted 1 September 2015

Academic Editor: Stefan K. Lhachimi

Background. Dietary exposure to high caffeine is a health risk for children. Governments are considering measures to restrict the sale of formulated caffeinated beverages (FCB) to children. *Objectives.* To investigate community concern about sales of high-caffeine drinks to children among Western Australian adults and describe Australian and New Zealand regulatory processes regarding FCB. *Methods.* Data from the 2009 and 2012 Department of Health's Nutrition Monitoring Survey Series of 2,832 Western Australians aged 18–64 years was pooled with descriptive and ordinal logistic regression analysis performed. Current regulatory processes for FCB are reported. *Results.* Most (85%) participants were concerned about the sale of high-caffeine drinks to children; 77.4% were very concerned in 2012 compared to 66.5% in 2009, $p < .008$. Females and those living with children had higher concern (odds ratio (OR) 2.11; 95% confidence interval (CI) 1.44–3.10; OR 2.16; 95% CI 1.51–3.09, resp., $p < .001$). Concern increased with each year of age (OR 1.04; 95% CI 1.02, 1.05, $p < .001$). *Conclusions.* Community concern regarding sale of high-caffeine energy drinks to children is high and increasing. Being female and living with children were associated with greater concern. These findings support the Australian and New Zealand regulatory controls of FCB, including labelling, promotion, and advertising to children.

1. Introduction

The growing concern regarding the consumption of caffeinated beverages such as energy drinks by children and adolescents has led to considerations of how to control or limit their intake by regulators and public health professionals [1–4]. Currently, there is no established safe level of consumption of caffeine for children and adolescents [5–7]. The National Health and Medical Research Council's Australian Dietary Guidelines, 2013, specify that high-caffeine beverages such as energy drinks are not suitable for children [8]. The American Academy of Paediatrics states "that caffeine and other stimulant substances contained in energy drinks have no place in the diet of children and adolescents" [9].

Reports of the adverse health effects of consuming an excess of energy drinks are increasing [10–13]. High acute consumption of caffeine increases the risk of toxic effects, particularly in children and adolescents who have not developed tolerance to caffeine [13, 14]. Clinical manifestations of caffeine toxicity include serious adverse cardiovascular effects, seizure, and deaths [14, 15]. Other effects include sleep disturbance, increased anxiety, nausea, palpitations, and headaches [6, 15, 16]. Childhood consumption may lead to habitual intakes in adult life [6]. Concomitant consumption of energy drinks and alcohol by youth has been identified as a public health issue of concern for health promoters, policy makers, and regulators [1, 12–14, 17, 18]. An additional concern is that many caffeinated products are essentially a soft drink, high in added sugar, nutrient poor, and energy dense, promoted for sale based on their stimulatory effects [19]. The displacement of nutrients from the diets of energy drink consumers and excess energy intake from discretionary foods

Table 1: Regulatory measures specific to the management of caffeine in food in Australia [22, 24, 25].

Regulatory measure	Key requirements
Standard 2.6.4 of the Code-Formulated Caffeinated Beverages	Must contain between 145 mg/L and 320 mg/L of caffeine and comply with labelling provisions disclosing nutrient composition including caffeine content (per serving and per 100 mL), along with daily usage and warnings that product is not suitable for children, pregnant, or lactating women
Standard 1.3.1 of the Code-Food Additives	Caffeine is approved as a food additive (flavour) in cola type drinks to a maximum amount of 145 mg/L
Standard 2.6.2 of the Code-Non-Alcoholic Beverages and Brewed Soft Drinks	Caffeine is prohibited as an ingredient in formulated beverage products
Standard 1.2.4 of the Code-Labelling of Ingredients	Caffeine must be included in the ingredient list where caffeine is added to the food
Standard 1.2.3 of the Code-Mandatory Warning and Advisory Statements and Declarations	Foods that contain guarana (rich in caffeine) are required to have an advisory statement on the label that the food contains caffeine. Other foods such as coffee, tea, and cocoa are not required to declare the presence of caffeine
New Zealand Food Safety Authority. *New Zealand Food (Supplemented Food) Standard 2013*	Foods that meet the definition of supplemented foods (excluding foods that meet the definition specified in Standard 2.6.4 of the Code) may contain caffeine for purposes other than as an additive. If containing more caffeine than is required to achieve a technological function under conditions of Good Manufacturing Practice the label on the package of supplemented food must include caffeine content (per serving and per 100 mL), along with daily usage and warnings that product is not suitable for children, pregnant, or lactating women Guarana may be added to a supplemented food, with the restriction that the label must include an advisory statement that the food contains caffeine

contribute to obesity and related chronic disease, including cardiometabolic disease [8, 19].

Health promotion uses a combination of strategies to improve health of the population including environmental changes through food regulation [20]. The Australian Beverage Council asserts that energy drinks in Australia are stringently regulated [21]. Energy drinks are classified and regulated as a food under Standard 2.6.4-Formulated Caffeinated Beverages (FCB) of the Australia New Zealand Food Standards Code (the Code). This standard specifies that energy drinks contain between 145 mg/L and 320 mg/L of caffeine; comply to labelling provisions disclosing nutrient composition including caffeine content (per serving and per 100 mL) along with daily usage; and display warnings that product is not suitable for children, pregnant, or lactating women [22]. The United States (US) and Europe do not have an upper limit for caffeinated beverages [2]. The industry claims that the products are not marketed or promoted to children; however, half of sales are through the supermarket where there is no restriction on who purchases them [21].

Standard 2.6.4 of the Code was developed in 2001, amidst community concern regarding the availability of these caffeinated beverages to children [2]. Energy drinks, termed as FCB, are defined as "a non-alcoholic water-based flavoured beverage which contains caffeine and may contain carbohydrates, amino acids, vitamins and other substances, including other foods, for the purpose of enhancing mental performance" [22]. Although the concentration of caffeine in FCB is specified, the volume of a retail unit and therefore the amount of caffeine consumed per retail unit are not regulated. In addition to Standard 2.6.4, other standards regulate caffeine in food (Table 1).

Community opinion can inform regulatory process with consultation with key stakeholders, including the community, identified as "an instrumental component of the joint food regulation system, and is fundamental to the development of good food regulation policy" [23]. Assessing community attitudes and perceptions toward regulatory issues helps inform decision making.

This paper explores the attitudes of adults living in Western Australia to the sale of caffeinated beverages to children aged 12 years or younger. The aim was to measure the current level of community concern; assess changes between 2009 and 2012; and explore factors associated with attitudes including demographics and behaviour. This paper also describes the Australian and NZ regulatory decision making process and outcomes regarding FCB and other foods containing caffeine.

2. Methods

2.1. Surveys. Data from the Department of Health in Western Australia's (WA) statewide 2009 and 2012 Nutrition Monitor Survey Series (NMSS) were pooled for this analysis. Computer assisted telephone interviews of WA adults aged 18 to 64 years were conducted in July to August 2009 and 2012. The samples were randomly drawn from the Electronic White Pages for WA and stratified according to area of residence. All sample households with an address were sent a primary approach letter explaining the purpose of the survey, how the sample was selected, who would be asked to do the survey, and about how long it would take. Every household in the initial sample was called and asked if someone aged 18–64 years was resident and, if so, which one had the most recent

Table 2: Sample demographics of Nutrition Monitoring Survey Series, Western Australia, 2009 and 2012.

	2009 $n = 1284$	2012 $n = 1548$	Total $n = 2832$	Weighted[a] %
Gender				
Female	830	1005	1,835	49.2
Male	454	543	997	50.8
Age groups				
18–24 years	71	66	137	15.6
25–34 years	180	144	324	22.5
35–44 years	340	377	717	22.7
45–54 years	356	466	822	21.6
55–64 years	337	495	832	17.7
Area of residence				
Metropolitan	965	1011	1976	79.3
Remote (Kimberley and Pilbara)	29	82	111	3.6
Rural	290	455	745	17.1

[a]Percentages were weighted for probability of selection and adjusted by age, sex, and geographic area to the 2011 Estimated Resident Population of Western Australia.

birthday. No substitutes were accepted. The response rate (completed/contacted) was 81.6% and 82.4% for 2009 and 2012, respectively. Surveys were granted approval from the WA Department of Human Research Ethics Committee.

2.2. Measures. Participants were asked to rate how concerned they were about the sale of high-caffeine drinks to children 12 years or younger using a five-point Likert scale from one "not very concerned"; "somewhat concerned"; "neither unconcerned nor concerned"; and "quite concerned" to five points "very concerned."

Their attitudes towards healthy eating were gauged by asking about the attention they paid to the health aspects of the food they eat, with the options of "pay a lot of attention," "take a bit of notice," or "don't really think about it."

Demographic data collected were age, gender, education level, household income, employment status, country of birth, residential area, and living arrangement. Participants' self-reported height and weight were used to derive their body mass index.

2.3. Statistical Analysis. Data were pooled and weighted to account for sample design and post adjusted for age, sex, and geographic area of 2011 Estimated Resident Population of WA as it was the most recent census year. Descriptive statistics were used to report the prevalence of participants' concern about sales of high-caffeine drink to children 12 years or younger. Ordinal logistic regression on participants' attitude toward sales of high-caffeine drink was performed. The direction of the rating in the regression went from one "not very concerned" to five "very concerned." A full model includes the following variables: survey year, demographics (gender, age group, education level, income, employment status, whether living with children, country of birth, and residential area), body mass index, and attention they paid to the health aspect of the food they eat. p values were derived from a survey design-based Pearson chi square test. Only variables with p value $< .05$ were retained in the final model and reported, with the exception of survey year which was retained in the model regardless of its significance. Survey module of Stata software version 12.0 (StataCorp LP, College Station, TX) was used for all analyses.

3. Results

A total of 2832 adults participated in the 2009 and 2012 surveys, and the sample details are shown in Table 2.

Table 3 shows that overall 85% of participants were "quite" or "very" concerned about the sales of high-caffeine drinks to children. Significantly more participants were "very concerned" in 2012 (77%) than in 2009 (67%), $p < .008$ (Table 3).

Regression analysis revealed female participants (OR 2.11) and those who live with children (OR 2.16) were twice as likely to be more concerned than their counterparts (Table 4, all p values $< .05$). Participants' concern also increased with age; for each incremental yearly increase in age participants were more likely to rate a high level of concern (OR 1.04). Participants residing in remote areas were significantly less likely to be very concerned than those living in other areas. There was no significant difference between participants' concern level across survey years when the other variables (body mass index and attention paid to the health aspect of food eaten) were included in the model.

4. Discussion

Community concern regarding the sale of high-caffeine drinks to children 12 years or younger remains high and increased between 2009 and 2012. As would be expected, females and participants living with children showed a higher level of concern. People in the community who are

Table 3: Prevalence of how concerned participants are about the sale of high-caffeine drinks to children 12 years or younger, NMSS, 2009 and 2012.

	Total %	2009 % [95% CI]	2012 % [95% CI]	p value[b]
Original categories (n = 2192)				.008
Not very concerned	5.3 [4.0, 7.1]	6.3 [4.4, 8.9]	3.5 [2.1, 5.7]	
Somewhat concerned	5.9 [4.5, 7.9]	6.1 [4.3, 8.4]	5.7 [3.3, 9.7]	
Neither unconcerned or concerned	2.9 [2.1, 4.0]	3.4 [2.3, 5.0]	2.0 [1.2, 3.4]	
Quite concerned	14.9 [12.9, 17.1]	17.1 [14.4, 20.2]	10.8 [8.2, 13.9]	
Very concerned	70.3 [67.7, 72.9]	66.5 [63.0, 69.8]	77.4 [73.0, 81.3]	
Do not know	0.6 [0.3, 1.1]	0.6 [0.3, 1.4]	0.5 [0.2, 1.4]	
Combined categories (n = 2122)[a]				.003
Not very concerned	5.5 [4.1, 7.3]	6.6 [4.6, 9.3]	3.6 [2.2, 5.8]	
Somewhat/quite concerned	21.7 [19.2, 24.3]	24.2 [21.0, 27.7]	17.1 [13.4, 21.5]	
Very concerned	72.8 [70.2, 75.4]	69.3 [65.7, 72.6]	79.4 [74.8, 83.3]	

[a]Excluded participants who said "neither unconcerned or concerned" and "don't know."
[b]p values were derived from a survey design-based Pearson chi square test.

Table 4: Factors related to how concerned the participants are about sale of high-caffeine food to children 12 years old or younger, NMSS, 2009 and 2012.

	How concerned the participants are about sales of high-caffeine food to children (from "not very" to "very" concerned) OR [95% CI]
Survey year	
2009	1.00
2012	1.38 [0.88, 2.17]
Gender	
Male	1.00
Female	2.11 [1.44, 3.10]***
Age (years)	1.04 [1.02, 1.05]***
Living with children	
No	1.00
Yes	2.16 [1.51, 3.09]***
Residential area	
Metro	1.00
Remote (Kimberley and Pilbara)	0.55 [0.31, 0.99]*
Rural	1.21 [0.82, 1.78]

*$p < .05$; ***$p < .001$. Results are odds ratio [95% confidence interval] from an ordinal logistic regression. The outcome variable is on a five-point Likert scale, from "not very concerned" (1) to "very concerned" (5).

responsible for caring for children may be more aware of the availability of high-caffeine energy drinks to children, influences on children's dietary choices, and adverse impacts of energy drink ingredients.

Our findings show general population concern about sales of high-caffeine drinks to children, broader than US research which found that consumers of high-caffeinated beverages were concerned about product safety. A market research company assessed attitudes of consumers and non-consumers of energy drinks in the US finding that 59% of those consuming energy drinks were concerned about possible adverse health effects [26]. Thirty-nine percent of people surveyed had reduced their intake due to perceived adverse health effects, and the majority supported the inclusion of maximum daily intake levels of caffeine information on the label (79% female; 71% male).

Our findings of high community concern are important as there has been a rapid growth in the number and types of caffeinated beverage products on the market. Between 2001 and 2010, energy drink sales in Australasia have quadrupled from 34.5 to 155.6 million litres and are predicted to reach 220 million litres by 2018 [2, 27]. The market leader, Frucor Beverages brand V, owned by Suntory Holdings, accounts for more than a third of sales volumes in Australasia. Red Bull (owned by Red Bull Australia Pty Ltd.) and Mother (owned by Coca-Cola Amatil) are also brand leaders, with

Mother and V Australasian market specific brands [27]. The marketing of energy drinks is designed to appeal to young people, for example, edgy campaigns incorporating extreme sports [2, 28] such as the Red Bull Stratos promotion of Felix Baumgartner 9.09-minute fall from the stratosphere back to earth [29].

Regulatory organisations, researchers, and health promoters across the world are investigating the risks of energy drink consumption in vulnerable populations and there is national and international discourse on the risk of caffeine of consumption by vulnerable groups amidst increasing sales in the energy drink market. These risk assessments on energy drinks may help inform regulatory decision making in Australia. In 2010, Health Canada created the expert panel on caffeinated energy drinks to review the safety of caffeinated energy drinks in the food supply which concluded that although the risk of adverse health effects following the consumption of energy drinks in the Canadian context was low, serious adverse event signals had occurred [30]. The EFSA safety assessment of caffeine concluded that there was a lack of evidence on which to set a safe level of caffeine consumption for either children or adolescents [31]. The French Agency for Food, Environmental and Occupational Health & Safety (ANSES) concluded that a causal relationship between energy drink intake and adverse symptoms was assessed likely or very likely in 25 out of 212 analysed cases, reported to the agency since 2008 [10]. ANSES recommended that "at risk" individuals including children and adolescents should avoid consuming energy drinks and that energy drinks should not be consumed with alcohol or during exercise [10].

Globally, the consumption of caffeinated beverages amongst children and younger adolescents appears to be increasing [10, 17, 32]; however, there is a lack of current Australian and NZ dietary consumption trend data. The current risk assessment of caffeinated beverages in children used NZ intake data that is over 10 years old to guide policy options deliberations [2]. The estimated caffeine consumption derived from the 2002 NZ Children's National Nutrition Survey with the addition of one retail unit of an energy drink estimated that 70% of children and 40% of teenagers would exceed an adverse effect level of 3 mg/kg bw/day [7]. The 2012 Australian Health Survey estimated that the usual daily caffeine consumption in children 9–13 years was 23 mg per day for boys and 18 mg/day for girls; and the intake at the 95th percentile was 81 mg per day and 63 mg per day, respectively [33].

There is no established dietary reference standard, such as an acceptable daily intake (ADI) level, for caffeine intake in children or adolescents; however, ≤2.5 mg/kg body weight/day level has been used in risk assessments [6, 7, 10, 34]. In 2000, in a review of safety of dietary caffeine consumption including toxicological/pharmacological effects, addictive effects, or other hazards at low doses, the authors concluded that the "no effect" dose response, along with the threshold dose for behavioural effects in children, has yet to be established [6]. A recent systematic review of the health effects of energy drinks concluded that a precautionary approach was warranted until sufficient scientific evidence is available to establish safe levels of dietary consumption [5].

Consumption needs to be explored beyond mean intakes as there is evidence of high chronic and high acute caffeinated beverage consumption among children and adolescents. A survey across 16 European Union countries found that 18% of children and 68% of adolescents consumed at least one energy drink in the past year (average exposures were 1.01 mg/kg bw/day and 0.38 mg/kg bw/day, resp.) and energy drinks contributed on average 43% and 13% to total caffeine intakes [35]. However, this does not identify high or chronic consumption patterns. The proportion of children and adolescents assessed as "high chronic consumers" (minimum intake of 4-5 energy drinks per week) was 16% (average intake 0.95 L/week) and 12% (average intake 7 L/month), respectively. Of note, 12% of adolescents were also assessed as "high acute consumers" (energy drink intake ≥1.065 L/session).

In Australia the number of calls to the New South Wales Poison Centre in Australia regarding toxic effects experienced following energy drink exposure increased from 12 in 2004 to 65 in 2010, highlighting this issue of high acute consumption, with adolescents identified as being particularly at risk [12]. This led to the recommendation that energy drinks include the national poison hotline number on the energy drink label.

The Australia and New Zealand Food Regulatory System and Response. In Australia and New Zealand (NZ), the system of food policy and laws is founded by two key agreements. Firstly, the Food Regulation Agreement (Australia) between the Commonwealth and the States and Territories brings into operation a consistent and cooperative approach to food regulation across Australia. This agreement established several policy bodies: the Australia and New Zealand Food Regulation Ministerial Council (Ministerial Council), which became the Legislative and Governance Forum on Food Regulation (the Forum) in February 2011, along with the Food Regulation Standing Committee (FRSC). Secondly, the Joint Food Standards Treaty between Australia and NZ principally facilitates trade between these nations. These agreements are supported by the Food Standards Australia New Zealand Act 1991 (FSANZ Act), which established FSANZ as the statutory organisation charged with developing and administering the Code. The Food Regulatory System (FRS) separates policy, standards development, and implementation and enforcement decision making processes [23].

Provision of safe food for all Australians and New Zealanders, especially for vulnerable subgroups in the population, is a key tenet of the FRS. Under the FSANZ Act, the objectives (in descending priority order) of the Authority in developing or reviewing food regulatory measures and variations of food regulatory measures are (a) protection of public health and safety; (b) provision of information relating to food to enable consumers to make informed choices; and (c) prevention of misleading or deceptive conduct [36]. The Act also requires FSANZ to consider trade and industry issues such as competitiveness and efficiency, along with alignment to both national and international food standards. Additionally, risk assessments need to be based on the best scientific evidence. During the review or development of the Code, FSANZ is required to have regard to policy guidelines [36].

In 2003, the Ministerial Council developed its Policy Guideline on the Addition of Caffeine to Foods (Policy Guideline) with the aim of minimising the risk of dietary exposure to caffeine of at risk individuals in the population [2, 37]. In addition to the high order principles, the policy guidance statements include specific policy principles relating to caffeine risks [37, p. Ministerial Council Policy Guideline on the Addition of Caffeine to Foods] as follows.

2003 Ministerial Council Policy Guideline on the Addition of Caffeine to Foods

"*Other Principles*

(1) Endeavour to limit the possible adverse effect of caffeine-containing foods on vulnerable subgroups of the population.

(2) Ensure that the effect of caffeine additions to individual foods is considered in the context of the total diet.

(3) Ensure the appropriate use of advisory statements on caffeine-containing foods in alignment with scientifically substantiated risk to vulnerable subgroups of the population.

Until further evidence becomes available, maintain the status quo (as currently in place in Australia) for caffeine regulation by

(i) Maintaining the current additive permissions for caffeine;

(ii) Restricting the use of new products containing non-traditional caffeine rich ingredients (including guarana) to boost the caffeine content in other foods, beyond the current provisions for caffeine.

Caffeinated cola drinks and formulated caffeinated beverages will be permitted in accordance with the current standards. Foods, which naturally contain caffeine and have a long history of use and consumer awareness/association with caffeine, such as tea, coffee, and cocoa, are to be exempted from the labelling provisions and the use of these foods naturally containing caffeine to be added to other foods will continue to be allowed. Guarana, as a non-traditional food containing caffeine, will continue to have special labelling provisions outlined in the Food Standards Code."

The review was ordered in response to ongoing public concern regarding the health implications of caffeinated beverage consumption, given the growth in the number and types of caffeinated products on the market [2]. The scope of the review included examination of current scientific evidence on the health effects of consuming caffeine (with a focus on the effects of dietary exposure of vulnerable population subgroups including children and adolescents); industry energy drink developments, for example, growth in the marketplace; and consideration of regulatory actions taken in other countries (e.g., for best practice and trade implications) [2]. Three policy options were proposed: to update; to maintain (keep the status quo); or to rescind the current Policy Guideline. After reviewing the evidence and following public consultation, a new Ministerial Policy Guideline-Regulatory Management of Caffeine in the Food Supply was endorsed on 27 June 2014 [38], with policy guidance statements, in addition to the high order principles [38, p. Ministerial Policy Guideline Regulatory Management of Caffeine in the Food Supply], as follows.

2014 Ministerial Policy Guideline Regulatory Management of Caffeine in the Food Supply

"*Specific Policy Principles*. The regulatory management of caffeine in the food supply should

(a) be based on risk analysis ensuring consideration of general population and taking into account vulnerable population groups including children, adolescents, pregnant, and lactating women and caffeine sensitive consumers;

(b) consider exposure to caffeine from all dietary sources;

(c) be informed by emerging evidence and the regulation of caffeine in overseas jurisdictions.

Additional Policy Advice

FSANZ is encouraged to work with research agencies to monitor caffeine consumption across the population, including consumption by vulnerable population groups.

Regulatory management of caffeine in the food supply may include regulatory and non-regulatory risk management approaches."

Despite the regulatory controls for FCB already in place, there is a high level of concern about sale of these beverages to children. In Australia, there have been calls to ban the sale of energy drinks to children and adolescents by the Country Women's Association with backing from the Australian Medical Association (AMA) [39]. Following the inaugural international energy drinks 2014 conference held in Australia, attendees released a statement highlighting evidence of adverse events associated with energy drink consumption and called for stronger regulatory approaches as the current measures had failed to protect the at risk subgroups, particularly children [40].

Measures proposed to reduce caffeine consumption in at risk subgroups include specifying a maximum retail unit volume, reducing the caffeine level in energy drinks, banning the sale of energy drinks to people under 18 years of age, restricting advertising and marketing of energy drinks to children and adolescents including events targeting this age group, and strengthening health warning advice to consumers [3, 13, 28, 30, 40, 41]. In response to concerns raised in Europe, Lithuania enacted a law banning the sale and advertising of energy drinks to people under eighteen years of age [42].

The findings of this current research suggest that the Australian and NZ government policy response is lagging behind public concern regarding the sale of caffeinated beverages to children. The Australian and NZ Food Regulation System identifies its stakeholders as including individual consumers, industry bodies, primary producers, food manufacturers, importers and retailers, public health organisations,

consumer advocacy organisations, and community groups. There are opportunities during the development of standards for stakeholders to respond to public consultations. From time to time consumer research is undertaken by FSANZ and regular monitoring of public concern and opinion regarding contemporary food regulation issues is a role of government, such as the NMSS reported here.

Minimising the adverse events related to energy drink consumption is the policy goal and there are increasing calls for urgent action [40, 43, 44]. Strength of community opinion is considered when developing regulatory policy options and currently, across the WA population, adults support the use of food regulatory control of food labelling and advertising [45]. It is important that governments set limits to control the supply and promotion of food products with adverse health implications, particularly for vulnerable population groups. The Forum endorsement of the revised policy guideline on the regulatory management of caffeine in the food supply was, in part, in response to public concerns raised about risks associated with dietary exposure. This policy guideline will be used to guide the development or review of food standards relating to this important issue.

A limitation of this current research is that it is a cross-sectional study based on self-reported opinion which may be influenced by perceived social desirability and as the survey was conducted on an Australian population, care should be taken in generalising the findings. Further monitoring and surveillance of the contribution of specific foods to total caffeine intake is recommended, including coffee, tea, soft drinks, and energy drinks. Monitoring should report on mean caffeine intake, high acute and high chronic consumption levels, and the intake of population groups who are sensitive to caffeine, particularly children and adolescents.

5. Conclusion

The findings of high and growing level of community concern in WA regarding consumption of high caffeinated beverages to children, particularly by females, those living with children, and with increasing age, coupled with the worldwide public and scientific community concerns regarding adverse health effects for children and adolescents, are of interest to health promoters and regulators. In the current context of limited dietary intake data and the lack of an established acceptable daily intake level and considering the potential adverse health effects of acute or high intake, the precautionary principle applies. It is important that government maintain regulatory controls of formulated caffeinated beverages, including labelling, promotion, advertising, and sale of these beverages to children. The research findings suggest that the Australian and New Zealand government policy response is lagging behind public concern regarding the sale of caffeinated beverages to children.

Conflict of Interests

The authors declare that there is no conflict of interests regarding publication of this paper.

Authors' Contribution

Christina Mary Pollard, Catrina Lisa McStay, and Xingqiong Meng contributed equally to this work.

References

[1] A. M. Arria, M. C. O'Brien, R. R. Griffiths et al., *The Use of Caffeine in Energy Drinks. Letter to The Honourable Margaret A Hamburg, M.D., Commissioner, Food and Drug Administration*, 2013, http://graphics8.nytimes.com/packages/pdf/business/BestofScienceLetter_v22.pdf.

[2] Food Regulations Standing Committee (FRSC) Working Group, "Food regulations policy options paper for the regulation of caffeine in foods," 2013, https://www.health.gov.au/internet/main/publishing.nsf/Content/A294B740C7928C3CCA257BF0001CFFF4/$File/The%20Regulation%20of%20Caffeine%20in%20Foods.pdf.

[3] J. J. Breda, S. H. Whiting, R. Encarnação et al., "Energy drink consumption in Europe: a review of the risks, adverse health effects, and policy options to respond," *Frontiers in Public Health*, vol. 2, article 134, 2014.

[4] C. M. Pollard, X. Meng, and C. McStay, "Community concern about the sale of high-caffeine drinks to children under 12 years of age: western Australia population survey results," *Health Promotion Journal of Australia*, vol. 24, no. 2, pp. 156–157, 2013.

[5] T. Burrows, K. Pursey, M. Neve, and P. Stanwell, "What are the health implications associated with the consumption of energy drinks? A systematic review," *Nutrition Reviews*, vol. 71, no. 3, pp. 135–148, 2013.

[6] P. F. S. A. Smith, J. Minors, J. McNeil, and A. Proudfoot, *Report from the Expert Working Group on the Safety Aspects of Dietary Caffeine*, Australia New Zealand Food Authority, 2000.

[7] B. Thomson and S. Schiess, "Risk Profile: caffeine in energy drinks and energy shots," 2010, http://www.foodsafety.govt.nz/elibrary/industry/Risk_Profile_Caffeine-Science_Research.pdf.

[8] National Health and Medical Research Council, *Australian Dietary Guidelines*, National Health and Medical Research Council, Canberra, Australia, 2013.

[9] Committee on Nutrition and the Council on Sports Medicine and Fitness, "Sports drinks and energy drinks for children and adolescents: are they appropriate?" *Pediatrics*, vol. 127, no. 6, pp. 1182–1189, 2011.

[10] French Agency for Food Environment and Occupational Health & Safety (ANSES), "Opinion of the French Agency for Food, Environmental and Occupational Health & Safety on the assessment of risks concerning the consumption of so-called 'energy drinks'," 2013, http://www.anses.fr/en/documents/NUT2012sa0212EN.pdf.

[11] M. Goldfarb, C. Tellier, and G. Thanassoulis, "Review of published cases of adverse cardiovascular events after ingestion of energy drinks," *American Journal of Cardiology*, vol. 113, no. 1, pp. 168–172, 2014.

[12] N. Gunja and J. A. Brown, "Energy drinks: health risks and toxicity," *Medical Journal of Australia*, vol. 196, no. 1, pp. 46–49, 2012.

[13] C. J. Reissig, E. C. Strain, and R. R. Griffiths, "Caffeinated energy drinks—a growing problem," *Drug and Alcohol Dependence*, vol. 99, no. 1–3, pp. 1–10, 2009.

[14] B. J. Wolk, M. Ganetsky, and K. M. Babu, "Toxicity of energy drinks," *Current Opinion in Pediatrics*, vol. 24, no. 2, pp. 243–251, 2012.

[15] K. M. Babu, R. J. Church, and W. Lewander, "Energy drinks: the new eye-opener for adolescents," *Clinical Pediatric Emergency Medicine*, vol. 9, no. 1, pp. 35–42, 2008.

[16] G. S. A. Trapp, K. Allen, T. A. O'Sullivan, M. Robinson, P. Jacoby, and W. H. Oddy, "Energy drink consumption is associated with anxiety in australian young adult males," *Depression and Anxiety*, vol. 31, no. 5, pp. 420–428, 2014.

[17] S. Azagba, D. Langille, and M. Asbridge, "An emerging adolescent health risk: caffeinated energy drink consumption patterns among high school students," *Preventive Medicine*, vol. 62, pp. 54–59, 2014.

[18] S. C. Jones, "'You wouldn't know it had alcohol in it until you read the can': adolescents and alcohol-energy drinks," *Australasian Marketing Journal*, vol. 19, no. 3, pp. 189–195, 2011.

[19] C. M. Brown, A. G. Dulloo, and J.-P. Montani, "Sugary drinks in the pathogenesis of obesity and cardiovascular diseases," *International Journal of Obesity*, vol. 32, no. 6, pp. S28–S34, 2008.

[20] T. Shilton, P. Howat, R. James, and T. Lower, "Review of competencies for Australian health promotion," *Promotion & Education*, vol. 10, no. 4, pp. 162–209, 2003.

[21] Australian Beverage Council, *Energy Drinks—An Industry Commitment*, Australian Beverage Council, 2014, http://australianbeverages.org/wp-content/uploads/2013/03/EnergyDrinks_An-IndustryCommitment.pdf.

[22] Food Standards Australia New Zealand, *Australia New Zealand Food Standards Code Standard 2.6.4 Formulated Caffeinated Beverages*, 2014, https://www.comlaw.gov.au/Details/F2013C00107.

[23] Australian Government Department of Health, *Engaging in the Australian and New Zealand Joint Food Regulatory System*, Australian Government Department of Health, 2014, https://www.health.gov.au/internet/main/publishing.nsf/Content/foodsecretariat-stakeholder-engagement.

[24] New Zealand Food Safety Authority, "New Zealand Food (Supplemented Food) Standard 2013," 2014, http://www.foodsafety.govt.nz/elibrary/industry/nzfood-supplementedfood-standard-2013.pdf.

[25] Food Standards Australia New Zealand, "Australia New Zealand Food Standards Code," 2014, http://www.foodstandards.gov.au/code/Pages/default.aspx.

[26] Mintel, "Nearly six in 10 US energy drink consumers worry about their safety," 2014, http://www.mintel.com/press-centre/food-and-drink/energy-drink-safety.

[27] Canadean, *Energy Drinks, An Antipodes Perspective*, Canadean, 2013, http://www.canadean.com/news/energy-drinks,-an-antipodes-perspective/.

[28] J. L. Pomeranz, C. R. Munsell, and J. L. Harris, "Energy drinks: an emerging public health hazard for youth," *Journal of Public Health Policy*, vol. 34, no. 2, pp. 254–271, 2013.

[29] Red Bull, "World record jump," 2014, http://www.redbullstratos.com/the-mission/world-record-jump/.

[30] N. Macdonald, R. Hamilton, P. Malloy, Y. Moride, and J. Shearer, *Report by the Expert Panel on Caffeinated Energy Drinks*, Government of Canada, 2014, http://www.hc-sc.gc.ca/dhp-mps/alt_formats/pdf/prodnatur/activit/groupe-expert-panel/report_rapport-eng.pdf.

[31] European Food Safety Authority Panel on Dietetic Products Nutrition and Allergies, *Scientific Opinion on the Safety of Caffeine*, 2015, http://www.efsa.europa.eu/en/consultations/call/150115.pdf?bcsi_scan_313cddce030931be=yIwqitE9o/ISyrpBzKjUCki1n3gBAAAAjORCAA==&bcsi_scan_filename=150115.pdf.

[32] L. Gallimberti, A. Buja, S. Chindamo et al., "Energy drink consumption in children and early adolescents," *European Journal of Pediatrics*, vol. 172, no. 10, pp. 1335–1340, 2013.

[33] Australian Bureau of Statistics, *Australian Health Survey: Usual Nutrient Intakes, 2011-12*, Australian Bureau of Statistics, Canberra, Australia, 2015.

[34] P. Nawrot, S. Jordan, J. Eastwood, J. Rotstein, A. Hugenholtz, and M. Feeley, "Effects of caffeine on human health," *Food Additives & Contaminants*, vol. 20, no. 1, pp. 1–30, 2003.

[35] S. Zucconi, C. Volpato, F. Adinolfi et al., "Gathering consumption data on specific consumer groups of energy drinks," Tech. Rep. EN-394, Supporting Publications, 2013, http:/www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/394e.pdf?bcsi_scan_c221d61a0ea4ff4c=vAA4PW+dzTyZwvqzYhn5EhHZ200DAAAAyM+MGA==&bcsi_scan_filename=394e.pdf.

[36] Food Standards Australia New Zealand (FSANZ) Act 1991, 1991.

[37] Australian Government Department of Health, *The Ministerial Council Policy Guideline on the Addition of Caffeine to Foods*, Australian Government Department of Health, 2003, http://www.foodstandards.gov.au/code/fofr/fofrpolicy/documents/Caffeine%20-%20Policy%20Guideline.pdf.

[38] Australian Government Department of Health, *Ministerial Policy Guideline Regulatory Management of Caffeine in the Food Supply*, 2015, http://www.health.gov.au/internet/main/publishing.nsf/Content/foodsecretariat-policy-guidelines#3.

[39] Country Women's Association of NSW, "CWA of NSW calls for ban on energy drinks to minors," 2014, https://cwaofnsw.org.au/article.php?group_id=461&id=6.

[40] Deakin University, "Time to take a stand against the sale of caffeinated energy drinks," 2015, http://www.deakin.edu.au/news/latest-media-releases/2015-media-releases-archives/time-to-take-a-stand-against-the-sale-of-caffeinated-energy-drinks.

[41] J. Thorlton, D. A. Colby, and P. Devine, "Proposed actions for the US food and drug administration to implement to minimize adverse effects associated with energy drink consumption," *American Journal of Public Health*, vol. 104, no. 7, pp. 1175–1180, 2014.

[42] AusFoodNews, "Lithuania ban on energy drink sales to under 18s comes in with broader restrictions and warnings," 2014, http://ausfoodnews.com.au/2014/11/05/lithuania-ban-on-energy-drink-sales-to-under-18s-comes-in-with-broader-restrictions-and-warnings.html.

[43] J. Thorlton, D. A. Colby, and P. Devine, "Proposed actions for the US Food and Drug Administration to implement to minimize adverse effects associated with energy drink consumption," *American Journal of Public Health*, vol. 104, no. 7, pp. 1175–1180, 2014.

[44] A. J. Budney and J. A. Emond, "Caffeine addiction? Caffeine for youth? Time to act!," *Addiction*, vol. 109, no. 11, pp. 1771–1772, 2015.

[45] C. M. Pollard, A. Daly, M. Moore, and C. W. Binns, "Public say food regulatory policies to improve health in Western Australia are important: population survey results," *Australian and New Zealand Journal of Public Health*, vol. 37, no. 5, pp. 475–482, 2013.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 812325, 14 pages
http://dx.doi.org/10.1155/2015/812325

Research Article

Health Impacts of Increased Physical Activity from Changes in Transportation Infrastructure: Quantitative Estimates for Three Communities

Theodore J. Mansfield and Jacqueline MacDonald Gibson

Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 148A Rosenau Hall, CB No. 7431, Chapel Hill, NC 27599, USA

Correspondence should be addressed to Theodore J. Mansfield; tedmans@live.unc.edu

Received 1 June 2015; Accepted 22 July 2015

Academic Editor: Stefan K. Lhachimi

Recently, two quantitative tools have emerged for predicting the health impacts of projects that change population physical activity: the Health Economic Assessment Tool (HEAT) and Dynamic Modeling for Health Impact Assessment (DYNAMO-HIA). HEAT has been used to support health impact assessments of transportation infrastructure projects, but DYNAMO-HIA has not been previously employed for this purpose nor have the two tools been compared. To demonstrate the use of DYNAMO-HIA for supporting health impact assessments of transportation infrastructure projects, we employed the model in three communities (urban, suburban, and rural) in North Carolina. We also compared DYNAMO-HIA and HEAT predictions in the urban community. Using DYNAMO-HIA, we estimated benefit-cost ratios of 20.2 (95% C.I.: 8.7–30.6), 0.6 (0.3–0.9), and 4.7 (2.1–7.1) for the urban, suburban, and rural projects, respectively. For a 40-year time period, the HEAT predictions of deaths avoided by the urban infrastructure project were three times as high as DYNAMO-HIA's predictions due to HEAT's inability to account for changing population health characteristics over time. Quantitative health impact assessment coupled with economic valuation is a powerful tool for integrating health considerations into transportation decision-making. However, to avoid overestimating benefits, such quantitative HIAs should use dynamic, rather than static, approaches.

1. Introduction

In the United States, approximately 234,000 premature deaths are associated with physical inactivity each year [1]. The built environment influences walking and biking for transportation and, in turn, total physical activity [2, 3]. Many communities in the United States are designed in ways that do not support walking and biking, thereby contributing to low levels of physical activity [4]. Recently, transportation agencies across the United States have sought to integrate health considerations into decision-making [5, 6]. Health impact assessment (HIA) has emerged as a systematic framework for considering how decisions, such as modifications to the built environment, may impact public health and has informed a variety of decisions in the transportation sector [7, 8]. However, most transportation HIAs conducted to date have provided qualitative rather than quantitative estimates of health benefits arising from changes in physical activity (e.g., indicating that physical activity is expected to increase, without estimating the magnitude of the increase) [9]. Existing research links the built environment to physical activity levels and health outcomes, but quantitative models to predict the health impacts of modifications to the built environment remain poorly developed [10–12].

Within the past four years, two new tools to support quantitative HIAs have emerged. The first tool, the Health Economic Assessment Tool (HEAT) for cycling and walking, was introduced by the World Health Organization in 2011 [13]. More recently, the European Union Health Programme released the Dynamic Model for Health Impact Assessment

(DYNAMO-HIA) [14]. These two tools employ fundamentally different methods; while DYNAMO-HIA is dynamic, capable of tracking changes in population health over many years, HEAT is static, providing health impact estimates for a single year. The HEAT method has been used in several HIAs of policies or projects to promote active transportation (walking or cycling instead of driving) [14]. DYNAMO-HIA has been applied to estimate the health impacts of a ban on alcohol imports in Sweden, smoking cessation in Great Britain, reduced salt intake in Europe, decreased smoking prevalence in Copenhagen, and body mass index reduction in Netherlands [15–18]. However, to our knowledge, DYNAMO-HIA has not yet been applied to predict the health impacts of increased physical activity arising from changes in the built environment. Further, the estimates from these two methods have not been compared.

To demonstrate the use of quantitative tools for estimating the health effects of physical activity in HIAs of the built environment, this paper describes quantitative HIAs of proposed changes to the built environment in three North Carolina communities. All three HIAs used DYNAMO-HIA to estimate the health effects of increased transportation walking time expected to arise due to modifications to the built environment. Changes in premature mortality, coronary heart disease (CHD), type 2 diabetes, hypertension, and stroke were estimated for each community. In addition, each HIA estimated the ratio of health benefits to expected project costs. For one of the case studies, we additionally compared results obtained from DYNAMO-HIA with those obtained from the HEAT model. Our objective in making this comparison was to determine whether the health impact estimates differ when using a dynamic approach (as in DYNAMO-HIA) as compared to a static approach (as in HEAT). We hypothesized that the static approach may overestimate health benefits by failing to account for overall improvements in population health from one year to the next and, as a result, estimating benefits in each year relative to a population for which no benefits have yet accrued. Our overall purpose was twofold: first, to demonstrate that quantitative tools in general may provide objective, evidence-based decision support within the HIA framework and, second, to provide insight into the advantages and disadvantages of emerging quantitative tools and methods to conduct HIAs.

The HIAs presented in this study were conducted as examples to support *WalkBikeNC*, a statewide bicycle and pedestrian plan developed by the North Carolina Department of Transportation (NCDOT) in 2013 [19]. *WalkBikeNC* presents a unified policy framework to support active travel statewide, but it does not propose projects. Instead, specific bicycle and pedestrian infrastructure projects are planned and implemented by local authorities in accordance with *WalkBikeNC*. Such projects may be included in a range of local plans, including small-area plans, comprehensive transportation plans, and bicycle and pedestrian master plans. The three HIAs described in this paper consider pedestrian infrastructure improvements aligned with the policy framework established in *WalkBikeNC* at three planning scales: a small-area plan, a comprehensive plan, and a streetscape plan.

2. Materials and Methods

All three case studies followed the six steps of HIA proposed by the US National Research Council: (1) screening; (2) scoping; (3) assessment; (4) recommendations; (5) reporting; and (6) monitoring and evaluation [7]. The first two steps of HIA, screening and scoping, focus on identifying and characterizing health concerns and disparities in the community. The third step, assessment, explores how the decision to be made influences these concerns and disparities through qualitative understanding and/or quantitative modeling of causal pathways as understood in the scientific literature. The conclusions from the assessment stage inform the fourth stage, recommendations. Finally, reporting and monitoring and evaluation aim to engage stakeholders, hold decision-makers accountable, and evaluate the effectiveness of the decision in addressing identified health concerns at some point in the future. Because this paper focuses on improving the assessment stage through the application of quantitative methods, details of steps 4–6 are not presented; these details can be found elsewhere [19, 20]. Details on the screening and scoping stages are provided below, because these steps influenced the scope of the assessment phase.

2.1. Site Selection (Screening). Case study sites were selected in coordination with NCDOT. In all three communities, the proposed changes to the built environment were included in adopted local plans but had not received funding as of October 2012 (when this project began). Projects were selected to provide variation across three dimensions: (1) development context (rural, suburban, and urban); (2) planning scale (corridor plan, small-area plan, and comprehensive plan); and (3) geographic region within North Carolina (Piedmont region, coastal region, mountain region). Table S1 and Figures S1–S3 in Supplementary Material available online at http://dx.doi.org/10.1155/2015/812325 provide maps, demographic data, and information about the changes to the built environment proposed for each project.

The first HIA is conducted on changes to the built environment proposed in the City of Raleigh's Blue Ridge Road Corridor (BRRC) small-area plan (*urban, small-area plan, Piedmont region*). The BRRC is located eight kilometers east of downtown Raleigh, the second-largest city in North Carolina and the state capital. The BRRC small-area plan is the result of a planning and visioning process to guide development in the corridor as it urbanizes. The plan includes dense, mixed-use land development, construction of a compact street network, and construction of additional pedestrian and bicycling facilities. We considered the effects on time spent walking for transportation and the resulting health outcomes if the plan were implemented in its entirety [21].

The second HIA is conducted on construction of new sidewalks in the town of Winterville as proposed in the Greenville Metropolitan Planning Organization's Bicycle and Pedestrian Master Plan (*suburban, comprehensive plan,* coastal *region*). This plan proposes both pedestrian and bicycle projects throughout the Greenville metropolitan area,

a mid-size community in eastern North Carolina. We estimated the health impacts of building all sidewalks proposed in the plan within the municipal boundaries of Winterville, a suburban community on the outskirts of the Greenville region [22].

The third HIA is conducted on streetscape improvements proposed in the Town of Sparta's Downtown Streetscape Master Plan (*rural, corridor plan, mountain region*). Sparta is a prototypical rural main-street community, with a small, walkable downtown containing shops and services surrounded by low-density development. We estimated the health impacts of proposed improvements to the downtown streetscape, including improved sidewalks and street crossings [23].

2.2. Selection of Health Outcomes (Scoping). Facilitated discussions with local decision-makers and residents in each community confirmed that existing transportation infrastructure (e.g., lack of sidewalks) and overall community design (e.g., lack of destinations within easy walking distance) limit opportunities for walking as a means of transportation. The potential health outcomes that could be affected if new, pedestrian-friendly infrastructure were in place and if, as a result, residents spent more time walking for transportation were then selected from a literature review. The literature review identified several health outcomes for which nonvigorous transportation physical activity has been shown to have a preventive effect: coronary heart disease (CHD), type 2 diabetes mellitus, hypertension, stroke, and premature mortality from all causes [24–27]. Additionally, these four diseases were identified as existing health concerns related to physical activity levels in each community.

2.3. Health Impacts Model (Assessment). We used DYNAMO-HIA to estimate the health impacts of increased transportation physical activity in all three communities. We then additionally used a modified version of the HEAT model, implemented in *Analytica 4.5* (Lumina Decision Systems, Los Gatos, CA) in the BRRC. These two models and their data requirements are described in turn below.

DYNAMO-HIA is a dynamic health impacts model that employs Markov Chain modeling to estimate the effects of a health intervention on a population over time [15]. Conceptually, Markov Chain models divide a system into distinct groups of risk factor states linked by transition probabilities, which define the likelihood that a member of one group will transition to another group over time (Figure 1). The model moves forward in discrete one-year time steps, estimating the population in each group at time step using the previous group populations and transition probabilities between groups. To estimate the health impacts of an intervention that changes health behaviors, an intervention scenario is specified in which the probabilities of transitioning from a healthy to a diseased state (represented in Figure 1 as P_1, P_2, P_4, and P_5) or from a healthy or diseased state to death (P_3 and P_6–P_9) are altered based on changes in the distribution of risk factors in the population (e.g., amount of time walking for transportation). As the model steps forward through time, changes in these transition probabilities affect the rate at which healthy individuals transition to diseased states and/or death. Alongside the intervention scenario, a baseline scenario is also specified in which transition probabilities are not affected by the intervention. Health impacts are estimated by comparing health outcomes between the two scenarios over time. DYNAMO-HIA requires a large amount of baseline health data: age- and sex-specific population distributions, mortality rates, disease prevalence, disease incidence rates, and risk factor prevalence. In the intervention scenario, a change in risk factor prevalence and/or a transition between risk factor states over time must also be specified. Finally, dose-response functions must be characterized for each health outcome of interest. DYNAMO-HIA is available free of charge (http://www.dynamo-hia.eu/) and may be installed on any Windows-based machine.

We developed DYNAMO-HIA models for each community. Each model included community-specific population and health data as described in Section 2.3.1. A baseline, "no-build" scenario and an intervention scenario were specified for each community. In the baseline scenarios, weekly time spent walking for transportation was taken from recent surveys as described in Section 2.3.3. In the intervention scenarios, studies linking proposed built environment changes in each community to increases in walking for transportation were used to estimate post-construction walking as described in Section 2.3.4. Relative risks linking time spent walking for transportation to modeled health outcomes were taken from epidemiological studies (Table 1). Health impacts were estimated by taking the difference in projected health outcomes between the two scenarios over time each year for 40 years.

To develop 95% confidence intervals for our health impact estimates, each model was run five times, changing relative risk parameters in the model to the upper and lower bound of the 95% confidence intervals reported in epidemiological studies in each iteration. The first model used central values for all relative risk parameters, the second model used the lower bound of the confidence interval for mortality and central values for all diseases, the third model used the upper bound of the confidence interval for mortality and central values for all diseases, the fourth model used lower bounds for all diseases and the central value for mortality, and the fifth model used upper bounds for all diseases and the central value for mortality. Varying each relative risk parameter in turn and rerunning each model enabled the construction of 95% confidence intervals for all of our results reflecting uncertainty in the relative risk parameters used; however, uncertainty in other model parameters (e.g., magnitude of changes in walking for transportation) is not reflected in these estimates. All confidence intervals reported throughout this paper were developed using this approach.

Unlike DYNAMO-HIA, the HEAT model is static: it estimates a fraction of cases of premature mortality that could be avoided if a population spent more time walking or cycling and assumes that this fraction is constant from year to year. That is, health benefits of increased activity do not accrue from year to year for a given individual. The WHO has made an online tool for automating these calculations (http://www.heatwalkingcycling.org/) available. In order to

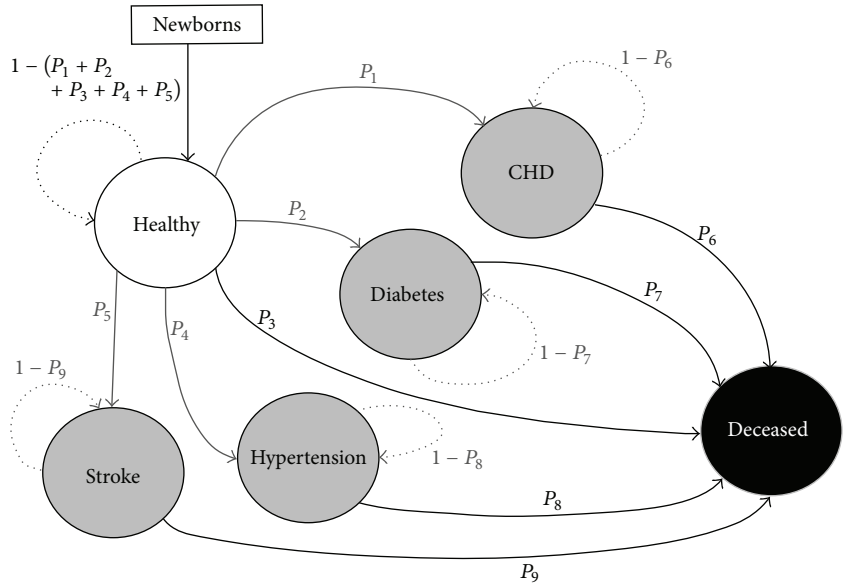


Figure 1: Schematic of DYNAMO-HIA model representing simulation of one time step for one scenario (reference or intervention). Each circle represents a population state. Solid lines represent possible transitions between states at each time step, whereas dotted lines represent staying in the same state during a time step. The variables P_1–P_9 represent transition probabilities between states.

Table 1: Relative risks.

Health outcome	Sex	Low category (1–149 minutes' walking for transportation per week)	High category (150+ minutes' walking for transportation per week)
All-cause mortality [23]	Combined	0.95 (0.98–0.92)[a]	0.90 (0.96–0.85)
CHD [24]	Male	0.99 (1.08–0.91)[c]	0.99 (1.10–0.90)[c]
	Female	0.95 (1.08–0.83)[c]	0.80 (0.92–0.69)[c]
Type 2 diabetes [26]	Combined	0.77 (1.02–0.58)[b]	0.69 (0.88–0.54)[b]
Hypertension [26]	Combined	0.76 (0.94–0.61)[b]	0.69 (0.83–0.58)[b]
Stroke [25]	Male	0.94 (1.06–0.83)[c]	0.88 (1.02–0.77)[c]
	Female	0.88 (1.01–0.77)[c]	0.87 (1.01–0.75)[c]

[a]95% confidence interval shown for all relative risks.
[b]Adjusted for race, education, income, and smoking status.
[c]Adjusted for education, smoking status, alcohol consumption, body mass index, systolic blood pressure, cholesterol, history of diabetes, and occupational and leisure-time physical activity.

compare the results obtained with DYNAMO-HIA with those obtained using the HEAT model approach, we reconstructed the HEAT tool using *Analytica*. This reconstruction additionally includes morbidity, which is not included in the base HEAT model. Details of this reconstruction are provided elsewhere [28].

Like DYNAMO-HIA, our reconstructed version of the HEAT model requires baseline data on population size by age and sex, baseline death rates, baseline disease prevalence and incidence rates for each health outcome of interest, and relative risks linking each health outcome to a risk factor (in this case, walking for transportation). In addition, information about the time spent walking for transportation under current conditions and under the intervention scenario is needed. Sources for these data, used in both the DYNAMO-HIA models the reconstructed HEAT model in the BRRC, are described below.

2.3.1. Baseline Population and Health Data. We estimated age- and sex-specific population distributions by applying county-level age and sex distributions to refine Census block-group data for each case study location (Figure S2) [29, 30]. Baseline death and birth rates were taken from county-level data obtained from the NC State Center for Health Statistics [31]. We developed age-specific prevalence functions for CHD, type 2 diabetes mellitus, hypertension, and stroke for each case study location by fitting second-order prevalence functions to data from the Behavioral Risk Factor Surveillance System (BRFSS) survey [32]. Disease prevalence data were not available stratified by both age and sex; thus, we stratified by age only and assumed identical prevalence functions for males and females. Incidence data are not available from the State Center for Health Statistics for the diseases considered in this study. Thus, incidence functions for each case study location were estimated using a

differential equation-based method described in Brinks (see Supplementary Material, Section 2.1 and Table S4) [33].

2.3.2. Relative Risks. Relative risks of each health outcome as a function of transportation walking were drawn from previous studies (summarized in Table 1). Categorical dose-response functions for type 2 diabetes mellitus and hypertension were taken from a study of US adults that used data from the National Health and Nutrition Examination Survey [26]. To our knowledge, no studies exist linking transportation physical activity levels to CHD or stroke risk in US adults; thus, relative risks were taken from two studies of a large cohort of Finnish adults [24, 25]. To estimate the relative risk of premature mortality as a function of time spent walking for transportation, a dose-response function derived in a recent meta-analysis was employed; this same function is used to calculate the relative risk of all-cause mortality in the HEAT model [13, 27]:

$$\mathrm{RR}_{\mathrm{mortality}} = 0.89^{(y/168)}, \tag{1}$$

where y is weekly minutes spent walking for transportation. We used (1) to estimate the relative risk of all-cause mortality for the same exposure categories used in studies linking walking for transportation to disease risk. Specifically, these studies grouped populations into three levels of time spent walking for transportation: a reference category (none), a low category (1–149 min/week), and a high category (150+ min/week). The high category reflects the Centers for Disease Control and Prevention (CDC) minimum recommendation for total adult physical activity [34]. Using (1), we calculated relative risks for all-cause mortality at the midpoint of the low transportation walking category (75 min/week) and at the low point of the high transportation walking category (150 min/week).

2.3.3. Baseline Active Transportation Behavior. In Winterville and Sparta, we estimated baseline transportation physical activity using data from the 2009 North Carolina BRFSS survey [32]. In the BRRC, we used an active transportation survey conducted within the neighborhood in 2012 utilizing a widely used and validated physical activity questionnaire [20, 28, 35]. Responses to these surveys were recategorized according to the CDC physical activity categories described above.

2.3.4. Estimating Changes in Active Transportation Behavior. Due to differences in data availability and the nature of the plans considered, different methods were used in each case study community to estimate how changes in the built environment are expected to affect transportation physical activity.

The method for estimating changes in walking time if the BRRC small-area plan were implemented is described in detail elsewhere [20, 28]. Briefly, because multiple built environment changes are proposed in addition to pedestrian infrastructure improvements, the net effect of all of these changes on transportation walking is estimated using a multidimensional walkability index that links intersection density, population density, land-use diversity, and retail floor area ratio to walking for transportation [36]. The walkability index is calculated from

$$\begin{aligned}\text{Walkability Score} = {} & \left(2 \times Z_{\text{intersetion}}\right) + \left(Z_{\text{residential}}\right) \\ & + \left(Z_{\text{FAR}}\right) + \left(Z_{\text{land-use}}\right),\end{aligned} \tag{2}$$

where Z variables represent normalized versions of intersection density ($Z_{\text{intersetion}}$), the number of intersections divided by land area; residential density ($Z_{\text{residential}}$), the number of housing units divided by the residential land area; retail floor area (Z_{FAR}), the square footage of retail floor area divided by the square footage of land devoted to retail use; and land-use diversity ($Z_{\text{land-use}}$), computed as described in Cervero and Kockelman [37]. Previous studies that have linked transportation walking time to the walkability score were then used to estimate the increase in time spent walking as a result of the increase in walkability score that would occur if the small-area plan were fully implemented [28, 38].

In Winterville, the proposed changes to the built environment consist solely of new sidewalk construction. Thus, a relationship linking sidewalk density to transportation walking was used to estimate changes in transportation physical activity. A 1 km/km^2 increase in sidewalk density is associated with an increase in the odds of an individual having taken a walking trip in the previous week by 2.3 percent [39]. Thus, the odds ratio of walking before and after construction may be expressed as:

$$\frac{O_{W,\text{after}}}{O_{W,\text{before}}} = 1.023^{(D_{s,\text{after}} - D_{s,\text{before}})}. \tag{3}$$

$O_{W,\text{before}}$ is the odds of walking given the density of sidewalks before construction, $D_{s,\text{before}}$ (km/km^2), and $O_{W,\text{after}}$ is the odds of walking given the density of sidewalks after construction, $D_{s,\text{after}}$ (km/km^2). Rearranging (3) and expressing in terms of probabilities, this becomes:

$$\frac{P_{W,\text{after}}}{\left(1 - P_{W,\text{after}}\right)} = \frac{P_{W,\text{before}} 1.023^{(D_{s,\text{after}} - D_{s,\text{before}})}}{\left(1 - P_{W,\text{before}}\right)}. \tag{4}$$

$P_{W,after}$ is the probability that an individual takes at least one walk trip per week after construction, and $P_{W,\text{before}}$ is the probability that an individual has taken a walking trip in the past week before construction, assumed to be equal to the proportion of the population reporting any walking in the BRFSS. We iteratively solved for $P_{W,\text{after}}$ and adjusted the proportion of non-walkers in the population accordingly. We assumed that new walkers were distributed between the low- and high-walk-time categories in the same manner as walkers were distributed between these two categories before construction.

In Sparta, we used changes in a composite pedestrian environment factor (PEF)—which includes sidewalk quality, ease of street crossings, topography, and density of the street grid—to estimate changes in average weekly walking distance [40]. Each subcategory is assessed on a 3-point scale; the PEF is calculated by adding these four subcategory scores

and transforming the result into an ordinal variable (low, medium, or high). After construction of streetscape improvement in Sparta, sidewalk quality and ease of street crossings would improve significantly while topography and the configuration of the street network would remain unchanged. Therefore, we assumed that the sidewalk quality and ease of street crossings subcategories would change from 1 (current conditions) to 3 (post-construction), while the topography and street grid density would remain unchanged. This change in subscores would change the PEF from low to medium. In turn, per-capita weekly walking distance would increase by 0.92 kilometers [40]. Assuming a typical walking speed of 4 kilometers per hour, per-capita transportation walking time would increase by 13.6 minutes per week, on average [41]. Because this relationship was derived in an urban setting using small geographies, while Sparta is a rural town, we assumed that only individuals living within a 0.4-kilometer buffer of the proposed improvements (25% of the population) would increase their walking. We increased the percentage of population in each walking time bin proportionally so that the average per-capita walking time for individuals living within 0.4 kilometers of the proposed improvements equaled to the preconstruction average plus 13.6 minutes.

2.3.5. Economic Valuation. To compare the benefits of estimated health impacts to project costs, we applied economic valuations to each health outcome considered. For mortality, we used the value of a statistical life suggested by the United States Department of Transportation (USDOT) in 2013, $9.1 M USD per avoided premature death [42]. For each disease, we used yearly disease costs estimated by the Milken Institute that combine treatment costs and indirect costs from productivity losses resulting from lost workdays and reduced presenteeism (in Supplementary Material, Table S7) [43]. For the BRRC and Winterville, we estimated project costs using average bid data for North Carolina ($89.57 per linear meter of sidewalk; $142.08 and $150.70 per square meter of poured concrete sidewalk and curb and gutter, resp.) [44]. For Sparta, we used the cost estimate provided in the plan, $686,157 USD [23]. Ongoing maintenance costs are not considered. Benefits and costs were discounted to the present using a 5% discount rate per USDOT guidance [45]. A sensitivity analysis was conducted using 3.5% and 7% discount rates based on guidance from the United States Office of Management and Budget and NCDOT, respectively (in Supplementary Material, Figure S5) [45, 46].

3. Results

3.1. Health Outcomes. To estimate the health impacts of built environment changes in each community, we used DYNAMO-HIA to predict changes in premature mortality and incidence of CHD, type 2 diabetes, hypertension, and stroke over 40 years due to increased walking for transportation. In the BRRC, DYNAMO-HIA estimates a significant reduction in premature all-cause mortality as well as significant preventive effects for hypertension, type 2 diabetes mellitus, and CHD (Figure 2). In Sparta, significant reductions in premature mortality, cases of hypertension, and cases of type 2 diabetes mellitus are estimated; however, estimated effects on avoided cases of CHD are minimal. In Winterville, DYNAMO-HIA estimates small, yet significant, reductions in premature mortality and cases of hypertension and minimal effects on type 2 diabetes and CHD. Across all sites, no significant reductions in cases of stroke are estimated. The total population benefits of avoided mortality and the prevention of hypertension and type 2 diabetes accrue over time but demonstrate diminishing returns (Figure 2, Table 2). For example, DYNAMO-HIA estimates that the cumulative number of premature deaths avoided in the BRRC will increase from 4.9 (1.8–7.7) ten years after construction to 14 (5.2–23) 40 years after construction (Table 2). Similarly, within ten years of construction, an estimated 12 (4.5–17) and 4.9 (2.6–7.6) cases of hypertension and type 2 diabetes will have been prevented, and these numbers are expected to increase to 32 (12–45) and 16 (8.3–24) within 40 years. Generally, health outcomes for which a strong preventive effect is demonstrated in the literature and for which baseline community prevalence is high (e.g., hypertension) are most influenced by increases in transportation physical activity.

Comparing across sites, DYNAMO-HIA estimates stronger preventive effects on a per-capita basis in the BRRC and Sparta than in Winterville (Figure 2). For example, the cumulative cases of premature mortality prevented by year 40 are 0.99 and 0.36 per 1,000 people in the BRRC and Sparta, respectively, as compared to 0.08 per 1,000 people in Winterville. This result occurs because the proposed changes to the built environment in the BRRC and Sparta are estimated to increase transportation walking more in the BRRC and in Sparta than in Winterville (Table 2). For example, the average time spent walking per week is expected to increase by 17 minutes in the BRRC and 2.2 minutes in Sparta, in comparison to a smaller increase of 0.7 minutes per week in Winterville (Table 2). Additionally, a preventive effect on CHD is only estimated in the BRRC. As shown in Table 1, the preventive effect of walking for transportation on CHD is strong only for females in the highest physical activity category. The population in the BRRC has a greater proportion of women compared to the other two sites (in Supplementary Material, Figure S4) and a greater predicted change in the proportion of the population walking more than 150 minutes per week for transportation (Table 2); thus, the effect of increased transportation walking on avoided cases of CHD is significant in the BRRC but not in the other two sites.

3.2. Economic Valuation. To estimate the economic value of health benefits in each community, we multiplied projected avoided deaths and avoided disease cases per year by their respective economic values. The economic value of estimated health benefits exceeds project construction costs within one year in the BRRC and within three years in Sparta (Table 2) assuming a 5% discount rate. Over the 40-year time period considered, the benefit-cost ratios in the BRRC and Sparta are 20.2 (8.7–30.6) and 4.7 (2.1–7.1), respectively. However, the present value of the health benefits in Winterville is less than the estimated project costs: the benefit-to-cost ratio in Winterville over 40 years is 0.6 (0.3–0.9) (Table 2). This latter

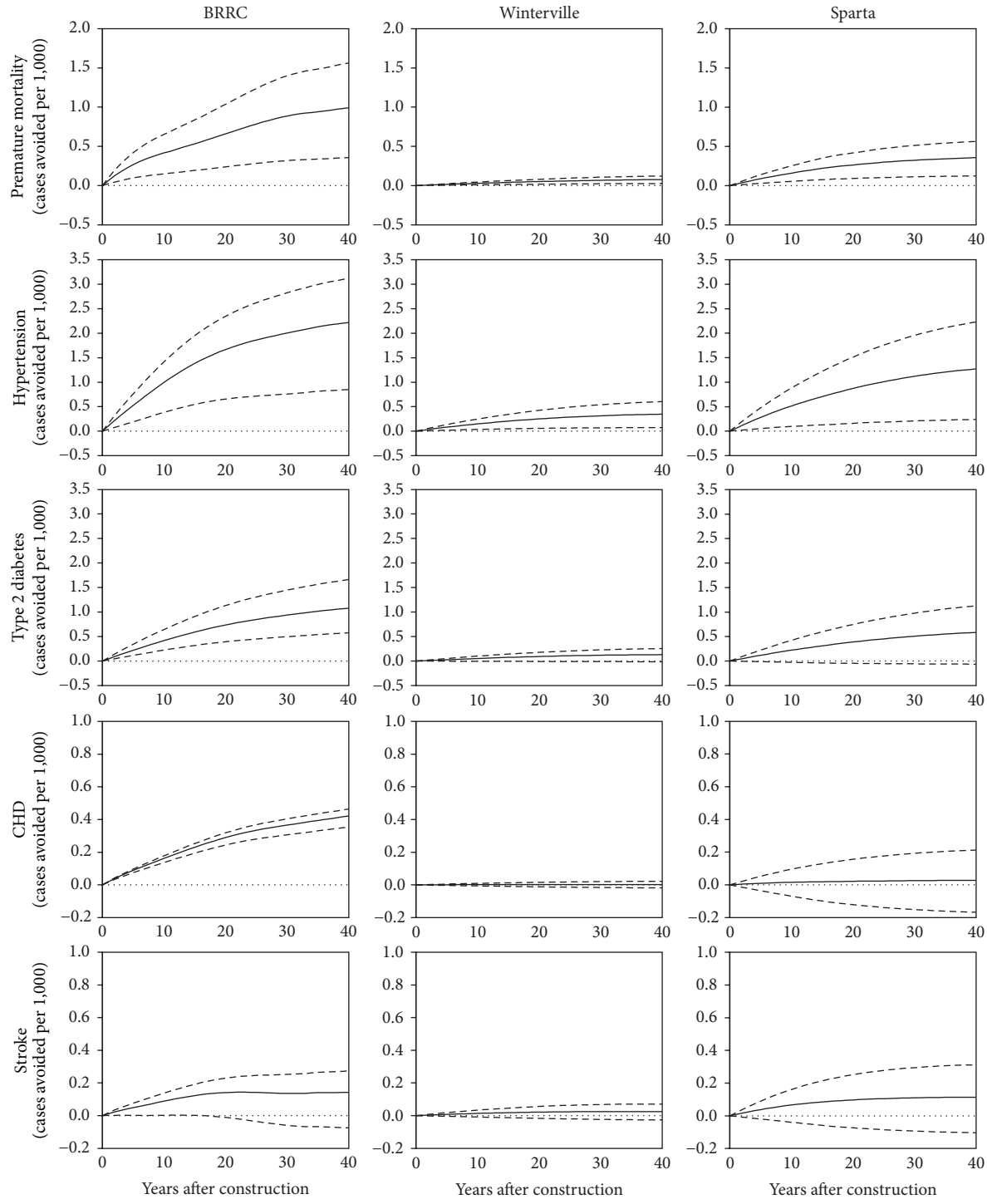


Figure 2: Estimated health impacts per 1,000 persons for each community (solid lines), with 95% confidence intervals reflecting uncertainty in relative risk parameters (dashed lines).

finding results from the design of the Winterville project and the population density in that community; while significant sidewalk construction is proposed, the new sidewalks will be spread over a very large area of relatively low population density, dampening the potential behavioral impact. The net present value of the BRRC and Sparta projects remains positive even when considering a higher discount rate (7%) and remains negative in Winterville even when considering a lower discount rate (3.5%) (in Supplementary Material, Figure S5).

In all communities, health benefits are overwhelmingly driven by avoided premature mortality (Figure S5). Avoided

Table 2: Summary of findings, with 95% confidence intervals based on uncertainty in relative risk parameters.

Built environment variables	BRRC			Winterville			Sparta		
	Before	After	Change	Before	After	Change	Before	After	Change
Walkability score	−3.61	0.96	**+4.57**	—	—	—	—	—	—
Sidewalk density (km/km^2)	—	—	—	0.8	3.8	**+3.0**	—	—	—
PEF (categorical)	—	—	—	—	—	—	Low	Medium	**+1**
Walking outcomes[a]	Before	After	Change	Before	After	Change	Before	After	Change
No walking (percent)	40.7%	40.7%	**0**%	84.3%	83.4%	**−0.9**%	85.4%	82.4%	**−3.0**%
1–149 min/week (percent)	41.5%	21.2%	**−20.3**%	12.3%	12.9%	**+0.6**%	12.1%	14.6%	**+2.5**%
150+ min/week (percent)	17.8%	38.1%	**+20.3**%	3.4%	3.6%	**+0.2**%	2.5%	3.0%	**+0.5**%
Ave. walk time (min/week)	13.1	30.4	**+17**	12.5	13.2	**+0.7**	10.4	12.6	**+2.2**
Health outcomes[a]	Years after construction			Years after construction			Years after construction		
	10	20	40	10	20	40	10	20	40
Avoided premature mortality	4.9 (1.8–7.7)	8.5 (3.1–13.3)	14.3 (5.2–22.6)	0.3 (0.1–0.5)	0.5 (0.2–0.9)	0.9 (0.3–1.4)	0.3 (0.1–0.4)	0.4 (0.2–0.7)	0.5 (0.2–0.8)
Avoided cases of CHD	1.9 (1.6–2.1)	3.7 (3.1–4.1)	6.1 (5.1–6.7)	0.0 (−0.1–0.1)	0.0 (−0.1–0.2)	0.0 (−0.2–0.3)	0.0 (−0.1–0.2)	0.0 (−0.2–0.3)	0.0 (−0.2–0.3)
Avoided cases of type 2 diabetes	4.9 (2.6–7.6)	9.4 (5.1–14.5)	15.6 (8.3–24.1)	0.5 (0.0–1.0)	1.0 (−0.1–1.9)	1.5 (−0.2–2.9)	0.4 (0.0–0.7)	0.6 (−0.1–1.2)	0.8 (−0.1–1.6)
Avoided cases of hypertension	11.8 (4.5–16.7)	21.4 (8.4–30.1)	32.1 (12.3–45.1)	1.5 (0.4–2.5)	2.7 (0.6–4.5)	4.0 (0.9–6.9)	0.9 (0.2–1.5)	1.4 (0.3–2.4)	1.8 (0.4–3.2)
Avoided cases of stroke	1.1 (0.0–1.6)	1.8 (−0.1–2.9)	2.1 (−1.1–4.0)	0.1 (−0.1–0.3)	0.2 (−0.2–0.6)	0.3 (−0.3–0.8)	0.1 (−0.1–0.3)	0.2 (−0.1–0.4)	0.2 (−0.2–0.5)
Economic outcomes[b,c]	Years after construction			Years after construction			Years after construction		
	10	20	40	10	20	40	10	20	40
Net present value (2012 USD)	33.4M (10.8–53.7)	50.4M (18.4–79.0)	66.8M (26.8–103)	−5.1M (−6.5–−3.9)	−3.9M (−5.9–−2.1)	−2.9M (−5.3–−0.6)	1.4M (0.1–2.5)	2.2M (0.5–3.7)	2.6M (0.7–4.2)
Benefit-cost ratio	10.6 (4.1–16.5)	15.5 (6.3–23.7)	20.2 (8.7–30.6)	0.3 (0.1–0.5)	0.5 (0.2–0.7)	0.6 (0.3–0.9)	3.0 (1.1–4.6)	4.1 (1.7–6.3)	4.7 (2.1–7.1)
Time for B : C to exceed 1	1 year (1-2 years)			Benefits do not exceed costs			3 years (2–9 years)		

[a]Estimates of walking for transportation after construction in Winterville do not add to 100% due to rounding.
[b]For all health and economic outcomes, 95% confidence intervals are estimated using the lower and upper bounds of the relative risk parameters as noted in Table 1.
[c]5% discount rate assumed.

premature mortality constitutes 92%, 86%, and 89% of the total net present value of health benefits over 40 years in the BRRC, Winterville, and Sparta, respectively. This result occurs due to the much higher value placed on an avoided premature death, in comparison to the value placed on avoided chronic disease cases (in Supplementary Material, Table S7).

3.3. Comparison of DYNAMO-HIA and HEAT. To compare the dynamic approach used in DYNAMO-HIA and the static approach used in the HEAT model, we re-estimated health impacts in the BRRC using our reconstructed HEAT model and compared these findings to impacts estimated by our DYNAMO-HIA model. For all health outcomes considered, the HEAT model estimates a higher number of avoided cases per year than the DYNAMO-HIA model (Figure 3). The difference between the two approaches increases with time (Figure 3). When considering the cumulative health impacts over multiple years, the differences in the two approaches become substantial (Figure 4). The reconstructed HEAT model estimates that 41 premature deaths would be prevented over 40 years—2.9 times as many deaths averted as predicted by the DYNAMO-HIA model. Similarly, central estimates of avoided hypertension, type 2 diabetes, CHD, and stroke increase by factors of 3.3, 1.6, 2.5, and 6.7 when using the static approach, in comparison to the dynamic approach (Figure 4).

The static approach overestimates health benefits by failing to account for changing disease prevalence over time. In the static model, avoided cases for each year are estimated for the population as a whole without accounting for population disease prevalence. In contrast, the dynamic

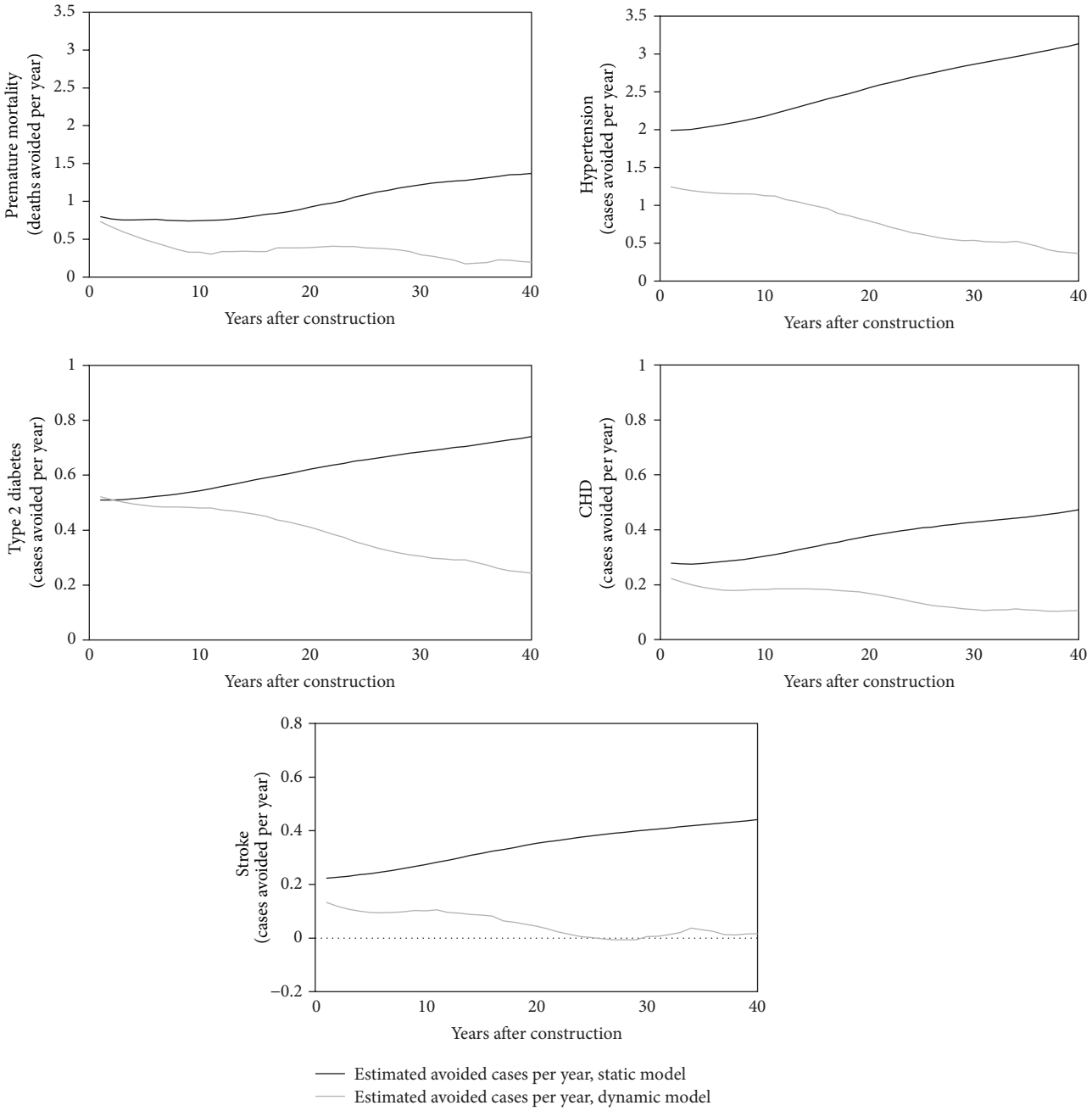


Figure 3: Estimated health impacts per year obtained using the HEAT (static) model (solid black lines) and DYNAMO-HIA (dynamic) model (solid grey lines) for the BRRC case study.

model removes individuals who develop a disease from the population that is able to avoid a new case in subsequent years (i.e., individuals who develop a disease transition to diseased states (Figure 1), after which they are not included in estimations of new avoided cases). Additionally, the dynamic model references data from the previous year in estimating benefits for a given year whereas the static model has no memory of population health data in the previous year. Thus, relative to the dynamic model, the static model overestimates benefits in the future because it fails to account for changes in disease prevalence over time. In other words, the dynamic model is able to incrementally approach a new steady state in which an intervention has shifted disease incidence functions downwards for a portion of the population; once this steady state is reached, new benefits no longer accrue as lower risk individuals delay the onset of disease but do not completely avoid disease over time. Once these individuals transition into a diseased state, they are no longer included in avoided cases calculations. Static models, however, do not approach a new steady state because benefits are always calculated relative to a population in which no benefits have been accrued and disease prevalence is not accounted for. Thus,

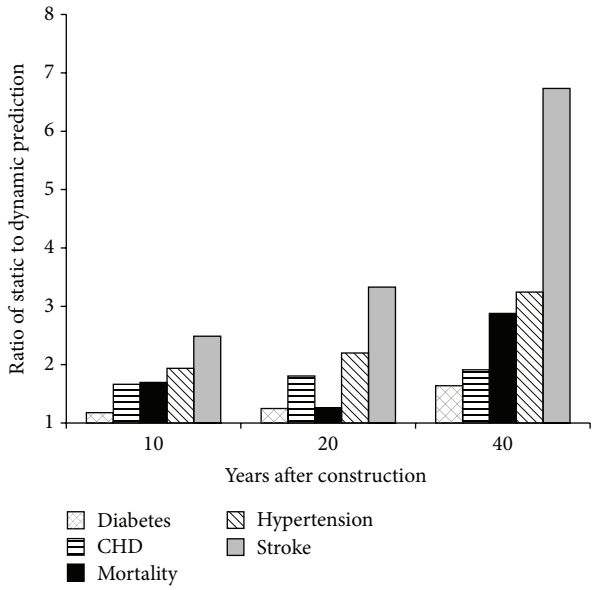


Figure 4: Ratio of cumulative health impact estimates from HEAT (static) and DYNAMO-HIA (dynamic) models at 10, 20, and 40 years after construction.

benefits will continue to accrue beyond the point at which the dynamic model reaches a new steady state. As a result, the static model increasingly overestimates benefits over time relative to the dynamic model. This behavior is illustrated in Figure 3; at each time step, the rate of change in avoided cases of type 2 diabetes stays relatively stable for the static model, increasing slightly as the population grows over time. In the dynamic model, the rate of change in the number of cases avoided decreases over time as the model approaches steady state in which all individuals who walk more have a decreased risk, but still some risk, for developing type 2 diabetes throughout their lifetimes (Figure 3).

4. Discussion

Using the dynamic DYNAMO-HIA tool, we predicted that the health benefits of changes to the built environment that support walking for transportation would exceed construction costs in two of the three case study communities. In the urban BRRC neighborhood, the benefit-cost ratio of changes to the built environment that would increase walkability was estimated to be 20 over 40 years. In the small rural town of Sparta, the benefit-cost ratio of proposed improvements to the downtown streetscape reached 4.7 over 40 years. In contrast, the benefit-cost ratio of constructing proposed sidewalks in suburban Winterville reached only 0.6 over 40 years. In addition, our comparison of estimates from the reconstructed HEAT model and estimates from the DYNAMO-HIA model showed that the static approach tends to over-predict benefits when considering effects over multiple years. Thus, if sufficient data and capacity exist, dynamic tools such as DYNAMO-HIA should be used rather than static tools to estimate the health impacts of policies and projects that increase transportation physical activity.

4.1. Comparison with Recent Active Transportation HIAs. A number of transportation HIAs using a range of modeling techniques to link changes in the built environment to health benefits from increased transportation physical activity have been completed in recent years [14]. To our knowledge, only one example of a dynamic model used to estimate the health benefits of built environment changes exists: a system dynamics model was used in an HIA of large-scale bicycle infrastructure construction in Auckland, New Zealand [47]. This model linked bicycle infrastructure investment scenarios to changes in the perceived safety of bicycling to work and resulting mode shifts to bicycle commuting. Health impacts were then estimated for resulting changes in bicycle crash risk, air pollution exposure, and physical activity levels. Bicycle mode shares were predicted for several investment scenarios, including a business-as-usual scenario. A relative risk function comparing cyclists to non-cyclists was used to estimate changes in mortality from increased physical activity for each scenario over time. Benefit-cost ratios ranged from 6 to 24, driven largely by the value of prevented premature mortality resulting from increased physical activity [47].

A number of HIAs using static models, including HEAT, have also recently been performed. A study in Dane County, Wisconsin, estimated a benefit-cost ratio of 1.7 for a hypothetical countywide sidewalk construction project. The study used a regression model to link sidewalk presence to time spent walking and biking for transportation. The results of this model were used to estimate transportation physical activity given sidewalk construction across the county. Increased physical activity was then linked to reduced weight gain and ultimately reduced costs associated with obesity using a static model [48]. An HIA of the construction of a bicycle path in Dublin, Ireland, estimated benefit-cost ratios ranging from 2.2 to 11.8. This HIA used a survey to estimate increased bicycling to work after construction and the HEAT model to estimate health and economic benefits [49]. Finally, an assessment in Portland, Oregon, used a traffic demand model to estimate increased bicycle commuting due to past and planned investments in bicycle infrastructure throughout the city. Using the HEAT model to estimate benefits from resulting increases in physical activity, benefit-cost ratios ranged from 20 to 53 [50]. As in our study, avoided premature mortality dominated the monetary value of the health benefits of increased physical activity (Figure S5).

Previous studies have found benefit-cost ratios for changes in the built environment that support walking and biking for transportation ranging 1.7 to 53. Our results are within this range for the BRRC and Sparta but not in Winterville. The population density in Winterville may be too low for the proposed improvements to be economically viable when considering health benefits alone. This finding demonstrates that the health benefits of changes in the built environment that increase physical activity may not always exceed project costs. Thus, quantitative HIA may be an important tool for prioritizing investments to maximize the overall value of health benefits.

As HIA for active transportation projects and policies is refined, it will be important to consider differential treatment effects for different age groups and to include social equity

considerations [14]. Physical activity may have a stronger preventive effect for older individuals, and many countries worldwide are seeing shifts in population distribution towards older age groups. The dynamic model used in this assessment is able to easily incorporate age-specific dose-response information, if available. The usefulness of such stratifications is demonstrated in our estimates for CHD: due to differences in population characteristics and predicted changes in behavior across sites, we estimate reduced incidence of CHD in the BRRC but not in Sparta or Winterville. This difference is driven by differential treatment effects at higher doses of transportation walking for men and women (Table 1). To increase the consideration of social equity in transportation HIA, scalable models are needed. Using the DYNAMO-HIA model at three different scales, we provide evidence that quantitative assessment methods are robust across scales. If modeling methods are robust at different scales, a series of neighborhood-scale models may be used to compare the health impacts of transportation decisions in neighborhoods with different socioeconomic conditions and may reveal disproportionate impacts. Such an application could better inform investments in active transportation infrastructure to address social equity concerns.

In sum, previous studies provide strong evidence that built environment changes meaningfully impact health outcomes and are often quite economically advantageous. Our application of a novel dynamic model yields findings consistent with the existing literature, building the robustness of the link between the built environment, physical activity, and health benefits. Further, we demonstrate that dynamic models may be applied across a variety of scales and are able to incorporate differential treatment effects for different age groups and for men and women. Thus, dynamic models may help address identified limitations of transportation HIA in practice.

4.2. Limitations. Our estimates of post-construction physical activity do not consider activity substitution (i.e., reducing other activities after increasing transportation physical activity) or self-selection (i.e., more active individuals may be more likely to increase transportation physical activity). However, longitudinal evidence suggests that activity substitution is minimal, and increases in physical activity remain when self-selection is accounted for [51–53]. In addition, our estimates exclude potential increases in physical activity from walking for leisure and from bicycling and, in this regard, could underestimate health benefits.

Additionally, we consider only one health pathway (physical activity), while transportation influences health in other ways, including exposure to air pollution and crash risk. Other health pathways may respond to built environment changes in opposite directions and with different magnitudes. For example, compact urban forms may increase physical activity but also increase exposure to air pollution [54]. A recent HIA in London found health benefits from increased physical activity but also negative health impacts from increased exposure to air pollution and elevated crash risk for active commuters [55]. However, recent HIAs of active transportation consistently find changes in physical activity to be the largest contributor to estimated health impacts [14].

While DYNAMO-HIA is able to use continuous relative risk functions, continuous prevalence data are also required when doing so and must be characterized using the mean, standard deviation, and skewness of the distribution. Baseline distributions of walking for transportation were noncontinuous (taken from categorical survey responses) and difficult to characterize as continuous distributions due to excess zeroes. Further, continuous dose-response functions were not available linking walking for transportation with CHD, type 2 diabetes, hypertension, or stroke. To overcome these difficulties, the model uses a discrete dose-response function that caps health benefits at 150 minutes of transportation physical activity per week. As a result, the model may underestimate benefits for those accruing more than 150 minutes of transportation physical activity per week. To analyze the potential magnitude of this underestimation, we recomputed the static (HEAT) model predicted mortality reduction using a continuous dose-response function combined with categorical prevalence data using smaller bins (i.e., divided into eleven categories of weekly time spent walking for transportation). The latter model estimates an additional 26 (+63%) avoided deaths after 40 years. However, since both these models are prone to overestimation, this difference may be artificially inflated.

This paper considered only three communities in North Carolina. While representing a range of urban development contexts (rural, suburban, and urban), all three communities had low baseline levels of transportation physical activity and limited public transit service. Further, community-specific disease prevalence and incidence may reflect population characteristics specific to North Carolina. Thus, our findings concerning the relative costs and benefits of the planned infrastructure investments in these three communities may not generalize to highly urban settings with higher baseline levels of transportation physical activity, higher levels of public transit usage, and/or different demographic characteristics than North Carolina. However, the differences revealed comparing estimates from DYNAMO-HIA and the HEAT model stem from the different structures of the modeling approaches themselves and thus may be generalizable across communities of many types.

Finally, disease prevalence and incidence are estimated using county data. However, these data are identical in the baseline and intervention scenarios so any resulting bias is likely minimal.

5. Conclusion

Using DYNAMO-HIA to conduct three quantitative HIAs, we demonstrated that investments in infrastructure that supports active transportation may have meaningful impacts on health outcomes via increased transportation physical activity. These health outcomes may also have considerable financial implications: in two of the three cases, the benefits of avoided disease and premature mortality alone exceeded construction costs.

Dynamic health impact models, such as DYNAMO-HIA, offer significant advantages over static models, such as HEAT. Static models may overestimate health benefits by failing to account for changing population health characteristics over time. However, it may be difficult to implement continuous relative risk functions using existing dynamic modeling tools if baseline exposure information is difficult to characterize as continuous distributions or if continuous dose-response information is available only for certain health outcomes. If continuous dose-response functions are discretized into just a few categories, the benefits of physical activity may be underestimated for individuals who are very physically active. Providing greater flexibility in characterizing exposure or allowing continuous dose-response functions to be used alongside categorical exposure data in existing tools would address this shortcoming in practice. Overall, the advantages of dynamic models outweigh the current limitations of available tools.

Quantitative HIA is a feasible tool for objective, evidence-based decision support linking health outcomes to increased—or decreased—physical activity resulting from changes in the built environment. Transportation decision-makers routinely use models to estimate congestion reduction and improvement in traffic safety and translate these outcomes into monetary benefits [56]. Thus, quantitative HIA combined with economic valuation enables the health benefits of increased transportation physical activity from changes in the built environment to be considered alongside traditional transportation metrics. As transportation agencies search for ways to better integrate health considerations into transportation decision-making, quantitative HIA fills a critical gap, translating investment in infrastructure that supports active travel into a metric that enables direct comparison with other types of projects. Further, quantitative assessments of competing built environment risks, such as physical activity, air pollution, and traffic fatalities, may help align larger planning efforts (e.g., comprehensive plans) with health goals by comparing the public health impacts of alternative future scenarios. Using three cases across North Carolina, we demonstrated that quantitative models linking built environment changes to physical activity and health impacts are feasible, provide meaningful results to decision-makers, and may help prioritize resources in pursuit of public health goals.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by the North Carolina Department of Transportation and by funds from the NIH (T32ES007018). The authors would like to acknowledge the valuable contributions of Philip Bors and Lauren Blackburn along with their advisory team: Julie Hunkins, Helen Chaney, Lori Rhew, Ruth Petersen, Monique Bethell, Tim Schwantes, Chuck Flink, Matt Hayes, Jennifer MacDougall, Jackie Epping, and Candace Rutt.

References

[1] US Burden of Disease Collaborators, "The state of US health, 1990–2010: burden of diseases, injuries, and risk factors," *Journal of the American Medical Association*, vol. 310, no. 6, pp. 591–606, 2013.

[2] R. Ewing and R. Cervero, "Travel and the built environment: a meta-analysis," *Journal of the American Planning Association*, vol. 76, no. 3, pp. 265–294, 2010.

[3] A. E. Bauman, R. S. Reis, J. F. Sallis, J. C. Wells, R. J. F. Loos, and B. W. Martin, "Correlates of physical activity: why are some people physically active and others not?" *The Lancet*, vol. 380, no. 9838, pp. 258–271, 2012.

[4] I.-M. Lee, R. Ewing, and H. D. Sesso, "The built environment and physical activity levels: the Harvard Alumni Health Study," *American Journal of Preventive Medicine*, vol. 37, no. 4, pp. 293–298, 2009.

[5] US Department of Transportation Research and Innovative Technology Administration and John A. Volpe National Transportation Systems Center, *Metropolitan Area Transportation Planning for Healthy Communizes*, Federal Highway Administration, Office of Planning, Washington, DC, USA, 2012.

[6] US Department of Transportation Research and Innovative Technology Administration and John A. Volpe National Transportation Systems Center, *Statewide Transportation Planning for Healthy Communizes*, Federal Highway Administration, Office of Planning, Washington, DC, USA, 2014.

[7] National Research Council, *Improving Health in the United States: The Role of Health Impact Assessment*, National Academies Press, Washington, DC, USA, 2011.

[8] A. Wernham, "HIA in the United States: state of practice, and future directions," in *Proceedings of the National Health Impact Assessment Meeting*, Washington, DC, USA, 2013.

[9] R. Bhatia and E. Seto, "Quantitative estimation in health impact assessment: opportunities and challenges," *Environmental Impact Assessment Review*, vol. 31, no. 3, pp. 301–309, 2011.

[10] G. R. McCormack and A. Shiell, "In search of causality: a systematic review of the relationship between the built environment and physical activity among adults," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, article 125, 2011.

[11] J. M. MacDonald, R. J. Stokes, D. A. Cohen, A. Kofner, and G. K. Ridgeway, "The effect of light rail transit on body mass index and physical activity," *American Journal of Preventive Medicine*, vol. 39, no. 2, pp. 105–112, 2010.

[12] D. B. Hess and J. K. Russell, "Influence of built environment and transportation access on body mass index of older adults: survey results from Erie County, New York," *Transport Policy*, vol. 20, pp. 128–137, 2012.

[13] S. Kahlmeier, P. Kelly, C. Foster et al., *Health Economic Assessment Tools (HEAT) for Walking and for Cycling: Methodology and User Guide, 2014 Update*, World Health Organization Regional Office for Europe, Copenhagen, Denmark, 2014.

[14] N. Mueller, D. Rojas-Rueda, T. Cole-Hunter et al., "Health impact assessment of active transportation: a systematic review," *Preventive Medicine*, vol. 76, pp. 103–114, 2015.

[15] S. K. Lhachimi, W. J. Nusselder, H. A. Smit et al., "DYNAMO-HIA: a dynamic modeling tool for generic health impact assessments," *PLoS ONE*, vol. 7, no. 5, Article ID e33317, 2012.

[16] M. A. Hendriksen, J. M. van Raaij, J. M. Geleijnse, J. Breda, H. C. Boshuizen, and O. Y. Gorlova, "Health gain by salt reduction in Europe: A Modelling Study," *PLOS ONE*, vol. 10, no. 3, Article ID e0118873, 2015.

[17] A. L. Holm, H. Brønnum-Hansen, K. M. Robinson, and F. Diderichsen, "Assessment of health impacts of decreased smoking prevalence in Copenhagen: application of the DYNAMO-HIA model," *Scandinavian Journal of Public Health*, vol. 42, no. 5, pp. 409–416, 2014.

[18] H. C. Boshuizen, S. K. Lhachimi, P. H. M. van Baal et al., "The DYNAMO-HIA model: an efficient implementation of a risk factor/chronic disease Markov model for use in Health Impact Assessment (HIA)," *Demography*, vol. 49, no. 4, pp. 1259–1283, 2012.

[19] North Carolina Department of Transpiration, *WalkBikeNC: North Carolina Statewide Pedestrian and Bicycle Plan*, North Carolina Department of Transportation, Raleigh, NC, USA, 2013, http://www.ncdot.gov/bikeped/download/WalkBikeNC-PlanChapterslowres.pdf.

[20] J. MacDonald Gibson, D. Rodriguez, T. Dennerlein, J. Mead, and S. Bevington, *Predicting Effects of Urban Design on Public Health: A Case Study in Raleigh, North Carolina*, University of North Carolina, Chapel Hill, NC, USA, 2014, http://www.pewtrusts.org/en/~/media/Assets/External-Sites/Health-Impact-Project/BlueRidgeCorridor.pdf.

[21] Urban Design Associates, JDavis Architects, M. A. Bryson, RCLCO, and Long Leaf Historic Resources, *Blue Ridge Road District Study*, City of Raleigh, Raleigh, NC, USA, 2012, http://www.raleighnc.gov/home/content/PlanUrbanDesign/Articles/BlueRidgeRoadDistrictStudy.html.

[22] Greenways Incorporated and Kimley-Horn and Associates, *Greenville Plan Greenville Bicycle and Pedestrian Master Plan*, 2011, http://www.greenways.com/downloads/GvilleNC_FullPlan_LowRes.pdf.

[23] Destination by Design Planning Group, *Downtown Sparta Streetscape Strategy*, 2012.

[24] G. Hu, P. Jousilahti, K. Borodulin et al., "Occupational, commuting and leisure-time physical activity in relation to coronary heart disease among middle-aged Finnish men and women," *Atherosclerosis*, vol. 194, no. 2, pp. 490–497, 2007.

[25] G. Hu, C. Sarti, P. Jousilahti, K. Silventoinen, N. C. Barengo, and J. Tuomilehto, "Leisure time, occupational, and commuting physical activity and the risk of stroke," *Stroke*, vol. 36, no. 9, pp. 1994–1999, 2005.

[26] G. L. Furie and M. M. Desai, "Active transportation and cardiovascular disease risk factors in U.S. adults," *American Journal of Preventive Medicine*, vol. 43, no. 6, pp. 621–628, 2012.

[27] P. Kelly, S. Kahlmeier, T. Götschi et al., "Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 11, no. 1, p. 132, 2014.

[28] J. M. Gibson, D. Rodriguez, T. Dennerlein et al., "Predicting urban design effects on physical activity and public health: a case study," *Health & Place*, vol. 35, pp. 79–84, 2015.

[29] US Census Bureau, "Census 2010, Summary File 1, Table DP-1," Generated by Theodore Mansfield using American FactFinder, 2013, http://factfinder2.census.gov.

[30] North Carolina State Center for Health Statistics, *North Carolina Population Estimates: 2009*, 2009, http://www.schs.state.nc.us/schs/data/population/sdpop.cfm.

[31] North Carolina State Center for Health Statistics, "North Carolina Selected Vital Statistics Volume 1–2009," 2009, http://www.schs.state.nc.us/schs/vitalstats/volume1/2009/.

[32] North Carolina State Center for Health Statistics, "Behavioral risk factor surveillance system survey calendar year 2009 results," 2009, http://www.schs.state.nc.us/schs/brfss/2009/index.html.

[33] R. Brinks, "A new method for deriving incidence rates from prevalence data and its application to dementia in Germany," http://arxiv.org/abs/1112.2720.

[34] United States Centers for Disease Control and Prevention, *Physical Activity Guidelines for Americans*, 2008, http://www.cdc.gov/physicalactivity/everyone/guidelines/index.html.

[35] C. L. Craig, A. L. Marshall, M. Sjöström et al., "International physical activity questionnaire: 12-country reliability and validity," *Medicine and Science in Sports and Exercise*, vol. 35, no. 8, pp. 1381–1395, 2003.

[36] L. D. Frank, J. F. Sallis, B. E. Saelens et al., "The development of a walkability index: application to the neighborhood quality of life study," *British Journal of Sports Medicine*, vol. 44, no. 13, pp. 924–933, 2010.

[37] R. Cervero and K. Kockelman, "Travel demand and the 3Ds: density, diversity, and design," *Transportation Research Part D: Transport and Environment*, vol. 2, no. 3, pp. 199–219, 1997.

[38] J. F. Sallis, B. E. Saelens, L. D. Frank et al., "Neighborhood built environment and income: examining multiple health outcomes," *Social Science & Medicine*, vol. 68, no. 7, pp. 1285–1293, 2009.

[39] Y. Fan, *The built environment, activity space, and time allocation: an activity-based framework for modeling the land use and travel connection [Doctoral dissertation]*, Department of City and Regional at the University of North Carolina, Chapel Hill, NC, USA, 2007.

[40] M. G. Boarnet, M. Greenwald, and T. E. McMillan, "Walking, urban design, and health: toward a cost-benefit analysis framework," *Journal of Planning Education and Research*, vol. 27, no. 3, pp. 341–358, 2008.

[41] R. C. Browning, E. A. Baker, J. A. Herron, and R. Kram, "Effects of obesity and sex on the energetic cost and preferred speed of walking," *Journal of Applied Physiology*, vol. 100, no. 2, pp. 390–398, 2006.

[42] US Department of Transportation, *Guidance on Treatment of the Economic Value of a Statistical Life in U.S. Department of Transportation Analyses*, US Department of Transportation, Washington, DC, USA, 2014, http://www.dot.gov/sites/dot.gov/files/docs/VSL_Guidance_2014.pdf.

[43] The Milken Institute, *An Unhealthy America: The Economic Impact of Chronic Disease*, 2015, http://www.chronicdisease-impact.com/ebcd.taf.

[44] North Carolina Department of Transportation, "2012 Bid Averages: Statewide [Data file]," 2013, https://connect.ncdot.gov/letting/LetCentral/2012%20Bid%20Average-English.xls.

[45] Office of Management and Budget, *Circular A-94: Guidelines for Discount Rates for Benefit-Cost Analysis of Federal Programs*, Office of Management and Budget, Washington, DC, USA, 1992, http://www.whitehouse.gov/sites/default/files/omb/assets/a94/a094.pdf.

[46] North Carolina Department of Transportation, *From Policy to Projects: 2040 North Carolina Statewide Transportation Plan*, North Carolina Department of Transportation, Raleigh, NC, USA, 2012, http://www.ncdot.gov/download/performance/NCDOT_2040TransportationPlan.pdf.

[47] A. Macmillan, J. Connor, K. Witten, R. Kearns, D. Rees, and A. Woodward, "The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling," *Environmental Health Perspectives*, vol. 122, no. 4, pp. 335–344, 2014.

[48] J. Y. Guo and S. Gandavarapu, "An economic evaluation of health-promotive built environment changes," *Preventive Medicine*, vol. 50, supplement 1, pp. S44–S49, 2010.

[49] G. Deenihan and B. Caulfield, "Estimating the health economic benefits of cycling," *Journal of Transport and Health*, vol. 1, no. 2, pp. 141–149, 2014.

[50] T. Gotschi, "Costs and benefits of bicycling investments in Portland, Oregon," *Journal of Physical Activity and Health*, vol. 8, supplement 1, pp. S49–S58, 2011.

[51] S. Sahlqvist, A. Goodman, A. R. Cooper, and D. Ogilvie, "Change in active travel and changes in recreational and total physical activity in adults: longitudinal findings from the iConnect study," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 10, article 28, 2013.

[52] A. Goodman, S. Sahlqvist, and D. Ogilvie, "New walking and cycling routes and increased physical activity: one- and 2-year findings from the UK iConnect study," *The American Journal of Public Health*, vol. 104, no. 9, pp. e38–e46, 2014.

[53] H. M. Badland, M. Oliver, R. A. Kearns et al., "Association of neighbourhood residence and preferences with the built environment, work-related travel behaviours, and health implications for employed adults: findings from the URBAN study," *Social Science & Medicine*, vol. 75, no. 8, pp. 1469–1476, 2012.

[54] T. J. Mansfield, D. A. Rodriguez, J. Huegy, and J. M. Gibson, "The effects of urban form on ambient air pollution and public health risk: a case study in raleigh, north Carolina," *Risk Analysis*, vol. 35, no. 5, pp. 901–918, 2015.

[55] J. Woodcock, M. Givoni, and A. S. Morgan, "Health impact modelling of active travel visions for England and Wales using an Integrated Transport and Health Impact Modelling Tool (ITHIM)," *PLoS ONE*, vol. 8, no. 1, Article ID e51462, 2013.

[56] E. Gwee, G. Currie, and J. Stanley, "International variation in cost-benefit analysis of urban rail projects: impact on outcomes," *Transportation Research Record*, vol. 2261, pp. 73–85, 2011.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 618640, 8 pages
http://dx.doi.org/10.1155/2015/618640

Research Article

Smoke-Free Workplaces Are Associated with Protection from Second-Hand Smoke at Homes in Nigeria: Evidence for Population-Level Decisions

Dorota Kaleta,[1] Kinga Polanska,[1] and Bukola Usidame[2]

[1]*Department of Tobacco Control, Preventive Medicine, Medical University of Łódź, 90-752 Łódź, Poland*
[2]*Department of Public Policy, University of Massachusetts, Boston, MA, USA*

Correspondence should be addressed to Dorota Kaleta; dkaleta@op.pl

Received 6 May 2015; Revised 12 June 2015; Accepted 2 July 2015

Academic Editor: Giedrius Vanagas

The evidence suggests that smoke-free workplace policies may change social norms towards exposing others to second-hand smoke at home. The aim of the study was to assess whether being employed in a smoke-free workplace (SFWP) is associated with living in a smoke-free home (SFH). We used the data from the Global Adult Tobacco Survey conducted in Nigeria in 2012, in which 9,765 individuals were interviewed including 1,856 persons who worked indoors. The percentage of Nigerians employed in SFWP that reported living in a SFH was higher compared to those employed in a workplace where smoking occurred (95% versus 73%). Working in a SFWP was associated with a significantly higher likelihood of living in a SFH (OR = 5.3; $p < 0.001$). Urban inhabitants indicated more frequently that they lived in SFH compared to rural residents (OR = 2.0; $p = 0.006$). The odds of living in a SFH were significantly higher among nonsmokers and nonsmokeless tobacco users compared to smokers and smokeless tobacco users (OR = 28.8; $p < 0.001$; OR = 7.0; $p < 0.001$). These findings support the need for implementation of comprehensive smoke-free policies in Nigeria that result in substantial health benefits.

1. Introduction

Taking into account the level of exposure and its health consequences, tobacco smoking and second-hand smoke (SHS) constitute one of the biggest public health threats [1–4]. According to the most recent data, about 40% of children and a third of nonsmoking adults are exposed to SHS [1, 5]. The estimates of worldwide burden of a disease indicated that approximately 600,000 deaths were attributed to SHS with 47% of deaths occurring in women, 28% in children, and 26% in men [5].

Over the years many policies have been implemented in order to improve the health of particular populations [6–9]. In 2008, the World Health Organization (WHO) identified six evidence-based tobacco control measures that are the most effective in reducing exposure to cigarette smoking and environmental tobacco smoke [6]. Among them creation of smoke-free public places and workplaces continues to be the most commonly established measure with the highest level of achievement [10].

Scientific evidence, unequivocally, indicates that there is no safe level of exposure to SHS and that an environment which is completely smoke-free and does not allow any exceptions is the only proven way to fully protect people from the harm of that exposure [2–4, 10]. The increase in the number of countries with comprehensive smoke-free legislation shows that effective laws are relatively easy to pass and enforce and involve little or no cost [10]. This policy measure has high levels of public support, causes no financial harm to businesses, and improves the health of both smokers and nonsmokers. Implementation of smoke-free environments including smoke-free workplaces has been associated with a reduction in tobacco consumption, an increase in smoking cessation, and consequently a reduction in hospital admissions due to cardiovascular and respiratory diseases [11–18]. Moreover, public places smoking bans and

home smoking bans are not isolated from each other and do not only protect people from the health risk of SHS but also reduce the likelihood that children will start smoking [11–20].

Despite the existing scientific evidence that creation of 100% smoke-free environments is an effective and inexpensive way of protecting residents from health and economic consequences of tobacco smoke exposure, Nigeria has still not introduced comprehensive smoke-free measures [21–26]. Under provisions of the smoking decree number 20 of 1990, reviewed in 2001, smoking in specific public places such as health-care, educational (except for universities), and governmental facilities is banned in Nigeria [21, 22]. On the other hand, smoke-free environments are still not created and reinforced in the indoor offices, restaurants, cafes, pubs, and bars.

The Global Adult Tobacco Survey (GATS) in Nigeria, as in the first country in Sub-Saharan Africa, provides a special insight into the country's tobacco use and control measures context. Based on the survey estimates in 2012, close to 5 million (5.6%) of Nigerian adults aged 15 years or older used tobacco products and 3 million of them (3.9% of the population) smoked tobacco [22]. Proportion of smokers is higher for men (7.3%) than it is for women (0.4%). In addition, estimated 17.3% of adults (2.7 million people) had been exposed to SHS in their workplaces and 6.6% (5.2 million people) at their homes.

The prevalence of active and passive smoking in Nigeria is not as high as in other low- and middle-income countries but given the high number of smokers and passive smokers together with an increasing trend in tobacco consumption exposure to tobacco smoke is becoming a growing public health threat [22]. The studies performed in Nigeria indicate that despite a high level of awareness of the dangers of SHS and positive attitudes to smoke-free laws, a high proportion of the population is still exposed to SHS in public places, which calls for policy level interventions to improve the implementation of the smoke-free law [23–26]. In addition, the analysis performed in Osun State in Nigeria indicated poor awareness of the existing law of prohibition of smoking in public places which generate the necessity to increase sensitization of the general public [26]. Evaluation of the association between smoke-free public places and SHS exposure at home might be crucial for strengthening implementation and enforcement of the comprehensive legislation prohibiting smoking in public and in workplaces in the country.

The aim of this study was to assess whether being employed in a smoke-free workplace is associated with living in a smoke-free home in Nigeria.

2. Materials and Methods

2.1. Study Design and Population. Data used for the current analysis is available from the Global Adult Tobacco Survey (GATS), which was conducted in Nigeria in 2012. The complete description of the methodological assumptions has been published elsewhere [22]. Briefly, the survey was designed to generate precise cross-sectional estimates at the national level. The final probability selection of the sample units was equivalent to that of being selected under the three-stage stratified-cluster sampling in order to produce key indicators by gender, for the country as a whole as well as classified by residence (urban or rural), and to allow for comparison of the estimates among the 6 geopolitical zones of Nigeria. Following the GATS sampling protocol, a sample of at least 8,000 respondents was required with 4,000 adults each from urban and rural areas. The household sample size was then adjusted to account for the potential sample size loss due to ineligibility and nonresponse. A total of 11,107 households were sampled, of which 5,776 were from urban areas and 5,331 were from rural areas. One eligible household member was randomly selected from each participating household, which resulted in 9,765 individuals completing the survey. GATS Nigeria included a household questionnaire and an individual questionnaire. The questionnaires were applied during face to face interviews with the persons who were 15 years of age or older and they were recorded on an electronic data collection device.

The overall response rate for GATS Nigeria was 89.1%. The household response rate was 90.3% (86.8% urban, 94.1% rural), while the individual response rate was 98.6% (98.0% urban, 99.2% rural) [22].

Current analysis is restricted to the GATS respondents working indoors or both indoors and outdoors but outside their home (2277 participants). After removing people with missing variables, the final analysis is focused on 1856 people (82% of the population that reported indoor work).

2.2. Variables. For the purpose of the current analysis the dependent variable was "living in a smoke-free home." The participant was classified as living in a smoke-free home if she/he answered "smoking is never allowed inside my home" to the question "Which of the following best describes the rules about smoking inside your home?" If the answer was "smoking is allowed inside my home," "smoking is generally not allowed inside my home but there are exceptions," or "there are no rules about smoking at home" he/she was considered as not living in a smoke-free environment. The independent variable was "working in a smoke-free environment" based on answer "no" to the question "During the past 30 days, did anyone smoke in the indoor areas where you work?"

Additional factors included in the analysis were as follows: age (15–29, 30–44, 45–59, 60 years, and above), gender (male, female), number of people in the household, residence (urban, rural), region (North East, North Central, North West, South East, South West, South-South), and education (no formal education, primary school completed, secondary school completed, higher secondary school completed, college/university, and above). Based on the question "Which activity best describes your main work status over the past 12 months?" the participants were divided into the employees (including government and nongovernment employees) and self-employed. If the study subjects indicated current smoking on a daily or less than a daily basis they were considered smokers. Similar classification was considered for smokeless tobacco use (yes, if the participant indicated usage of these products daily or occasionally).

2.3. Statistical Analysis. The STATISTICA Windows XP version 10.0 program was used to carry out the statistical analysis. Initially, a descriptive analysis for all the variables involved in the analysis was completed. Univariate and multivariate logistic regression analyses were run to estimate the odds ratio (OR) and 95% confidence intervals (95% CI) of living in a smoke-free home if employed in a smoke-free workplace compared with being employed in a workplace where smoking occurred. The logistic regression model was adjusted for all the covariates (significant at a 0.1 level) to reduce the risk of confounding. Age and number of household members were treated as continuous variables. Analysis is performed for total study population as well as for smokers and nonsmokers separately (see S1–S3 in Supplementary Material available online at http://dx.doi.org/10.1155/2015/618640). We tested for multicollinearity for covariates that were controlled for in the analysis. The multicollinearity diagnostics variance inflation factors (VIF) were all less than five, which indicates that the assumption of reasonable independence among predictor variables was met.

3. Results

3.1. Characteristics of the Study Participants. Most of the study subjects who worked indoors but out of their homes were younger than 45 (Table 1). Males constituted 60.6% of the population included in analysis. About 66% of the respondents indicated that they completed a higher secondary school or have a college/university degree. Similar percentages of the study participants lived in urban areas. 5.8% of the respondents were current smokers, whereas of the people working indoors, 14.6% indicated SHS exposure in the workplace and 7.9% at home. Among the self-employed nonsmokers, 15.8% indicated exposure to SHS in the workplace, whereas among the nonsmoking employees, 8.2% reported the same (Table 2). For the smokers those percentages were 63.6% among the self-employed and 26.4% among the employees.

3.2. Predictors of Smoke-Free Home. Among the Nigerians who reported smoke-free workplaces 95.4% lived in smoke-free homes, whereas among those who indicated SHS exposure in their workplaces smoke-free home was declared by only 73.1% (Table 1). Additional analysis, which was performed for the smokers, indicated that among those who had a smoke-free work environment about half also lived in smoke-free homes (54.2%). The percentage of smoke-free homes was indicated much less frequently by smokers who declared SHS exposure in their workplace (22.4%) (Supplementary Table S1). Among the nonsmokers, smoke-free home was declared by 84.2% of those exposed to SHS in the workplace and 97.0% of those who declared a smoke-free workplace (Supplementary Table S2). The highest proportion of the people who lived in smoke-free homes was observed in the South West region (95.4%) and among the participants with college/university degrees (94.9%). The people who lived in urban areas were more likely to indicate smoke-free homes compared to those from rural areas (94.7% versus 87.1%). The current smokers and smokeless tobacco users were less likely to live in smoke-free homes compared to those who did not indicate any of these habits (39.8% versus 95.4% and 63.6% versus 92.5%, resp.).

Table 3 presents the results of the unadjusted and adjusted logistic regression analyses of the predictors of a smoke-free home. In the univariate model, working in a smoke-free environment was associated with a significantly higher likelihood of living in a smoke-free home (OR = 7.6; $p < 0.001$). This association persisted after including a variety of covariates in the model (OR = 5.3; $p < 0.001$). The analysis performed separately for the smokers and nonsmokers indicated a similar, more than 4 times higher chance of living in a smoke-free home in the case of those working in a smoke-free workplace comparing to the people who declared SHS exposure in the environment of work (among the smokers adjusted OR = 4.4; $p = 0.005$ among the nonsmokers adjusted OR = 4.9; $p < 0.001$) (Supplementary Table S3). Women were significantly more likely to live in a smoke-free home than men (OR = 2.2; $p < 0.001$) in the univariate analysis but not in the multivariate assessment ($p = 0.8$). The people living in urban areas indicated significantly more frequently that they lived in smoke-free homes than those from rural areas in the unadjusted (OR = 2.6; $p < 0.001$) and similarly in the adjusted analyses (OR = 2.0; $p = 0.006$). In the univariate analysis the chance of having smoke-free homes was higher in the South West region (OR = 2.5; $p = 0.002$), but higher in the North West Nigeria in the case of the multivariate logistic regression (OR = 2.3; $p = 0.05$). In the univariate analysis, more respondents with a higher level of education indicated that smoking was never allowed inside their homes compared to those that have completed secondary school education (higher secondary school OR = 2.4; $p = 0.007$; college/university OR = 3.0; $p = 0.003$). However, in the multivariate analysis, those that have completed a primary school (OR = 3.9; $p = 0.003$) and a higher secondary school (OR = 2.9; $p = 0.005$) had a higher chance of living in smoke-free homes. The odds of living in a smoke-free home were significantly higher for the nonsmokers and the nonsmokeless tobacco users relative to those who indicated they were current smokers and smokeless tobacco users (OR = 31.1; $p < 0.001$; OR = 7.0; $p < 0.001$ resp.). The multivariate results confirmed the figures observed in the univariate analysis (OR = 28.8; $p < 0.001$; OR = 7.0; $p < 0.001$, resp.).

4. Discussion

The study indicated that working in a smoke-free workplace was associated with a significantly higher likelihood of living in a smoke-free home after adjusting for a variety of confounders. In addition, the people living in urban areas as well as the nonsmokers and nonsmokeless tobacco users indicated significantly more frequently that they lived in smoke-free homes than those from rural areas, current smokers, and smokeless tobacco users.

The analysis, which utilized data from the GATS conducted in 15 low- and middle-income countries between 2008 and 2011, indicated, similarly as our assessment, positive associations between being employed in a smoke-free workplace with living in a smoke-free home (for 13 countries

Table 1: Descriptive statistics for the respondent characteristics.

	Respondents (employee or self-employed) who worked indoors (out of home) $N = 1856$			
	Total respondents who worked indoors		Smoke-free at home	
	N	%	N	%
Smoke-free in the workplace	1585	85.4	1512	95.4
SHS in the workplace	271	14.6	198	73.1
Age				
15–29	478	25.7	439	91.8
30–44	895	48.2	820	91.6
45–59	365	19.7	342	93.7
60 and above	118	6.4	109	92.4
Gender				
Male	1125	60.6	1014	90.1
Female	731	39.4	696	95.2
Residence				
Rural	622	33.5	542	87.1
Urban	1234	66.5	1168	94.7
Geographical regions				
North East	141	7.6	122	86.5
North Central	206	11.1	184	89.3
North West	253	13.6	225	88.9
South East	266	14.3	239	89.9
South West	718	38.7	685	95.4
South-South	272	14.7	255	93.8
Education				
No formal education	326	17.6	284	87.1
Primary completed	209	11.3	193	92.3
Secondary school completed	100	5.4	86	86.0
Higher secondary school completed	852	45.9	797	93.5
College/university and above	369	19.9	350	94.9
Occupation				
Employee	770	41.5	719	93.4
Self-employed	1086	58.5	991	91.3
Current smoking				
Yes	108	5.8	43	39.8
No	1748	94.2	1667	95.4
Smokeless tobacco use				
Yes	22	1.2	14	63.6
No	1834	98.8	1696	92.5

SHS: second-hand smoke.

Table 2: Second-hand smoke (SHS) exposure among the self-employed and employees depending on their smoking status.

SHS exposure in the workplace	Self-employed $N = 1086$		Employee $N = 770$	
	Smokers $n = 55$	Nonsmokers $n = 1031$	Smokers $n = 53$	Nonsmokers $n = 717$
No	20 (36.4%)	868 (84.2%)	39 (73.6%)	658 (91.8%)
Yes	35 (63.6)	163 (15.8%)	14 (26.4%)	59 (8.2%)

Table 3: Predictors of a smoke-free home. Crude and adjusted odds ratio (95% CI).

	Crude OR (95% CI)	p	Adjusted OR (95% CI)	p
Smoke-free in the workplace				
Yes	7.6 (5.3–10.9)	$p < 0.001$	5.3 (3.4–8.5)	$p < 0.001$
No	1 (ref.)		1 (ref.)	
Age (years)	1.0 (0.99–1.0)	$p = 0.4$		
Gender				
Male	1 (ref.)		1 (ref.)	
Female	2.2 (1.5–3.2)	$p < 0.001$	0.9 (0.6–1.5)	$p = 0.8$
Residence				
Rural	1 (ref.)		1 (ref.)	
Urban	2.6 (1.9–3.7)	$p < 0.001$	2.0 (1.2–3.2)	$p = 0.006$
Regions				
North East	0.8 (0.4–1.5)	$p = 0.4$	0.7 (0.3–1.5)	$p = 0.3$
North Central	1 (ref.)		1 (ref.)	
North West	1.0 (0.5–1.7)	$p = 0.9$	2.3 (1.0–5.2)	$p = 0.05$
South East	1.1 (0.6–1.9)	$p = 0.9$	0.9 (0.4–1.9)	$p = 0.8$
South West	2.5 (1.4–4.4)	$p = 0.002$	1.2 (0.6–2.5)	$p = 0.7$
South-South	1.8 (0.9–3.5)	$p = 0.08$	1.1 (0.5–2.6)	$p = 0.8$
Education				
No formal education	1.1 (0.6–2.1)	$p = 0.8$	2.2 (1.0–4.9)	$p = 0.05$
Primary completed	2.0 (0.9–4.2)	$p = 0.08$	3.9 (1.6–9.7)	$p = 0.003$
Secondary school completed	1 (ref.)		1 (ref.)	
Higher secondary school completed	2.4 (1.3–4.4)	$p = 0.007$	2.9 (1.4–6.1)	$p = 0.005$
College/university and above	3.0 (1.5–6.2)	$p = 0.003$	2.3 (0.9–5.6)	$p = 0.07$
Occupation				
Employee	1.4 (1.0–1.9)	$p = 0.1$	1.3 (0.8–2.2)	$p = 0.9$
Self-employed	1 (ref.)		1 (ref.)	
Current smoking				
Yes	1 (ref.)		1 (ref.)	
No	31.1 (19.9–48.6)	$p < 0.001$	28.8 (16.8–49.5)	$p < 0.001$
Smokeless tobacco use				
Yes	1 (ref.)		1 (ref.)	
No	7.0 (2.9–17.0)	$p < 0.001$	7.0 (2.5–19.3)	$p < 0.001$
Number of people in household	0.99 (0.9–1.0)	$p = 0.7$		

such associations were statistically significant) [11]. However, substantial differences in the percentage of people indicating both smoke-free homes and smoke-free workplaces have been observed between the countries (varied from 21% in China to 75% in Mexico). In Nigeria, such percentages were even higher than those observed in Mexico, Thailand (73%), and Ukraine (71%). This can result from the low prevalence of smoking observed in Nigeria compared to the other countries. In addition, some cultural and religious norms or high level of awareness about the dangers of exposure to SHS, as well as evidence of support for tobacco control laws, could be responsible for the observed results [11, 12, 23–26]. Similar results (as observed in low- or middle-income countries) are indicated in longitudinal studies performed in the US, where living in a country fully covered by a 100% clean indoor air law in workplaces or restaurants/bars was associated with an increased likelihood of having a voluntary 100% smoke-free home rule both for people living with smokers (OR = 7.8, 95% CI = 5.3–11.4) and not living with smokers (OR = 4.1, 95% CI = 3.3–5.2) [27, 28]. Comparably, significant reduction in smoking at home after implementation of comprehensive smoke-free policies has been observed in Ireland and in the UK [29]. Another evaluation by Edwards et al. (2008) has indicated that, in New Zealand between 2003 and 2006, SHS exposure in workplaces decreased from 20% to 8% and proportion of smoke-free homes increased from 64% to 70% [30].

These results provide evidence against arguments that smoke-free legislation may displace smoking from public to private places and can be used as the tool for implementing 100% smoke-free public places in Nigeria. It needs to be stressed that in 2004 Nigeria signed and in 2005 ratified the WHO FCTC, which highlighted the importance of accelerating the implementation of comprehensive tobacco control legislation in this country [21]. The other aspect, which needs

to be considered, is that although the percentages of active and passive smokers in this country are not as high as in other low/middle-income countries, taking into account the extent of the population this constitutes a significant public health problem. This means that implementing comprehensive policy measures might result in significant benefits.

A smoke-free workplace is a cost-effective, public health approach that encourages the important long-term goal of eliminating tobacco use and SHS exposure [7, 9]. The existing evidence indicates that creation of public and private policies to restrict smoking has been found to be an effective approach to promoting cessation including reduction of the average daily consumption of cigarettes, increasing the percentage of smokers contemplating quitting, and increasing the percentage of successful quitting [13–18]. This public health approach can affect large numbers of individuals at minimal cost and thus is an essential component of any successful strategy to promote smoking cessation.

Our results, similar to most low- or middle-income countries, indicate that the people in urban settings were more likely to live in a smoke-free home environment than those from rural areas. This can be explained by different types of dwellings observed in these two areas (of enclosed structure in urban and open space in rural settings) [11]. In addition, the rural areas have the highest level of illiteracy in the country, which might also explain the higher prevalence of tobacco use and SHS exposure [31].

Our results show that there is a higher chance of having smoke-free homes in the North West compared to the North Central region of Nigeria. Data from the GATS Nigeria (for the whole population included in the survey) indicated that the residents of the North Central region (12.6%; 1.3 million) had the highest and those from the North West (3.6%; 0.6 million) the lowest SHS exposure at home among all regions [11]. The differences between the regions can result from sociodemographic, cultural/religious, and economic determinants as well as from the implementation of policy measures and public awareness about the active and passive smoking and their consequences [11].

In the current analysis the self-employed individuals constitute about 60% of the population and in this group, among smokers as well as nonsmokers, the exposure to SHS in the workplace was indicated more frequently than among the employees. The case of a self-employed person still follows the logic that even if a person works away from home, as a one-man business, he or she has no motivation to set a smoke-free policy only for him/herself. It can be assumed that if the study had included the average number of employees in the workplace in the data collection, it would have provided a clearer understanding of the issue of limited smoke-free workplace among the self-employed. This requires further attention and indicates that such a group of workers constitutes the target group for antismoking and policy interventions.

Not surprisingly, the nonsmokers and nonsmokeless tobacco users indicated significantly more frequently that they lived in smoke-free homes than those who declared current smoking status and smokeless tobacco use.

The study has several strengths. The Global Adult Tobacco Survey (GATS) is a cross-sectional, nationally representative survey and covers a large number of respondents obtained from a general population framework, and so it ensures the reliability and validity of the results. In addition, the data obtained for the current analysis are based on similar questions as those used for the assessment performed in other low- and middle-income countries, which guarantees direct comparability of the results [11, 12].

The limitations of the study also need to be considered. Firstly, the lack of verification of self-reported smoking status and SHS exposure at home/workplace can create misclassification. The verification of active or passive smoking by biomarkers or environmental measurements is generally not feasible for the large cross-sectional studies. However, studies indicate that validated self-reported smoking status and SHS exposure are accurate in most studies and correlate well with the biomarker measurements [32]. Secondly, the cross-sectional study in which both variables (smoke-free workplace and smoke-free home) are measured simultaneously limits causal interpretation of our findings. However, the longitudinal studies conducted in other countries have demonstrated that people who worked in smoke-free workplaces are more likely to live in smoke-free homes [27–30]. Poor surveillance of tobacco use in Nigeria means that more robust prepost or longitudinal study designs applied to explain the association cannot yet be employed.

In Nigeria, in the future studies, questions on the most effective and appropriate interventions for different sectors of the workforce (such as men and women, younger and older workers, temporary or casual workers) need to be addressed. Moreover, there is a need to identify the most effective ways of encouraging employee compliance with a smoke-free policy and resource needs of the large, medium, and small enterprises in implementing smoke-free legislation. Based on the policy recommendations for Nigeria the adaptation and implementation of the law must be a collaborative effort between federal, state, and local governments [33].

5. Conclusions

Our results support the evidence that smoke-free workplaces have the important additional effect of stimulating smoke-free homes in Nigeria. Since home remains a major source of SHS exposure for children, this work clearly indicates additional justification for enacting smoke-free workplaces as the motivation for voluntary smoke-free home rules. The results from the GATS can also be used against arguments that a smoke-free legislation may displace smoking from public to private places and strengthen implementation of the 100% smoke-free legislation in Nigeria.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgments

This survey was funded by Bloomberg Initiative to Reduce Tobacco Use, a program of Bloomberg Philanthropies, and implemented by the National Bureau for Statistics (NBS) and under the coordination of the Federal Ministry of Health (FMOH). The World Health Organization (WHO) and Centers for Disease Control and Prevention (CDC) provided technical assistance.

References

[1] M. Eriksen, J. Mackay, and H. Ross, *The Tobacco Atlas*, American Cancer Society, Atlanta, Ga, USA; World Lung Foundation, New York, NY, USA, 4th edition, 2012.

[2] US Department of Health and Human Services, *The Health Consequences of Smoking: 50 Years of Progress. A Report of the Surgeon General*, US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, Ga, USA, 2014.

[3] IARC, *Tobacco Smoke and Involuntary Smoking*, vol. 83 of *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, IARC, Lyon, France, 2004.

[4] IARC, *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Personal Habits and Indoor Combustions Volume 100 E. A Review of Human Carcinogens*, IARC, Lyon, France, 2012.

[5] M. Öberg, M. S. Jaakkola, A. Woodward, A. Peruga, and A. Prüss-Ustün, "Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries," *The Lancet*, vol. 377, no. 9760, pp. 139–146, 2011.

[6] WHO, *O Report on the Global Tobacco Epidemic, 2008. The MPOWER Package*, World Health Organization, Geneva, Switzerland, 2008, http://www.who.int/tobacco/mpower/mpower_report_full_2008.pdf.

[7] K. Cahill and T. Lancaster, "Is the workplace an effective setting for helping people to stop smoking?" Cochrane Database of Systematic Reviews, 2014.

[8] M. Bala, L. Strzeszynski, R. Topor-Madry, and K. Cahill, "Can tobacco control programmes that include a mass media campaign help to reduce levels of smoking among adults?" Cochrane Database of Systematic Reviews, 2013.

[9] J. E. Callinan, A. Clarke, K. Doherty, and C. Kelleher, "Does legislation to ban smoking reduce exposure to secondhand smoke and smoking behaviour?" Cochrane Database of Systematic Reviews, 2010.

[10] WHO, *WHO Report on the Global Tobacco Epidemic, 2013. Enforcing Bans on Tobacco Advertising, Promotion and Sponsorship*, World Health Organization, Geneva, Switzerland, 2013, http://www.who.int/tobacco/global_report/2013/en/index.html.

[11] G. P. Nazar, J. T. Lee, S. A. Glantz, M. Arora, N. Pearce, and C. Millett, "Association between being employed in a smoke-free workplace and living in a smoke-free home: evidence from 15 low and middle income countries," *Preventive Medicine*, vol. 59, no. 100, pp. 47–53, 2014.

[12] J. T. Lee, S. Agrawal, S. Basu, S. A. Glantz, and C. Millett, "Association between smoke-free workplace and second-hand smoke exposure at home in India," *Tobacco Control*, vol. 23, no. 4, pp. 308–312, 2014.

[13] C. M. Fichtenberg and S. A. Glantz, "Effect of smoke-free workplaces on smoking behaviour: systematic review," *British Medical Journal*, vol. 325, no. 7357, pp. 188–191, 2002.

[14] J. M. Moskowitz, Z. Lin, and E. S. Hudes, "The impact of workplace smoking ordinances in California on smoking cessation," *American Journal of Public Health*, vol. 90, no. 5, pp. 757–761, 2000.

[15] A. B. Naiman, R. H. Glazier, and R. Moineddin, "Is there an impact of public smoking bans on self-reported smoking status and exposure to secondhand smoke?" *BMC Public Health*, vol. 11, article 146, 2011.

[16] R. Borland, H.-H. Yong, K. M. Cummings, A. Hyland, S. Anderson, and G. T. Fong, "Determinants and consequences of smoke-free homes: findings from the International Tobacco Control (ITC) Four Country Survey," *Tobacco Control*, vol. 15, supplement 3, pp. iii42–iii50, 2006.

[17] U. Mons, G. E. Nagelhout, S. Allwright et al., "Impact of national smoke-free legislation on home smoking bans: findings from the International Tobacco Control Policy Evaluation Project Europe Surveys," *Tobacco Control*, vol. 22, no. e1, pp. e2–e9, 2013.

[18] R. W. Zablocki, S. D. Edland, M. G. Myers, D. R. Strong, R. Hofstetter, and W. K. Al-Delaimy, "Smoking ban policies and their influence on smoking behaviors among current California smokers: a population-based study," *Preventive Medicine*, vol. 59, pp. 73–78, 2014.

[19] A. J. Farkas, E. A. Gilpin, M. M. White, and J. P. Pierce, "Association between household and workplace smoking restrictions and adolescent smoking," *Journal of the American Medical Association*, vol. 284, no. 6, pp. 717–722, 2000.

[20] P. C. Akhtar, D. B. Currie, C. E. Currie, and S. J. Haw, "Changes in child exposure to environmental tobacco smoke (CHETS) study after implementation of smoke-free legislation in Scotland: national cross sectional survey," *British Medical Journal*, vol. 335, no. 7619, pp. 545–549, 2007.

[21] World Health Organization, *Framework Convention on Tobacco Control*, World Health Organization, Geneva, Switzerland, 2015, http://www.who.int/fctc/about/en.

[22] GATS Nigeria, *Global Adult Tobacco Survey: Country Report*, Federal Ministry of Health, Abuja, Nigeria, 2012.

[23] O. O. Onigbogi, O. Odukoya, M. Onigbogi, and O. Sekoni, "Knowledge and attitude toward smoke-free legislation and second-hand smoking exposure among workers in indoor bars, beer parlors and discotheques in Osun State of Nigeria," *International Journal of Health Policy and Management*, vol. 4, no. 4, pp. 229–234, 2015.

[24] O. O. Desalu, C. C. Onyedum, O. O. Adewole, A. E. Fawibe, and A. K. Salami, "Secondhand smoke exposure among nonsmoking adults in two Nigerian cities," *Annals of African Medicine*, vol. 10, no. 2, pp. 103–111, 2011.

[25] E. O. Poluyi, O. O. Odukoya, B. A. Aina, and B. Faseru, "Tobacco related knowledge and support for smoke-free policies among community pharmacists in Lagos state, Nigeria," *Pharmacy Practice*, vol. 13, no. 1, pp. 486–493, 2015.

[26] S. A. Olowookere, E. G. Adepoju, and O. O. Gbolahan, "Awareness and attitude to the law banning smoking in public places in Osun State, Nigeria," *Tobacco Induced Diseases*, vol. 12, article 6, 2014.

[27] K.-W. Cheng, S. A. Glantz, and J. M. Lightwood, "Association between smokefree laws and voluntary smokefree-home rules," *The American Journal of Preventive Medicine*, vol. 41, no. 6, pp. 566–572, 2011.

[28] K.-W. Cheng, C. A. Okechukwu, R. McMillen, and S. A. Glantz, "Association between clean indoor air laws and voluntary smokefree rules in homes and cars," *Tobacco Control*, vol. 24, no. 2, pp. 168–174, 2015.

[29] G. T. Fong, A. Hyland, R. Borland et al., "Reductions in tobacco smoke pollution and increases in support for smoke-free public places following the implementation of comprehensive smoke-free workplace legislation in the Republic of Ireland: findings from the ITC Ireland/UK Survey," *Tobacco Control*, vol. 15, supplement 3, pp. iii51–iii58, 2006.

[30] R. Edwards, G. Thomson, N. Wilson et al., "After the smoke has cleared: evaluation of the impact of a new national smoke-free law in New Zealand," *Tobacco Control*, vol. 17, article e2, 2008.

[31] R. E. Uyanga, "The indigenization policy and educational advancement in Nigeria," *The International Journal of Diversity in Organisations, Communities and Nations*, vol. 10, no. 6, pp. 199–211, 2011.

[32] A. Florescu, R. Ferrence, T. Einarson, P. Selby, O. Soldin, and G. Koren, "Methods for quantification of exposure to cigarette smoking and environmental tobacco smoke: focus on developmental toxicology," *Therapeutic Drug Monitoring*, vol. 31, no. 1, pp. 14–30, 2009.

[33] I. Agaku, A. Akinyele, and A. Oluwafemi, "Tobacco control in Nigeria—policy recommendations," *Tobacco Induced Diseases*, vol. 10, article 8, 2012.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 136505, 10 pages
http://dx.doi.org/10.1155/2015/136505

Research Article

Worksite Tobacco Prevention: A Randomized, Controlled Trial of Adoption, Dissemination Strategies, and Aggregated Health-Related Outcomes across Companies

Verena Friedrich,[1] Adrian Brügger,[2] and Georg F. Bauer[3]

[1]*Centre of Continuing Education, University of Bern, Schanzeneckstrasse 1, 3001 Bern, Switzerland*
[2]*Institute of Marketing and Management, Department of Consumer Behavior, University of Bern, 3012 Bern, Switzerland*
[3]*Epidemiology, Biostatistics and Prevention Institute, University of Zurich, 8001 Zurich, Switzerland*

Correspondence should be addressed to Georg F. Bauer; georg.bauer@uzh.ch

Received 5 June 2015; Revised 20 August 2015; Accepted 27 August 2015

Academic Editor: Stefan K. Lhachimi

Evidence based public health requires knowledge about successful dissemination of public health measures. This study analyses (a) the changes in worksite tobacco prevention (TP) in the Canton of Zurich, Switzerland, between 2007 and 2009; (b1) the results of a multistep versus a "brochure only" dissemination strategy; (b2) the results of a monothematic versus a comprehensive dissemination strategy that aim to get companies to adopt TP measures; and (c) whether worksite TP is associated with health-related outcomes. A longitudinal design with randomized control groups was applied. Data on worksite TP and health-related outcomes were gathered by a written questionnaire (baseline $n = 1627$; follow-up $n = 1452$) and analysed using descriptive statistics, nonparametric procedures, and ordinal regression models. TP measures at worksites improved slightly between 2007 and 2009. The multistep dissemination was superior to the "brochure only" condition. No significant differences between the monothematic and the comprehensive dissemination strategies were observed. However, improvements in TP measures at worksites were associated with improvements in health-related outcomes. Although dissemination was approached at a mass scale, little change in the advocated adoption of TP measures was observed, suggesting the need for even more aggressive outreach or an acceptance that these channels do not seem to be sufficiently effective.

1. Introduction

Measures for tobacco prevention (TP) at the worksite are a key strategy of tobacco control [1], since many people spend a great part of their day at work, where second-hand smoke (SHS) is still common. As Kramer et al. [2] showed in their synopsis of the literature, structural TP measures (especially smoke-free policies) and behavioural prevention measures (e.g., group interventions and consultations for smokers) at the worksite contribute to reducing the prevalence of smoking and cigarette consumption among staff [3], to improved air quality [4], and reduced absenteeism [5].

However, for these TP measures to have a large public health impact, they must be disseminated across organizations and adopted by them [6, 7]. Laws or regulations for the protection of nonsmokers, which have been introduced in many countries in recent years, have contributed significantly to the dissemination of smoke-free policies. However, as Radtke et al. [8] showed for Switzerland, 28% of the working population were still exposed to SHS at the worksite in 2010, even though the Swiss Federal law for protection against passive smoking was implemented in that year. It is therefore an ongoing challenge for public health professionals to sensitize decision-makers and support companies in the implementation of TP measures, both in countries with and without workplace smoking policies. Although several studies have addressed the dissemination of health promotion interventions associated with different health topics and settings [9], no study has yet addressed the dissemination of TP across worksites.

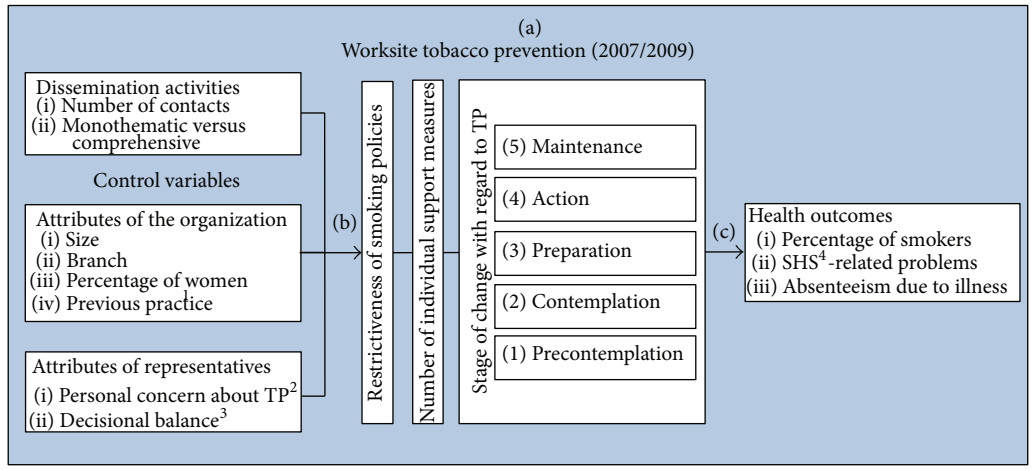


Figure 1: Variables to be analysed with regard to changes of worksite TP (research question a), predictors (question b), and outcomes (question c) of worksite TP. *Notes.* [1]Other workplace health promotion measures; [2]tobacco prevention; [3]pros and cons regarding WHP; [4]second-hand smoke.

Accordingly, this study aims to contribute to the knowledge of dissemination and adoption of TP across companies by addressing the following research questions (Figure 1). It presents follow-up data to a 2007 survey of more than 1400 companies in the Canton of Zurich, Switzerland (for the baseline results see [10]).

(a) In what way did the overall adoption of TP by companies in the Canton of Zurich, Switzerland, change between 2007 and 2009?

(b1) Does the implementation of a multistep dissemination strategy to promote TP in companies produce larger effects regarding the adoption of TP than only sending out an information brochure to companies?

(b2) Does embedding the topic of TP in a comprehensive workplace health promotion (WHP) dissemination strategy lead to better adoption of worksite TP than promoting TP in isolation (monothematic dissemination strategy)?

(c) Does improved worksite TP lead to improved health-related outcomes aggregated at company level?

(a) Firstly, regarding the overall adoption of worksite TP, we examine how its prevalence changed across companies between 2007 and 2009. In this period, the Swiss Federal law for protection against passive smoking was not yet in force, so companies were relatively free in their application of nonsmoking regulations. Worksite TP includes smoking restriction policies (spatial restrictions) and a number of individual support measures offered to smokers (e.g., cessation courses). Besides looking at the full adoption of these preventive measures by companies, we also use the transtheoretical model of change to analyse the five stages of change (SoC) of companies in the adoption process [11]: precontemplation, contemplation, preparation, action, and maintenance. This is a more differentiated, sensitive measure covering both attitudinal and behavioural aspects associated with future behaviour change and clearly distinguishes between short-term behaviour change (action) and long-term behaviour change (maintenance).

(b1) Secondly, we analyse if the implementation of a multistep dissemination strategy in companies to promote TP is more effective for the adoption of TP than only sending out an information brochure. A systematic review of dissemination studies concluded that more active, multimodal dissemination and implementation strategies are more likely to be effective [9]. In comparing the "brochure only" to the multistep strategy, we control for variables that are known to be associated with TP at worksites [10]: the characteristics of the organization (e.g., size, industry type, and previous WHP practice) and the characteristics of the decision-makers within the organization (personal concern for TP, perceived advantages of TP).

(b2) Thirdly, we examine whether an approach to dissemination presenting TP as an integral part of a comprehensive WHP strategy yields better results than an approach presenting TP in isolation. The rationale for comprehensive WHP has been widely discussed [12, 13]. Regarding worksite TP, the aim of integrating it into a broader WHP strategy is to discuss factors associated with smoking cessation (e.g., body weight and diet; assumed stress relief) or with the introduction of smoke-free worksites (e.g., corporate communications, team climate) in the more positively connoted context of promoting health at work. In comparing the comprehensive and monothematic dissemination strategies, we control for the above mentioned variables (characteristics of the organisation and the representative).

(c) Fourthly, we examine whether improved worksite TP is associated with improved health outcomes aggregated at the company level. We included outcomes reported to be positively affected by workplace TP: the percentage of smokers in the workforce [3], SHS-exposure and related complaints [4], and the absenteeism rate due to illness [5].

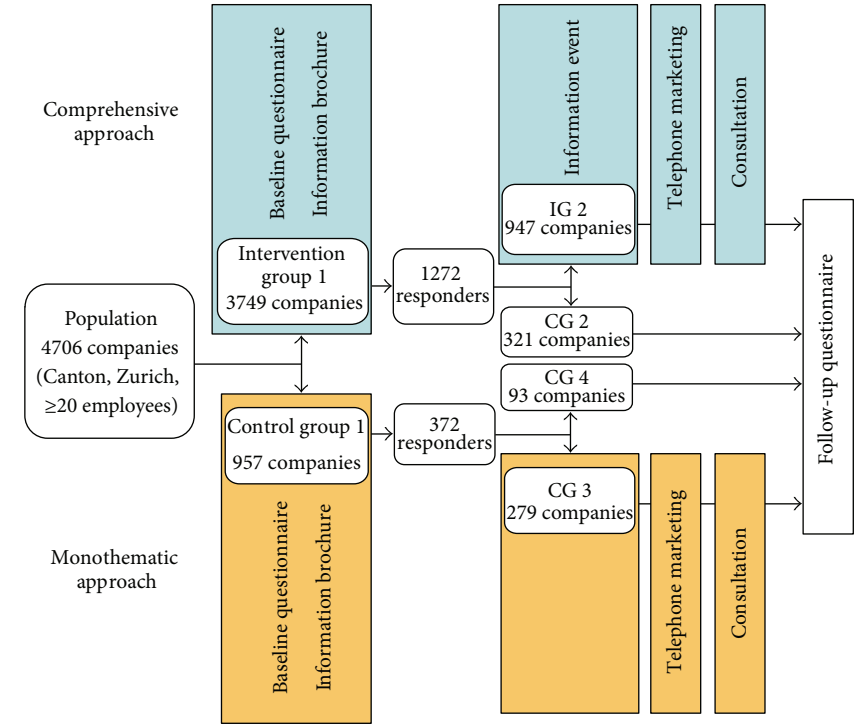


Figure 2: Intervention and study design.

2. Material and Methods

2.1. Study Population and Study Design. The study included all companies in the Canton of Zurich, Switzerland, with 20 or more employees ($N = 4706$), varying in their baseline levels of worksite TP (see Table 1). This study population was randomly divided into intervention and control groups (simple random allocation using SPSS random number allocator; see Figure 2). Intervention group 1 received a "comprehensive" dissemination strategy in which TP was part of a comprehensive approach to WHP. Control group 1 received a "monothematic" dissemination strategy with TP as a singular topic. The comprehensive dissemination strategy was carried out by Zurich University's WHP consulting team. The monothematic dissemination strategy was implemented by TP specialists from the Swiss Lung League, a national organization for the prevention and control of lung diseases.

In a first step, both intervention group 1 and control group 1 received bulk mailings with information brochures and the baseline questionnaire. The "comprehensive" brochure explained the benefits of worksite TP and WHP and of their integration, showed how to proceed in practice, and gave information about the consulting agency's offers. The "monothematic" brochure provided information about the benefits of worksite TP, about related support services, and invited companies to register on a website as smoke-free workplaces. As the subsequent dissemination activities could not be offered to all companies due to limited resources, the companies who had answered the baseline questionnaire were further randomly subdivided into intervention group 2 (IG 2; comprehensive intervention), control group 3 (CG 3; monothematic intervention), and control groups 2 and 4 (CG 2 and 4, no further intervention).

Next, intervention group 2 companies were invited to information events. As only 29 of 947 companies participated, this step was omitted for control group 3. The subsequent steps were telephone marketing (selectively calling companies that had expressed interest in WHP services in the baseline questionnaire; intervention group 2: 133 companies reached; control group 3: 80 companies reached) and free initial consultations (intervention group 2: 92 companies; control group 3: 49 companies. These numbers were a subsample of those reached by the previous step). The dissemination activities are described in detail in [14].

2.2. Data Collection. Company addresses and information about company size (number of employees) and branch were obtained from the Federal Statistics Office. Data on dissemination activities implemented by the dissemination teams across companies were recorded in a customer database. The following measures of dissemination activities were derived: presence or absence of any dissemination activity apart from the brochure (1/0); number of contacts between dissemination team and company; and type of dissemination activity (0 = brochure only, 1 = monothematic intervention,

Table 1: Changes in worksite TP between 2007 (n = 1627) and 2009 (n = 1452); cross-sectional data in percent (scale coding in brackets).

	There is no policy (0)	Smoking is allowed anywhere (1)	Smoking is allowed outside and in certain indoor areas (2)	Smoking is allowed outside, not in buildings (3)	Smoking is not allowed anywhere (4)
			Prevalence and restrictiveness of smoking policies		
2007	5.8	4.1	47.6	40.0	2.5
2009	4.9	2.8	41.4	48.6	2.3
			Number of individual support measures*		
	0	1	2	3	
2007	85.0	11.6	2.8	0.6	—
2009	84.2	12.9	2.3	0.6	—
	Precontemplation (1)	Contemplation (2)	Preparation (3)	Action (4)	Maintenance (5)
			Stage of change "smoke-free policy"		
2007	20.8	21.9	3.9	5.9	47.5
2009	14.5	15.8	3.7	5.5	60.5
			Stage of change "cessation courses"		
2007	67.9	24.2	1.5	3.1	3.3
2009	66.9	22.1	1.8	5.7	3.5

*For example, cessation courses, information material, or individual counseling for smokers.

and 2 = comprehensive intervention). The other variables (Figure 1) were measured using a written questionnaire addressed to the human resources or occupational health managers as representatives of the companies who should have the best overview of the relevant variables. The baseline measurement was made in June 2007 ($t1$); a follow-up measurement was made in March 2009 after the dissemination activities had been implemented ($t2$). Both questionnaires were sent to all companies in the Canton of Zurich, Switzerland, with 20 or more employees ($t1$ N = 4706; $t2$ N = 4472). The questionnaire was the same for the baseline and follow-up levels. Details of the variables (number of items, response format) are described in the Appendix and in [10].

2.3. Participating Sample. Of the questionnaires sent to 4706 companies at baseline level, 1648 were returned (response rate after excluding undeliverable questionnaires: 36.5%). Of those, 1627 could be analyzed. For the follow-up, undeliverables were excluded from the addressees, resulting in a population of N = 4472. A total of 1502 questionnaires were completed at follow-up, of which 1452 were suitable for analyses. A total of 244 questionnaires were returned blank because the company had fewer than 20 employees, moved away, or had ceased to exist. After subtracting these from the population, the response rate was 35.5%. A total of 827 companies returned both the baseline and follow-up questionnaires. Compared to data from the Federal Statistics Office, the survey samples are representative of worksites in the Canton of Zurich as regards company size and the four major branches (construction, hospitality, healthcare/welfare, and trading/maintenance/repair), except that healthcare and welfare organizations are slightly overrepresented. Regarding the characteristics of the organizations and representatives, the follow-up sample is comparable to the baseline sample (see Appendix and [10]). At follow-up, 64% of the participating companies had 20–49 employees, 30% 50–250 employees, and 5% more than 250 employees at baseline. Of the representatives who completed the questionnaire at follow-up, 72% were nonsmokers, 48% were CEOs, and 44% were human resource managers. Irrespective of their position, 49% reported that they were authorized to decide on occupational health measures in their companies.

2.4. Data Analyses. Descriptive statistics and Wilcoxon tests were used to examine the changes in worksite TP (policy restrictiveness, number of individual support measures) and stages of change regarding smoke-free policy and cessation course between 2007 and 2009 (research question a).

To answer research questions (b1) and (b2), t-tests and ordinal regression analyses were performed, with control variables (attributes of the organization and the representatives; measured at $t1$) and dissemination strategies as factors ((b1) brochure only versus brochure and other, number of contacts; (b2) monothematic versus comprehensive strategy), and differences in worksite TP (policy restrictiveness, individual support measures) and stages of change (smoke-free policy, cessation course) between 2009 and 2007 ($t2$ minus $t1$ data) as dependent variables.

Ordinal regression analyses were performed to test whether changes in TP (policy restrictiveness, individual support measures) and the organizations' stage of change (smoke-free policy, cessation course) predict improvements in health-related outcomes (percent of smokers, second-hand smoke related problems, and absenteeism; research question c). We used categorized changes in TP (−1 = decrease, 0 = no change, and 1 = increase) as factors and differences ($t2$ minus $t1$ data) in health-related outcomes as dependent variables.

Table 2: Longitudinal changes with regard to worksite TP and health-related outcomes aggregated at company level (Wilcoxon tests).

	Mdn_{t1}[a]	Mdn_{t2}[a]	Decrease	No change	Increase	n	Z	p	r[b]
Worksite TP[1]									
Restrictiveness of smoking policy	2	3	9.4%	68.5%	22.1%	809	−5.941	<.001	−0.21
Number of individual support measures	0	0	9.1%	78.8%	12.1%	827	−1.609	.054	−0.06
SOC[2] smoke-free policy	4	5	12.6%	54.8%	32.6%	786	−7.666	<.001	−0.27
SOC[2] cessation courses	1	1	16.1%	65.6%	18.3%	771	−2.376	.009	−0.09
Health-related outcomes aggregated at company level									
% smokers in the workforce[3]	2	2	17.7%	56.0%	26.3%	723	−3.254	<.001	−0.12
SHS[4]-related problems[5]	1	1	31.4%	49.0%	19.6%	816	−6.195	<.001	−0.22
Absenteeism due to illness[5]	2	2	27.0%	46.3%	26.8%	800	−0.790	.215	−0.03

Note. [a]Median.
[b]Effect size.
[1]For scale coding, see Table 1.
[2]Stage of change.
[3]1 = 0–20%, 5 = 80–100%.
[4]Environmental tobacco smoke.
[5]Answers were given on a five-point scale from 1 ("does not apply") to 5 ("applies").

3. Results

3.1. Changes in Adoption of Worksite TP between 2007 and 2009. A comparison of cross-sectional data in 2007 and 2009 showed that the proportion of organizations with no smoking regulations (or smoking allowed anywhere) was slightly reduced (see Table 1). The largest change was found in tightening the policy from "smoking is allowed outside and in certain indoors areas" to "allowed outside, not in buildings." The number of individual support measures (e.g., cessation courses) did not differ between the two years. In 2009, more companies were in a higher stage of change (SOC) concerning the introduction of smoke-free policies, with a particularly large increase of the maintenance stage. The SOC regarding cessation courses remained low.

Longitudinal analyses showed that the changes within companies participating in the baseline and follow-up surveys ($n = 827$) were also rather small, but conformed to the cross-sectional data. Wilcoxon tests analysed whether changes between $t2$ and $t1$ data in worksite TP (−1 = decrease, 0 = no change, and 1 = increase) were statistically significant (Table 2). The results indicated that more companies tightened their smoking policies than those adopting a looser policy. There were also more companies with an SOC increase in smoke-free policy and cessation courses than those with an SOC decrease. However, effect sizes for these differences were rather small [15].

3.2. Effect of Dissemination Strategies on the Adoption of Worksite TP. In a first step, t-tests assessed whether control groups 2 (comprehensive brochure only) and 4 (monothematic brochure only) differed with regard to changes in worksite TP ($t2$ minus $t1$ data). As shown in Table 3, there were no significant differences between the two groups. Therefore, CG 2 and CG 4 data were pooled for the following analyses.

To analyse the relationship between control variables, interventions, and the adoption of worksite TP, we performed ordinal regression analyses (Table 4). The results showed that construction companies improved more than companies in other branches with regard to the restrictiveness of their TP policy, and hospitality venues improved less. However, construction companies improved less with regard to the number of individual support measures. Health and welfare organizations improved more in their stage of change regarding cessation courses than other sectors. Companies that had adopted many other WHP measures at $t1$ improved less in their SoC regarding a smoke-free policy. The same is true for companies that reported high personal concern and perceived advantages at $t1$.

As to the effect of the dissemination strategies, the results showed that the "brochure only" intervention (control groups 2 and 4) was associated with less improvement in individual support measures than any additional (monothematic or comprehensive) interventions. However, the number of additional contacts were not predictive of the outcomes. In comparing the dissemination strategies, we found no significant differences between intervention group 2 (comprehensive dissemination strategy) and control group 3 (monothematic dissemination strategy). However, the results at least show a statistically not significant tendency for the comprehensive intervention to be superior as regards policy restrictiveness and the related SoC, whereas the monothematic intervention was superior as regards outcomes relating to smoking cessation.

3.3. Effect of Worksite TP on Health-Related Outcomes. In the entire sample, more companies reported an increased percentage of smokers in their workforce than a decreased one (Table 2), whereas problems related to environmental tobacco smoke decreased. Absenteeism due to illness did not change significantly. As regards the question of whether changes in worksite TP predict changes in health-related outcomes, ordinal regression analyses with categorized changes in TP as factors showed that less restrictive policies and a decreased SoC regarding smoke-free policies were associated with an increased percentage of smokers in the workforce and increased SHS-related problems between 2007 and 2009

Table 3: Differences between control group 2 (comprehensive brochure only) and control group 4 (monothematic brochure only) with respect to advances ($t2 - t1$ data) in worksite TP (t-tests).

	Control gr. 2 ($n = 151$[a])		Control gr. 4 ($n = 40$[a])		
	Mean	SD[b]	Mean	SD[b]	t
Differences[1] in					
Restrictiveness of policy	.155	.8628	.244	.4889	−.632
Number of individual support measures	.055	.5325	−.023	.5112	.860
SOC[2] smoke-free policy	.49	1.687	.45	2.062	.139
SOC[2] cessation courses	.05	.889	.10	.810	−.345

Note. [a]Companies who participated in both surveys and answered the respective questions both times.
[b]Standard deviation.
[1]$t2$ data minus $t1$ data; for scale coding, see Table 1.
[2]Stage of change.

(Table 5). A decreased SoC regarding smoke-free policies is also associated with more absenteeism due to illness. Decreased and unchanged numbers of individual support measures are associated with more SHS-related problems, whereas a decreased SoC regarding cessation courses is associated with a higher percentage of smokers in the workforce.

4. Discussion

This study shows that in 2009 only half of the worksites in the Canton of Zurich, Switzerland, had adopted an effective smoking policy (48% banned smoking indoors, 2.3% anywhere). Only few offered individual support measures for smokers. Longitudinal analysis (research question a) showed slight but significant improvements in worksite TP, notably regarding smoking policies. Companies in the construction sector improved more than other companies, probably because they started from a considerably lower level of TP in 2007. In contrast, hospitality venues improved less. Protection from SHS for employees in hospitality venues still remains inconsistent and suboptimal, as the Swiss Federal law for protection against passive smoking from 2010 still allows exemptions from the smoking ban in this sector [16].

The other control variables (previous practice, personal concern, and perceived advantages) that predicted the state of worksite TP at baseline (cf. [10]) are not or negatively associated with improvements in worksite TP, measured as $t2$ minus $t1$ data. This may be due to the fact that companies with high levels of these control variables already had higher levels of TP at $t1$ [10] and thus less room for improvement.

Regarding research question (b1), the results showed that, compared to the "brochure only" groups, additional (monothematic or comprehensive) dissemination activities had a positive impact on the number of individual support measures and the related SoC of companies. However, this effect could not be found for the outcomes relating to smoking policy. This can be explained by the societal context of this study: it took place during the run-up to the Federal law for protection against passive smoking at worksites, which was accompanied by vigorous public debates over the benefit of smoke-free policies, and as part of a general trend towards such policies (e.g., a nonsmoking policy in Swiss trains had been introduced in 2005). Thus, the dissemination activities probably could not generate any effect on worksite smoking policies in *addition* to the historical developments that were strong enough to produce significant results in the longitudinal analysis. Also, the number of contacts between the intervention teams and the companies (in addition to brochure) did not predict improvement in any of the worksite TP measures. Research question (b2) focused on a comparison of the monothematic and comprehensive dissemination strategy. Although the data (bottom of Table 3) suggest that, compared to the monothematic dissemination strategy, the comprehensive strategy led to greater improvements in policy restrictiveness, whereas the monothematic dissemination strategy was superior as regards the improvement of individual support measures, no statistically significant differences between the two dissemination strategies were found.

With respect to research question (c), this study showed that changes in worksite TP are associated with changes in relevant health-related outcomes aggregated on the company level. In particular, a less restrictive policy, less individual support measures, and a decreased stage of change regarding smoke-free policy predict an increase in SHS-related problems. That is, maintaining positive changes in organizations seems to pay off. However, these findings should be interpreted with caution as they rely on self-reporting from a single representative within an organization.

The strengths of the study lie in its longitudinal, randomized controlled trial design, the large sample size, and the real-world setting in which the interventions were studied, as urged by Rabin et al. [9]. The heterogeneity of companies included in this broad field study assures that the results are generalizable to contexts where legal regulations are debated but not yet in force.

However, this field setting also caused some limitations: first, the compared dissemination strategies differed not only with respect to the implementation teams but also with respect to the steps taken: the information events in the intervention group were omitted for the control groups. Second, for feasibility reasons, baseline questionnaires and information brochures were sent in the same mailing; to

Table 4: Bivariate odds ratios for predictors ($t1$ data) of differences in worksite tobacco prevention (ordinal regressions).

	Differences ($t2$ minus $t1$ data) in			
	Policy restrictiveness[a]	Number of individual support measures[a]	SOC[b] smoke-free policy[a]	SOC[b] cessation course[a]
Organizational attributes[1]				
Company size				
<50 employees	1.62	1.62	0.98	1.21
50–250 employees	1.82	1.90	0.97	0.98
>250 employees	ref.	ref.	ref.	ref.
Branch				
Construction	2.27**	0.41**	0.94	0.80
Hospitality	0.46**	1.02	1.18	0.89
Health and welfare	0.83	0.91	0.77	1.58*
Other	ref.	ref.	ref.	ref.
% women in workforce				
<20%	1.60+	0.77	0.83	0.92
20–39%	1.01	0.67	0.83	0.79
40–59%	1.16	0.81	1.14	0.79
60–79%	0.78	0.81	0.68	1.18
>80%	ref.	ref.	ref.	ref.
Other WHP[2]-measures	1.04	0.97	0.86*	0.91
Attributes of the representative[1]				
Personal concern	0.89+	0.90	0.80**	0.88+
Perceived advantages[3]	0.79*	0.94	0.59**	0.94
Dissemination strategies				
Brochure only	0.97	0.53*	0.87	0.56*
Brochure and other[4]	ref.	ref	ref.	ref.
Number of contacts[5]	0.99	1.04	1.02	1.05
Type of intervention				
Monothematic intervention[6]	0.51	1.42	0.80	1.63
Comprehensive intervention[7]	ref.	ref.	ref.	ref.

Note. [a]For scale coding, see Table 1.
[b]Stage of change.
+ $p < .10$; * $p < .05$; ** $p < .01$.
[1]Measured at $t1$.
[2]Workplace health promotion (0 up to 6 measures, as stated in the questionnaire).
[3]Pros and cons for the respective measure, cons recoded.
[4]Information event, telephone marketing, and free initial consultation.
[5]In addition to brochure; as listed in the customer database.
[6]Control group 3 companies with intervention.
[7]Intervention group 2 companies with intervention.

achieve a sound baseline, the questionnaires should have been sent *before* the brochures, also allowing for a no-brochure control group. Third, self-selection in survey participation in this study probably led to an overestimation of the actual prevalence of TP, since more advanced companies are more likely to participate. Also, the companies that participated in the information events and in initial consultations presumably attach greater importance to the wellbeing of their employees and have the resources to invest in TP or broader WHP activities. This does not limit the external validity of the results, as this selection bias will also apply to future efforts to disseminate voluntary TP measures outside a study context. However, this is problematic from a public health equity perspective, because less privileged groups of employees remain disadvantaged and have less chance of benefiting from voluntary TP and WHP. Thus, the importance of consistent legal regulations regarding tobacco control reaching all companies independently of existing resources and motivation for TP becomes even clearer.

5. Conclusions

This is the first large-scale, longitudinal field study examining the active dissemination and adoption of TP across companies from all economic sectors in a region. Over a period

Table 5: Bivariate odds ratios for predictors of differences ($t2$ data minus $t1$ data) in health-related outcomes aggregated at the company level (ordinal regressions).

	% Smokers	Differences ($t2$ minus $t1$ data) in SHS[a]-related problems	Absenteeism
Change in policy restrictiveness			
Decrease	1.73*	4.95**	1.32
No change	0.91	2.91**	1.21
Increase	ref.	ref.	ref.
Change in individual support measures			
Decrease	0.83	1.64+	1.14
No change	0.96	1.53*	0.86
Increase	ref.	ref.	ref.
Change in SOC[1] policy			
Decrease	1.77*	2.47**	1.53*
No change	1.12	2.30**	1.12
Increase	ref.	ref.	ref.
Change in SOC[1] course			
Decrease	1.82*	1.01	1.00
No change	1.12	0.99	0.98
Increase	ref.	ref.	ref.

Note. [a]Environmental tobacco smoke.
+ $p < .10$; * $p < .05$; ** $p < .01$.
[1]SOC = Stage of change.

of two years (2007–2009), it showed slight improvements in worksite TP, specifically with respect to smoking policies; the companies' activities with respect to individual support measures for smokers remained low. Although dissemination was approached at a mass scale, it had modest reach and we observed only slight improvements in the advocated adoption of policies and programs. On the one hand, this might be explained by the societal context of this study: the dissemination activities probably could not generate any effect in *addition* to the historical trend. On the other hand, the fact that we did not observe clear benefits of the dissemination activities might also suggest the need for even more aggressive outreach and communication to target companies, or an acceptance that these channels do not seem to be sufficiently effective to merit investment. One might argue that legal regulations will be the adequate measure to ensure TP at worksites; however, as Radtke et al. [8] showed, a considerable proportion of the working population were still exposed to SHS at the worksite in 2010, even though the Swiss Federal law for protection against passive smoking was then in force. Therefore, further research is needed to explore more promising channels of dissemination in order to promote worksite TP. The finding of our study, indicating that improvements in TP measures at worksites are associated with improvements in aggregated health-related outcomes, shows that this will be worthwhile.

Appendix

A. Questionnaire Measures

The present research assessed the following variables at baseline and follow-up via a written questionnaire (Table 6).

(1) Attributes of the Organisation. The percentage of female employees was assessed by one item with given answer options in percent (see Table 4). To analyse the companies' previous practice with regard to workplace health promotion, we used four items to assess the extent to which health promotion measures (e.g., courses for general health behaviours) were already implemented at the workplace. Answers were given on a five-point scale (1 = not interested, 3 = intention to implement in the next months, and 5 = systematically implemented).

(2) Attributes of the Representative. Representatives indicated the extent to which they were personally concerned about tobacco prevention (e.g., "smoking is a private matter and none of the company's business"; 1 = disagree, 5 = agree) and how they evaluated a selection of possible pros and cons with regard to smoke-free workplaces and smoking cessation courses (health and economic benefits, rejection by employees and investments; 1 = disagree, 5 = agree).

(3) Worksite Tobacco Prevention. Participants used a single forced-choice item to indicate the extent to which tobacco prevention policies were already in place (0 = "there is no policy," 1 = "smoking is allowed anywhere except a few nonsmoking areas," 2 = "smoking is allowed outside and in certain indoor areas," 3 = "smoking is allowed outside, but not in buildings," and 4 = "smoking is not allowed anywhere"). To assess the prevalence and number (0–3) of individual support measures, we asked whether companies offered smoking cessation courses for their personnel, individual consultancy for smokers, or information material. To assess the stages of change with regard to worksite tobacco prevention, we used one item focussing on smoke-free policies and another

Table 6: Sample composition of the baseline and follow-up studies.

	Baseline (n = 1627)		Follow-up (n = 1452)	
	n	%	n	%
Characteristics of the representatives				
Smoking status				
Nonsmoker	1198	74.0	1046	72.0
Occasional smoker	209	12.9	165	11.4
Smoker	212	13.1	184	12.7
Function				
CEO	820	49.7	703	48.4
Human resource manager	747	45.3	649	44.7
Health and safety manager	209	12.7	181	12.5
Authority to decide upon WHP measures				
Not authorised	762	47.6	741	51
Authorised	839	52.4	711	49
Characteristics of the organisations				
Size				
20–49 employees	1035	63.0	927	63.8
50–250 employees	528	32.2	435	30.0
>250 employees	79	4.8	76	5.2
Branch				
Construction	140	8.5	117	8.1
Hospitality	131	8.0	127	8.7
Healthcare and welfare	247	15.0	201	13.8
Other	1127	68.5	996	69.2
% Women				
<20%	472	29.6	414	28.5
20–39%	292	18.3	266	18.3
40–59%	414	25.9	370	25.5
60–79%	232	14.5	232	16.0
>80%	187	11.7	153	10.5

focussing on cessation courses. Answers were given on a five-point scale with the following answer options: "We are not interested in adopting a smoke-free policy/cessation courses" (1, precontemplation); "We have not yet implemented a smoke-free policy/cessation courses but are interested in doing so" (2, contemplation); "We intend to implement a smoke-free policy/cessation courses in the next few months" (3, preparation); "Until now we have implemented a smoke-free policy/cessation courses only erratically" (4, action); "We have systematically implemented a smoke-free policy/cessation courses" (5, maintenance).

(4) Health Outcomes. The percentage of smokers was assessed by one item with given answer options in percent (1 =< 20%, 2 = 20–39%, 3 = 40–59%, 4 = 60–79%, and 5 => 80%). Problems related to environmental tobacco smoke (i.e., exposure to and complaints about environmental tobacco smoke) were assessed using a five-point rating scale (1 = disagree, 5 = agree). Absenteeism was assessed using a single item: "In our company we have a high level of absenteeism due to illness" (1 = disagree, 5 = agree).

Apart from the questions relating to previous practice and personal concern, all items included the response option "I don't know" in case an answer was not possible.

Conflict of Interests

The authors declare that there is no conflict of interests.

Acknowledgment

This work was funded by the Tobacco Prevention Fund (Swiss Federal Office of Public Health).

References

[1] World Health Organization, *WHO Framework Convention on Tobacco Control*, World Health Organization (WHO), Geneva, Switzerland, 2003.

[2] I. Kramer, I. Sockoll, and W. Bödeker, "Die evidenzbasis für betriebliche gesundheitsförderung und prävention—eine synopse des wissenschaftlichen kenntnisstandes," in *Fehlzeiten-Report 2008—Betriebliches Gesundheitsmanagement: Kosten und Nutzen*, B. Badura, H. Schröder, and C. Vetter, Eds., vol. 2008 of *Fehlzeiten-Report*, pp. 65–76, Springer, Berlin, Germany, 2009.

[3] D. T. Levy, F. Chaloupka, and J. Gitchell, "The effects of tobacco control policies on smoking rates: a tobacco control scorecard," *Journal of Public Health Management and Practice*, vol. 10, no. 4, pp. 338–353, 2004.

[4] J. L. Repace, J. N. Hyde, and D. Brugge, "Air pollution in Boston bars before and after a smoking ban," *BMC Public Health*, vol. 6, article 266, 2006.

[5] E. Kelloway, J. Barling, and C. Weber, "Smoking and absence at work: a quantitative review," in *Voluntary Employee Withdrawal and Inattendance: A Current Perspective*, M. Koslowsky and M. Krausz, Eds., pp. 167–178, Plenum Publishing, New York, NY, USA, 2002.

[6] R. E. Glasgow, T. M. Vogt, and S. M. Boles, "Evaluating the public health impact of health promotion interventions: the RE-AIM framework," *American Journal of Public Health*, vol. 89, no. 9, pp. 1322–1327, 1999.

[7] G. Parcel, C. Perry, and W. Taylor, "Beyond demonstration: diffusion of health promotion innovations," in *Health Promotion at the Community Level*, N. Bracht, Ed., pp. 229–251, Sage Publications, Newbury Park, Calif, USA, 1990.

[8] T. Radtke, R. Keller, H. Krebs, and R. Hornung, *Passivrauchen in der Schweizer Bevölkerung 2010. Zusammenfassung des Passivrauchberichts 2011*, Universität Zürich, Zürich, Switzerland, 2011.

[9] B. A. Rabin, R. E. Glasgow, J. F. Kerner, M. P. Klump, and R. C. Brownson, "Dissemination and implementation research on community-based cancer prevention. A systematic review," *American Journal of Preventive Medicine*, vol. 38, no. 4, pp. 443–456, 2010.

[10] V. Friedrich, A. Brügger, and G. Bauer, "Worksite tobacco prevention in the Canton of Zurich: stages of change, predictors, and outcomes," *International Journal of Public Health*, vol. 54, no. 6, pp. 427–438, 2009.

[11] J. M. Prochaska, "A transtheoretical model for assessing organizational change: a study of family service agencies' movement

to time-limited therapy," *Families in Society*, vol. 81, no. 1, pp. 76–84, 2000.

[12] D. Stokols, "Translating social ecological theory into guidelines for community health promotion," *American Journal of Health Promotion*, vol. 10, no. 4, pp. 282–298, 1996.

[13] B. J. Weiner, M. A. Lewis, and L. A. Linnan, "Using organization theory to understand the determinants of effective implementation of worksite health promotion programs," *Health Education Research*, vol. 24, no. 2, pp. 292–305, 2009.

[14] V. Friedrich, S. Hofmann, and G. Bauer, "Strategies of active dissemination of workplace health promotion," *International Journal of Workplace Health Management*, vol. 8, no. 1, pp. 3–14, 2015.

[15] J. Cohen, *Statistical Power Analysis for the Behavioural Sciences*, Lawrence Erlbaum Associates, Hillsdale, Mich, USA, 2nd edition, 1988.

[16] S. Rajkumar, S. Hoffmann, M. Röösli, and G. F. Bauer, "Evaluation of implementation, compliance and acceptance of partial smoking bans among hospitality workers before and after the Swiss Tobacco Control Act," *Journal of Public Health*, vol. 37, no. 1, pp. 89–96, 2015.

Hindawi Publishing Corporation
BioMed Research International
Volume 2015, Article ID 240240, 9 pages
http://dx.doi.org/10.1155/2015/240240

Research Article

Evaluation of the Effectiveness and Implementation of an Adapted Evidence-Based Mammography Intervention for African American Women

Linda Highfield,[1] Marieke A. Hartman,[2] L. Kay Bartholomew,[2] Philomene Balihe,[3] and Valerie A. Ausborn[3]

[1]*Department of Management, Policy and Community Health, School of Public Health, University of Texas, 1200 Pressler Street, Houston, TX 77030, USA*
[2]*Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas, 7000 Fannin Street, Houston, TX 77030, USA*
[3]*Episcopal Health Foundation, 500 Fannin Street, Suite 300, Houston, TX 77002, USA*

Correspondence should be addressed to Linda Highfield; linda.d.highfield@uth.tmc.edu

Received 23 June 2015; Revised 10 August 2015; Accepted 18 August 2015

Academic Editor: Giedrius Vanagas

Breast cancer mortality disparities continue, particularly for uninsured and minority women. A number of effective evidence-based interventions (EBIs) exist for addressing barriers to mammography screening; however, their uptake and use in community has been limited. Few cancer-specific studies have evaluated adapted EBIs in new contexts, and fewer still have considered implementation. This study sought to (1) evaluate the effectiveness of an adapted mammography EBI in improving appointment keeping in African American women and (2) describe processes of implementation in a new practice setting. We used the type 1 hybrid design to test effectiveness and implementation using a quasi-experimental design. Logistic regression and intent-to-treat analysis were used to evaluate mammography appointment attendance. The no-show rate was 44% (comparison) versus 19% (intervention). The adjusted odds of a woman in the intervention group attending her appointment were 3.88 ($p < 0.001$). The adjusted odds of a woman attending her appointment in the intent-to-treat analysis were 2.31 ($p < 0.05$). Adapted EBI effectiveness was 3.88 (adjusted OR) versus 2.10 (OR) for the original program, indicating enhanced program effect. A number of implementation barriers and facilitators were identified. Our findings support previous studies noting that sequentially measuring EBI efficacy and effectiveness, followed by implementation, may be missing important contextual information.

1. Background

Breast cancer is the most common cancer in the United States and is the second leading cause of cancer mortality in women [1, 2], with lower incidence in African American women but higher stage at diagnosis and greater mortality as compared to non-Hispanic white women [2, 3]. Enhancing guideline adherent mammography routines among these women may be important to address this disparity [3]. While a number of effective evidence-based interventions (EBIs) exist for addressing barriers to mammography screening, like other EBIs, their uptake and use in community settings have been limited [4–7]. Reasons for lack of uptake include cancer planners' anticipation of a misfit between interventions tested in controlled efficacy trials and the needs of their settings [8–10]. Both the perception of lack of fit and the possibility of real deficits in an EBI fit for a new community can be addressed by judicious and systematic adaptation of EBIs by research-practice partnerships and consultation with the community to improve fit [8–15].

Planners face the challenge of striking a balance between program fidelity, that is, implementation of an EBI as intended, and adaption to the needs of the adopting site [15]. Some efforts to promote use of evidence-based programs suggest that the primary concern should be fidelity rather than adaptation because of the lack of data to suggest that

adaptation improves program effectiveness [16]. However, in a review of over 500 studies that demonstrated that program implementation affected outcomes of prevention programs, Durlak and DuPre point out that while higher levels of fidelity were closely tied to improved program outcomes, levels of fidelity were well below 100% across interventions [17]. Therefore some adaptation occurred and might have been seen as necessary for program implementation. Elliott and Mihalic have outlined four ways that programs are typically adapted: adding or deleting program components; changing program components or content; changing the process or intensity of implementation; and making cultural modifications [16]. Barrera Jr. and colleagues found that behavioral interventions were more effective when adapted for a new cultural group than usual care and other control conditions and that most planners agreed that adaptation begins with data collection, to inform the need for adaptation, and ends with testing in the new setting [18].

Best practice is to always evaluate an EBI used in a new setting, however, particularly one that has been adapted. Evaluation of adapted EBIs is recommended, since adaptation may harm the effective elements of an EBI (i.e., core elements) [11]. Besides this need for impact evaluation, there is a need to evaluate the feasibility and fidelity of intervention implementation in the new population and setting [10]. However, few cancer-specific studies have evaluated effectiveness of adapted evidence-based interventions in new contexts, and fewer still have evaluated implementation in real-world contexts specifically [19, 20]. Of the few studies that have evaluated implementation of cancer-specific EBIs, facilitators for implementation and fidelity included the use and enthusiasm of program champions, academic detailing, and training (a higher degree of control) and team involvement/communication. Barriers included lack of attendance at training sessions, incomplete exposure to EBI tools/components, and competing demands at the practice level [20]. The authors could find no published studies that discussed real-world implementation of mammography EBIs in particular.

Therefore, the objectives of this study were to (1) evaluate the effectiveness of an adapted mammography EBI in improving appointment keeping for mammography in African American women and (2) describe processes of implementation of the EBI in a practice setting. Study results will test the hypothesis in which the effectiveness of the original EBI will be retained after adaptation and provide lessons learned for future intervention implementation in the real-world setting of mammography screening.

2. Methods

2.1. Evidence-Based Intervention. For this study, we adapted the intervention "Breast Cancer Screening among Nonadherent Women," originally developed by Duke University and Kaiser Foundation Health Plan [21]. The intervention is a tailored telephone counseling reminder based on the Transtheoretical Model of Change [22]. The program assessed a woman's stage of readiness to attend her appointment through a series of survey questions and counseled her through barriers to attendance. Following the Transtheoretical Model, the five stages were as follows: precontemplation, no intention to attend appointment; contemplation, intends to attend appointment; preparation, intends to attend appointment and is making preparations for taking action; action, has attended the appointment; maintenance, keeps attending appointments [22]. In the original trial, women who were off schedule with screening were more than twice as likely to get a mammogram if they received the telephone counseling (OR = 2.10).

We adapted the intervention using Int Map Adapt, a modified version of intervention mapping (for full details of the intervention adaptation, please see Highfield et al. also in this issue) in the following ways: (1) performing needs assessment among local African American women to identify salient barriers and include them in the barrier scripts; (2) developing a foundational communication process based on active listening to make it easier for the patient navigator to hold a real world rather than research conversation (when not dealing with a specific barrier) and to develop rapport with the patient; (3) changing assessment of stage of readiness to include only two categories precontemplation/contemplation or preparation/action and then matching the script to whether the women intended to keep her appointment or is unsure; (4) pretesting the changes with local women to assess acceptability and fine-tune scripts; and (5) developing an implementation protocol and training the navigator [14]. The adapted intervention aimed to increase scheduled mobile mammography screening appointment attendance rates among low-income African American women with care provided by a mobile mammography provider which was the largest nonprofit breast cancer screening organization in the greater Houston area. The systematic and collaborative adaptation process of the original EBI for use in local practice is reported elsewhere (see Highfield et al., this issue).

2.2. Study Design. We used the type 1 hybrid design to test the intervention's effectiveness and to gather information on the implementation [23, 24]. This type of design focuses on effectiveness evaluation and answers questions such as "what are possible facilitators and barriers to real-world implementation of an EBI?" and "what potential modifications could be made to maximize implementation?" in addition. We originally planned a randomized controlled trial but found that the navigator could not alternate between usual care and the adapted intervention. Therefore, we changed to a quasi-experimental, sequential recruitment design in which we assigned contacted women to usual care or adapted intervention in sequential groups of 50 patients. See enrollment and study limitations for further detail. The time period for enrollment and collection of patient data was predetermined based on funding and availability of the clinical partner and took place from February to December 2012. We sought to contact as many patients as possible within this time window. This study operated under Institutional Review Board approval from St. Luke's Episcopal Hospital Institutional Review Board.

2.3. Study Setting. A local mobile mammography partner served as the site for implementation of the intervention (including recruitment and data collection). In 2011, the organization provided 33,784 screening and diagnostic procedures for those able to pay; 19,369 screening and diagnostic procedures at no charge to low-income, uninsured women; and 8,857 free patient navigation services to patients without insurance. Mobile screening mammography services are provided to over 7,000 women a year, covering a 15-county region centered on Harris County, TX. Services are provided in a variety of settings, including schools, worksites, federally qualified health centers, churches, and other community settings. The mobile mammography provider in this study serves a diverse population including Caucasians, Hispanics, Asians, African Americans, and immigrant populations. Approximately 20% of the low-income, uninsured patient population at the time of study was African American (2,200 women). The baseline expected no-show rate for uninsured, low-income African American women was 38% (unpublished data).

2.4. Patient Enrollment. Inclusion criteria were as follows: African American, female, age 35–64, uninsured, income of ≤200% of the federal poverty level (FPL), and an upcoming appointment for a mobile screening mammogram at a program partner site. We identified eligible patients from the electronic patient scheduling records. The patient navigator made three calls to reach all eligible patients including calls at different times of the day and weekends for those who were not reached in the initial attempt. Reached individual patients received one phone call from the patient navigator in order to deliver the intervention. We expected intervention calls would take on average 6–10 minutes. Reached individual patients were initially enrolled into each group by randomization (using a randomized controlled trial (RCT) design) from February to April 2012; however, we ran into implementation issues with the patient navigator (see Section 3), so we adjusted to a sequential enrollment procedure from May to December 2012. The navigator called patients in the comparison group and provided them with a standard appointment reminder which included the date, time, and location of their upcoming appointment. If a patient did not answer the phone, the navigator left a voicemail message containing the reminder. The navigator read to patients in the intervention group an oral consent over the phone and after consent asked the following staging question: "How confident are you that you will keep your upcoming appointment?" The navigator then counseled as needed for any barriers uncovered in the phone call per the intervention protocol. No blinding was used in this study.

2.5. Measures and Data Tracking. The primary outcome of appointment keeping was ascertained from mobile mammography clinical records (nonattendance = 0, attendance = 1). In addition, we collected information about sites, age of patient, number of days between reminder call and appointment, and study stage (i.e., design coded as 0/1 for randomized controlled trial versus quasi-experimental one). Appointments were scheduled to 41 different sites across 8 counties. The sites were divided into 2 categories, community sites (local nonprofit organizations or government agencies, community initiatives, schools, health fairs, or other community organizations) and hospital/clinic sites (local hospital, federally qualified health center, or charity clinic). Age and days to appointment were categorized in the following categories: 35–39, 40–49, and 50–64 years old and 0 days, 1 day, 2 days, 3-4 days, and 5 days or more between phone call and appointment.

We evaluated the secondary outcome of implementation fidelity by monitoring of intervention phone calls and comparing them to the protocol, making site visits to the mammography site, and meeting with implementation staff (researchers and practitioners). A series of three phone calls made by the patient navigator were recorded at the beginning of implementation. These recordings were evaluated by the research team for compliance in asking the staging question and using active listening and scripted responses to patient identified barriers during the phone call. Following review, the navigator received feedback on the staging question and active listening. During implementation, phone calls were periodically monitored on-site by a member of the research team for the same compliance issues. In addition, we made postintervention follow-up phone calls to a randomly selected subset of intervention patients ($n = 50$) to assess their perception of the EBI calls and systems barriers encountered (see Topic-List).

Topic-List for Follow-Up Calls and Implementation Evaluation

(1) Is there anything you want to tell me about your mammogram appointment so we can make the experience better?

(2) Do you remember talking to anyone from [mobile mammography program name] before your mammogram appointment? Do you remember who?

(3) What was the conversation about?

(4) What stands out about your experience talking to (navigator name)?

(5) Please tell me what caused you to keep this mammogram appointment? (Probe: was there anything else?)

(6) Do you remember anything specific about the conversation with (navigator name) that helped you keep your appointment? If yes, what about the conversation helped you keep your appointment?

(7) Are there any other reasons you kept your mammogram appointment?

(8) On a scale from 1 to 5, with 5 being most helpful, how helpful did you find your phone call with (insert navigator name)?

(9) Can you think of things that would have been helpful to hear from (navigator name) that would have made the phone call better?

(10) Was there anything more we could have done to help you keep your appointment?

(11) Do you have any other comments regarding our study of how to get women to keep their mammogram appointments?

We tracked all data for the pilot either in an Access database or in paper data collection forms. The database included fields for a unique identifier for each patient, date and time of attempted call(s) with outcome of each (reached, not reached, left message, and bad number), barriers, and systems barriers encountered during the session, such as the patient was not aware they needed a doctor's order to receive a mammogram. We also included an open text field for the patient navigator to record notes during the call. Data available from the mammography partner's data system included age, sponsored status (lack of insurance and ≤200% FPL), site of screening, date and time of appointment, and contact information including phone number. The research assistant cleaned the data by comparing the Access database with the paper forms and existing records from the mammography partner. Any inconsistencies between the database and paper forms were investigated with the site and patient navigator for clarification. Data from both databases were combined into one Access database and exported to Stata for analysis.

2.6. Data Analysis. We used Stata (Stata Corp., College Station, TX, USA) for statistical analysis. We calculated descriptive statistics and then conducted logistic regression analysis to report attendance in the intervention group as compared to the comparison group. Chi-square tests (and Fisher's exact tests when cell sizes were less than five) were used to evaluate group differences between potential confounding variables, including age, days between reminder calls, mammography site (community versus clinical setting) and appointment time, and the study stage (i.e., design change). Both unadjusted and adjusted logistic regression models were fitted to determine intervention's effectiveness in improving mammography attendance. Factors in the adjusted model included, besides group (intervention; control), mammography site, age of patient, number of days between reminder call and appointment, and navigator making the reminder calls. Study stage (design) was not included in the model as it was highly collinear with navigator as only one navigator made calls during each phase. Following the basic analysis, we further evaluated the effectiveness of the EBI using intent-to-treat analysis [25–28]. In this study, we used intent-to-treat analysis which considered the outcomes (appointment attendance) for all women based on their group designation at the time of phone call attempt (intervention or control) and not just those who were reached and treated following protocol by the patient navigator. Intent-to-treat analysis ignores deviations in protocol, noncompliance, and anything that may happen after group assignment [25–28]. We conducted power analysis using a two-tailed two-sample frequencies Fisher's exact test with $\alpha = 0.05$ and adjusted for unequal sample sizes to evaluate ability to detect a difference between the groups.

3. Results

Figure 1 shows the CONSORT/TREND diagram with the total number of enrolled patients per study stage (randomization and sequential enrolment stage), those assigned, allocated, exposed to the intervention, followed-up, and analyzed, both in the basic effectiveness ($n = 151$) and intent-to-treat analysis ($n = 198$). The intervention and comparison groups were similar with regard to age and number of days between reminder call and appointment as shown in Table 1. The average and median age for patients in both groups was 51 years (range: 36–64). The average and median number of days between reminder call and patient appointment was 3 days for both groups (range: 0–13 days). No effect was observed for the study stage (design change) ($\chi^2 = 0.292$). The groups were different in regard to the type of mammography site, with women in the intervention group being screened in community settings more frequently than the control group in both the basic analysis and intent-to-treat analysis (see Table 1). The no-show rate for patients in the comparison group was 44%. The no-show rate for patients in the intervention group was 19% meaning that the EBI in this study led to a 57% reduction in the no-show rate in the basic analysis (calculated as percent change).

3.1. Effectiveness Results. The unadjusted and adjusted results are presented in Table 2. The unadjusted odds of a woman in the intervention group attending her appointment was 3.38 times higher than for a woman in the control group ($p < 0.001$) in the basic analysis. The adjusted odds of a woman in the intervention group attending her appointment were 3.88 as compared to the control group ($p < 0.001$). No effect was found for the change in study design. In the intent-to-treat analysis, the unadjusted odds of a woman attending her appointment if she was in the intervention group were 1.84 ($p < 0.05$). The adjusted odds of a woman attending her appointment in the intent-to-treat analysis were 2.31 as compared with the control group ($p < 0.05$). With the no-show rate of 44% observed in the comparison group, using a two-tailed test and $\alpha = 0.05$, there was 87% power to detect a change in the no-show rate to 19% in the intervention group for this study in the basic analysis.

3.2. Implementation Results. We encountered a number of systems barriers to implementation. These included the following: confusion about responsibility for implementation of usual care reminder calls; lack of clear communication about the prerequisites of a doctor's order and clinical exam prior to screening; and inconsistent notification about costs associated with screening.

Fourteen out of 96 (15%) patients in the intervention group reported encountering systems barriers, including the fact that they were unaware of their upcoming appointments, unaware of the need for a doctor's order to obtain a mammogram, and unaware of the out-of-pocket cost of the mammogram. Additionally, some sites reported issues with the mobile units going to the wrong location, sites being cancelled with short notice due to mechanical issues (mobile

Table 1: Descriptive statistics for patients in the control versus intervention groups.

Patient characteristics	Basic analysis		Intent-to-treat analysis	
	Intervention group n (%)	Usual care group n (%)	Intervention group n (%)	Usual care group n (%)
Age group				
35–39	5 (8%)	4 (5%)	8 (8%)	5 (5%)
40–49	25 (40%)	38 (43%)	38 (40%)	41 (41%)
50–64	33 (52%)	46 (52%)	49 (52%)	53 (54%)
	$\chi^2 = 0.8162$, $p = 0.691^{\wedge}$		$\chi^2 = 0.8810$, $p = 0.670^{\wedge}$	
Mobile site				
Clinic	36 (60%)	71 (81%)	49 (57%)	73 (74%)
Community	24 (40%)	17 (19%)	37 (43%)	26 (26%)
	$\chi^2 = 7.6191$, $p = 0.006^{*}$		$\chi^2 = 5.769$, $p = 0.016^{*}$	
Days from call to appointment				
0	4 (5%)	4 (4%)	8 (8%)	4 (4%)
1	25 (29%)	28 (26%)	36 (38%)	29 (28%)
2	12 (14%)	20 (19%)	18 (19%)	21 (21%)
3-4	33 (39%)	44 (41%)	15 (16%)	25 (25%)
≥5	11 (19%)	12 (11%)	19 (20%)	23 (23%)
	$\chi^2 = 2.9846$, $p = 0.558^{\wedge}$		$\chi^2 = 4.948$, $p = 0.298^{\wedge}$	
Screening outcome				
Attendance	51 (81%)	49 (56%)	70 (73%)	60 (59%)
Nonattendance	12 (19%)	39 (44%)	26 (27%)	41 (41%)
	$\chi^2 = 10.2484$, $p = 0.001^{**}$		$\chi^2 = 4.003$, $p = 0.045^{*}$	

*Statistically significant at $p < 0.05$.
**Statistically significant at $p = 0.001$.
$^{\wedge}$Fisher's exact test used for p value.

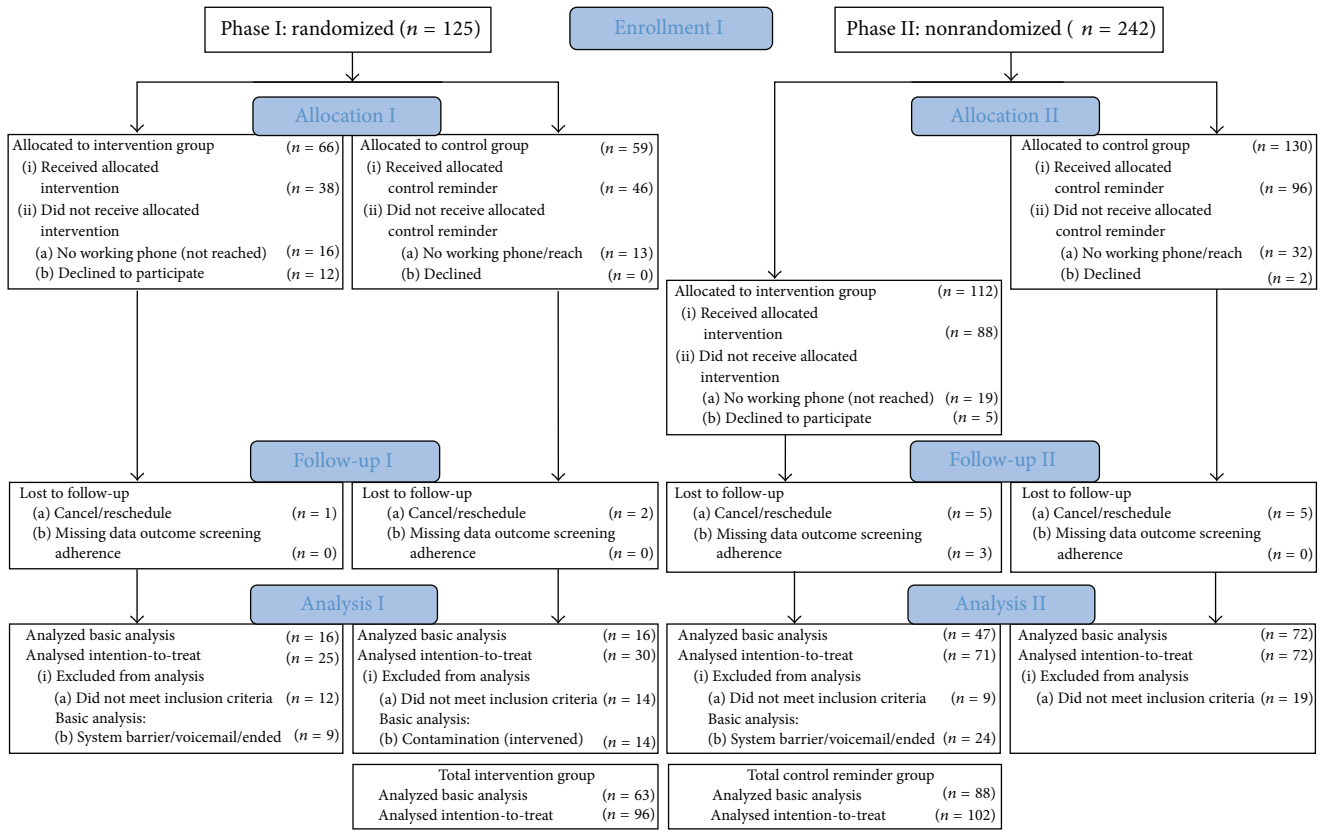


Figure 1: Flow chart of patient enrolment, follow-up, and basic and intention-to-treat analyses.

Table 2: Unadjusted and adjusted logistic regression results for mammography appointment attendance.

	Crude OR (95% CI)	Adjusted OR (95% CI)
Basic analysis		
Group (control versus intervention)	3.38** (1.59–7.21)	3.88** (1.70–8.86)
Age	0.997 (0.95–1.04)	0.903 (0.492–1.66)
Days between call and appointment	1.001 (0.753–1.33)	1.11 (0.799–1.54)
Navigator	1.33 (0.875–2.01)	1.25 (0.773–2.04)
Mobile site	0.843 (0.397–1.79)	0.562 (0.241–1.31)
Intent-to-treat analysis		
Group (control versus intervention)	1.84* (1.01–3.35)	2.31* (1.09–4.93)
Age	1.04 (0.642–1.68)	1.07 (0.637–1.78)
Days between call and appointment	1.03 (0.805–1.31)	1.11 (0.856–1.43)
Navigator	1.27 (0.872–1.84)	1.07 (0.684–1.67)
Mobile site	1.24 (0.665–2.33)	1.64 (0.831–3.25)

* $p < 0.05$.
** $p < 0.001$.

unit broke down or mammography machine needed service), and unclear communication about scheduling procedures, such as how many patients could be seen and at what time for scheduled mobile screening dates.

Of the 50 randomly selected intervention patients for follow-up phone calls to assess patient perception of the EBI 42 completed the interview (84%). In these calls, we found that 34 patients remembered their reminder phone call from the navigator (81%). The patients who remembered their call and conversation reported positive interactions with her. They reported, for example, "*She was warm, friendly, helpful, sweet, supportive and sincere*." When asked if there was something about the phone call from the navigator that helped them to keep their appointment, 18 patients reported positive impact, such as "*The encouragement from her [the navigator] went beyond a reminder call*," "*she cared*," "*put me first*," "*helped me overcome my misconceptions*," and "*was nice*." When asked how helpful they found the phone call, all patients that remembered the conversation ($n = 34$) rated it as 5 out of 5, except one patient who rated it 4 out of 5. Finally, the patients who attended their appointment were asked to share their thoughts of the reminder phone call program. Patients reported that "*It's important. Catch it (breast cancer) early to have a chance*," "*They (women) need to go and have it done!*," "*Taking care of yourself is a major point to bring up*," "*It's a wonderful program*," "*It's a must*," and "*I think it's great that we are talking to women to let them know that mammograms are important*," and recommended "*Use media and marketing to reach women without insurance*," "*Transportation is a very big deal and would be a help. Maybe find a church with a van that could help out*," "*Spread the news about breast health–put it in churches and schools. I was telling people at my church about the mammography program and they had never heard about it*."

Seven of the 42 patients we conducted follow-up calls with did not attend their appointment. When asked if there was anything we could have done to have helped them keep their appointment, three reported they were sick on their appointment day, three had last minute transportation issues, and one reported that she did not have the money for the copay. Due to the nature of the mobile program, in many cases our patient navigator was not able to reschedule patients directly if during the intervention call they indicated a desire to change their appointment. Patients had to be routed through the mobile program coordinators or the site coordinators in the community in order to reschedule. This meant a loss of continuity with the patient and in some cases patients reported having difficulty reaching the coordinators to reschedule.

4. Discussion

This study used a hybrid type 1 design to evaluate both effectiveness of an adapted EBI in a practice setting and the implementation process. The effectiveness of the adapted EBI was 3.88 (adjusted OR) versus 2.10 (OR) for the original program [21]. This is consistent with the findings of Barrera Jr. et al., which systematically adapted EBIs improvement program effectiveness when compared to a control reminder [18]. The adapted EBI in this study reduced appointment no-shows by 57 percent from baseline in the clinical practice and would be suitable for scale-up.

Few published studies provide a detailed description of EBI effectiveness with implementation outcomes in a single study, particularly for cancer-specific EBIs [19, 20, 23]. The "Communicating Health Options Through Information and Cancer Education" (CHOICE), "Improving Systems for CRC Screening at Harvard Vanguard Medical Associates" (HVMA), and "Improving CRC Screening and Follow-up in the Veterans Health Administration" (VHA) programs all considered implementation context during their evaluations [20]. These studies encountered some of the same implementation barriers and facilitators we found in this study [20]. We found that monitoring of implementation was valuable and that the study approach needed flexibility to deal with evolving implementation issues, such as the lack of consistent standard care reminder calls and the navigator struggling with the simultaneous implementation of the control and

intervention process. This was consistent with the recommendations from the Cool Pool trial where they noted that continuous monitoring of implementation was critical. Other authors have also noted the importance of continuous monitoring of implementation [20]. Additionally, the ability to identify and measure all implementation issues at the beginning of a study is limited and has been noted as a barrier in previous studies [20]. For example, in this study, we did not know prior to implementation monitoring whether reminder calls were implemented consistently. Implementation issues like these are likely to arise only once monitoring begins and may also appear over time, requiring subsequent intervention or changes in protocol to address them.

Our findings further highlight themes from previous studies which have noted that the predominant research paradigm of sequentially measuring EBI efficacy and effectiveness, followed by implementation studies, may be missing important contextual information [7, 23, 29]. We were able to find and correct problems with implementation based on the results of our process evaluation which we monitored continuously throughout the study. The process evaluation also allowed us to find problems with fidelity early in the study and correct them. The major correction was to change the design so that the navigator did not have to conduct two different interventions in the same period of time. We also added plans for an intent-to-treat analysis to increase our confidence in the validity of our results. After an unsuccessful attempt to train the original navigator to adhere to protocol, we replaced the navigator and conducted repeated trainings and monitoring of calls more frequently with the new navigator. This monitoring process may have contributed to the increased effectiveness of the EBI that we observed in this study in addition to the systematic adaptations made to the EBI.

4.1. Strengths and Weaknesses of the Study. This study has a number of strengths and limitations which should be considered. This study is one of only a very few studies to evaluate both EBI effectiveness and implementation in a community context and provides critical insights for the future translation of EBIs, particularly for mammography interventions. A major strength of this study was using a systematic process for adapting an EBI [11]. The process included working in a research-community partnership with an advisory board comprised of researchers and practitioners who worked together to perform a community needs assessment, select an EBI, adapt the EBI based on the needs assessment, pretest it, implement it, and evaluate it [13].

A number of limitations must also be considered for this practice-based evaluation study. First, this study was conducted in a mobile mammography practice and the findings may not be generalizable to other implementation contexts. Best practice indicates that anytime an intervention is being considered for a new population or context, that needs assessment and evaluation is needed (see Introduction and Highfield et al., 2014 [13]). Our findings show that by retaining core elements of an intervention, such as stage-based telephone counseling on barriers for mammography appointments that effectiveness can be maintained or even improved in a new population. We have no reason to believe this would not be the case when extending our intervention into a broader population context. For instance, many of the barriers to mammography screening we found among African American women are complementary to barriers to screening faced by all underserved women [30]. The most significant weakness of this study was our initial inability to train the navigator to keep the control and adapted intervention groups separate in the first study design. However, the process evaluation in this study enabled us to correct the behavior of the navigator and redesign the study from a RCT to a quasi-experimental design that we expected would be more feasible in practice. In the quasi-experimental design, we enrolled 47 patients in the intervention. The main threats to validity from this type of enrollment were selection bias, where patients reached during the sequential enrollment time period may not have been representative of the larger patient population in the clinic. We dealt with this validity threat by comparing patient demographics between the randomized and sequential design process, by including a design change variable in the regression analysis and by conducting an intent-to-treat analysis, which is useful for dealing with deviations in protocol. Also, even though this study lacked a true nonintervened group, the baseline no-show rate of 38% serves as a proxy control group since reminder calls were made only rarely. Additionally, the effect of using a control group receiving standard reminder calls as opposed to a true no-contact control group could have downward biased the results of the study. In other words, this may have made it harder to find an intervention effect in our study. Our control and intervention groups in both the basic analysis and intent-to-treat analysis differed in where they received their screening, with women in the intervention group being more likely to be screened in a community as opposed to a clinical setting. However, all women in the study were required to have a doctor's order to obtain screening and the differences were consistent in both the basic analysis and intent-to-treat analysis, so we believe the effect would be minimal on the results. Further, no significant difference was observed in the regression model between screening sites when controlling for other factors. Finally, we evaluated the EBI using patient data available from the clinical provider in this study. There may have been important factors such as educational level and occupation which we were not able to evaluate due to a lack of data availability from the provider. While these factors may be important, it is important to note that these are nonmodifiable factors and have been shown to have limited value when designing and evaluating EBI programs [30]. Lastly, women enrolled in our study were low-income and uninsured, two factors that generally correlate with education and occupation, so while we did not measure those directly, we believe their effect would have been minimal on the outcomes.

4.2. Lessons Learned. Interventions are rarely implemented with complete fidelity and in this study the navigator struggled to implement either intervention protocol, but especially the usual care group with fidelity. The navigator stated that she wanted to help all women attend their appointment and seemed not to be able to adhere to protocol. Even when we hired a second navigator, protocol adherence continued

to be somewhat difficult. We believe the important learning from this is about staffing in a research study versus staffing in a clinical or other professional setting. In the original evaluation of an EBI in a research setting, research assistants (usually students) do not have a particularly strong professional identity or habitual way of doing tasks closely approximating the research protocol. In contrast, in a practice implementation, new protocol driven tasks are given to professional care providers who may be unable to divert from their normal practice. It is important for future researchers to consider issues of fidelity when adapting, training, monitoring, and measuring EBIs in community contexts.

Additionally, best practice indicates the need for EBI testing in new contexts (e.g., effectiveness testing in the new setting) [11, 31]; however measuring and addressing context specific implementation issues remain challenging [32–34]. Currently, there is a lack of standardized and validated measures that can be used to assess implementation [35]. Additionally, studies have noted the need for multilevel interventions that consider implementation context; however there currently is no packaged approach to implementation available in the published literature [36]. Future studies should consider creating a packaged implementation intervention which could be tested and evaluated in the context of EBI implementation in the community.

5. Conclusion

This study provides an example of the real-world implementation of an adapted EBI. It demonstrates best practice for adaptation and evaluation of an EBI using a hybrid type 1 design and can be used for a model of blending research and practice to increase the uptake of EBIs and to make sure that they show effectiveness in new settings.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to acknowledge the contributions of our community partners. Collaboration took place with The Rose, The Breast Health Collaborative of Texas and Episcopal Health Foundation. Without their efforts, this project would not have been possible. The authors would also like to acknowledge the Avon Foundation for funding this study (Project nos. 05-2010-004 and 05-2011-008). Marieke A. Hartman was supported by the postdoctoral fellowship, University of Texas School of Public Health Cancer Education and Career Development Program, National Cancer Institute (NIH Grant R25 CA57712).

References

[1] National Cancer Institute, *SEER Cancer Statistics Review, 1975–2003*, National Cancer Institute, Bethesda, Md, USA, 2006.

[2] C. Desantis, R. Siegel, P. Bandi, and A. Jemal, "Breast cancer statistics, 2011," *CA—Cancer Journal for Clinicians*, vol. 61, no. 6, pp. 409–418, 2011.

[3] R. F. Young, K. Schwartz, and J. Booza, "Medical barriers to mammography screening of african american women in a high cancer mortality area: implications for cancer educators and health providers," *Journal of Cancer Education*, vol. 26, no. 2, pp. 262–269, 2011.

[4] J. R. Harris, A. Cheadle, P. A. Hannon et al., "A framework for disseminating evidence-based health promotion practices," *Preventing Chronic Disease*, vol. 9, no. E22, Article ID 110081, 2012.

[5] J. Kerner, B. Rimer, and K. Emmons, "Introduction to the special section on dissemination—dissemination research and research dissemination: how can we close the gap?" *Health Psychology*, vol. 24, no. 5, pp. 443–446, 2005.

[6] E. M. Rogers, *Diffusion of Innovations*, Free Press, New York, NY, USA, 2003.

[7] R. E. Glasgow, E. Lichtenstein, and A. C. Marcus, "Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition," *American Journal of Public Health*, vol. 93, no. 8, pp. 1261–1267, 2003.

[8] L. N. Krivitsky, S. J. Parker, A. Pal, L. Meckler, R. Shengelia, and M. C. Reid, "A systematic review of health promotion and disease prevention program adaptations: how are programs adapted?" in *Research for the Public Good: Applying the Methods of Translational Research to Improve Human Health and Well-Being*, E. Wethington and R. E. Dunifon, Eds., pp. 73–99, American Psychological Association, Washington, DC, USA, 2012.

[9] F. G. Castro, M. Barrera Jr., and C. R. Martinez Jr., "The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit," *Prevention Science*, vol. 5, no. 1, pp. 41–45, 2004.

[10] V. S. McKleroy, J. S. Galbraith, B. Cummings et al., "Adapting evidence-based behavioral interventions for new settings and target populations," *AIDS Education & Prevention*, vol. 18, pp. 59–73, 2006.

[11] L. K. Bartholomew, G. S. Parcel, G. Kok, and N. H. Gottlieb, *Planning Health Promotion Programs: An Intervention Mapping Approach*, Jossey-Bass, San Francisco, Calif, USA, 2011.

[12] S. J. Lee, I. Altschul, and C. T. Mowbray, "Using planned adaptation to implement evidence-based programs with new populations," *American Journal of Community Psychology*, vol. 41, no. 3-4, pp. 290–303, 2008.

[13] L. Highfield, L. K. Bartholomew, M. A. Hartman, M. M. Ford, and P. Balihe, "Grounding evidence-based approaches to cancer prevention in the community: a case study of mammography barriers in underserved African American women," *Health Promotion Practice*, vol. 15, no. 6, pp. 904–914, 2014.

[14] S. M. Ahmed and A.-G. S. Palermo, "Community engagement in research: frameworks for education and peer review," *American Journal of Public Health*, vol. 100, no. 8, pp. 1380–1387, 2010.

[15] T. E. Backer, *Finding the Balance: Program Fidelity and Adaptation in Substance Abuse Prevention: A State of the Art Review*, Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Prevention, Rockville, Md, USA, 2002.

[16] D. S. Elliott and S. Mihalic, "Issues in disseminating and replicating effective prevention programs," *Prevention Science*, vol. 5, no. 1, pp. 47–53, 2004.

[17] J. A. Durlak and E. P. DuPre, "Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation," *American Journal of Community Psychology*, vol. 41, no. 3-4, pp. 327–350, 2008.

[18] M. Barrera Jr., F. G. Castro, L. A. Strycker, and D. J. Toobert, "Cultural adaptations of behavioral health interventions: a progress report," *Journal of Consulting and Clinical Psychology*, vol. 81, no. 2, pp. 196–205, 2013.

[19] S. H. Taplin, R. A. Price, H. M. Edwards et al., "Introduction: understanding and influencing multilevel factors across the cancer care continuum," *Journal of the National Cancer Institute Monographs*, vol. 2012, no. 44, pp. 2–10, 2012.

[20] E. M. Yano, L. W. Green, K. Glanz et al., "Implementation and spread of interventions into the multilevel context of routine practice and policy: implications for the cancer care continuum," *Journal of the National Cancer Institute—Monographs*, no. 44, pp. 86–99, 2012.

[21] I. M. Lipkus, B. K. Rimer, S. Halabi, and T. S. Strigo, "Can tailored interventions increase mammography use among HMO women?" *American Journal of Preventive Medicine*, vol. 18, no. 1, pp. 1–10, 2000.

[22] C. C. DiClemente and J. O. Prochaska, "Toward a comprehensive, transtheoretical model of change: stages of change and addictive behaviorss," in *Treating Addictive Behaviors*, R. William and N. Heather, Eds., pp. 3–24, Plenum Press, New York, NY, USA, 1998.

[23] G. M. Curran, M. Bauer, B. Mittman, J. M. Pyne, and C. Stetler, "Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact," *Medical Care*, vol. 50, no. 3, pp. 217–226, 2012.

[24] A. C. Bernet, D. E. Willens, and M. S. Bauer, "Effectiveness-implementation hybrid designs: implications for quality improvement science," *Implementation Science*, vol. 8, supplement 1, article S2, 2013.

[25] S. K. Gupta, "Intention-to-treat concept: a review," *Perspectives in Clinical Research*, vol. 2, no. 3, pp. 109–112, 2011.

[26] D. J. Newell, "Intention-to-treat analysis: implications for quantitative and qualitative research," *International Journal of Epidemiology*, vol. 21, no. 5, pp. 837–841, 1992.

[27] R. Wertz, "Intention to treat: once randomized, always analyzed," in *Proceedings of the Clinical Aphasiology Conference*, vol. 23, pp. 57–64, 1995.

[28] S. R. Heritier, V. J. Gebski, and A. C. Keech, "Inclusion of patients in clinical trial analysis: the intention-to-treat principle," *Medical Journal of Australia*, vol. 179, no. 8, pp. 438–440, 2003.

[29] S. R. Tunis, D. B. Stryer, and C. M. Clancy, "Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy," *The Journal of the American Medical Association*, vol. 290, no. 12, pp. 1624–1632, 2003.

[30] O. M. Fayanju, S. Kraenzle, B. F. Drake, M. Oka, and M. S. Goodman, "Perceived barriers to mammography among underserved women in a Breast Health Center Outreach Program," *The American Journal of Surgery*, vol. 208, no. 3, pp. 425–434, 2014.

[31] A. S. Lau, "Making the case for selective and directed cultural adaptations of evidence-based treatments: examples from parent training," *Clinical Psychology: Science and Practice*, vol. 13, no. 4, pp. 295–310, 2006.

[32] R. P. T. M. Grol, M. C. Bosch, M. E. J. L. Hulscher, M. P. Eccles, and M. Wensing, "Planning and studying improvement in patient care: the use of theoretical perspectives," *Milbank Quarterly*, vol. 85, no. 1, pp. 93–138, 2007.

[33] E. K. Proctor and R. C. Brownson, "Measurement issues in dissemination and implementation research," in *Dissemination and Implementation Research in Health: Translating Research to Practice*, R. C. Brownson, G. A. Colditz, and E. K. Proctor, Eds., pp. 261–280, Oxford University Press, New York, NY, USA, 2012.

[34] E. Proctor, H. Silmere, R. Raghavan et al., "Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda," *Administration and Policy in Mental Health and Mental Health Services Research*, vol. 38, no. 2, pp. 65–76, 2011.

[35] S. R. Chaudoir, A. G. Dugan, and C. H. Barr, "Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures," *Implementation Science*, vol. 8, no. 1, article 22, 2013.

[36] S. R. Kirsh, R. H. Lawrence, and D. C. Aron, "Tailoring an intervention to the context and system redesign related to the intervention: a case study of implementing shared medical appointments for diabetes," *Implementation Science*, vol. 3, no. 1, article 34, 2008.