

Distribution Agreement

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

Signature:

Akshay Sharma

Date

New approaches for increasing HIV testing among men who have sex with men in the
United States

By

Akshay Sharma
Doctor of Philosophy

Epidemiology

Patrick S. Sullivan, D.V.M., Ph.D.
Advisor

Brent A. Johnson, Ph.D.
Committee Member

A.D. McNaghten, Ph.D.
Committee Member

Rob B. Stephenson, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

New approaches for increasing HIV testing among men who have sex with men in the
United States

By

Akshay Sharma

M.P.H., Emory University, 2010
M.B.B.S., Bangalore Medical College and Research Institute, 2006

Advisor: Patrick S. Sullivan, D.V.M., Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Epidemiology
2015

Abstract

New approaches for increasing HIV testing among men who have sex with men in the United States

By Akshay Sharma

Men who have sex with men (MSM) remain disproportionately affected by HIV in the United States. Testing is the cornerstone of comprehensive prevention efforts and the gateway to treatment. However, many MSM do not test annually, and a high proportion are unaware they are infected. This dissertation comprises three studies focusing on the acceptability and evaluation of new strategies to promote frequent HIV testing.

In the first study, we determined collective attitudes towards using six testing options among 973 MSM recruited online in 2012: Testing at a physician's office; Individual voluntary counseling and testing; Couples' HIV counseling and testing; Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) self-collection for laboratory testing. Most participants indicated being likely to use all modalities except DBS self-collection. Rapid home self-testing and testing at a physician's office were consistently top ranked.

In the second study, we surveyed 840 internet-using MSM in 2014 regarding past experiences with HIV self-testing, and future intentions of distributing free rapid home test kits and testing with their social or sexual network associates. Although 80% were aware, only 9% reported using home tests before. MSM self-identifying as bisexual had reduced odds of potentially distributing kits. Unprotected anal intercourse with one man in the past year, previous HIV testing, and willingness to distribute kits were associated with higher odds of potentially testing with others.

In the third study, we simulated several randomized trials evaluating a new HIV prevention strategy, formulated three analytical outcomes (screening a certain number of times during the entire trial; screening at least once within a clinically meaningful interval; instantaneous rate of screening), and compared them with regard to interpreting results and estimating power. Methods utilizing detailed prospective data, while also accounting for within-participant correlations, were less likely to miss the actual underlying intervention benefit compared to those relying on dichotomous measures derived from aggregating events.

Our findings demonstrate the potential for combining multiple testing options as part of packages to increase HIV testing among MSM. Rapid home tests hold immense promise and research assessing their public health impact should be prioritized.

New approaches for increasing HIV testing among men who have sex with men in the
United States

By

Akshay Sharma

M.P.H., Emory University, 2010
M.B.B.S., Bangalore Medical College and Research Institute, 2006

Advisor: Patrick S. Sullivan, D.V.M., Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Epidemiology
2015

TABLE OF CONTENTS

CHAPTER 1: Background and significance	1
SECTION 1.1: MSM and HIV in the United States	1
SECTION 1.2: Role of HIV testing in prevention efforts.....	2
SECTION 1.3: Inadequate HIV testing among MSM	3
SECTION 1.4: Menu of available HIV testing options	4
SECTION 1.5: Internet-based HIV prevention strategies.....	14
SECTION 1.6: Motivation for dissertation studies.....	15
SECTION 1.7: Specific objectives for dissertation studies	17
SECTION 1.8: Structure of this dissertation.....	19
CHAPTER 2: Acceptability and intended usage preferences for six HIV testing options among internet-using men who have sex with men	20
SECTION 2.1: Abstract	22
SECTION 2.2: Introduction	24
SECTION 2.3: Methods	28
SECTION 2.4: Results	31
SECTION 2.5: Discussion	34
CHAPTER 3: Willingness to potentially self-test, distribute and test with social or sexual network associates using rapid home HIV test kits and associated factors among men who have sex with men	46
SECTION 3.1: Abstract	48
SECTION 3.2: Introduction	50
SECTION 3.3: Methods	55
SECTION 3.4: Results	60

SECTION 3.5: Discussion	65
CHAPTER 4: Evaluating interventions to promote routine preventive screenings: A comparison of analytical outcomes.....	92
SECTION 4.1: Abstract	94
SECTION 4.2: Introduction	96
SECTION 4.3: Methods	99
SECTION 4.4: Results	103
SECTION 4.5: Discussion	105
SECTION 4.6: Supplementary content	116
CHAPTER 5: Relevance and Contributions.....	144
SECTION 5.1: Review of major findings.....	144
SECTION 5.2: Current relevance of dissertation studies	147
SECTION 5.3: Public health contributions of this dissertation	150
SECTION 5.4: Future directions emerging from this dissertation.....	153
REFERENCES.....	155
APPENDICES	181
A1: Survey questions on the acceptability and ranking of six HIV testing options....	181
A2: Press copy of Chapter 2 (<i>SpringerPlus</i> 2014, 3:109).....	184
A3: KnowAtHome.org landing page	194
A4: Survey questions on the awareness and previous use of home HIV tests, and willingness to potentially self-test, distribute and test with friends or sex partners using rapid home test kits	195
A5: SAS code for simulations and subsequent analyses.....	204
A6: Press copy of Chapter 4 (<i>Contemporary Clinical Trials</i> 2015, 41:152-159).....	321

LIST OF TABLES

TABLE 2.1: Demographic and behavioral characteristics of 973 HIV-negative or unknown status MSM respondents to a national online health survey, United States, 2012.....	41
TABLE 2.2: HIV testing characteristics of 795 HIV-negative or unknown status MSM respondents to a national online health survey who had previously tested, United States, 2012.	43
TABLE 2.3: Stated likelihood of using different HIV testing options if offered free of charge by selected demographic and behavioral characteristics of 973 HIV-negative or unknown status MSM respondents to a national online health survey, United States, 2012.....	44
TABLE 3.1: Demographic and behavioral characteristics of 840 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	73
TABLE 3.2: Baseline testing characteristics of 840 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	76
TABLE 3.3: Reasons and main reason for never being tested for HIV cited by 153 HIV unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	78
TABLE 3.4: Reasons and main reason for not previously using home HIV tests cited by 596 HIV-negative and unknown status MSM participants who had heard about such tests before in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	79
TABLE 3.5: Associations between demographic and behavioral characteristics and the stated likelihood of potentially distributing free oral fluid rapid home tests among 840	

HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	80
TABLE 3.6: Associations between demographic and behavioral characteristics and the stated likelihood of potentially distributing free finger-stick blood rapid home tests among 840 HIV- negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	82
TABLE 3.7: Associations between demographic and behavioral characteristics and the stated likelihood of potentially testing with friends or sex partners using free rapid home tests among 840 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	84
TABLE 3.8: Reasons and main reason for being unlikely to self-test using free rapid home tests cited by 79 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	86
TABLE 3.9: Reasons and main reason for being unlikely to potentially distribute free rapid home tests cited by 179 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	87
TABLE 3.10: Reasons and main reason for being unlikely to potentially test with friends or sex partners using free rapid home tests cited by 232 HIV-negative or unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	88
TABLE 3.S1: Derivation of the analytic sample comprising 840 HIV-negative and unknown status MSM participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	89
TABLE 3.S2: Comparison of 840 HIV-negative and unknown status MSM participants included in the final analytic sample with 152 excluded participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	91

TABLE 4.1: Illustrative demographic characteristics of 1,200 simulated participants from a single replication of a randomized trial.	113
TABLE 4.2: Screening characteristics from a single replication of a randomized trial with 1,200 simulated participants assuming the first parameter specification under Scenario A.	114
TABLE 4.3: Point estimates from different analytical approaches on 10 replications of a randomized trial each with 1,200 simulated participants assuming the first parameter specification under Scenario A.	115
TABLE 4.S1: Data layout example for the first analytical outcome definition.....	122
TABLE 4.S2: Data layout example for the second analytical outcome definition.....	124
TABLE 4.S3: Data layout example for the third analytical outcome definition	130

LIST OF FIGURES

FIGURE 2.1: Stated likelihood of using different currently available HIV testing options if offered free of charge to 973 HIV-negative or unknown status men who have sex with men in a national online health survey, United States, 2012.	39
FIGURE 2.2: Modified Borda Count ranking of different currently available HIV testing options if offered free of charge to 973 HIV-negative or unknown status men who have sex with men in a national online health survey, United States, 2012.	40
FIGURE 3.1: Stated likelihood of potentially self-testing, distributing and testing with friends or sex partners using free rapid home test kits among 840 HIV-negative or unknown status men who have sex with men participating in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.	72
FIGURE 4.1: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario A.	110
FIGURE 4.2: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario B.	111
FIGURE 4.3: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario C.	112
FIGURE 4.S1: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario A.	132
FIGURE 4.S2: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario B.	133
FIGURE 4.S3: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario C.	134

FIGURE 4.S4: Retention at 3, 6, 9 and 12 months for 1,200 simulated participants by study arm from a single trial replication under three censoring scenarios.....	135
FIGURE 4.S5: Frequency distribution of point estimates generated from SLR and GEE for correlated data on 1,000 trial replications each with 1,200 simulated participants assuming the first parameter specification under Scenario A.	136
FIGURE 4.S6: Frequency distribution of point estimates generated from the total follow-up time and gap time models for stratified Cox PH regression on 1,000 trial replications each with 1,200 simulated participants assuming the first parameter specification under Scenario A.	137
FIGURE 4.S7: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario A.	138
FIGURE 4.S8: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario B.	139
FIGURE 4.S9: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario C.	140
FIGURE 4.S10: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario A.	141
FIGURE 4.S11: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario B.	142
FIGURE 4.S12: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario C.	143

CHAPTER 1: Background and significance

SECTION 1.1: MSM and HIV in the United States

Gay, bisexual and other men who have sex with men (collectively referred to as MSM) represent approximately 2% of the United States (US) population[1], but have been a disproportionately affected risk group since early days of the Human Immunodeficiency Virus (HIV) epidemic. Recent estimates released by the Centers for Disease Control and Prevention (CDC) indicate that among adult and adolescent men (aged 13 years and older), the annual number of HIV diagnoses attributed to male-male sex increased from 2009 through 2013[2]. Incident infections among MSM have been rising over the past decade[3], with most new diagnoses occurring among young, black MSM relative to any other age or racial category[4]. 2008 National HIV Behavioral Surveillance (NHBS) data from 21 US cities from MSM aged 18-24 years estimated an annual HIV incidence density of 5.1% among blacks, higher than the incidence among men who were white (1.6%), Hispanic (1.9%) and of other races (2.9%)[5]. According to the CDC, more than half of all new infections per year from 2006 to 2009 in the US occurred among MSM (56% in 2006, 58% in 2007, 56% in 2008, and 61% in 2009)[6]. In 2010, this subgroup accounted for more than three fourths (78%) of incident HIV infections among males, and almost two thirds (63%) of all new infections in the US[7]. By the end of 2012, more than half (51%) of the estimated 880,440 persons living with an HIV diagnosis were gay

and bisexual men[8]. The high prevalence of infection (19%) among MSM[9] implies members of this community have an increased likelihood of being exposed to HIV.

SECTION 1.2: Role of HIV testing in prevention efforts

Novel prevention strategies that incorporate efficacious behavioral, biomedical and structural interventions are urgently needed to help reverse current trends among MSM[10]. Interestingly, testing for HIV is the cornerstone of almost all comprehensive risk reduction efforts. Besides the fact that testing is the essential first step in the process of offering services appropriately tailored to an individual's situation, awareness of one's positive serostatus may confer a prevention benefit. Research has shown that persons who know that they are infected take steps to protect their own health and are motivated to interrupt onward transmission[11]. Given that the estimated HIV transmission rate in the US among persons who are unaware of their infection is 3.5 times that among persons who know that they are positive[12], scaling up testing efforts is critical. Theoretically, if all infected individuals could learn their serostatus and subsequently adopt changes in behavior similar to those adopted by individuals already aware of their infection, incident sexual HIV infections could be reduced by over 30% per year in the US[12]. Increasing the proportion of infected persons who know their serostatus is an established national HIV prevention priority[13]. Recent agent-based models for South African MSM demonstrate that increases in testing can incrementally prevent the proportion of new infections, depending upon which other biomedical and behavioral approaches (increases in antiretroviral treatment [ART] coverage, decreases in unprotected anal intercourse [UAI], increases in pre-exposure prophylaxis [PrEP] acceptance) are combined with HIV

testing as part of prevention packages[14]. Finally, testing is also the gateway to early engagement of HIV-positive persons in treatment and care[15].

SECTION 1.3: Inadequate HIV testing among MSM

Opt-out screening, wherein patients (regardless of behavioral risks) are informed that an HIV test will be performed unless they decline, has been recommended by the CDC for all individuals aged 13-64 in US healthcare settings[16]. The CDC's guidelines for sexually active MSM state that they should be tested for HIV at least once a year, and that higher risk MSM who have multiple and/or anonymous sex partners or use illicit drugs concurrent with sexual activity should be screened for sexually transmitted infections (STIs) at 3 to 6 month intervals[17]. However, recently there have been calls to revisit these recommendations and consider HIV testing at 3 to 6 month intervals for all sexually active MSM, regardless of their self-reported risk behaviors[18]. The United States Preventive Services Task Force (USPSTF) acknowledges that although current evidence is insufficient to determine optimum time intervals for HIV screening, MSM are at very high risk and would likely benefit from frequent screening[19]. Testing within the past year among MSM participants in NHBS reportedly increased from 63% in 2008 to 67% in 2011[20], but the fact that more than a third (34%) of HIV-positive men did not know they were infected at the time of their 2011 interview[21] is concerning.

Commonly cited barriers to testing among MSM include stigma, low risk perception, denial of risk factors, fear of a positive diagnosis, concerns regarding disclosure of serostatus, transportation issues, and the lack of access to healthcare[22-29]. Exploring

new approaches to increase the coverage and frequency of HIV testing in this heavily impacted population is warranted.

SECTION 1.4: Menu of available HIV testing options

Several testing options are currently available to MSM in the US. Depending upon their preferences or circumstances, men can choose from traditional approaches such as testing at their physician's office and individual voluntary counseling and testing (VCT) offered at community-based organizations (CBOs), to newer modalities such as rapid home self-testing with a recently introduced oral fluid test (OraQuick® In-Home HIV Test) and couples' HIV counseling and testing (CHCT). Other options include expedited/express testing which could potentially be offered as part of street outreach programs, and home dried blood spot (DBS) specimen self-collection for laboratory testing (HomeAccess® HIV-1 Test System). Given this menu of testing services, how willing are MSM to use each approach and which of these modalities rank higher than others in terms of intended usage preference? Understanding the acceptability of each option is an initial step in developing effective public health interventions to increase HIV testing among MSM.

Testing at a physician's office

The American College of Physicians (ACP) recommends that clinicians adopt routine screening for HIV, and encourage testing of all adult and adolescent patients in healthcare settings[30]. Blanket screening has an advantage of reducing stigma often associated with

testing that requires an initial assessment of risk behaviors[31, 32]. Here, a patient's oral fluid or blood sample is tested for HIV in a laboratory, and the result is provided either in person or over the phone usually within 7 days. Some facilities conduct rapid HIV tests and deliver results at the appointment itself. HIV testing in clinical settings is currently covered by Medicare (one screening test per year)[33] and Medicaid in some US states[34]. Under the Affordable Care Act (ACA), most new private health insurance plans must now cover testing for everyone aged 15 to 65 years without additional cost-sharing such as copays or deductibles[35]. Some healthcare providers may offer HIV testing on a sliding scale depending upon their patient's financial status.

The ACP's recommendations also state that clinicians should determine the need for repeat HIV screening on an individual basis, testing higher-risk patients including MSM more frequently than lower-risk patients[30]. These guidelines have been derived from an evaluation of screening recommendations previously developed by the United States Preventive Services Task Force (USPSTF) and the CDC. However, a comprehensive review found that physicians in the US often face policy level barriers (consent process, pre-test counseling, inadequate reimbursement), logistical barriers (insufficient time, competing priorities, language issues), and educational barriers (lack of patient acceptance, lack of knowledge/training) in the process of implementing routine HIV testing[36]. Patient level barriers traditionally encountered by MSM include social and economic factors such as homophobia, potential discrimination on receiving a positive test result, and lack of financial resources to access healthcare services[22-29].

Despite these impediments, physician's offices have been one of the most common venues where HIV tests are administered, often as part of general physical exams.

Historically, 65% of the 2,441 persons diagnosed with Acquired Immune Deficiency Syndrome (AIDS) between 1990-1992 residing in 11 US states and cities were tested for HIV in acute healthcare settings: 33% in hospitals, 28% in physician's offices, and 4% in emergency departments[37]. Depending upon the kind of relationship and levels of trust that MSM share with their physicians, some men may feel more comfortable getting tested at their doctor's office. NHBS data indicate that 36% of the 7,057 MSM interviewed between 2003-2005[26] and 26% of the 5,082 MSM interviewed in 2008[27] who reported having been tested for HIV within the past year, indicated doing so at their private doctor's office.

Individual voluntary counseling and testing (VCT)

Previous research has demonstrated that VCT is an efficacious, cost-effective approach to promote behavior change and reduce HIV transmission[38-40]. VCT is usually provided at CBOs, public health clinics and testing centers across the US. Here, individuals first receive education and prevention counseling in a one-on-one session with a counselor, who also discusses their HIV-related risks (pre-test counseling). Next, they are tested by a trained professional, usually using either a rapid oral fluid or finger-stick blood test. Finally, they receive their result in person from the counselor and develop a plan for the future. HIV-negative individuals are advised on risk reducing techniques to ensure they remain uninfected. Persons who test preliminary positive are provided emotional support, referred for confirmatory testing, linked to treatment programs, encouraged to disclose their serostatus to their sex partners, and counseled to prevent onward transmission (post-

test counseling). VCT therefore constitutes an opportunity for both primary prevention (i.e. preventing uninfected people from contracting HIV) and secondary prevention (i.e. avoiding the progression of the disease in infected people by linking them to medical care). This testing option usually takes approximately 30-45 minutes if a rapid test is used, and is frequently offered free or on a sliding scale depending upon one's income.

Meta-analytic evidence from studies among MSM in developed countries suggests that when HIV counseling with testing was linked to sexual behavior change, it appeared to be more effective in reducing risk behaviors among HIV-positive participants; VCT as implemented in the reviewed studies had little effect on HIV-negative participants[41, 42]. However, results from randomized controlled trials (RCTs) indicate that the nature and duration of prevention counseling might influence its effectiveness. Project RESPECT, a multicenter trial conducted with 5,758 heterosexual HIV-negative patients in five public sexually transmitted disease (STD) clinics in the US, demonstrated that short counseling interventions using personalized risk reduction plans decreased participants' risk behaviors, compared to receiving only didactic prevention messages from health-care providers[43]. The EXPLORE RCT comparing an intensive intervention of 10 one-on-one counseling sessions followed by maintenance sessions every 3 months with a standard of twice-yearly individualized prevention counseling sessions among 4,295 HIV-negative MSM, resulted in UAI reductions with HIV-positive and unknown status partners, but not in HIV incidence[44]. NHBS data from 2008 estimate that 42% of the 5,082 MSM who reportedly tested within the past year, did so at HIV counseling and testing sites, public health clinics and community health centers, settings primarily offering VCT[27].

Couples' HIV counseling and testing (CHCT)

CHCT, in which both partners participate in the whole cycle of counseling and testing together, was introduced more than 20 years ago in Africa and has been described by the CDC as a “high-leverage HIV prevention intervention” for heterosexual couples in that region[45]. Here, couples first receive pre-test information, risk ascertainment and counseling together, and then get tested for HIV by a trained professional. Next, the counselor presents two sets of results to both partners at the same time: their personal results (negative or positive) and their results as a couple (sero-concordant negative, sero-concordant positive, or sero-discordant). Finally, couples receive post-test risk reduction counseling tailored to their couple serostatus, and jointly develop a plan for the future. The entire process usually takes approximately 45 minutes to an hour. CHCT can address issues such as failure to disclose one's serostatus to their partner, enables the formulation of risk-reduction plans based on partner serostatus, and has been shown to decrease high-risk behavior and HIV transmission in heterosexual couples[38, 46-48]. Modeling has demonstrated that CHCT has the potential to avert more than two-thirds of incident infections among urban African men and women[49], and this approach may prove effective in reducing transmission among serodiscordant MSM couples.

The American Community Survey estimated that the number of same-sex couples living in the US increased by 30% from 2000 to 2005[50]. In addition to knowing one's own HIV status, knowledge of partner serostatus is important in decreasing transmission risk. Recent studies indicate that between one third and two thirds of new HIV infections

among MSM are likely attributable to sex with their main partners[51, 52]. Given these trends, preventive efforts targeting MSM couples may help reduce the spread of HIV. Many men already have informal agreements with their partners relating to rules and boundaries around sexual behavior within and outside their relationship[53]. However, interventions targeting MSM couples are relatively new, and sustained prevention services which address issues with partner disclosure are needed[11].

CHCT is currently being offered free of charge at some CBOs and HIV testing centers in select US cities (New York, Los Angeles, Miami, Chicago, Washington DC, Dallas, Houston, Philadelphia, Atlanta, San Francisco, Baltimore, and Orlando). CDC's national diffusion plan for CHCT was initiated in 2012 focusing initial roll-out in the 12 jurisdictions with highest HIV prevalence and incident diagnoses, and this service will be expanded to other areas as demand increases and resources become available[54]. Initial studies have found that MSM's attitudes towards this testing option are generally favorable. In a recent national online survey, young men of color and those with main sex partners expressed a high level of willingness to participate in couples' counseling and testing with a male partner within the next year[55]. Focus groups with MSM in Atlanta, Chicago and Seattle also revealed high acceptability for CHCT, the approach being regarded as a possible mechanism through which partners could disclose their HIV status and plan for sexual risk reduction based on their test results[56].

Expedited/express testing

HIV prevention counseling might not be appropriate or feasible in all settings[16], and could be a barrier to getting tested for some MSM. Expedited/express testing is a new service that could potentially be offered at large social events such as gay pride, as well as through street outreach programs. Unlike VCT and CHCT, this approach excludes the standard HIV education and prevention counseling components. Here, individuals first provide their contact information electronically, after which a trained professional obtains an oral fluid or finger-stick blood specimen for rapid testing. Depending upon the type of test used, results are available within a few minutes or even seconds (e.g. INSTI™ HIV-1 Antibody Test Kit can provide results within 60 seconds[57]). Individuals could choose to receive results by text message or email, or retrieve them online using a confidential PIN whenever they are ready. Preliminary positive persons would receive their results by phone or in person by a trained HIV counselor, who would provide them support and referrals for confirmatory testing and linkage to care.

Previous studies have indicated equivocal benefits of providing prevention counseling in conjunction with HIV testing to reduce sexual risk behaviors[41]. Rapid testing without an initial risk assessment could help reduce the stigma associated with HIV testing modalities which first require a discussion about recent behaviors[31, 32]. Additionally, this approach could reduce the length of time associated with the testing process, and offers the potential advantage of reaching large numbers of individuals concentrated at specific geographic locations. Findings from a demonstration project in seven US cities indicated that rapid HIV testing in outreach and community settings was a feasible approach to reach members of minority subgroups and people who are at risk for HIV, including MSM[58]. An RCT conducted at a needle exchange and two bathhouses

frequented by MSM in Seattle found that rapid oral fluid and blood testing resulted in significantly more people receiving HIV test results compared with traditional testing, and that making counseling optional increased testing at the needle exchange[59].

Evaluating attitudes towards expedited/express testing is an initial step in determining whether such an approach should be considered when conceptualizing new packages of HIV prevention interventions for MSM.

Rapid home self-testing using an oral fluid test

In July 2012, the US Food and Drug Administration (FDA) approved the OraQuick® In-Home HIV Test, the first over-the-counter (OTC) rapid test that can detect the presence of antibodies in oral fluid within 20 minutes[60]. After more than two decades of politicking and lawsuits over its potential approval[61], individuals can now buy this test online or at a pharmacy and learn their HIV status in the privacy of their home almost instantly. Here, one first collects an oral fluid sample by swabbing their upper and lower gum, and then puts that swab inside a developer vial provided as part of the kit. HIV test results can be read by persons on their own within 20 minutes, by comparing the lines on their test device with color images printed in the instruction booklet. In addition to the step-by-step instructions included with the kit, individuals can watch an online video for guidance[62]. If one has any questions or concerns, or if one receives a preliminary positive test result, they can call a 24x7 support center free of charge. According to the company's website, the OraQuick In-Home HIV Test currently costs about \$40[63].

Research conducted years before this test became commercially available indicated high levels of interest in rapid home self-testing. Over a third (37%) of the 2,964 respondents to the 1999 California Behavioral Risk Factor Surveillance System (BRFSS) telephone survey, administered by the California Department of Health Services through random-digit dialing of residences, indicated they would consider using instant home HIV tests, the acceptability of this approach being higher among individuals who had never been tested before[64]. In another survey conducted among public clinic attendees being tested for HIV in San Francisco, almost a quarter (24%) of the 354 participants indicated that an instant home test, if available, would be their most preferred choice[65]. A study among clients of a needle exchange, an STD clinic, and three sex venues for MSM in Seattle found that participants preferred rapid home self-testing to clinic-based testing as they considered it to be more private, convenient and faster[24]. Furthermore, participants in that study who preferred oral fluid testing reported a dislike of venipuncture[24]. In a recent meta-analysis, wherein majority (14/21, 67%) of studies evaluated oral fluid HIV tests, their preference was attributed to non-invasiveness, convenience, and ease of specimen collection[66].

Potential alternative uses of rapid home HIV self-tests include distribution among one's social or sexual networks, and screening with partners in the same location before having sex. Peer-driven prevention approaches have been effective in identifying high risk persons with undiagnosed infections[67], and making such test kits available to interested MSM could offer a potential public health benefit. Recent work has indicated that using rapid home tests as a partner screening device may be a useful risk reduction strategy for MSM who engage in UAI[68]. Despite concerns raised by some researchers that such

tests may not have a profound impact on either the HIV public health crisis or among populations in greatest need[69], the historical enthusiasm among members of high risk groups suggests that research evaluating their efficacy is warranted, especially now that these are readily available in the US.

Home dried blood spot (DBS) specimen self-collection for laboratory testing

The Home Access® HIV-1 Test System is a DBS specimen collection kit, which was approved by the FDA for OTC sale in July 1996[70]. Even this test can be purchased online or at a pharmacy and performed in the privacy of one's home without medical supervision, but it differs from rapid home HIV self-testing in some aspects. Here, individuals first register their kit by calling a toll free number and complete pre-test counseling by phone. Next, they prick their finger and collect a few drops of blood on a specimen collection card. After letting the card dry overnight, they package their sample in a weather resistant protective pouch and return that to the company's accredited laboratory in a prepaid shipping envelope. Finally, they access their test results by phone and receive post-test counseling within 7 business days from when they return their specimen. The Home Access® HIV-1 Test System currently costs \$44[71].

DBS specimen collection kits for home HIV testing have been unevenly adopted in the US. During the first year of availability, a comparison of kit users with individuals who tested at publicly funded testing locations suggested that these were used by persons at high risk for HIV and by those who did not seek other testing[72]. Few years later, data from seven states included in the HIV Testing Survey (HITS) of 2000 revealed that the

overall awareness and use of these tests were limited: 54% of the 2,836 respondents were aware of the home collection kit, and only 4% reported using them[73]. However, in a research study on the feasibility and acceptability of bimonthly specimen collection for HIV testing among high-risk individuals, which included MSM recruited from four sites in the HIV Network for Prevention Trials (HIVNET) cohort, 90% of 357 expected DBS specimens were returned to the laboratory[74]. Our previous survey among 6,163 HIV-negative internet-using MSM indicated that 62% of men reported being very likely and 20% of men reported being somewhat likely to accept free home specimen collection kits for HIV testing as part of an online prevention study[75]. In our follow-up online study, 79% of the 895 MSM who were sent DBS specimen collection kits returned their samples for laboratory testing[76]. How this testing modality compares with more recent technologies, such as the oral fluid rapid home HIV test, in terms of acceptability and preference is yet to be determined.

SECTION 1.5: Internet-based HIV prevention strategies

Emerging technologies have the potential to target hard-to-reach MSM populations, and engage individuals in easily accessible areas who are being missed in existing outreach efforts[77, 78]. Internet-based and mobile phone-based HIV prevention strategies are demonstrating feasibility, acceptability and efficacy for both consumers as well as providers[79, 80]. Results from a recent RCT provide support for effectively using the internet in delivering HIV prevention messages to rural MSM[81]. Additionally, meta-analytic evidence has suggested that online and offline computer-based HIV prevention

interventions are just as efficacious as many commonly utilized human-delivered interventions, especially when they employ individualized content tailoring [82]. In the context of public health research, online surveys enable data collection in a time and resource efficient manner from a large number of subjects who could be residing in widely scattered geographical locations. For example, the European MSM Internet Survey (EMIS) on HIV-related male homosexual behaviors and prevention needs gathered 184,469 responses from 38 countries between June and August 2010[83]. According to the American Community Survey Reports, 74% of all households reported internet use and 73% indicated having access to a high speed connection in 2013[84]. Given that MSM in the US are increasingly using the internet to meet sex partners[85, 86], and that online negotiations of both high-risk and safe sex have become increasingly prevalent among members of this community[86-90], this medium holds immense promise for the streamlined delivery of new prevention strategies.

SECTION 1.6: Motivation for dissertation studies

The high burden of HIV among MSM in the US warrants the development of novel approaches to increase the frequency of testing. Depending upon one's needs and personal circumstances, gay, bisexual and other MSM residing in the US can choose from a menu of testing options. We believe that exploring current attitudes towards established modalities (e.g. testing at a physician's office) and newer concepts (e.g. expedited/express testing) is an initial step in formulating effective public health strategies targeting this disproportionately impacted population. Although previous

research has examined selected testing approaches in isolation[55, 75, 91], we are not aware of any study that has investigated MSM's acceptability of different modalities presented collectively. Combination testing packages could help individuals in putting together annual testing strategies tailored to their unique situations and risk perceptions. For the purpose of informing the preliminary development of such an intervention, we sought to explore the likelihood of using as many as six different HIV testing approaches among internet-using MSM, and identify their most preferred options stratified by demographic and behavioral characteristics (*Study 1*).

Data regarding uptake of the Home Access® HIV-1 Test System after its OTC release in 1996[70] have been previously reported, both in real world and research settings[73, 92]. However, information about MSM's adoption of the OraQuick® In-Home HIV Test after it became commercially available in 2012[60] is currently lacking. The Sure Check® HIV 1/2 Assay is another rapid test which can detect antibodies in finger-stick blood in 15 minutes[93], and at the time of conducting our research its manufacturers were seeking an FDA premarket approval for a home use version. Alternative uses of this relatively new technology, such as kit distribution and testing with friends or sex partners in the same place, could facilitate the penetration of HIV testing into high risk networks. Given the limited literature on MSM's intentions of employing such strategies, we planned on studying different aspects of rapid home testing with oral fluid and finger-stick blood tests including self-perceived barriers among MSM recruited online as part of KnowAtHome[94], a CDC-sponsored research project (*Study 2*).

The proliferation of new preventive health screening technologies is indeed an exciting development in recent years, but an important step after exploring attitudes towards their

adoption is their rigorous field evaluation. RCTs are widely considered to be the “gold-standard” scientific tool for testing new interventions. Researchers conducting trials to promote regular health exams need to determine *a priori* what constitutes a meaningful outcome. For example, consider a hypothetical RCT seeking to determine whether a newly conceptualized combination HIV testing package can improve testing behaviors in a 6-month period among high risk MSM. How should we define our study outcome? Possible options include a dichotomous measure (e.g. testing ≥ 2 times over the entire duration of the study versus not), an interval-specific measure (e.g. testing at least once in a 3-month interval versus not) and a survival time measure (e.g. “hazard” or instantaneous rate of testing). Depending upon the choice of our analytical outcome, different statistical approaches which appropriately account for recurrent events would be needed to figure out whether the intervention truly had an underlying effect, directly impacting adequate sample size estimation. In an effort to understand the efficiency of multiple approaches for evaluating a prevention strategy aimed at increasing routine health screenings, we simulated several replications of a “successful” RCT, formulated at least three outcome definitions for analyses and compared them with regard to interpreting results and estimating power (*Study 3*).

SECTION 1.7: Specific objectives for dissertation studies

The following is a list of specific objectives for each of our three studies focusing on the evaluation of new strategies to promote frequent HIV testing among MSM:

Study 1

- Determine the acceptability of the following six HIV testing approaches presented collectively to internet-using MSM in the US when hypothetically offered free of charge: Testing at a physician's office; VCT; CHCT; Expedited/express testing; Rapid home self-testing using an oral fluid test; Home DBS specimen self-collection for laboratory testing
- Identify which HIV testing options rank higher than others in terms of intended usage preference, overall and within selected demographic and behavioral strata

Study 2

- Establish the levels of awareness and previous use of OTC home HIV tests among internet-using MSM residing in the US
- Identify factors associated with their reported likelihood of potentially distributing free rapid home HIV test kits and potentially testing with their social or sexual network associates in the same place
- Describe self-perceived barriers to self-testing, kit distribution and testing with friends or sex partners

Study 3

- Simulate several replications of a "successful" RCT (i.e. one with a positive intervention effect) under various censoring scenarios

- Formulate three outcome definitions (screening a certain number of times during the entire study period versus not, screening at least once within a clinically meaningful interval versus not, “hazard” or instantaneous rate of screening) and perform corresponding analyses
- Compare them with regard to interpreting results and estimating power at different sample sizes

SECTION 1.8: Structure of this dissertation

The next chapter (Chapter 2) contains an original scientific manuscript that addresses the objectives of our first dissertation study on HIV testing preferences among MSM in the US. This is followed by a chapter (Chapter 3) in which we address the objectives of our second dissertation study on MSM’s past experiences and future intentions regarding various aspects of home HIV testing. The following chapter (Chapter 4) contains another original scientific manuscript that addresses the objectives of our third dissertation study on comparing multiple analytical outcomes for an RCT seeking to evaluate a new public health intervention. In the final chapter (Chapter 5), we discuss the current relevance and contributions of studies included in this dissertation to the field of HIV prevention.

CHAPTER 2: Acceptability and intended usage preferences for six HIV testing options among internet-using men who have sex with men

[Chapter published as a manuscript in *SpringerPlus* 2014, 3:109. Designated ‘Highly accessed’. Available at: <http://www.springerplus.com/content/3/1/109>. Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium.]

Authors and affiliations: Akshay Sharma^{a, b, *}, Rob B Stephenson^c, Darcy White^c, Patrick S Sullivan^{a, b}.

^a Department of Epidemiology, Emory University Laney Graduate School, 1518 Clifton Road NE, Atlanta GA 30322, USA.

^b Department of Epidemiology, Emory University Rollins School of Public Health, 1518 Clifton Road NE, Atlanta GA 30322, USA.

^c Hubert Department of Global Health, Emory University Rollins School of Public Health, 1518 Clifton Road NE, Atlanta GA 30322, USA.

***Correspondence:** ashar24@emory.edu

Running head: HIV testing preferences among MSM.

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: AS participated in designing the study and collecting data, performed the statistical analyses, and drafted the manuscript. RBS participated in conceiving and designing the study, contributed to data interpretation, and helped draft the manuscript. DW participated in coordinating the study and collecting data, and helped draft the manuscript. PSS participated in conceiving and designing the study, contributed to data interpretation, and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements: This work was facilitated by the Center for AIDS Research at Emory University (P30 AI050409).

Abbreviations: CDC, Centers for Disease Control and Prevention; CHCT, Couples' HIV Counseling and Testing; DBS, Dried Blood Spot; FDA, Food and Drug Administration; GED, General Educational Development; HIV, Human Immunodeficiency Virus; MBC, Modified Borda Count; MSM, Men who have Sex with Men; STI, Sexually Transmitted Infection; UAI, Unprotected Anal Intercourse; US, United States; VCT, Individual Voluntary Counseling and Testing.

SECTION 2.1: Abstract

Background: Men who have sex with men (MSM) continue to be disproportionately impacted by the Human Immunodeficiency Virus (HIV) epidemic in the United States (US). Testing for HIV is the cornerstone of comprehensive prevention efforts and the gateway to early engagement of infected individuals in medical care. We sought to determine attitudes towards six different HIV testing modalities presented collectively to internet-using MSM and identify which options rank higher than others in terms of intended usage preference.

Methods: Between October and November 2012, we surveyed 973 HIV-negative or -unknown status MSM and assessed their acceptability of each of the following services hypothetically offered free of charge: Testing at a physician's office; Voluntary counseling and testing (VCT); Couples' HIV counseling and testing (CHCT); Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) specimen self-collection for laboratory testing. Kruskal-Wallis tests were used to determine whether the stated likelihood of using each of these modalities differed by selected respondent characteristics. Men were also asked to rank these options in order of intended usage preference, and consensual rankings were determined using the modified Borda count (MBC) method.

Results: Most participants reported being extremely likely or somewhat likely to use all HIV testing modalities except DBS self-collection for laboratory testing. Younger MSM indicated greater acceptability for expedited/express testing ($P < 0.001$), and MSM with lower educational levels reported being more likely to use CHCT ($P < 0.001$). Non-

Hispanic black MSM indicated lower acceptability for VCT ($P < 0.001$). Rapid home self-testing using an oral fluid test and testing at a physician's office were the two most preferred options across all demographic and behavioral strata.

Conclusions: Novel approaches to increase the frequency of HIV testing among US MSM are urgently needed. Combination testing packages could enable high risk MSM in putting together annual testing strategies personalized to their circumstances, and warrant due consideration as an element of combination HIV prevention packages.

Key words: HIV testing preferences; internet-using men who have sex with men; combination prevention approaches; rapid home HIV self-testing.

SECTION 2.2: Introduction

Men who have sex with men (MSM) comprise approximately 4% of the adult male United States (US) population[1], but are the most heavily impacted risk group for Human Immunodeficiency Virus (HIV) infection. Since 2000, incident infections among MSM have been increasing annually[3], with the rate of new diagnoses in this group being at least 44 times that of other men[1]. In 2010, MSM accounted for more than three fourths (78%) of new HIV infections among males, and almost two thirds (63%) of all new infections in the US[7]. Most incident diagnoses occurred among young (ages 13-24), black MSM relative to any other age or racial category[7]. Better prevention strategies are needed to help reverse current trends.

Behavioral interventions, such as risk reduction counseling, and biomedical approaches, such as condoms and pre-exposure prophylaxis, have complementary roles in HIV prevention. Modeling experiments have shown that offering packages of currently available interventions can avert at least 25% of new infections among MSM over a decade[10]. Testing for HIV is not just a critical first step in developing client-specific recommendations regarding the adoption of these approaches, but can be considered an important prevention activity in itself. Meta-analytic evidence suggests that seropositive individuals aware of their status are motivated to interrupt onward transmission and reduce risky behaviors including unprotected anal intercourse (UAI)[11, 95]. HIV testing is also the gateway to early engagement of infected individuals in treatment and care[15], wherein resulting viral load reductions are known to offer substantial prevention benefits[96].

The US Centers for Disease Control and Prevention (CDC) currently recommends that sexually active MSM should be tested for HIV annually, and that higher risk MSM who have multiple partners or use illicit drugs concurrent with sexual activity should be screened for sexually transmitted infections (STI) at 3-6 month intervals[17]. Although the nationwide prevalence of lifetime testing among MSM is high[27, 97], many men report not being tested within the past year[27] and a high proportion of seropositive MSM are unaware that they are infected[25, 98]. The estimated HIV transmission rate from persons who are unaware of their infection is 3.5 times that from serostatus-aware individuals[12]. MSM therefore remain a key risk group for expanded testing efforts. Increasing the percentage of infected individuals who know their serostatus is one of the goals of the National HIV/AIDS Strategy[13] and a Healthy People 2020 objective[99].

Depending upon their preferences or circumstances, MSM in the US can choose from several HIV testing approaches ranging from the traditional to the contemporary. Physician's offices, frequently offering screening as part of routine general physical exams, have remained one of the most common testing venues[26, 27]. Individual voluntary counseling and testing (VCT) is usually provided at community-based organizations, and involves one-on-one sessions comprising of pre-test risk assessments and post-test counseling. VCT has proven efficacious in promoting behavior change among high risk persons who learn they are living with HIV, and constitutes an opportunity for both primary and secondary prevention[42, 44]. In 2012, the CDC initiated a national diffusion plan for couples' HIV counseling and testing (CHCT) targeting same-sex couples in 12 US jurisdictions with the highest HIV burden[54]. Here, partners participate in the whole cycle of counseling and testing together, and receive risk

reduction messages tailored to their couple serostatus (sero-concordant negative, sero-concordant positive, or sero-discordant)[100].

Although prevention counseling is desirable for high risk individuals, the CDC recognizes that such counseling might not be appropriate or feasible in all settings[16], and it could pose a barrier to testing. States such as New York that have streamlined regulations regarding pre-test counseling have seen increases in HIV testing[101]. Given the recent licensure of rapid tests with processing times as little as 60 seconds[57], an expedited/express testing approach that excludes prevention counseling sessions could be provided through street outreach programs at large events such as gay pride. Individuals could choose to receive their results by text message or email, or retrieve them online using a confidential personal identification number (PIN) whenever ready. Preliminary positive persons would receive their results either by phone or in person by a trained counselor. In addition to saving time, this approach can help reduce stigma associated with HIV testing options requiring an assessment of risk behaviors[31, 32]. Rapid home self-testing with a recently US Food and Drug Administration (FDA) approved oral fluid test[60], is another testing modality offering privacy and convenience[66]. Individuals can read their own test results within 20 minutes, and have the option of calling a 24x7 support center toll free if they have questions or receive a preliminary positive result. This non-invasive approach differs from home dried blood spot (DBS) self-collection wherein specimens need to be returned for laboratory HIV testing, and results are available by phone within 7 days[70].

Considering this menu of available options, we believe that analogous to combination prevention approaches, combination testing packages need due consideration as an

element in continuing efforts to increase HIV testing frequencies among high risk populations. Such an intervention could enable individuals in putting together annual personalized testing strategies tailored to their needs and risk perceptions. Previous online and in-person research studies among MSM, each focusing on selected testing modalities in isolation, have found generally favorable attitudes towards their adoption[28, 55, 56, 64, 74, 75, 91, 102]. We sought to explore the acceptability of six HIV testing approaches presented collectively to internet-using MSM in the US when hypothetically offered free of charge. Another objective of our study was determining which testing options rank higher than others in terms of intended usage preference. Identifying variations in ranking orders within demographic and behavioral strata of MSM represents an initial step in developing comprehensive packages of HIV testing services targeting specific subgroups.

SECTION 2.3: Methods

MSM were recruited online through selective placement of banner advertisements displayed on a social networking website (Facebook.com) from October to November 2012. Recruitment was targeted only towards internet users in the US who indicated in their Facebook profile that they were male, 18 years of age or older and interested in men. Individuals who clicked through the banner advertisements were directed to an online informed consent module, and those who consented were screened to determine eligibility before being administered an internet-based survey. Eligibility criteria included being reportedly male, 18 years of age or older, currently residing within the US, and having at least one male sex partner in the past 6 months.

Demographic information collected from participants included age, race/ethnicity, state of residence, education, employment, self-identified sexual orientation and whether they had a main partner. Questions pertaining to the participants' behaviors included whether they had engaged in UAI with male sex partners in the past 6 months, and HIV testing characteristics detailing the timing, location and type of their most recent test. Men who reported being previously tested were asked to indicate one or more motivations for their decision to test from a list of pre-specified options based on subject area expertise, and provided with the choice of typing in an open-ended response.

Participants who reported not being infected with HIV were provided brief descriptions about different testing approaches, and then asked about their likelihood of using each option hypothetically offered free of charge. Acceptability was assessed by the question: "On a scale from one to five, how likely would you be to use this service?" Six questions

of this form were asked, one for each of the following approaches: Testing at a physician's office; VCT; CHCT; Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) specimen self-collection for laboratory testing. Responses were collected in the following Likert item format: 1=Extremely unlikely; 2=Somewhat unlikely; 3=Neutral; 4=Somewhat likely; 5=Extremely likely. Finally, men were asked to rank these options in order of intended usage preference from the one they were most likely to use (assigned Rank 1) to the one they were least likely to use (assigned Rank 6).

Statistical analyses were performed using SAS version 9.3[103]. The analytic sample only included self-reported HIV-negative or unknown status MSM who answered at least one of six questions on the acceptability of various testing approaches. Fisher's exact tests were used to compare these respondents with men who were excluded.

Demographic and behavioral characteristics of all study participants and HIV testing characteristics of men who reported being previously tested were tabulated. Responses for their decisions to test were summarized, and open-ended comments were manually reviewed and reassigned to appropriate pre-specified options.

The acceptability of various HIV testing approaches stratified by selected demographic and behavioral characteristics was summarized by finding the median and mean of participants' five-point Likert item responses. Given the ordinal nature of these data, non-parametric tests are preferable for statistical inferences[104]. The Kruskal-Wallis ANOVA, a generalized form of the Wilcoxon-Mann-Whitney test, was used to determine whether the intended usage likelihood of a particular testing option differed by the following respondent characteristics: age; race/ethnicity; education; whether they had a

main partner; whether they had engaged in UAI with male sex partners in the past 6 months; HIV testing history. Because of our *a priori* decision to conduct 36 independent tests (6 testing approaches times 6 participant characteristics), the alpha level was corrected using the Sidak equation to limit the overall risk of making at least one Type I error to 0.05[105]. Each Kruskal-Wallis test was considered statistically significant only if its associated probability was smaller than 0.001. Additional analyses were performed to examine whether participants' stated likelihood of using any testing option differed by geographic region of residence.

The modified Borda count (MBC) method was used to identify the relative orders of preferences for the various testing modalities overall, as well as stratified by selected participant characteristics. The original system invented by Jean-Charles de Borda in 1770 was intended for use in elections with a single winner[106]. Each testing approach was assigned a certain number of points corresponding to the position in which it was ranked by individual respondents. The number of points given for a participant's first and subsequent choices was determined by the total number of options he actually ranked, rather than the total number of options available. Points for each approach were summed to determine ranking orders representing the collective best compromise within each stratum. This method effectively penalized respondents who did not rank all six testing approaches, by diminishing the number of points their ranking distributed among these options, thereby favoring approaches supported by a broad consensus.

SECTION 2.4: Results

Overall, 432,632 advertising impressions on Facebook resulted in 4,638 click-throughs to the survey over a 10-day period; 1,739 (38% of click-throughs) consented and were asked questions used to determine eligibility. Of these, 15 identified their gender other than male, 37 were less than 18 years of age, 15 did not reside within the US, 335 did not self-report sex with a man in the past 6 months, and 86 did not respond to one or more of the eligibility questions, yielding a sample of 1,285 (74% of respondents to eligibility questions) who could potentially complete the survey. We restricted our analyses to 973 (81% of 1,204 HIV-negative or unknown status participants) who answered at least one of the six acceptability questions. Compared to these participants, excluded men were more likely to be non-Hispanic black, but similar with respect to all other characteristics (data not shown in table).

Table 2.1 summarizes the demographic and behavioral characteristics of respondents included in our analyses. Majority of the participants were young (mean age in years = 31; median = 26) non-Hispanic white men with some college education or higher. More than one third had a main partner for ≥ 1 year, one fifth reported having UAI with ≥ 2 men within the past 6 months, and almost one fifth had never been tested for HIV.

The HIV testing characteristics of 795 participants who reported being previously tested are described in Table 2.2. More than two fifths of the men had their most recent test > 1 year prior to the survey. Among the 56% who tested within the past year, almost a quarter indicated testing routinely every 12 months, approximately one third tested routinely every 6 months and almost one fifth tested routinely every 3 months. Private doctors'

offices and public health clinics were the most commonly reported testing locations, followed by HIV counseling and testing sites.

Regarding participants' decisions to test for HIV, 55% of the 795 indicated they got tested routinely, 25% before they started having sex with a new partner, 21% whenever they had the opportunity, 15% after they had UAI with someone whose HIV status they did not know, 10% whenever someone they had sex with told them they had an STI, 10% if they started to notice or feel symptoms of an STI, 3% after they had UAI with someone they knew to be HIV-positive, and 2% whenever they felt the need to test. Respondents could have indicated more than one motivation for their decision to previously test for HIV: 23% chose multiple reasons, 72% chose a single reason and 5% did not specify a reason (data not shown in table).

Figure 2.1 depicts MSM's stated likelihood of using each of the six testing approaches hypothetically offered free of charge. Overall, majority of the men reported being extremely likely or somewhat likely to use different options. DBS self-collection for laboratory testing was the only approach that appealed to less than half the participants.

The intended usage likelihood for each testing approach stratified by selected demographic and behavioral characteristics is summarized in Table 2.3. On adjusting for multiple comparisons, younger participants were significantly more likely to use expedited/express testing ($P < 0.001$), non-Hispanic black participants reported lower acceptability for VCT ($P < 0.001$), and participants with lower educational levels were more likely to use CHCT ($P < 0.001$). The stated likelihood of using any particular option

did not significantly differ by the behavioral characteristics of respondents or by their region of residence (data not shown in table).

The MBC ranking orders for all six HIV testing approaches are presented in Figure 2.2. Overall, rapid home self-testing using an oral fluid test and testing at a physician's office were the two most preferred options. Expedited/express testing and VCT were next, followed by DBS self-collection for laboratory testing and CHCT. Similar patterns were observed on stratifying by HIV testing history, relationship status, and history of UAI with a male sex partner within the past 6 months.

SECTION 2.5: Discussion

Our study sought to explore attitudes towards using long established and newer HIV testing modalities available in the US. Specifically, we were interested in determining the acceptability and intended usage preferences for six testing options hypothetically offered free of charge to internet-using MSM. Based on self-reported likelihood of using each approach, our results indicate high overall acceptability, demonstrating the potential for combining multiple options as part of comprehensive packages to promote regular testing in this disproportionately impacted population. Motivations for our participants' decisions to test are comparable to MSM undergoing HIV testing at a community-based program in Seattle[107]. Clear preferences for test types emerged across subgroups, revealing which approaches men would potentially employ in developing personalized testing strategies.

Across demographic and behavioral participant characteristics, MSM in our study generally reported being extremely or somewhat likely to use various testing modalities. Younger men significantly preferred expedited/express testing compared to older men. Possible explanations include the reduced time associated with this approach, not having to receive counseling, and the likely appeal of receiving results through text messages if so desired. Young US adults are avid users of text messaging, current statistics indicating that cell phone owners between 18-24 years exchange an average of 109.5 messages per day[108]. Previous research with MSM has found higher acceptability for rapid versus traditional testing in outreach settings, and an increase in testing when counseling was made optional[22, 59]. Compared to men of other racial/ethnic categories, non-Hispanic black men reported being less likely to use VCT. Black MSM's experiences with societal

and institutional racism, coupled with a general distrust of the medical community and heightened perceptions of stigma, have posed personal and systemic barriers to them accessing HIV prevention resources[109]. Although we agree with advocates of providing culturally competent counseling and testing services[17, 110, 111], this result needs to be interpreted with caution due to the underrepresentation of black MSM in our study. Finally, lower educational levels were significantly associated with a higher stated likelihood of using CHCT, the direction of this result being consistent with a recent study among internet-using MSM in seven countries[112]. Another online study found that South African MSM with more schooling were significantly less likely to express willingness to utilize CHCT services[113]. Possible explanations could include greater financial resources enabling more access to health care providers[114], and a lower perceived or actual risk of acquiring HIV among better educated individuals[115].

Given that rapid home self-testing and testing at a physician's office consensually emerged as the top ranked choices, these warrant consideration as key components of potential future combination HIV testing packages for MSM. Privacy, convenience, ease of specimen collection, almost instantaneous results, and not having to visit a testing facility have been reported as favorable attributes of rapid home oral fluid self-testing by high risk populations globally[24, 64, 65, 91, 102, 116]. Depending upon the kind of relationship and levels of trust MSM share with their physicians, some men may feel more comfortable getting tested at their doctor's office. Favorable attributes of this option that distinguish it from rapid home self-testing include the availability of in-person post-test counseling for newly diagnosed positives as well as the potential for early initiation of treatment. Expedited/express testing was consistently ranked as the third choice across

risk groups in our study, higher than both VCT and CHCT, suggesting a niche role for this modality in frequent testing strategies. Previous studies among MSM have reported mixed reactions towards pre-test counseling, ranging from generally positive attitudes[117] to considering it ‘repetitive’ and ‘unnecessary’[22]. Despite only a quarter of our participants’ negative perceptions about CHCT, comparable to online research with MSM in the US[55], Australia and the United Kingdom[112], this approach ranked low in terms of intended usage preference when presented in conjunction with other testing alternatives. Limited awareness about the intentions and content of this relatively new intervention for MSM[54] could explain why even men in main partner relationships for longer than a year preferred other alternatives. DBS specimen self-collection kits for laboratory HIV testing have been unevenly adopted in the US. Although their acceptability and use in the context of research studies have been high[74, 75], this approach has had minimal impact on the testing behavior of high risk individuals due to concerns regarding privacy and accuracy[92].

Strengths of our study include the evaluation of attitudes towards using six different HIV testing approaches presented collectively to a group of MSM recruited through the internet in a time, cost and resource efficient manner. Considering that online negotiations of both high-risk and safe sex have become increasingly prevalent among members of this community[86-89], we believe that understanding their testing preferences is critical in advancing internet-based HIV prevention efforts. Participants could only enter the online survey by clicking on banner advertisements displayed on Facebook, and because multiple surveys could not be completed from the same browser, it is unlikely that the same individual completed the survey more than once. People tend

to be more open and honest while reporting sensitive risk behavior information using computer-based technologies compared to traditional questionnaires[118], thereby improving data accuracy and reducing the possibility of social desirability bias[119].

Limitations of our study include not being able to generalize to all MSM users of Facebook, users of other online social networks, or MSM in the general population. Because our banner advertisements were displayed only to men who had reported being interested in men in their Facebook profile, MSM who did not disclose their interest in men in their online profile were systematically underrepresented. One limitation of collecting data online is the inability to verify participants' self-reported demographic characteristics. Non-Hispanic black men comprised a smaller proportion of our sample relative to the general US population prevalence, an unfortunate reality that has plagued online research studies[120]. Reduced access to and use of both basic and high-speed internet among black Americans compared to white or Hispanic individuals may explain this disparity[121]. Because questions in our survey involving a 6-month recall period were answered based on memory, our results could be subjected to recall bias. Although we collected data on participants' geographic region of residence, the lack of information regarding whether they lived in urban versus rural areas limited our ability to explore preference patterns within these strata. Additionally, usage intentions do not always translate into actions[92], and the extent to which newer modalities will be adopted by MSM in research as well as real world settings is yet to be determined.

Despite these concerns, we believe that our results have important implications for future HIV prevention research. In this time of great challenge and opportunity, we envision an approach of combination testing packages to enable individuals form personalized HIV

testing strategies. The fact that MSM belonging to all demographic and behavioral risk strata in our study were willing to use most testing approaches is encouraging. Moreover, their order of intended usage preferences suggest that newer options such as rapid home self-testing could be incorporated as key components of comprehensive interventions to promote testing and increase serostatus awareness. Further research, especially among black MSM, is needed to explain the relative ranks assigned to these options and explore how different modalities can be packaged together. Given the challenges with recruiting racial and ethnic MSM online[120], additional in-person surveys or qualitative work with black MSM may be required to fully capture the perspectives of this critical population. Understanding circumstances in which men would use particular approaches, and how they would combine multiple options to test in a year is imperative. To this end, we are conducting qualitative research with MSM using online focus group discussions, the results of which we hope will provide an in-depth understanding of these issues. The efficacy of each approach in increasing HIV testing frequencies should be a high priority as part of developing comprehensive prevention strategies for MSM in the US.

FIGURE 2.1: Stated likelihood of using different currently available HIV testing options if offered free of charge to 973 HIV-negative or unknown status men who have sex with men in a national online health survey, United States, 2012.

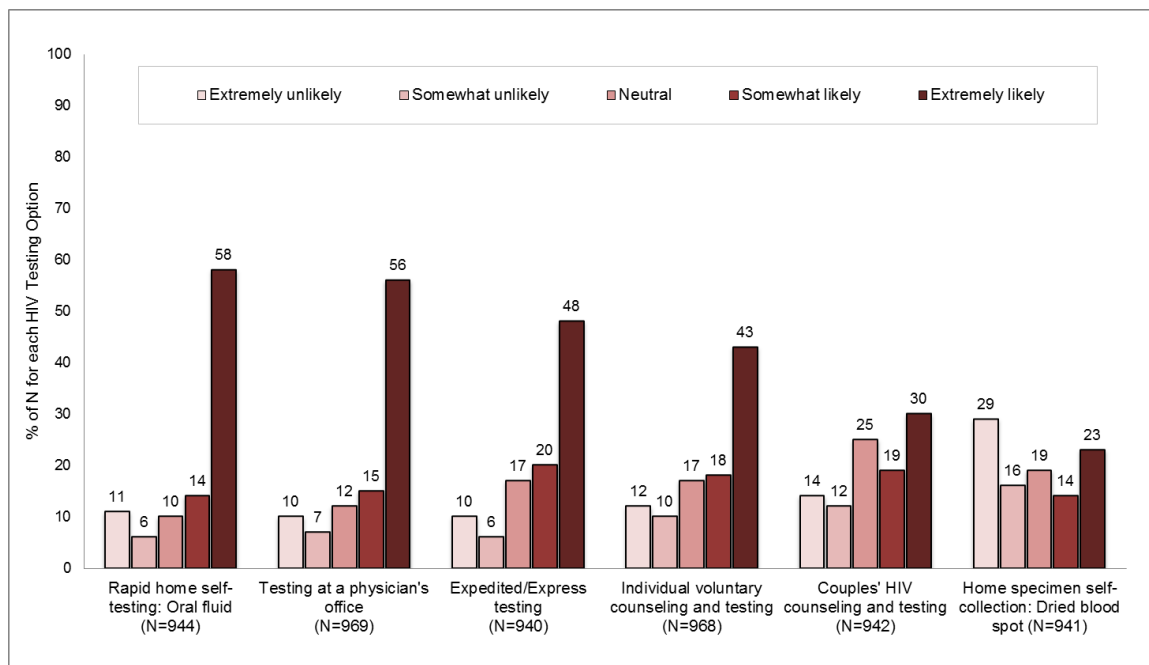


FIGURE 2.2: Modified Borda Count ranking of different currently available HIV testing options if offered free of charge to 973 HIV-negative or unknown status men who have sex with men in a national online health survey, United States, 2012.

Ranking	Overall preferences	Stratified by demographic and behavioral characteristics								
		HIV testing history		Had a main partner			Had unprotected anal intercourse with a male sex partner in the past 6 months			
		Never tested	Tested at least once	Yes, for ≥ 1 year	Yes, for < 1 year	No	Yes, with ≥ 2 men	Yes, with 1 man	No	
1	Home	Home	Physician	Home	Physician	Home	Home	Physician	Physician	
2	Physician	Physician	Home	Physician	Home	Physician	Physician	Home	Home	
3	Express	Express	Express	Express	Express	Express	Express	Express	Express	
4	VCT	DBS	VCT	VCT	VCT	VCT	VCT	VCT	VCT	
5	DBS	VCT	DBS	DBS	CHCT	DBS	DBS	CHCT	DBS	
6	CHCT	CHCT	CHCT	CHCT	DBS	CHCT	CHCT	DBS	CHCT	

Home Rapid home self-testing: Oral fluid	VCT Individual voluntary counseling and testing
Physician Testing at a physician's office	DBS Home specimen self-collection: Dried blood spot
Express Expedited/Express testing	CHCT Couples' HIV counseling and testing

TABLE 2.1: Demographic and behavioral characteristics of 973 HIV-negative or unknown status MSM^a respondents to a national online health survey, United States, 2012.

Characteristic	n ^b	%
Age group (years) ^c :		
18-24	410	42
25-34	269	28
35-44	123	13
≥ 45	171	18
Race/Ethnicity:		
White, non-Hispanic	751	77
Black, non-Hispanic	14	1
Hispanic	117	12
Other ^d	91	9
Census region:		
West	279	29
Midwest	218	22
Northeast	190	20
South	285	29
Unknown	1	0
Education:		
College, Post graduate, or Professional school	384	39
Some college, Associate's degree, and/or Technical school	391	40
High school, GED ^e or less	193	20
Unknown	5	1
Employment:		
Employed full-time	515	53
Employed part-time	221	23
Unemployed	191	20
Retired	39	4
Unknown	7	1
Self-identified sexual orientation:		
Homosexual/Gay	945	97
Bisexual	19	2
Other ^f	9	1
Had a main partner ^g :		
Yes, for ≥ 1 year	366	38
Yes, for < 1 year	175	18
No	430	44
Unknown	2	0

Had UAI ^h with a male sex partner in the past 6 months:		
Yes, with ≥ 2 men	196	20
Yes, with 1 man	409	42
No	333	34
Unknown	35	4
HIV testing history:		
Never tested	160	16
Tested at least once	795	82
Unknown	18	2
HIV status (Result of most recent HIV test):		
Negative	773	79
Unknown ⁱ	200	21

^a MSM: Men who have sex with men.

^b Sample size (N) = 973.

^c Age: Mean = 31, Median = 26, Range = 18-77.

^d Includes 31 Asian/Pacific Islander, 12 Native American/Alaska Native, 36 multiracial, 9 other and 3 unknown.

^e GED: General educational development.

^f Includes 2 who indicated they were “Heterosexual/Straight”, 2 who indicated they were “Questioning/Unsure”, 4 who indicated “Other” as their response and 1 unknown.

^g Defined as “Someone you feel committed to above all others. You might call this person your boyfriend, partner, significant other, spouse, or husband”.

^h UAI: Unprotected anal intercourse. Neither the respondent nor his partner used a condom.

ⁱ Includes 160 who never tested, 9 who tested but did not receive a result, 1 who tested and received an indeterminate result, and 30 who declined to answer.

TABLE 2.2: HIV testing characteristics of 795 HIV-negative or unknown status MSM^a respondents to a national online health survey who had previously tested, United States, 2012.

Characteristic	n ^b	%
Time of most recent HIV test:		
More than 2 years ago	151	19
Between the past 1 - 2 years	187	24
Within the past 1 year		
Test routinely every year	99	12
Test routinely every 6 months	141	18
Test routinely every 3 months	71	9
Test routinely at other intervals	10	1
Do not test routinely	128	16
Unknown	8	1
Location of most recent HIV test:		
Private doctor's office (including HMO ^c) ^d	325	41
Public health clinic/Community health center/STD ^e clinic	242	30
HIV counseling and testing site	106	13
Street outreach program/Mobile unit	41	5
Home or other private location	21	3
Other ^f	60	8
Most recent HIV test type:		
Test that required drawing blood with a syringe	427	54
Finger-stick blood rapid test	171	22
Oral fluid rapid test	154	19
Unknown	43	5

^a MSM: Men who have sex with men.

^b Sample size (N) = 795.

^c HMO: Health maintenance organization.

^d Includes 12 who tested in the Emergency Room and 19 who tested as an inpatient.

^e Sexually transmitted disease.

^f Includes 8 who tested in the military, 3 who tested at a correctional facility (jail or prison), 41 other and 8 unknown.

TABLE 2.3: Stated likelihood of using different HIV testing options if offered free of charge by selected demographic and behavioral characteristics of 973 HIV-negative or unknown status MSM^a respondents to a national online health survey, United States, 2012.

Characteristic	HIV testing option					
	Testing at a physician's office	Individual voluntary counseling and testing (VCT)	Couples' HIV counseling and testing (CHCT)	Expedited / express testing	Rapid home self-testing: Oral fluid	Home specimen self-collection: DBS ^b
	Median ^c (Mean ^c)					
Age group (years) ^d :						
18-24	5 (4.2)	4 (3.8)	4 (3.5)	4 (4.0)	5 (4.0)	3 (2.8)
25-34	5 (4.0)	4 (3.8)	3 (3.4)	5 (4.1)	5 (4.1)	3 (3.1)
35-44	5 (3.8)	4 (3.6)	3 (3.3)	4 (3.8)	5 (4.0)	3 (2.8)
≥ 45	5 (3.7)	4 (3.3)	3 (3.1)	4 (3.5)	5 (3.9)	3 (2.7)
Race/Ethnicity ^e :						
White, non-Hispanic	5 (4.0)	4 (3.6)	3 (3.3)	4 (3.9)	5 (4.0)	3 (2.9)
Black, non-Hispanic	5 (4.1)	3 (3.2)	4 (3.7)	5 (4.2)	2 (2.6)	1 (2.6)
Hispanic	5 (4.0)	5 (3.9)	4 (3.6)	4 (3.9)	5 (4.1)	3 (2.7)
Other ^f	5 (3.8)	5 (4.1)	4 (3.5)	5 (4.1)	5 (3.9)	3 (2.9)
Education ^g :						
College, Post graduate, or Professional school	5 (3.8)	4 (3.6)	3 (3.2)	4 (4.0)	5 (4.0)	3 (2.9)
Some college, Associate's degree, and/or Technical school	5 (4.1)	4 (3.8)	4 (3.5)	4 (3.9)	5 (4.0)	3 (2.8)
High school, GED ^h or less	5 (4.2)	4 (3.7)	4 (3.6)	4 (3.9)	5 (4.1)	3 (3.2)
Had a main partner ⁱ :						
Yes, for ≥ 1 year	5 (3.9)	4 (3.5)	3 (3.2)	4 (3.8)	5 (3.9)	3 (2.9)

Yes, for < 1 year	5 (4.1)	4 (3.7)	4 (3.6)	4 (3.9)	5 (4.0)	3 (2.7)
No	5 (4.0)	4 (3.9)	3 (3.4)	4 (4.0)	5 (4.1)	3 (3.0)
Had UAI ^j with a male sex partner in the past 6 months:						
Yes, with ≥ 2 men	5 (4.0)	4 (3.7)	3 (3.3)	5 (4.1)	5 (4.2)	3 (3.0)
Yes, with 1 man	5 (4.0)	4 (3.6)	3 (3.4)	4 (3.8)	5 (3.9)	3 (2.8)
No	5 (3.9)	4 (3.8)	4 (3.4)	4 (3.9)	5 (4.0)	3 (2.8)
HIV testing history:						
Never tested	5 (4.1)	4 (3.5)	3 (3.2)	4 (3.9)	5 (4.2)	3 (3.2)
Tested at least once	5 (4.0)	4 (3.7)	4 (3.4)	4 (3.9)	5 (4.0)	3 (2.8)

^a MSM: Men who have sex with men.

^b DBS: Dried blood spot.

^c Five-point Likert item format: 1=Extremely unlikely, 2=Somewhat unlikely, 3=Neutral, 4=Somewhat likely, 5=Extremely likely.

^d Kruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using expedited/express testing differed by age group was significant ($P < 0.001$).

^e Kruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using VCT differed by race/ethnicity was significant ($P < 0.001$).

^f Includes 31 Asian/Pacific Islander, 12 Native American/Alaska Native, 36 multiracial, 9 other and 3 unknown.

^g Kruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using CHCT differed by education was significant ($P < 0.001$).

^h GED: General educational development.

ⁱ Defined as “Someone you feel committed to above all others. You might call this person your boyfriend, partner, significant other, spouse, or husband”.

^j UAI: Unprotected anal intercourse. Neither the respondent nor his partner used a condom.

CHAPTER 3: Willingness to potentially self-test, distribute and test with social or sexual network associates using rapid home HIV test kits and associated factors among men who have sex with men

[Chapter will be reorganized into a manuscript and necessary clearances will be obtained from the Centers for Disease Control and Prevention before peer-reviewed journal submission in accordance with their data publication guidelines and agreement.]

Authors and affiliations: Akshay Sharma^{a, b, *}, Patrick S Sullivan^{a, b}, A D McNaghten^{a, b}, Laura Gravens^b, Robin J MacGowan^c, Pollyanna R Chavez^c, Arin Freeman^c, and GROUP^c. [Final authorship order for manuscript will be decided in consultation with the Centers for Disease Control and Prevention.]

^a Department of Epidemiology, Emory University Laney Graduate School, Atlanta GA 30322, USA.

^b Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta GA 30322, USA.

^c Centers for Disease Control and Prevention, Atlanta GA 30329, USA.

***Correspondence:** ashar24@emory.edu

Running head: Rapid HIV self-testing, kit distribution and testing with others among MSM.

Acknowledgements: Data for this work were collected as part of “Evaluation of rapid HIV self-testing among MSM in high prevalence cities”, a research study funded under contract #200201141989 through the Centers for Disease Control and Prevention.

Abbreviations: AIDS, Acquired Immune Deficiency Syndrome; aOR, Adjusted Odds Ratio; BRFSS, Behavioral Risk Factor Surveillance System; CDC, Centers for Disease Control and Prevention; CI, Confidence Interval; cOR, Crude Odds Ratio; FDA, Food and Drug Administration; GED, General Educational Development; HITS, HIV Testing Survey; HIV, Human Immunodeficiency Virus; MSM, Men who have Sex with Men; NHBS, National HIV Behavioral Surveillance System; OTC, Over-The-Counter; STD, Sexually Transmitted Disease; UAI, Unprotected Anal Intercourse; US, United States.

SECTION 3.1: Abstract

Background: Rapid home testing for Human Immunodeficiency Virus (HIV) is now a reality in the United States (US). We sought to (a) determine the awareness and previous use of over-the-counter home tests among internet-using men who have sex with men (MSM), (b) identify factors associated with their reported likelihood of potentially distributing free rapid home test kits and potentially testing with their social or sexual network associates, and (c) describe barriers to self-testing, kit distribution and testing with others.

Methods: Between May and October 2014, we surveyed 840 HIV-negative or unknown status MSM online about their past experiences with and future intentions regarding various aspects of home HIV testing. Multivariable logistic regression was used to model associations with their willingness to potentially distribute and test with friends or sex partners using hypothetically offered free rapid home tests. Reasons for being unlikely to self-test, distribute kits and test with others in the future were summarized.

Results: Majority (80%) of participants knew about home HIV tests, but only 9% had used them within the past year because of concerns regarding cost and accuracy. More men indicated being likely to potentially distribute oral fluid tests compared to finger-stick blood tests (90% versus 79%). MSM self-identifying as bisexual had reduced odds of being willing to distribute oral fluid tests, and those with lower educational attainment had greater odds of being willing to distribute finger-stick blood tests. Stated likelihood of self-testing with free rapid home HIV tests was positively associated with potential kit distribution. Almost three fourths (72%) of participants reported being likely to potentially test with others, the odds being significantly higher for MSM who engaged in

unprotected anal intercourse (UAI) with a single male partner within the past year, had previously tested for HIV, and who stated being likely to distribute such kits. Possible negative perceptions and reactions among kit recipients were the most common self-perceived barriers to potential distribution, and participants unwilling to potentially test with others were concerned about loss of privacy and confidentiality.

Conclusions: MSM's high acceptability of rapid home HIV testing is encouraging, and studying the efficacy of alternative uses of this new technology should be prioritized. Engaging men who are willing to distribute kits and test with others in future prevention efforts could potentially help target hard-to-reach and hidden high risk networks of MSM.

Key words: Rapid home HIV testing; men who have sex with men; alternative testing strategies; online prevention research.

SECTION 3.2: Introduction

Human Immunodeficiency Virus (HIV) testing is an important prevention activity as knowledge of one's positive serostatus has been shown to decrease high-risk sexual behaviors[11], and subsequent initiation of antiretroviral therapy can reduce onward transmission[96]. Given the reemergence of the HIV epidemic among men who have sex with men (MSM) in high income countries including the United States (US)[122], novel approaches to increase their frequency of testing are warranted. The Centers for Disease Control and Prevention (CDC) recommends at least annual HIV screening for all sexually active MSM[17], but recently there have been calls to reexamine these guidelines and consider HIV testing at 3 to 6 month intervals regardless of their self-reported risk behaviors[18]. Despite the benefits of regular screening, 2011 estimates from the National HIV Behavioral Surveillance System (NHBS) indicate that 33% of 7,312 MSM reported not having been tested in the past year[123]. Further, more than a third (34%) of 1,556 HIV-positive MSM did not know they were infected at the time of interview[21]. Increasing the proportion of infected persons who are aware of their serostatus is a national HIV prevention priority[13].

Historically, many reasons exist for MSM's reluctance to test for HIV. One study among men recruited from gay bars and dance clubs in five cities participating in NHBS from 2004-2005 found that majority of MSM with unrecognized infections did not test during the preceding year because they were scared of learning their status, afraid of losing their job, insurance and family, or worried that others would find out their result[25]. Principal reasons to avoid testing among high-risk individuals from six US states participating in the HIV Testing Surveys (HITS-I and HITS-II) included the fear of a positive diagnosis

and the denial of risk factors[23]. Besides these individual barriers, MSM patrons of three bathhouses in Seattle, Washington had cited counseling and testing barriers before rapid point-of-care testing became widely available in the US[22]. Men were concerned about having to schedule a test, and found it difficult to tolerate feelings of anxiety experienced during the long wait for results. Participants also frequently mentioned venipuncture as a dissatisfaction with traditional HIV testing methodologies[22].

Exploring attitudes towards newer technologies such as rapid home HIV tests is an initial step in developing effective prevention strategies targeting the MSM community. In July 2012, the US Food and Drug Administration (FDA) approved the OraQuick® In-Home HIV Test, the first over-the-counter (OTC) rapid test which can detect the presence of HIV antibodies in oral fluid within 20 minutes[60]. The Sure Check® HIV 1/2 Assay has the ability to detect antibodies in finger-stick blood within just 15 minutes[93], and at the time of our study its manufacturers were seeking an FDA premarket approval for a home use version. How the widespread availability of such test kits will influence the demographics of HIV testing in the US is still unclear, but they have the potential to reduce social and logistical barriers associated with traditional counseling and testing approaches faced by MSM including stigma, homophobia, lack of access to healthcare, transportation issues, and concerns regarding privacy and confidentiality[22-29].

Long before rapid HIV self-testing became commercially available in the US, there was a high level of interest in this concept, both among the general population and in high-risk groups. Over a third (37%) of 2,964 respondents in the 1999 California Behavioral Risk Factor Surveillance System (BRFSS) survey mentioned they would consider using instant home HIV tests, their acceptability being higher among individuals who reported never

having been tested[64]. Clients of a needle exchange, a sexually transmitted disease (STD) clinic, and three sex venues for MSM in Seattle, Washington who had not previously tested were more likely to prefer home HIV self-testing compared to clinic-based or standard blood testing as they opined this would be a private, convenient and faster approach[24]. Using conjoint analysis, a method employed to measure preferences for HIV testing strategies, a study conducted from 1999-2000 at four publicly-funded testing locations in California involving a large number of gay men found that participants favored tests that were accurate/timely and private/anonymous[124].

Results of a systematic review including 21 studies from 2000-2012, two thirds of which evaluated oral fluid self-testing and a third of which evaluated blood-based self-testing, lend support to these newer screening modalities[66]. Majority (89%) of the data came from research undertaken in high-income settings, with the study samples varying from low-risk general populations to high-risk subgroups including MSM. Both supervised (i.e. self-testing and counseling aided by a healthcare professional) and unsupervised (i.e. performed by a self-tester with access to phone/internet counseling) strategies were found to be highly acceptable and preferred to facility-based testing[66]. Even in countries where rapid home HIV testing has not yet been approved, MSM have indicated favorable attitudes towards this approach. Mixed-methods research with 172 Australian MSM regarding their acceptability of rapid oral fluid tests found that 63% would be likely to self-test and 61% would test more frequently if such a kit became available for home use as they thought it would be “quick, easy, confidential”[116]. Preferences for not having to see a physician and wanting immediate results were independent predictors of potentially increased testing frequency with rapid HIV self-tests in another recent

Australian survey among 2,018 gay men[125]. Separate studies in France[126] and Spain[127] found that 3% of 2,370 and 12% of 75 MSM who knew about the unauthorized online sale of HIV self-tests declared having used them respectively.

Although US MSM have been able to legally purchase the OraQuick® In-Home HIV Test since 2012, little is known about their actual uptake of this product. Two years after the Home Access® HIV-1 Test System was approved in 1996[70], a survey among high-risk populations revealed that almost half (46%) of 1,788 respondents had never heard of home testing, and only 2% of the 963 aware persons reported actual use of that kit[92]. Unlike rapid HIV tests, here individuals first collect a few drops of blood on a specimen collection card, mail that back to the company's laboratory for testing, and call to receive their results within 7 days. Even four years after its approval, more than a quarter (27%) of 1,017 MSM reported being unaware of this option, and only 6% of the 743 who had heard about the test reported ever using it before[73]. A preference for in-person counseling and concerns about its accuracy, privacy and cost were cited as reasons for not using the home collection kit[73, 92]. Evaluating whether MSM's keen interest in rapid home HIV testing over the past decade has translated into awareness and use of the commercially available product, and identifying self-perceived barriers to this relatively new technology are important.

Besides self-testing, other potential uses of OTC rapid HIV tests are to distribute them among one's social or sexual network associates and to test with others in the same place. Peer-driven HIV prevention strategies are effective in identifying high risk persons with undiagnosed infections, as has been established by CDC's Social Networks Demonstration Project[67, 128]. Modeling experiments have indicated that using home

HIV tests as a partner screening device may be a useful risk reduction strategy for MSM who do not use condoms, the advantage increasing with higher disease prevalence[68]. However, data on MSM's intentions to distribute rapid home tests or to test with friends or sex partners using such kits are limited. One study in New York City conducted before the OraQuick® In-Home HIV Test was approved by the FDA found that 80% of 60 high-risk MSM were willing to use such tests with sex partners[129]. In another study by the same research group, 27 non-monogamous MSM who regularly engaged in unprotected anal intercourse (UAI) were given oral fluid rapid HIV tests to screen sex partners[130]. Participants proposed using them with 89% of 140 partners over a 3-month period, and approximately 82% of those 124 accepted. Ten of the 101 men who agreed to test were identified as positive; seven were potential sex partners and three were acquaintances of study participants[130].

The acceptability and effectiveness of distributing rapid home HIV tests and using them to screen with prospective partners as a harm-reduction approach needs to be investigated in larger studies of MSM. Our objectives were to (a) determine the levels of awareness and previous use of OTC home HIV tests among internet-using MSM residing in the US, (b) identify factors associated with their reported likelihood of potentially distributing free rapid home test kits and potentially testing with their social or sexual network associates in the same place, and (c) describe self-perceived barriers to self-testing, kit distribution and testing with others. Understanding these issues can directly inform researchers and public health practitioners who plan on employing such tests in future prevention efforts.

SECTION 3.3: Methods

KnowAtHome is a CDC-sponsored research project seeking to determine the potential benefits of providing free rapid home HIV test kits to MSM residing in the US[94]. Data for this study were collected in a formative phase of KnowAtHome before participants were shipped a package containing an oral fluid test (OraQuick® In-Home HIV Test), a finger-stick blood test (Sure Check® HIV 1/2 Assay) and a dried blood spot specimen collection kit for confirmatory testing.

MSM were enrolled online from May to October 2014 through targeted banner advertising on a social networking website (Facebook.com), websites featuring MSM content selected by the Google Display Network, and some dating and sex-seeking websites identified by a marketing agency (Creaxion®). Individuals who clicked through the banner advertisements were directed to the KnowAtHome website where they were presented with some study information and an online consent form. Men who consented to participate were administered a short eligibility screener to evaluate whether they met the inclusion criteria. Eligible respondents were requested to provide their contact information (email address, mobile phone number, shipping address, and preferred name) as part of an online registration process. Men who successfully registered were prompted to complete a baseline survey hosted on secure web servers. Inclusion criteria were being reportedly male, 18 years of age or older, currently residing within the US, self-reporting anal sex with at least one man in the past 12 months, never having been diagnosed with a bleeding disorder, never having been part of an HIV vaccine trial, not currently taking any antiretroviral medications to prevent HIV, and not known to be HIV-positive.

Demographic and behavioral information collected from participants included age, race/ethnicity, state of residence, education, employment, annual household income, health insurance status, whether they had visited a health care provider in the past 12 months, self-identified sexual orientation and whether they had engaged in UAI with male partners in the past 12 months. Men were asked about their HIV testing history including the location, type and result of their most recent test. Because of our focus on home HIV testing, participants were questioned in detail about their awareness and previous use of currently available OTC kits. Information about the types of tests used within the past year (OraQuick® In-Home HIV Test, Home Access® HIV-1 Test System, or some other test), how these kits were obtained, and their annual frequency of home testing was collected. Men who reported never having been tested before, and those who knew about home HIV tests but had never used them, were asked to select their reasons and main reason from a list of pre-specified options based on subject area expertise and previous literature[22, 23, 25-28, 73, 92]. MSM who reported previously using commercially available home HIV tests were asked to indicate one or more motivations for their decision.

Participants were then provided brief descriptions about two different types of rapid home HIV test kits: a 20-minute oral fluid test and a 15-minute finger-stick blood test, and asked about their likelihood of using each option being provided to them free of charge as part of KnowAtHome. Men were also asked to report their likelihood of distributing each type of kit to their friends or sex partners if these were offered at no cost in the future, and their likelihood of potentially testing with others in the same place using rapid home HIV tests. Responses were collected as an ordinal variable in a five-

point Likert item format (“Very likely”, “Somewhat likely”, “Not sure”, “Somewhat unlikely” and “Very unlikely”). Participants who indicated they were likely to potentially distribute test kits were asked about whom they would give such tests, and those who indicated they were likely to potentially test with others were asked about whom they would test with and their preferred location of testing. Men who reported being unsure or unlikely to self-test, distribute or test with friends or sex partners in the future were asked to select their reasons and main reason from lists of pre-specified options.

Statistical analyses were performed using SAS version 9.3[103]. Our analytic sample only included self-reported HIV-negative or unknown status MSM who responded to the baseline survey. Fisher’s exact tests were used to compare them with men who registered for the study online but excluded because they did not begin the baseline survey[131].

Descriptive statistics were used to tabulate the demographic, behavioral and HIV testing characteristics of all participants. Reasons for never having been tested, reasons for not using commercially available home HIV tests, and motivations for the previous use of such kits were summarized among appropriate subgroups.

Our three primary analytical outcomes for regression modeling were: (i) the stated likelihood of potentially distributing a free oral fluid rapid home HIV test, (ii) the stated likelihood of potentially distributing a free finger-stick blood rapid home HIV test, and (iii) the stated likelihood of potentially testing with friends or sex partners using free rapid home HIV tests. Due to small numbers of responses in some categories, these measures were dichotomized into “Likely” (by combining “Very likely” and “Somewhat likely”), and “Unlikely” (by combining “Not sure”, “Somewhat unlikely” and “Very unlikely”). Several explanatory variables were considered for these analyses because of

their known associations with HIV testing behaviors and the differential burden of disease among MSM. These included age[22, 23, 28, 132, 133], race/ethnicity[24, 28, 75, 92, 133], census region[2, 134], educational attainment[5, 75, 92, 110, 133], health insurance status[26, 27, 97, 135, 136], self-identified sexual orientation[29, 137], having UAI with male sex partners[40, 95, 102, 126, 138], and HIV testing history[5, 24, 28, 98, 110]. Participants' reported likelihood of self-testing with free oral fluid and finger-stick blood rapid home HIV tests were also considered while modeling their stated likelihood of potentially distributing and testing with others using such kits.

Bivariate analyses were conducted to examine unadjusted associations with each of the three outcomes, and reported as crude odds ratios (cOR) with 95% confidence intervals (CIs). Estimated logit plots were produced to determine whether age, collected as a continuous measure, should be treated as a continuous or categorical variable. Because age demonstrated a non-linear relationship with our outcomes of interest, it was treated as a categorical measure. Only those variables found to be significant in Wald Chi-square tests[139] ($P < 0.05$) in the bivariate analyses were included in the initial multivariable model, along with all two-way interactions of first order factors. Backwards elimination was used to reduce initial logistic models until no additional effects met the 0.05 significance level for removal. Retained variables were assessed by generating Condition Indices and Variance Decomposition Proportions to detect possible issues with collinearity[140]. None of the interaction terms in either the model for stated likelihood of potentially distributing an oral fluid test or the model for stated likelihood of potentially distributing a finger-stick blood test were significant, and we did not detect any problems with collinearity. However, the only significant interaction term between

health insurance status and the reported likelihood of potentially distributing an oral fluid test had to be dropped from the model for stated likelihood of potentially testing with friends or sex partners due to a collinearity issue. Final models that provided associations with each of the three outcomes while controlling for other factors, reported as adjusted odds ratios (aOR) with 95% CIs, were evaluated for goodness of fit using the Hosmer-Lemeshow test[141].

Finally, we summarized reasons for why some men reported being unsure or unlikely to self-test with free rapid home HIV tests being provided to them as part of KnowAtHome, to potentially distribute such kits to individuals within their social or sexual networks if they were offered at no cost, and to potentially test with friends or sex partners in the future. Main self-identified barriers to each of these aspects chosen by men who indicated more than one reason were also tabulated.

SECTION 3.4: Results

Overall, 40,289,337 advertising impressions resulted in 155,945 click-throughs to the KnowAtHome website over a 20-week period; 3,725 (2% of click-throughs) consented and were asked questions used to determine eligibility; 3,112 (84% of respondents who consented) completed the eligibility screener; 1,154 (37% of respondents to eligibility questions) met the inclusion criteria; 1,056 (92% of eligible respondents) successfully registered online by providing their contact information. Careful scrutiny of these registrants revealed 56 potentially fraudulent cases and 4 observations erroneously retained from pre-launch website functionality testing. Removing them yielded a sample of 996 legitimate participants; 4 of these men withdrew from the study and we elected to not use their data. Our final analytic sample included 840 (85% of 992 self-reported HIV-negative or unknown status MSM) who responded to the baseline survey (Table 3.S1). Excluded participants did not differ with respect to age, race/ethnicity, census region, HIV testing history and serostatus (Table 3.S2).

Table 3.1 summarizes the demographic and behavioral characteristics of respondents included in our analyses. Majority were young, non-Hispanic white men residing in the South with some college education or higher. More than half (56%) reported having UAI with ≥ 2 male sex partners within the past year. Almost everyone reported being likely to self-test with an oral fluid rapid home HIV test, and 91% reported being likely to self-test with a finger-stick blood rapid home HIV test being provided to them as part of KnowAtHome. A greater proportion of men indicated being likely to distribute oral fluid tests compared to finger-stick blood tests (90% versus 79%) if these were offered free of

charge in the future, and almost three fourths reported being likely to potentially test with friends or sex partners using free rapid home HIV tests (Figure 3.1).

Participants' HIV testing histories and their awareness and previous use of commercially available home tests are presented in Table 3.2. Majority (82%) of our participants reported having been tested at least once in their lifetime. Private doctors' offices, public health clinics and HIV counseling and testing sites were the most commonly cited testing locations. Although eight in ten MSM had heard about OTC home HIV tests, only one in ten reported having used them within the past year. Among the 76 men who previously used such kits, almost everyone indicated testing with the OraQuick® In-Home HIV Test, most of them being reportedly purchased from pharmacies and online vendors. Reasons for use included the greater privacy (65%) and convenience (63%) these offered compared to clinic-based testing, self-testing after having sex (24%), not wanting others to know they were testing for HIV (22%), self-testing before having sex (18%), testing together with someone before sex (7%), being asked by a sex partner to take the test (1%), and some other reason (7%). Respondents could have indicated more than one motivation for previously using a home HIV test: 55% chose multiple reasons and 45% chose a single reason (data not shown in table).

Table 3.3 summarizes the reasons and main reason for never being tested for HIV before among 153 participants. Majority (43%) were afraid of finding out that they could be infected, and more than a third (37%) perceived themselves to be at low risk for HIV infection. The reasons and main reason for not previously using home HIV tests cited by 596 MSM who were aware about such tests are presented in Table 3.4. Concerns

regarding their cost (54%) and accuracy (43%) dominated the list of self-identified barriers, and a quarter indicated not knowing where they could get such test kits.

Tables 3.5 and 3.6 each show results from the bivariate and multivariable analyses of factors associated with MSM's stated likelihood of potentially distributing free oral fluid and finger-stick blood rapid home HIV tests respectively. Age and race/ethnicity were not associated with either of these analytical outcomes. Men with lower levels of education had significantly greater odds of reportedly being likely to distribute a finger-stick blood test compared to those with higher educational attainment. Participants who self-identified as bisexual had significantly reduced odds of reportedly being likely to distribute an oral fluid test compared to homosexual/gay identifying men. Stated likelihood of self-testing with an oral fluid test and a finger-stick blood test were both independently associated with increased odds of potentially distributing an oral fluid test. Stated likelihood of self-testing with a finger-stick blood test was positively associated with potentially distributing that type of test kit.

Crude and adjusted associations from logistic regression modeling of characteristics associated with the stated likelihood of potentially testing with friends or sex partners using hypothetically offered free rapid home HIV tests are presented in Table 3.7. Similar to our analyses on MSM's intentions of potentially distributing kits, age and race/ethnicity were not related to this outcome. Engaging in UAI with a single male partner within the past year was associated with increased odds of reportedly being likely to test with others compared to not having UAI. Participants who indicated having been previously tested for HIV had greater odds of potentially testing with others compared to those who had never been tested. Stated likelihood of potentially distributing an oral fluid

test and a finger-stick blood test were both associated with increased odds of reporting being likely to test with others in the same place.

Table 3.8 summarizes the reasons and main reason cited by 79 participants who indicated being unlikely to self-test using free oral fluid or finger-stick blood rapid home HIV tests being provided to them as part of KnowAtHome. More than half were reluctant to collect a blood sample for the finger-stick test, and a quarter were afraid of finding out they could be infected. Concerns about the accuracy of such tests were cited by more than one fifth of this subgroup. Very few men reported preferences for getting tested by a trained professional or for in-person counseling as barriers to self-testing. Overall, 19% of the 79 participants had never heard about OTC home HIV tests before, 72% were aware but had never used them before and 8% reported being aware and using the OraQuick® In-Home HIV Test within the past year (data not shown in table).

The reasons and main reason among 179 men who reported being unlikely to potentially distribute free oral fluid or finger-stick blood rapid home HIV tests are presented in Table 3.9. More than one third were afraid that kit recipients would think they had HIV, and a similar proportion thought recipients would get upset or angry. Almost one third were concerned that giving a test kit to someone would affect their friendship or sexual relationship. Among 759 men who indicated positive attitudes towards potentially distributing either type of rapid home HIV test kit, 63% would most likely give them to a main sex partner (someone they feel committed to above all others), 61% to a friend (who is not a sex partner), 54% to a casual sex partner (someone they do not feel committed to above all others), 33% to a one-time sex partner (someone they would have sex with only once and probably never again), 23% to a family member (who is not a sex partner), 22%

to an acquaintance (who is not a sex partner), 11% to a stranger (who is not a sex partner), and 7% to an exchange sex partner (someone they would have sex with in exchange for money, food or drugs). Participants could have indicated more than one individual to whom they would most likely distribute such test kits: 68% chose multiple persons, 18% chose a single person and 14% did not respond (data not shown in table).

Table 3.10 summarizes the reasons and main reason cited by 232 MSM who indicated being unlikely to potentially test with friends or sex partners using hypothetically offered free rapid home HIV tests. More than half did not want others to find out their result, and almost one third did not want to find out their partner's result. Respondents were also concerned that testing with others might affect their friendship, or that the person they were testing with could disclose their result to someone else. Among 603 participants who reported positive attitudes towards testing with others using such kits, 76% would most likely test with a main sex partner, 52% with a casual sex partner, 44% with a friend, 28% with a one-time sex partner, 13% with a family member, 7% with an acquaintance, 5% with an exchange sex partner, and 4% with a stranger. Men could have indicated more than one person with whom they would most likely test: 63% chose multiple individuals and 37% chose a single individual (data not shown in table). Further, 94% indicated they would most likely test with others at their own house, 59% at their partner's house, 9% in a car, 9% at a private sex party, 6% in a bathhouse or sex club, 5% at a bar or dance club, and 2% at a circuit party. Participants could have chosen more than one venue where they would most likely test with others: 56% indicated multiple locations and 44% indicated a single location (data not shown in table).

SECTION 3.5: Discussion

Our study sought to determine the levels of awareness and previous use of commercially available home HIV tests among internet-using MSM, identify factors associated with their willingness to potentially distribute and test with social or sexual network associates using free test kits in the future, and describe self-perceived barriers to multiple aspects of rapid home HIV testing. Despite the fact that majority of our participants had heard about OTC home tests before, only a small proportion reported having used them within the past year, primarily due to monetary concerns. However, almost everyone indicated being likely to self-test with an oral fluid rapid test, and most men were willing to use a finger-stick blood rapid test being provided to them free of charge as part of KnowAtHome. Distributing rapid home HIV tests and testing with friends or sex partners using kits hypothetically offered at no cost were highly acceptable alternative uses among our study participants. These findings are important as they demonstrate the potential role such technologies could play in facilitating the expansion and penetration of HIV testing into high risk networks.

Nearly two decades after the Home Access® HIV-1 Test System[70] and exactly two years after the OraQuick® In-Home HIV Test[60] received FDA approval, one fifth of our study population had never heard about home HIV testing. Although this is an improvement over awareness levels documented in the past[73, 92], expanding efforts to increase MSM's knowledge about this screening modality might be beneficial, particularly among those who are concerned about privacy and confidentiality. Majority of our participants who reported previously using home HIV tests employed the OraQuick® In-Home HIV Test, but their small numbers suggest that the historical

enthusiasm for such a product among high-risk individuals[24, 64, 66] has not yet translated into widespread uptake in real world settings. More than half our participants cited cost as a barrier to previously self-testing and one third indicated this to be their main reason. The OraQuick® In-Home HIV Test is currently available for purchase online and in pharmacies for approximately \$40[63]. Although research involving gay men from California conducted 12 years before its OTC approval found that participants were willing to pay up to \$39 for a highly accurate, instant home test[124], the current price might be problematic, especially for men of lower socioeconomic status. Post-hoc analyses revealed that compared to study participants whose annual household income was more than \$75,000, men whose incomes were less than \$20,000 (cOR: 2.0; 95% CI: 1.2, 3.4), between \$20,000 and \$39,999 (cOR: 1.7; 95% CI: 1.0, 2.8), and between \$40,000 and 74,999 (cOR: 2.1; 95% CI: 1.3, 3.4) had significantly greater odds of citing a concern regarding cost. The reportedly high willingness to self-test with free rapid home HIV tests being provided as part of KnowAtHome suggests that the availability of such kits at a lower price point has the potential to increase testing frequencies among these subgroups of MSM. The second most commonly cited reason for not using home HIV tests in the past was a concern regarding their accuracy. Educating individuals about the high sensitivities and specificities of such tests[142, 143], and advising them regarding the window period could help mitigate this barrier.

Regardless of age or race/ethnicity, majority of our participants indicated being likely to potentially distribute oral fluid and finger-stick blood rapid home HIV tests to individuals within their social or sexual networks if these were offered free of charge. Multivariable analyses found that men who were reportedly willing to self-test with either type of test

were significantly more likely to indicate potential kit distribution. Marketing and psychology literature has documented that consumers are more inclined to recommend services and products to friends when they perceive them positively based on personal experiences[144], and that word-of-mouth communication is highly influential in shaping recipients' attitudes and behaviors towards their adoption[145]. Furthermore, strong-tie recommendation sources, such as persons within one's social or familial networks, are advantageous because they can evaluate individuals who would potentially receive the usage recommendation in addition to the product itself[146]. Our findings are encouraging especially because recent research has established that peer-driven, network-oriented strategies are highly effective in identifying people with undiagnosed HIV infections in high risk communities including MSM[67, 147]. However, differences in sexual orientation might need to be addressed in order to maximize the effectiveness of future HIV prevention efforts[148]. We found that participants self-identifying as bisexual were reportedly less likely to distribute an oral fluid test compared to those identifying as homosexual/gay. Reduced risk perceptions and lower levels of discussion around HIV and Acquired Immune Deficiency Syndrome (AIDS) among bisexual men[149] may help explain this disparity. Why men with lower educational levels in our study were reportedly more likely to distribute a finger-stick blood test compared to those with higher educational attainment needs further exploration. We believe that investigating the feasibility and extent to which providing MSM with free rapid home HIV test kits for distribution among their network associates could increase testing and serostatus awareness is a critical next step in assessing their potential public health impact. To this end, men who will be enrolled in the randomized trial phase of

KnowAtHome would be allowed to order additional tests to replenish the ones they use or give away[94].

Recent qualitative research has suggested that testing with individuals belonging to one's social or sexual networks might be an acceptable alternative use for rapid home HIV tests among MSM[129]. Given the relative simplicity of this option combined with the availability of test results within minutes, such kits could be used as a partner screening tool before engaging in high risk sexual activity[130]. In our study, 72% of 835 men indicated favorable attitudes towards testing with friends or sex partners in the same place if such tests were made available free of charge. Participants who stated they were likely to potentially distribute both oral fluid and finger-stick blood rapid HIV tests were significantly more likely to report being willing to potentially test with others. Although not surprising, this is an important result, especially when considered from the perspective of a test kit recipient. Testing with a partner could provide a sense of emotional support and increase levels of trust, as has been mentioned by MSM in interviews exploring the acceptability of couples-based prevention interventions[56]. Men who indicated having been previously tested for HIV were reportedly more likely to potentially test with others compared to those who had never been tested, possibly because they felt confident about their negative status. Interestingly, participants who engaged in UAI with a single male sex partner within the past year had almost double the odds of potentially testing with others compared to those who did not report having UAI. More than three fourths of 603 men who were willing to potentially test with others using rapid home HIV test kits indicated they would do so most likely with a main sex partner. Given the documented high burden of HIV transmission within steady same-sex

relationships[51, 52] in conjunction with recent reports that MSM who are in primary partnerships have lower levels of routine testing than the general MSM population[150], our finding suggests that rapid home HIV tests could play an important role in facilitating screening at regular intervals within this subgroup. More than half of our participants who endorsed potentially testing with others reported they would test along with casual partners. Besides stimulating greater honesty in serostatus disclosure[129], perceived risks and rapid HIV test results could lead non-monogamous MSM to alter their partner-specific sexual behavior. Eleven of the 21 men in a small study from New York City ended sexual encounters when their potential partners refused screening with an oral fluid rapid test because they assumed that those men were HIV- positive or too high risk, and none of the five participants whose partners tested positive proceeded with the sexual encounter[151].

One tenth of our participants indicated being unwilling to distribute oral fluid rapid home HIV tests, and one fifth reported being unwilling to distribute finger-stick blood rapid home HIV tests in the future. Rather than having concerns about their accuracy or conduct, these men were worried about possible negative perceptions or reactions among test kit recipients. These findings are important from a prevention research perspective, because efforts can be made to educate MSM about effectively handling a potentially confrontational situation in future studies of actual kit distribution. Providing information on how peers and sex partners could be engaged in meaningful dialogue before giving them rapid home HIV tests might help alleviate concerns regarding how such an act might affect their interpersonal relationships. With respect to testing with others, more than half of our unwilling participants did not want their partners to find out their test

result, and more than a third were concerned that their partners would tell someone their test result. Given that loss of privacy and confidentiality have consistently been cited by MSM as reasons for not testing using traditional approaches[26-29], it is not surprising that some men would hold negative opinions about testing with a non-professional.

Strengths of our study include evaluating attitudes towards alternative uses of rapid home HIV tests among a large diverse sample of internet-using MSM from across the country.

Almost three fourths of all US households now have access to a high speed internet connection[84], and this has become one of the most popular venues for MSM to access gay-oriented information[85] and meet prospective sex partners[86]. Given our greater understanding of the advantages, challenges and unique possibilities offered by this medium to the field of HIV prevention quantitative research[80, 152], we believe that exploring preferences for newer testing technologies in online samples of high risk individuals is important. Determining the acceptability of and identifying potential barriers to self-testing, rapid HIV test kit distribution and testing with persons within one's social or sexual networks is a critical first step in developing new internet-based risk reduction strategies targeting MSM. Participants in our study could only respond to the baseline survey after completing a registration process, which allowed us to detect duplicate attempts and remove possible fraudulent cases. However, ours was a convenience sample and caution must be exercised in generalizing results to all internet-using MSM or those in the general US population. Targeted banner advertising would have missed men who did not disclose their interest in men on Facebook.com, and those who did not visit websites featuring MSM content that displayed our advertisements. Because some questions involved a 12-month recall, our results could be subjected to

recall bias. The sensitive nature of our survey might have precluded an honest disclosure of high risk behaviors on account of social desirability[153], but we do not expect this was a problem. Previous studies have suggested that MSM might be more forthcoming in an online survey due to its perceived anonymity compared to a pen-and-paper questionnaire[154]. Finally, although participants' willingness to self-test with kits being provided to them as part of KnowAtHome were likely overestimates, we do not think this is a major concern with respect to their intentions of future kit distribution and testing with others using hypothetically offered free kits. Follow-up data regarding actual self-testing among our study participants are in the process of being compiled and will be reported elsewhere. We can also confirm that the next phase of KnowAtHome will evaluate alternative uses of such test kits with social or sexual network associates[94].

HIV prevention among MSM in the US requires a multifaceted approach. Exploring preferences for new testing strategies and understanding barriers to their adoption can guide the potential development of subgroup-specific testing packages. Engaging MSM willing to distribute rapid home HIV tests to their social or sexual network associates in future prevention efforts could enable us target hard-to-reach and hidden populations at high risk. Further, if the high reported likelihood of testing with others translates into actually using such kits with partners, possible misperceptions about serostatus could be eliminated, and both individuals would be able to make subsequent informed decisions regarding their sexual practices and seek preventive or therapeutic treatment if warranted. MSM's acceptability for rapid home HIV testing suggests that this option holds immense promise and research evaluating its public health impact should be prioritized.

FIGURE 3.1: Stated likelihood of potentially self-testing, distributing and testing with friends or sex partners using free rapid home test kits among 840 HIV-negative or unknown status men who have sex with men participating in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

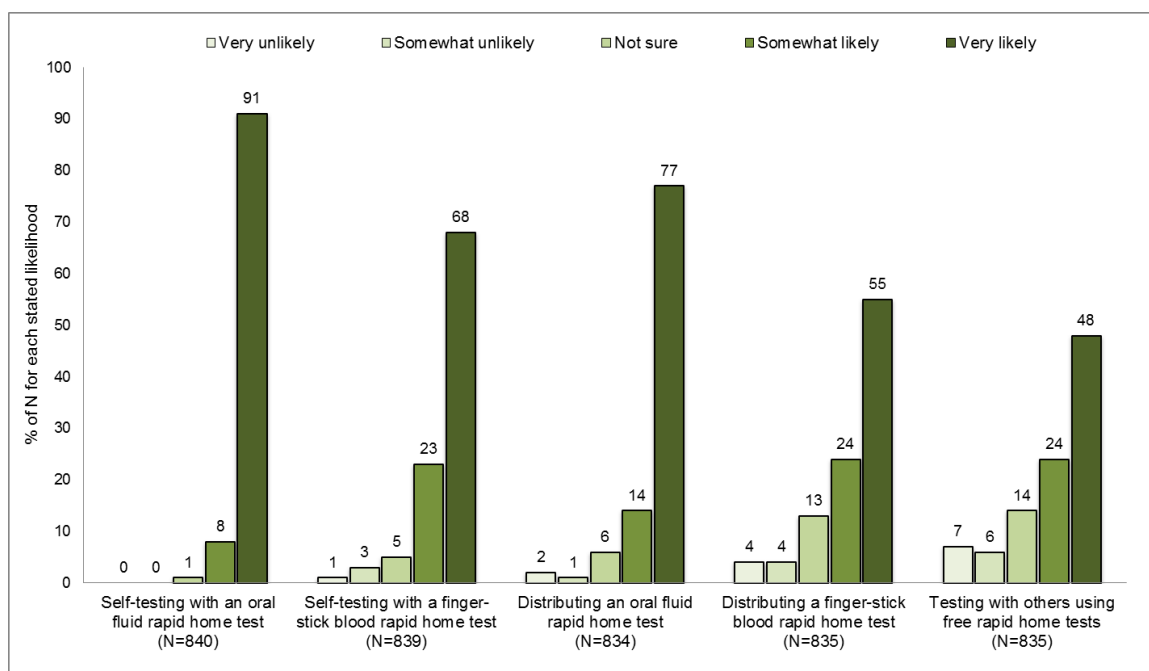


TABLE 3.1: Demographic and behavioral characteristics of 840 HIV-negative or unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	n ^b	(%)
Age group (years) ^c :		
18-24	251	(30)
25-34	290	(35)
35-44	104	(12)
≥ 45	195	(23)
Race/Ethnicity:		
White, non-Hispanic	485	(58)
Black, non-Hispanic	86	(10)
Hispanic	187	(22)
Other ^d	82	(10)
Census region ^e :		
South	311	(37)
West	209	(25)
Midwest	163	(19)
Northeast	156	(19)
Education:		
College, Post graduate, or Professional school	427	(51)
Some college, Associate's degree, or Technical degree	287	(34)
High school, GED ^f or less	122	(15)
Unknown	4	(0)
Employment:		
Employed full-time	536	(64)
Employed part-time	157	(19)
Unemployed	133	(16)
Unknown	14	(2)
Annual household income:		
\$0 to \$19,999	168	(20)
\$20,000 to \$39,999	191	(23)
\$40,000 to \$74,999	180	(21)
\$75,000 or more	169	(20)
Unknown	132	(16)
Had health insurance:		
No	158	(19)
Yes ^g	668	(80)
Unknown	14	(2)
Visited health care provider in the past 12 months:		
No	184	(22)

Yes	651	(78)
Unknown	5	(1)
Self-identified sexual orientation:		
Homosexual/Gay	709	(84)
Bisexual	116	(14)
Heterosexual/Straight	6	(1)
Unknown	9	(1)
Had UAI ^h with a male sex partner in the past 12 months:		
No	67	(8)
Yes, with 1 man	299	(36)
Yes, with 2 men	165	(20)
Yes, with ≥ 3 men	305	(36)
Unknown	4	(0)
Reported likelihood ⁱ of self-testing with a free oral fluid rapid home test:		
Unlikely	11	(1)
Likely	829	(99)
Reported likelihood ^j of self-testing with a free finger-stick blood rapid home test:		
Unlikely	74	(9)
Likely	765	(91)
Unknown	1	(0)
Reported likelihood ^k of potentially distributing a free oral fluid rapid home test:		
Unlikely	79	(9)
Likely	755	(90)
Unknown	6	(1)
Reported likelihood ^l of potentially distributing a free finger-stick blood rapid home test:		
Unlikely	175	(21)
Likely	660	(79)
Unknown	5	(1)
Reported likelihood ^m of potentially testing with others using free rapid home tests:		
Unlikely	232	(28)
Likely	603	(72)
Unknown	5	(1)

^a MSM: Men who have sex with men.

^b Sample size (N) = 840.

^c Age: Mean = 34, Median = 29, Range = 18-79.

^d Includes 51 Asian, 3 American Indian/Alaska Native, 2 Native Hawaiian/Other Pacific Islander, 19 multiracial and 7 other.

^e Does not include 1 participant from Puerto Rico.

^f GED: General educational development.

^g Includes 382 who were insured through their job, 86 who were insured through someone else's job, 115 who indicated they or someone else paid for their insurance and 85 who were insured through Medicare/Medicaid.

^h UAI: Unprotected anal intercourse. Respondent had insertive or receptive anal sex without using a

condom or not using one the whole time.

ⁱ Unlikely includes 1 “Very unlikely”, 3 “Somewhat unlikely” and 7 “Not sure”. Likely includes 765 “Very likely” and 64 “Somewhat likely”.

^j Unlikely includes 11 “Very unlikely”, 24 “Somewhat unlikely” and 39 “Not sure”. Likely includes 574 “Very likely” and 191 “Somewhat likely”.

^k Unlikely includes 16 “Very unlikely”, 10 “Somewhat unlikely” and 53 “Not sure”. Likely includes 638 “Very likely” and 117 “Somewhat likely”.

^l Unlikely includes 32 “Very unlikely”, 33 “Somewhat unlikely” and 110 “Not sure”. Likely includes 456 “Very likely” and 204 “Somewhat likely”.

^m Unlikely includes 60 “Very unlikely”, 52 “Somewhat unlikely” and 120 “Not sure”. Likely includes 400 “Very likely” and 203 “Somewhat likely”.

TABLE 3.2: Baseline testing characteristics of 840 HIV-negative or unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	n ^b	(%)
HIV testing history:		
Never tested	153	(18)
Tested at least once	687	(82)
Location of most recent HIV test:		
Private doctor's office	197	(29)
Public health clinic/Community health clinic/STD ^c clinic	183	(27)
HIV counseling and testing site	139	(20)
Home or other private location	58	(8)
Street outreach program/Mobile unit	47	(7)
Hospital (inpatient)/Emergency Room	13	(2)
Correctional facility (jail or prison)	1	(0)
Unknown	49	(7)
Most recent HIV test type:		
Test that required drawing blood with a syringe	334	(49)
Oral fluid rapid test	186	(27)
Finger-stick blood rapid test	160	(23)
Unknown	7	(1)
HIV status (Result of most recent HIV test) ^d :		
Negative	676	(98)
Unknown ^e	11	(2)
Awareness and previous use of home HIV tests:		
Never heard about before	163	(19)
Heard about but never used before	596	(71)
Heard about and used in the past 12 months	76	(9)
Type of test used:		
OraQuick® In-Home HIV Test	72	(95)
HomeAccess® HIV-1 Test System	2	(3)
Unknown	2	(3)
Obtained test from:		
Pharmacy	47	(62)
Online vendor	17	(22)
Another research study	4	(5)
Friend	2	(3)
HIV counseling and testing site	1	(1)
Unknown	5	(7)
Annual frequency of home HIV testing:		
Once	60	(79)
Twice	8	(11)
Three or more times	8	(11)
Unknown	5	(1)

^a MSM: Men who have sex with men.

^b Sample size (N) = 840.

^c STD: Sexually transmitted disease.

^d Study only included participants who did not report being HIV-positive.

^e Includes 10 who tested but did not receive a result and 1 who tested and received an indeterminate result.

TABLE 3.3: Reasons and main reason for never being tested for HIV cited by 153 HIV unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Pre-specified options for never being tested for HIV	Any reason ^b (N=153)		Main reason ^c (N=146)	
	n	(%)	n	(%)
I am afraid to find out I have HIV	65	(43)	40	(26)
I think I'm at low risk for HIV infection	56	(37)	30	(20)
I don't want my friends and family to know that I got tested	42	(28)	7	(5)
No particular reason	34	(22)	25	(16)
I don't know where to go for a test	34	(22)	9	(6)
If I test positive I will be rejected by my friends and family	32	(21)	10	(7)
I don't have enough money or insurance for a test	28	(18)	8	(5)
If I test positive I won't be able to get treatment	22	(14)	3	(2)
I don't have time	20	(13)	7	(5)
I do not want my result to be reported to the government	19	(12)	7	(5)
I don't want my sex partners to know that I got tested	7	(5)	0	(0)

^a MSM: Men who have sex with men.

^b Numbers do not add to total because respondents could select one or more reasons for never being tested for HIV.

^c Includes 94 who indicated their main reason from more than one of their reasons and 52 who indicated only one reason. Does not include 7 unknown.

TABLE 3.4: Reasons and main reason for not previously using home HIV tests cited by 596 HIV-negative and unknown status MSM^a participants who had heard about such tests before in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Pre-specified options for not previously using home HIV tests	Any reason ^b (N=596)		Main reason ^c (N=554)	
	n	(%)	n	(%)
I'm concerned about the cost of a home HIV test	324	(54)	198	(33)
I'm concerned about the accuracy of such a test	253	(43)	104	(17)
I do not know where to get a home HIV test kit	149	(25)	67	(11)
I got tested at a different location such as my doctor's office	129	(22)	85	(14)
I'm afraid of finding out that I have HIV	90	(15)	49	(8)
I'm concerned I would not be able to perform the test correctly	81	(14)	14	(2)
I would rather talk to a counselor when I get an HIV test	66	(11)	23	(4)
I'm concerned I would not be able to read the result properly	54	(9)	4	(1)
I do not want to mail my blood sample to a lab	36	(6)	5	(1)
I do not want to stick my finger to get a drop of blood	25	(4)	5	(1)

^a MSM: Men who have sex with men.

^b Numbers do not add to total because respondents could select one or more reasons for not previously using home HIV tests.

^c Includes 344 who indicated their main reason from more than one of their reasons and 210 who indicated only one reason. Does not include 42 unknown.

TABLE 3.5: Associations between demographic and behavioral characteristics and the stated likelihood^a of potentially distributing free oral fluid rapid home tests among 840 HIV-negative or unknown status MSM^b participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	Potentially distributing a free oral fluid rapid home test				Bivariate associations		Adjusted associations	
	Likely		Unlikely		cOR ^c	(95% CI ^d)	aOR ^e	(95% CI ^d)
	n	(%)	n	(%)				
Age group (years) ^f :								
18-24	228	(91)	22	(9)	1.0	-	-	-
25-34	266	(92)	23	(8)	1.1	(0.6, 2.1)	-	-
35-44	92	(90)	10	(10)	0.9	(0.4, 1.9)	-	-
≥ 45	169	(88)	24	(12)	0.7	(0.4, 1.3)	-	-
Race/Ethnicity:								
White, non-Hispanic	430	(89)	51	(11)	1.0	-	-	-
Black, non-Hispanic	79	(93)	6	(7)	1.6	(0.6, 3.8)	-	-
Hispanic	175	(94)	12	(6)	1.7	(0.9, 3.3)	-	-
Other ^g	71	(88)	10	(12)	0.8	(0.4, 1.7)	-	-
Census region ^h :								
South	282	(91)	27	(9)	1.0	-	-	-
West	195	(94)	13	(6)	1.4	(0.7, 2.9)	-	-
Midwest	137	(85)	24	(15)	0.5	(0.3, 1.0)ⁱ	-	-
Northeast	140	(90)	15	(10)	0.9	(0.5, 1.7)	-	-
Self-identified sexual orientation:								
Homosexual / Gay	649	(92)	57	(8)	1.0	-	1.0	-
Bisexual	93	(82)	20	(18)	0.4	(0.2, 0.7)	0.4	(0.2, 0.7)
Heterosexual / Straight	6	(100)	0	(0)	-	-	-	-
HIV testing history:								
Never tested	132	(86)	21	(14)	1.0	-	-	-
Tested at least once	623	(91)	58	(9)	1.7	(1.0, 2.9)^j	-	-
Reported likelihood ^k of self-testing with a free oral fluid rapid home test:								
Unlikely	4	(40)	6	(60)	1.0	-	1.0	-
Likely	751	(91)	73	(9)	15.4	(4.3, 55.9)	13.2	(3.4, 50.8)
Reported likelihood ^l of self-testing with a free finger-stick blood rapid								

home test:					
Unlikely	58 (81)	14 (19)	1.0 -	1.0 -	
Likely	696 (91)	65 (9)	2.6 (1.4, 4.9)	2.0 (1.0, 4.1)^m	

^a Likely to potentially distribute a free oral fluid rapid home test, N = 755 (includes 638 “Very likely” and 117 “Somewhat likely”). Unlikely to potentially distribute a free oral fluid rapid home test, N = 79 (includes 16 “Very unlikely”, 10 “Somewhat unlikely” and 53 “Not sure”).

^b MSM: Men who have sex with men.

^c cOR: Crude odds ratio.

^d CI: Confidence interval.

^e aOR: Adjusted odds ratio.

^f Age: Mean = 34, Median = 29, Range = 18-79.

^g Includes 51 Asian, 3 American Indian/Alaska Native, 2 Native Hawaiian/Other Pacific Islander, 19 multiracial and 7 other.

^h Does not include 1 participant from Puerto Rico.

ⁱ Result was significant. Upper limit of 95% CI rounded up from 0.983 to 1.0.

^j Result was significant. Lower limit of 95% CI rounded down from 1.003 to 1.0.

^k Unlikely includes 1 “Very unlikely”, 3 “Somewhat unlikely” and 7 “Not sure”. Likely includes 765 “Very likely” and 64 “Somewhat likely”.

^l Unlikely includes 11 “Very unlikely”, 24 “Somewhat unlikely” and 39 “Not sure”. Likely includes 574 “Very likely” and 191 “Somewhat likely”.

^m Result was significant. Lower limit of 95% CI rounded down from 1.021 to 1.0.

TABLE 3.6: Associations between demographic and behavioral characteristics and the stated likelihood^a of potentially distributing free finger-stick blood rapid home tests among 840 HIV- negative or unknown status MSM^b participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	Potentially distributing a free finger-stick blood rapid home test				Bivariate associations		Adjusted associations	
	Likely		Unlikely		cOR ^c	(95% CI ^d)	aOR ^e	(95% CI ^d)
	n	(%)	n	(%)				
Age group (years) ^f :								
18-24	202	(80)	49	(20)	1.0	-	-	-
25-34	225	(78)	63	(22)	0.9	(0.6, 1.3)	-	-
35-44	82	(80)	20	(20)	1.0	(0.6, 1.8)	-	-
≥ 45	151	(78)	43	(22)	0.9	(0.5, 1.4)	-	-
Race/Ethnicity:								
White, non-Hispanic	372	(77)	109	(23)	1.0	-	-	-
Black, non-Hispanic	67	(79)	18	(21)	1.1	(0.6, 1.9)	-	-
Hispanic	155	(83)	32	(17)	1.4	(0.9, 2.2)	-	-
Other ^g	66	(80)	16	(20)	1.2	(0.7, 2.2)	-	-
Education:								
College, Post graduate, or Professional school	317	(75)	105	(25)	1.0	-	1.0	-
Some college, Associate's degree, or Technical degree	241	(84)	46	(16)	1.7	(1.2, 2.5)	1.7	(1.1, 2.5)
High school, GED ^h or less	100	(82)	22	(18)	1.5	(0.9, 2.5)	1.7	(1.0, 2.9)ⁱ
Self-identified sexual orientation:								
Homosexual / Gay	568	(81)	137	(19)	1.0	-	-	-
Bisexual	80	(70)	35	(30)	0.6	(0.4, 0.9)	-	-
Heterosexual / Straight	6	(100)	0	(0)	-	-	-	-
Reported likelihood ^j of self-testing with a free oral fluid rapid home test:								
Unlikely	4	(36)	7	(64)	1.0	-	-	-
Likely	656	(80)	168	(20)	6.8	(2.0, 23.6)	-	-
Reported likelihood ^k of self-testing with a free								

finger-stick blood rapid
home test:

Unlikely	34 (47)	38 (53)	1.0 -	1.0 -
Likely	625 (82)	137 (18)	5.1 (3.1, 8.4)	5.1 (3.1, 8.5)

^a Likely to potentially distribute a free finger-stick blood rapid home test, N = 660 (includes 456 “Very likely” and 204 “Somewhat likely”). Unlikely to potentially distribute a free finger-stick blood rapid home test, N = 175 (includes 32 “Very unlikely”, 33 “Somewhat unlikely” and 110 “Not sure”).

^b MSM: Men who have sex with men.

^c cOR: Crude odds ratio.

^d CI: Confidence interval.

^e aOR: Adjusted odds ratio.

^f Age: Mean = 34, Median = 29, Range = 18-79.

^g Includes 51 Asian, 3 American Indian/Alaska Native, 2 Native Hawaiian/Other Pacific Islander, 19 multiracial and 7 other.

^h GED: General educational development.

ⁱ Result was significant. Lower limit of 95% CI rounded down from 1.006 to 1.0.

^j Unlikely includes 1 “Very unlikely”, 3 “Somewhat unlikely” and 7 “Not sure”. Likely includes 765 “Very likely” and 64 “Somewhat likely”.

^k Unlikely includes 11 “Very unlikely”, 24 “Somewhat unlikely” and 39 “Not sure”. Likely includes 574 “Very likely” and 191 “Somewhat likely”.

TABLE 3.7: Associations between demographic and behavioral characteristics and the stated likelihood^a of potentially testing with friends or sex partners using free rapid home tests among 840 HIV-negative or unknown status MSM^b participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	Potentially testing with others using free rapid home tests		Bivariate associations		Adjusted associations	
	Likely	Unlikely	cOR ^c	(95% CI ^d)	aOR ^e	(95% CI ^d)
	n (%)	n (%)				
Age group (years) ^f :						
18-24	182 (73)	68 (27)	1.0	-	-	-
25-34	210 (73)	79 (27)	1.0	(0.7, 1.5)	-	-
35-44	76 (74)	27 (26)	1.1	(0.6, 1.8)	-	-
≥ 45	135 (70)	58 (30)	0.9	(0.6, 1.3)	-	-
Race/Ethnicity:						
White, non-Hispanic	344 (72)	137 (28)	1.0	-	-	-
Black, non-Hispanic	68 (79)	18 (21)	1.5	(0.9, 2.6)	-	-
Hispanic	140 (75)	47 (25)	1.2	(0.8, 1.7)	-	-
Other ^g	51 (63)	30 (37)	0.7	(0.4, 1.1)	-	-
Had health insurance:						
No	124 (79)	33 (21)	1.0	-	-	-
Yes ^h	472 (71)	193 (29)	0.7	(0.4, 1.0)ⁱ	-	-
Had UAI ^j with a male sex partner in the past 12 months:						
No	43 (65)	23 (35)	1.0	-	1.0	-
Yes, with 1 man	227 (76)	70 (24)	1.7	(1.0, 3.1)	1.9	(1.1, 3.5)
Yes, with 2 men	125 (76)	40 (24)	1.7	(0.9, 3.1)	1.6	(0.8, 3.1)
Yes, with ≥ 3 men	206 (68)	97 (32)	1.1	(0.6, 2.0)	1.1	(0.6, 2.1)
HIV testing history:						
Never tested	98 (64)	54 (36)	1.0	-	1.0	-
Tested at least once	505 (74)	178 (26)	1.6	(1.1, 2.3)	1.6	(1.0, 2.3)^k
Reported likelihood ^l of self-testing with a free oral fluid rapid home test:						
Unlikely	4 (40)	6 (60)	1.0	-	-	-
Likely	599 (73)	226 (27)	4.0	(1.1, 14.2)	-	-
Reported likelihood ^m of						

self-testing with a free finger-stick blood rapid home test:								
Unlikely	45	(62)	28	(38)	1.0	-	-	-
Likely	557	(73)	204	(27)	1.7	(1.0, 2.8)ⁿ	-	-
Reported likelihood ^o of potentially distributing a free oral fluid rapid home test:								
Unlikely	30	(38)	48	(62)	1.0	-	1.0	-
Likely	572	(76)	182	(24)	5.0	(3.1, 8.2)	2.8	(1.5, 5.2)
Reported likelihood ^p of potentially distributing a free finger-stick blood rapid home test:								
Unlikely	90	(52)	83	(48)	1.0	-	1.0	-
Likely	511	(78)	148	(22)	3.2	(2.2, 4.5)	2.1	(1.3, 3.3)

^a Likely to potentially test with others using free rapid home tests, N = 603 (includes 400 “Very likely” and 203 “Somewhat likely”). Unlikely to potentially test with others using free rapid home tests, N = 232 (includes 60 “Very unlikely”, 52 “Somewhat unlikely” and 120 “Not sure”).

^b MSM: Men who have sex with men.

^c cOR: Crude odds ratio.

^d CI: Confidence interval.

^e aOR: Adjusted odds ratio.

^f Age: Mean = 34, Median = 29, Range = 18-79.

^g Includes 51 Asian, 3 American Indian/Alaska Native, 2 Native Hawaiian/Other Pacific Islander, 19 multiracial and 7 other.

^h Includes 382 who were insured through their job, 86 who were insured through someone else’s job, 115 who indicated they or someone else paid for their insurance and 85 who were insured through Medicare/Medicaid.

ⁱ Result was significant. Upper limit of 95% CI rounded up from 0.990 to 1.0.

^j UAI: Unprotected anal intercourse. Respondent had insertive or receptive anal sex without using a condom or not using one the whole time.

^k Result was significant. Lower limit of 95% CI rounded down from 1.048 to 1.0.

^l Unlikely includes 1 “Very unlikely”, 3 “Somewhat unlikely” and 7 “Not sure”. Likely includes 765 “Very likely” and 64 “Somewhat likely”.

^m Unlikely includes 11 “Very unlikely”, 24 “Somewhat unlikely” and 39 “Not sure”. Likely includes 574 “Very likely” and 191 “Somewhat likely”.

ⁿ Result was significant. Lower limit of 95% CI rounded down from 1.032 to 1.0.

^o Unlikely includes 16 “Very unlikely”, 10 “Somewhat unlikely” and 53 “Not sure”. Likely includes 638 “Very likely” and 117 “Somewhat likely”.

^p Unlikely includes 32 “Very unlikely”, 33 “Somewhat unlikely” and 110 “Not sure”. Likely includes 456 “Very likely” and 204 “Somewhat likely”.

TABLE 3.8: Reasons and main reason for being unlikely to self-test using free rapid home tests cited by 79 HIV-negative or unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Pre-specified options for being unlikely ^b to self-test using free rapid home HIV tests	Any reason ^c (N=79)		Main reason ^d (N=63)	
	n	(%)	n	(%)
I do not want to prick my finger to collect a blood sample	44	(56)	32	(40)
I'm afraid of finding out that I have HIV	20	(25)	11	(14)
I'm concerned about the accuracy of an oral fluid rapid HIV test	19	(24)	4	(5)
I'm concerned about the accuracy of a finger-stick blood rapid HIV test	15	(19)	3	(4)
I live with people who might see the test kits arrive by mail	12	(15)	6	(8)
I'm concerned I would not be able to perform the test correctly	7	(9)	4	(5)
I would rather be tested by someone who is trained to conduct the test	6	(8)	2	(3)
I would rather talk to a counselor when I get an HIV test	2	(3)	1	(1)
I live with people who might see me performing an HIV test	2	(3)	0	(0)
I do not want to swab my mouth to collect an oral fluid sample	1	(1)	0	(0)
I'm concerned I would not be able to read the result properly	1	(1)	0	(0)

^a MSM: Men who have sex with men.

^b Includes 1 “Very unlikely”, 3 “Somewhat unlikely” and 7 “Not sure” about self-testing with a free oral fluid rapid home test, and 11 “Very unlikely”, 24 “Somewhat unlikely” and 39 “Not sure” about self-testing with a free finger-stick blood rapid home test.

^c Numbers do not add to total because respondents could select one or more reasons for being unlikely to self-test using free rapid home tests.

^d Includes 25 who indicated their main reason from more than one of their reasons and 38 who indicated only one reason. Does not include 16 unknown.

TABLE 3.9: Reasons and main reason for being unlikely to potentially distribute free rapid home tests cited by 179 HIV-negative or unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Pre-specified options for being unlikely ^b to potentially distribute free rapid home HIV tests	Any reason ^c (N=179)		Main reason ^d (N=153)	
	n	(%)	n	(%)
I'm afraid they would think I have HIV	65	(36)	27	(15)
I think they would get upset or angry	64	(36)	35	(20)
I'm concerned this might affect our friendship	58	(32)	24	(13)
I'm concerned this might affect our sexual relationship	53	(30)	15	(8)
I'm concerned they would not be able to perform the test correctly	27	(15)	16	(9)
I would rather they get tested by someone who is trained to conduct the test	25	(14)	18	(10)
I would rather they talk to a counselor when they get an HIV test	19	(11)	10	(6)
I'm concerned about the accuracy of an oral fluid rapid HIV test	16	(9)	5	(3)
I'm concerned about the accuracy of a finger-stick blood rapid HIV test	11	(6)	3	(2)
I'm concerned they would not be able to read the result properly	5	(3)	0	(0)

^a MSM: Men who have sex with men.

^b Includes 16 “Very unlikely”, 10 “Somewhat unlikely” and 53 “Not sure” about potentially distributing a free oral fluid rapid home test, and 32 “Very unlikely”, 33 “Somewhat unlikely” and 110 “Not sure” about potentially distributing a free finger-stick blood rapid home test.

^c Numbers do not add to total because respondents could select one or more reasons for being unlikely to potentially distribute free rapid home tests.

^d Includes 94 who indicated their main reason from more than one of their reasons and 59 who indicated only one reason. Does not include 26 unknown.

TABLE 3.10: Reasons and main reason for being unlikely to potentially test with friends or sex partners using free rapid home tests cited by 232 HIV-negative or unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Pre-specified options for being unlikely ^b to potentially test with others using free rapid home HIV tests	Any reason ^c (N=232)		Main reason ^d (N=192)	
	n	(%)	n	(%)
I do not want them to find out my test result	118	(51)	77	(33)
I'm concerned this might affect our friendship	97	(42)	43	(19)
I'm concerned they would tell someone my test result	80	(35)	30	(13)
I do not want to find out their test result	68	(29)	15	(6)
I'm concerned this might affect our sexual relationship	55	(24)	8	(3)
I think they would get upset or angry	51	(22)	19	(8)
I'm concerned we might be drunk or high on drugs	4	(2)	0	(0)

^a MSM: Men who have sex with men.

^b Includes 60 "Very unlikely", 52 "Somewhat unlikely" and 120 "Not sure" about potentially testing with others using free rapid home tests.

^c Numbers do not add to total because respondents could select one or more reasons for being unlikely to potentially test with friends or sex partners using free rapid home tests.

^d Includes 132 who indicated their main reason from more than one of their reasons and 60 who indicated only one reason. Does not include 40 unknown.

TABLE 3.S1: Derivation of the analytic sample comprising 840 HIV-negative and unknown status MSM^a participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Phase	Description
	<p>Targeted banner advertisements displayed on:</p> <ul style="list-style-type: none"> - Social networking website (Facebook.com) - Websites featuring MSM content selected by the Google Display Network - Dating and sex-seeking websites identified by a marketing agency (Creaxion®)
Advertising	<p>From May 19-August 25, 2014 (15 weeks): 23,179,392 advertising impressions resulted in 115,344 click-throughs to the KnowAtHome website</p> <p>From September 22-October 27, 2014 (5 weeks): 17,109,945 advertising impressions resulted in 40,601 click-throughs to the KnowAtHome website</p>
	<p>3,725 (2% of 155,945 click-throughs) consented and were asked questions used to determine eligibility</p>
Recruitment	<p>3,112 (84% of respondents who consented) completed the eligibility screener</p> <p>1,154 (37% of respondents to eligibility questions) met the inclusion criteria</p> <p>1056 (92% of eligible respondents) successfully registered online by providing their contact information</p>
	<p>Removed 56 potentially fraudulent cases:</p> <ul style="list-style-type: none"> - 42 on account of duplicated contact information - 4 who had multiple observations^b in the eligibility screener - 10 with discrepant responses to questions repeated in the eligibility screener and baseline survey
Data Cleaning	<p>Removed 4 erroneously retained from website functionality testing before study launch</p> <p>996 legitimate participants in study population</p>
	<p>Removed 4 who withdrew from the study after completing the baseline survey</p>
Analysis	<p>Final sample included 840 (85% of 992 self-reported HIV-negative or unknown status MSM) who responded to the baseline survey and excluded 152 who did not respond</p>

^a MSM: Men who have sex with men.

^b Each participant ID had 4 observations with different responses in the eligibility screener.

TABLE 3.S2: Comparison of 840 HIV-negative and unknown status MSM^a participants included in the final analytic sample with 152 excluded participants in a national online research study on rapid HIV self-testing (KnowAtHome), United States, 2014.

Characteristic	Included (N=840)		Excluded (N=152)		P-value ^b
	n	(%)	n	(%)	
Age group (years) ^c :					0.49
18-24	251	(30)	40	(26)	
25-34	290	(35)	51	(34)	
35-44	104	(12)	17	(11)	
≥ 45	195	(23)	44	(29)	
Race/Ethnicity:					0.12
White, non-Hispanic	485	(58)	78	(51)	
Black, non-Hispanic	86	(10)	21	(14)	
Hispanic	187	(22)	43	(28)	
Other ^d	82	(10)	10	(7)	
Census region ^e :					0.31
South	311	(37)	61	(40)	
West	209	(25)	42	(28)	
Midwest	163	(19)	20	(13)	
Northeast	156	(19)	29	(19)	
HIV testing history:					0.25
Never tested	153	(18)	34	(22)	
Tested at least once	687	(82)	118	(78)	
HIV status (Result of most recent HIV test):					1.0
Negative	676	(98)	116	(98)	
Unknown	11	(2)	2	(2)	

^a MSM: Men who have sex with men.

^b Fisher's exact tests used to compare groups.

^c Age in "Included" group: Mean = 34, Median = 29, Range = 18-79. Age in "Excluded" group: Mean = 36, Median = 31, Range = 18-68.

^d "Included" group includes 51 Asian, 3 American Indian/Alaska Native, 2 Native Hawaiian/Other Pacific Islander, 19 multiracial and 7 other. "Excluded" group includes 2 Asian, 7 multiracial and 1 other.

^e "Included" group does not include 1 participant from Puerto Rico.

CHAPTER 4: Evaluating interventions to promote routine preventive screenings: A comparison of analytical outcomes

[Chapter published as a manuscript in *Contemporary Clinical Trials* 2015, 41:152-159.

Available at: [doi:10.1016/j.cct.2015.01.014](https://doi.org/10.1016/j.cct.2015.01.014). Permission for reuse of the full article in this dissertation (both print and electronic formats) obtained from Elsevier on April 3, 2015.

(License number: 3601380699356)]

Authors and affiliations: Akshay Sharma^{a,*}, Brent A Johnson^b, Patrick S Sullivan^a

^aDepartment of Epidemiology, Emory University Laney Graduate School and Rollins School of Public Health, 1518 Clifton Road NE, Atlanta GA 30322, United States.

^bDepartment of Biostatistics and Computational Biology, University of Rochester, 601 Elmwood Avenue, Rochester NY 14642, United States.

* **Corresponding author at:** Department of Epidemiology, Emory University, 1518 Clifton Road NE, Atlanta GA 30322, United States. Tel.: +1 404 857 8460. E-mail address: ashar24@emory.edu

Running head: Analytical outcomes in preventive screening trials.

Funding acknowledgement: This work was facilitated by the Center for AIDS Research at Emory University (P30 AI050409).

Abbreviations: AR1, Autoregressive first order; CDC, Centers for Disease Control and Prevention; GED, General Educational Development; GEE, Generalized Estimating Equations; HIV, Human Immunodeficiency Virus; HR, Hazard Ratio; IQR, Inter-quartile Range; MSM, Men who have Sex with Men; OR, Odds Ratio; PH, Proportional Hazards; RCT, Randomized Controlled Trial; SLR, Standard Logistic Regression; USPSTF, United States Preventive Services Task Force.

SECTION 4.1: Abstract

Background: Often in public health, we are interested in promoting routine preventive screenings (e.g., blood glucose monitoring, hypertension screening, or mammography). Evaluating novel interventions to encourage frequent screenings using randomized controlled trials can help inform evidence-based health promotion programs. When the desired behavior change is a recurrent event, specifying the most meaningful study outcomes may prove challenging.

Methods: To understand the efficiency of multiple approaches for evaluating an intervention seeking to increase regular health screenings we (a) simulated several replications of a trial with a positive intervention effect under various censoring scenarios, (b) formulated three different analytical outcome definitions (screening a certain number of times during the entire study period versus not, screening at least once within a clinically meaningful time period versus not, “hazard” or instantaneous rate of screening), and (c) compared them with regard to interpreting results and estimating power at different sample sizes.

Results: Approaches which better utilize detailed prospective data, while also accounting for within-participant correlations, are less likely to miss the actual underlying benefits conferred by a new prevention strategy compared to relying on a dichotomous measure derived from aggregating events over the study duration. Such approaches are also more powerful in realistic scenarios wherein some participants are lost to follow-up over time.

Conclusions: Researchers should carefully consider the choice of analytical outcomes and strive to employ more efficient approaches that model comprehensive event-specific

information, rather than summarizing repeated measures into less-informative dichotomous responses, while designing and conducting trials with recurrent preventive screenings.

Key words: Preventive screenings; analytical outcome choices; randomized trials with recurrent events; rapid home HIV self-testing; men who have sex with men.

SECTION 4.2: Introduction

Preventive screenings are an important component of health promotion efforts which can potentially reduce the significant economic burden of diseases[155]. Subgroup-specific recommendations have been developed to guide the timely identification and treatment for numerous morbidities. For example, the United States Preventive Services Task Force (USPSTF) recommends screening for lipid disorders in all men aged ≥ 35 years and women at an increased risk for coronary heart disease aged ≥ 45 years every five years[156], screening for type 2 diabetes in adults with hypertension or hyperlipidemia at 3-year intervals[157], and screening for Hepatitis B infection in pregnant women at the time of their first prenatal visit[158].

Given the emergence of new screening technologies, it is important to evaluate public health strategies to promote regular health exams. Parallel group randomized controlled trials (RCTs), reported according to established standards[159], are considered to be the most rigorous scientific tool for testing new interventions. Despite available guidance for variations in trial design[160, 161], intervention content[162, 163] and mode of delivery[164], limited discussion exists regarding design and methodological aspects unique to RCTs with recurrent events during follow-up[165-167]. Examples of such events include episodes of healthcare utilization, screening mammography, self-monitoring blood glucose and cholesterol levels, and screening for human immunodeficiency virus (HIV) infection.

Researchers conducting trials to promote preventive screening behaviors need to determine *a priori* what constitutes a meaningful outcome. Imagine a situation where we would like all adults get examined for hypertension annually, but only 20% actually

follow this recommendation. Suppose we are studying a new intervention aimed at increasing the frequency of screening. How should we define our outcome? One option is a dichotomous measure, such as checking for high blood pressure ≥ 5 times versus not over a 5-year period. However, this definition might misclassify meaningful behavioral changes (e.g. increasing from one to four screenings) as failures, and a promising intervention could be wrongly described as being ineffective. Further, someone could screen five times in a short period (e.g. within a year), but not again for the remaining 5 years, and still be counted as a success. Potential alternative outcomes are screening at least once within a 1-year interval or the rate of screening. Depending upon the choice of our outcome, different analytical approaches are needed to answer the primary question “Does the intervention work?” directly impacting adequate sample size estimation[168].

Despite the availability of different techniques to analyze recurrent events data[169-171], researchers often adopt naive approaches which either ignore the existence of multiple events, their timing during follow-up, or the correlation between repeated measures. For example, a recent systematic review of 83 RCTs evaluating interventions to prevent falls among the elderly indicated that more than half the studies inappropriately employed proportions/odds-ratio (OR) based approaches[172]. Another review of 105 longitudinal studies examining hospitalization data among heart failure patients found that 70% based their analyses on outcomes incorporating only the first admission, and almost one-third compared proportions of individuals experiencing one or more hospitalizations using either a chi-squared test or standard logistic regression (SLR)[173].

Given that researchers continue to summarize repeated measures into less-informative dichotomous responses, we sought to demonstrate how different choices of analytical

outcomes impact the ability to detect true underlying intervention effects. To understand the efficiency of multiple approaches for evaluating an intervention seeking to increase routine preventive screenings we (a) simulated several replications of a “successful” RCT (i.e. one with a positive intervention effect) under various censoring scenarios, (b) formulated three outcome definitions (screening a certain number of times during the entire study period versus not, screening at least once within a clinically meaningful time period versus not, “hazard” or instantaneous rate of screening) and performed corresponding analyses, and (c) compared them with regard to interpreting results and estimating power at different sample sizes. For demonstration purposes and our own scientific research interests, we are using the rationale of a randomized trial which seeks to determine the effectiveness of rapid HIV self-test kits in increasing testing among men who have sex with men (MSM) in the United States[94].

SECTION 4.3: Methods

Simulation strategy

Consider an RCT among HIV-negative or unknown status MSM prospectively followed for one year. The intervention to be evaluated is one to increase the frequency of HIV screening by distributing rapid HIV self-test kits that can be used at home. Intervention arm participants are given self-test kits and comparison arm participants are provided resources for identifying local HIV testing services. Men can report their test results online at the time of screening or during quarterly surveys. Participants are censored either because they are newly diagnosed as HIV-positive or because they are lost to follow-up.

SAS version 9.3[174] was used to simulate 360,000 iterations of such a trial under different assumptions and perform all subsequent analyses. Hypothetical participants were assigned demographic characteristics based on a previous study of behavioral risks involving voluntary HIV testing with a home specimen collection kit[76] and randomized to either the intervention or comparison arm. In that prospective study, 1% of participants had tested for HIV six times within a year, 1% had tested five times, 3% had tested four times, 8% had tested three times, 17% had tested twice, 31% had tested once and 39% had not tested even once within a year. Screening frequencies for simulated participants were generated using different assumptions for men in either trial arm. Annual HIV testing frequencies for men in the intervention arm were simulated such that the intervention was effective and that participants could screen for a maximum of six times.

Testing days were uniformly generated on the interval 1 to 365, assuming that all days of the year were equally likely to be selected, and the specific days of screening were separated to obtain HIV testing behavior within four 3-month time intervals.

First, two variations of a “successful” RCT (i.e. one with a positive intervention effect) were simulated assuming that 13% of comparison arm men screened (a) ≥ 3 times annually (as previously observed[76]) and (b) ≥ 2 times annually, with the intervention truly doubling these odds (i.e., assuming an OR of 2). Essentially, in the first parameter specification the odds of testing ≥ 3 times per year among intervention arm participants were twice the odds of testing ≥ 3 times per year among men in the comparison arm, and in the second parameter specification the odds of testing ≥ 2 times per year among intervention arm participants were twice the odds of testing ≥ 2 times per year among men in the comparison arm. For sensitivity analyses, we considered four additional parameter specifications such that 13% of the comparison arm screened either ≥ 3 or ≥ 2 times annually with an OR of 2.5, and 5% of the comparison arm screened either ≥ 3 or ≥ 2 times annually with an OR of 3.

Datasets were simulated by increasing starting sample sizes from 100 to 2000 in increments of 100 for each specification under the following retention and incidence scenarios:

Scenario A (Ideal retention) – No loss to follow-up and no incident HIV diagnoses

Scenario B (Modest retention) – 40% loss to follow-up and 5% incident HIV diagnoses

Scenario C (Poor retention) – 70% loss to follow-up and 10% incident HIV diagnoses

Censoring in Scenarios B and C was implemented by randomly designating some testers from Scenario A as incident positives, followed by selecting men from the remaining pool to be lost over time, and uniformly generating their last day of follow-up on the interval 1 to 365. Consult appendix for further details regarding the simulation strategy.

Analytical outcome definitions

Outcome definition 1 – Trial arms could naively be compared using a dichotomous outcome of screening a certain number of times during the entire study period versus not, e.g. testing ≥ 3 times versus fewer within a year, and performing SLR[140].

Outcome definition 2 – Alternatively, one could use a binary outcome of screening at least once within a clinically meaningful time period versus not, e.g. testing at least once versus not within a 3-month interval. Because screening behaviors for each individual may be related across consecutive time intervals, Generalized Estimating Equations (GEE) with an autoregressive first order (AR1) correlation structure could be employed for analyses[140].

Outcome definition 3 – Finally, one could examine whether the intervention increases the “hazard” or instantaneous rate of screening. Stratified Cox proportional hazards (PH) regression for recurrent events could be used to model survival times between screening tests. The total follow-up time approach uses “event times”, i.e. the actual time of two tests from baseline[166, 175], whereas the gap time approach uses “gap times”, i.e. the interval length between subsequent tests[166, 176]. Both models adjust for the fact that screening events may be correlated within individuals.

Power and sample size estimation

SLR, GEE for correlated data, and stratified Cox PH regression for recurrent events were used to analyze simulated datasets. For SLR, we excluded all those participants who were newly diagnosed but retained available screening data for those lost to follow-up. For GEE and stratified Cox PH regression, we used available information from all participants (including incident positives) until they were censored. Power for different approaches was estimated from 1,000 trial replications for a fixed sample size by determining the proportion for which the intervention effect magnitudes were significantly greater than the null. This process was sequentially repeated with 20 starting sample sizes under each censoring scenario for all six parameter specifications. Consult appendix for data layouts, model specifications and sample SAS code.

SECTION 4.4: Results

Preliminaries

Table 4.1 illustrates our randomization process with sample data on 1,200 simulated participants from a single trial replication. Study arms are balanced with respect to race/ethnicity, age, education, income, urban or rural residence, and self-identified sexual orientation. Annual screening characteristics under the first set of parameters in Scenario A are presented in Table 4.2. As specified, men in the intervention arm had higher testing frequencies versus those in the comparison arm, resulting in shorter inter-test intervals. Consult appendix for an illustration of screening patterns under different censoring scenarios (Figures 4.S1, 4.S2 and 4.S3).

Table 4.3 presents point estimates from different analytical approaches on 10 trial replications under the first set of parameters in Scenario A, each with 1,200 participants. By design, the SLR ORs comparing trial arms with respect to testing ≥ 3 times annually are close to 2. Although we present results from different methods in a single table for convenience, their interpretations differ and their magnitudes are not directly comparable. We also performed a subset of analyses wherein we adjusted for demographic covariates, and our findings remained unchanged (data not shown in table). Frequency distributions of estimates from 1,000 replications are presented in the appendix (Figures 4.S5 and 4.S6).

Graphical comparison of analytical outcomes

Figure 4.1 depicts two sets of power curves generated under the first and second parameter specifications in Scenario A. Each panel includes multiple analytical approaches corresponding to different outcome choices on 20,000 trial replications. Compared to other methods, SLR (using the first outcome definition of testing ≥ 3 times versus fewer within a year) was more sensitive to variations in the underlying benefit conferred by the intervention. Despite being powerful when the intervention actually increased the odds of screening ≥ 3 times annually (top panel), this approach performed poorly when the intervention increased the odds of screening ≥ 2 times annually (bottom panel). Expressed alternatively, sample size estimation in order to attain a desired level of statistical power was subject to more uncertainty when using the first outcome definition compared to both other choices. For example, depending upon the actual success of the intervention in improving screening behaviors, recruiting either 500 participants would suffice (top panel) or 1,200 would be needed (bottom panel) to detect a difference between trial arms with the conventional 80% power.

Figures 4.2 and 4.3 demonstrate that compared to SLR, models for GEE for correlated data (using the second outcome definition of testing at least once versus not within a 3-month interval) and stratified Cox PH regression for recurrent events (using the third outcome definition of the instantaneous rate of testing) performed increasingly better in detecting the intervention effect as greater proportions of participants were censored over time (Scenarios B and C respectively). Similar patterns were observed on comparing results from sensitivity analyses with four additional parameter specifications (Figures 4.S7, 4.S8, 4.S9, 4.S10, 4.S11 and 4.S12 presented in the appendix).

SECTION 4.5: Discussion

Researchers designing trials with recurrent screening events are often challenged with defining appropriate study endpoints. Our goal was to demonstrate how different analytical outcome choices can impact the ability to detect true underlying intervention effects. Therefore, we used simulations of varyingly “successful” RCTs to examine the efficiency of three outcome definition alternatives and their corresponding analyses. We conclude that approaches that model detailed prospective data while accounting for within-participant correlations are better than those which depend upon appropriately specifying a dichotomous outcome measure. In particular, when the dichotomous outcome is defined as number of tests exceeding a certain cut-point, SLR may fail in correctly assessing the effect of the intervention unless the theoretical and user-defined cut-point align closely; GEE for correlated data and survival analysis methods for recurrent events do not suffer this limitation. Besides being less likely to miss the actual benefits of a new prevention strategy, they are also more powerful in realistic scenarios encompassing attrition over time. Although we based our simulations on a trial of rapid HIV self-testing among MSM[94], our findings are more widely applicable to longitudinal studies with repeated health screenings. Depending upon the condition for which someone is getting evaluated, a positive result could lead them to either test more often or to cease screening altogether. For example, pregnant women might increase their frequency of screening for gestational diabetes if their glucose levels are not under control, but someone who receives a positive HIV test result would most likely stop screening, seek confirmatory testing and initiate treatment. This highlights the importance of employing statistical methods which utilize all available information to the fullest

extent. Next, we focus on practical issues while choosing relevant outcome measures and some methodological caveats of corresponding analytical approaches.

Suppose an intervention arm participant from our example self-tested three times on the first day of follow-up, and then did not test during the remainder of the study. Using the first outcome definition he would be considered a success at the end of a year, even though the practical value may only have been that of a single screening test. Contrast this with a comparison arm participant who tested twice annually at regular 6-month intervals (say first with his physician and later at a community-based organization).

Despite having a more reasonable inter-test interval, he would not be considered a success. Simply aggregating the number of screenings per individual at study conclusion and dichotomizing for SLR while ignoring the key aspect of when they were performed can therefore be problematic. Further, our simulations demonstrated that if the choice of cut-point is not reflective of the actual prevention implications of a novel screening approach, researchers may fail to identify a genuinely superior intervention. Depending upon the disease for which one is screening (e.g. HIV, certain cancers), analyzing data in this manner could also necessitate excluding all new diagnoses, even if they had tested a few times before identification. From our example, someone who was identified as HIV-positive before his third test should not be considered a failure with regard to screening ≥ 3 times per year because he might have continued to screen for HIV, and eventually be counted as a success had his previous test result been negative. Yet another statistical decision involves handling missing data due to incomplete follow-up (i.e., should such observations be retained or removed from the dataset?)

Choosing an analytical outcome of any screening within shorter and perhaps more clinically meaningful intervals can help resolve some of these issues. Consider someone from the intervention arm in our example who decided to test because he recently had unprotected sex. Suppose his first test result was negative but he self-tested again on the same day “just to be sure”. He later learned about the approximately 3-month long window period (interval between the time of infection and the production of measurable antibodies to HIV), and decided to test again after 3 months. Using the second outcome definition, his first two tests would be appropriately counted as a single success in one interval and his third test would be counted as a single success in the next interval.

Although GEE for correlated data allows for an assessment of the intervention effect in shorter time periods, even this approach does not take into account the actual timing of screening tests. However unlike SLR, this option is less sensitive to variations in the true underlying intervention effects, newly diagnosed individuals need not be excluded and one can better utilize information from individuals who are lost to follow-up.

Researchers who plan on collecting data about the days on which each screening test was performed should consider using the third outcome definition. Although several analytical methods exist to compare instantaneous rates between trial arms[166], we employed the total follow-up time and gap time models for stratified Cox PH regression for recurrent events. Given the level of detail in our data, we consider these choices to be more appropriate than count models. Both utilize all available information including the order and actual timing of subsequent screening tests. Unlike the previous options considered, one does not have to pre-specify a binary outcome such as testing certain number of times versus not within a particular time period. Power calculations assuming

different parameter specifications also showed how both models for this approach were least affected by the manner and extent to which the intervention truly increased preventive screenings in the study population.

Our simulations also demonstrated that methods which more fully utilized available data (stratified Cox PH regression and GEE) versus aggregating information (SLR) performed better in detecting the intervention effect as greater proportions of participants were censored over time. These results were in the expected direction because men who dropped out during later months before their third screening test were coded as failures when using the first outcome definition, and all new diagnoses had to be excluded. Although 40% (Scenario B) and 70% (Scenario C) losses to follow-up by year-end seem unusually high, more than 80% of the sample was retained in each arm at the first follow-up assessment in each scenario (Figure 4.S4 presented in the appendix). This exceeds the 70% retention threshold set by the Centers for Disease Control and Prevention (CDC) to support a designation of “best-evidence” HIV prevention interventions[177]. Loss to follow-up might not be problematic in short-term studies, but trials of a longer duration (e.g. an RCT evaluating an intervention to increase routine screenings where the recommended interval is a year) may benefit from analytical approaches that utilize maximum information.

Limitations of our study include making relatively simplistic assumptions while choosing probability distributions and parameters for the simulation process. HIV testing frequencies for comparison arm participants in three of the six parameter specifications were based on an actual behavioral study among MSM[76], but these might not be representative of current trends. However, similar results were observed on performing

sensitivity analyses with different specifications in each of the three censoring scenarios and we are confident regarding the validity of our findings. Although the assumption that nobody could test more than six times annually is arbitrary, we believe that counts higher than this would be extremely rare. Because we simulated screening patterns within a year-long trial, the median inter-test intervals of our hypothetical participants were shorter than what have been reported in a recent study following MSM over eight years[178]. We also acknowledge we have not considered all analytical options for the third outcome definition (e.g. marginal approach for stratified Cox PH regression[171], independent-increment model for the rate of recurrent events[179]). The well-established counting process approach was not employed as this assumes independence between different screening tests on the same individual[170, 176]. Negative binomial and zero-inflated models were not considered as these necessitate summarizing repeated outcomes on each participant with a single, one-dimensional measure of the total count of screening tests[180]. However, if the study design is such that only aggregated data rather than actual event times are collected, these alternatives could be used to estimate intervention effects without losing much efficiency[181].

Ultimately, the research question of interest and goals of the RCT should dictate the manner in which recurrent events data are measured and analyzed. Simple and familiar approaches are often appealing, but aggregating and dichotomizing information on repeated screening tests without regard to their timing could miss what are otherwise significant implications of a novel prevention strategy to increase routine health monitoring.

FIGURE 4.1: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario A.

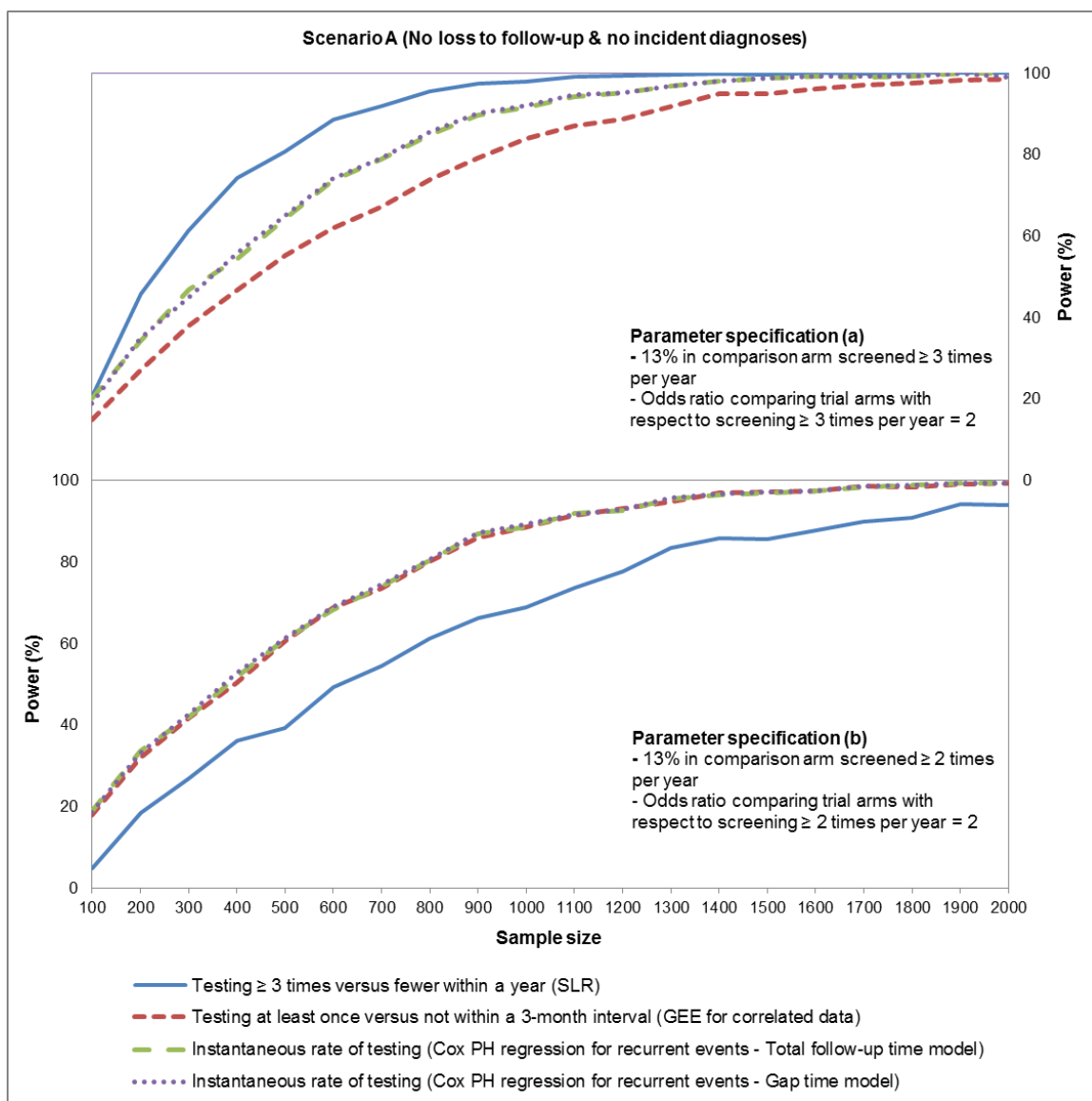


FIGURE 4.2: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario B.

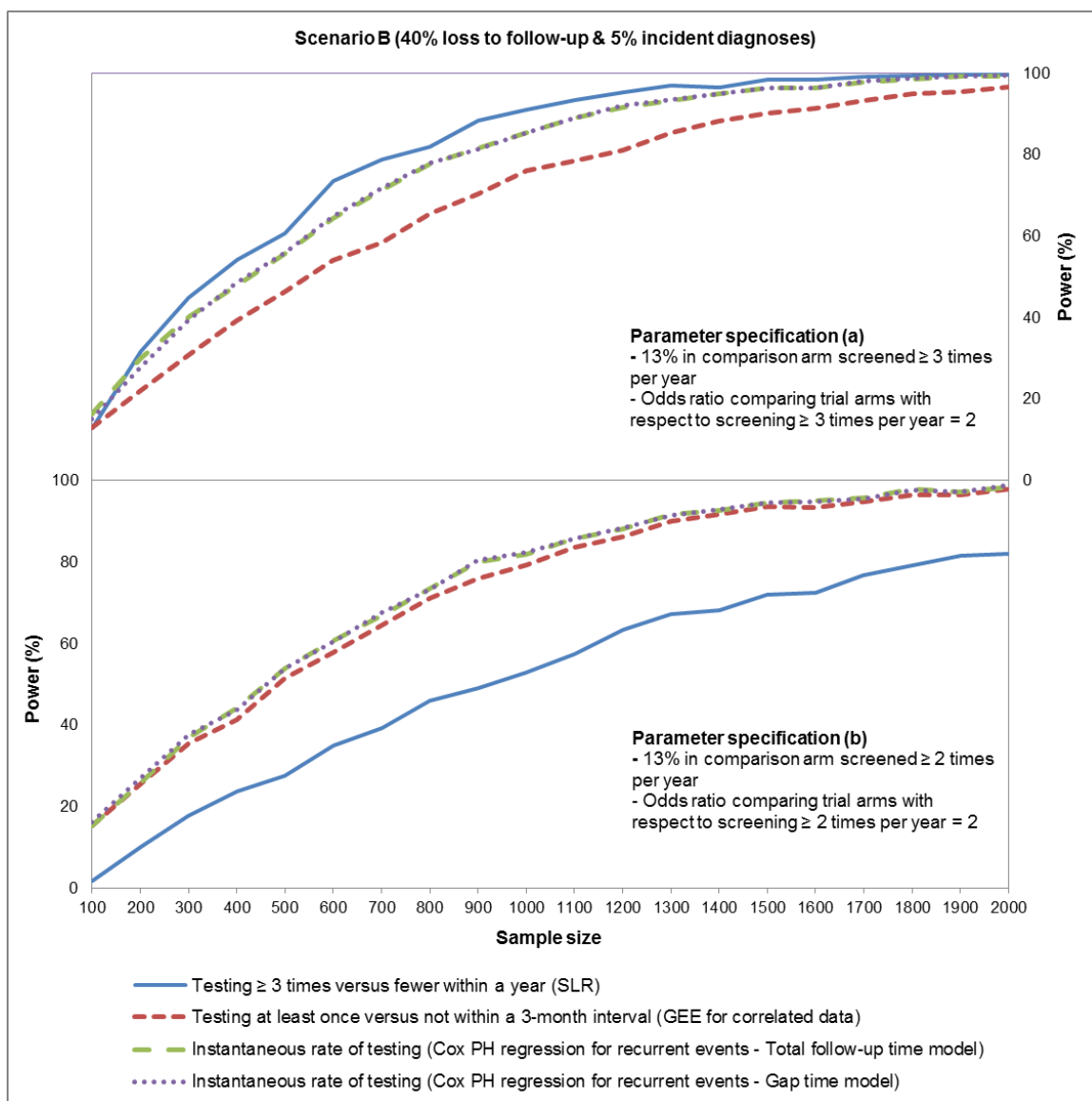


FIGURE 4.3: Power curves from 1,000 trial replications assuming the first and second parameter specifications under Scenario C.

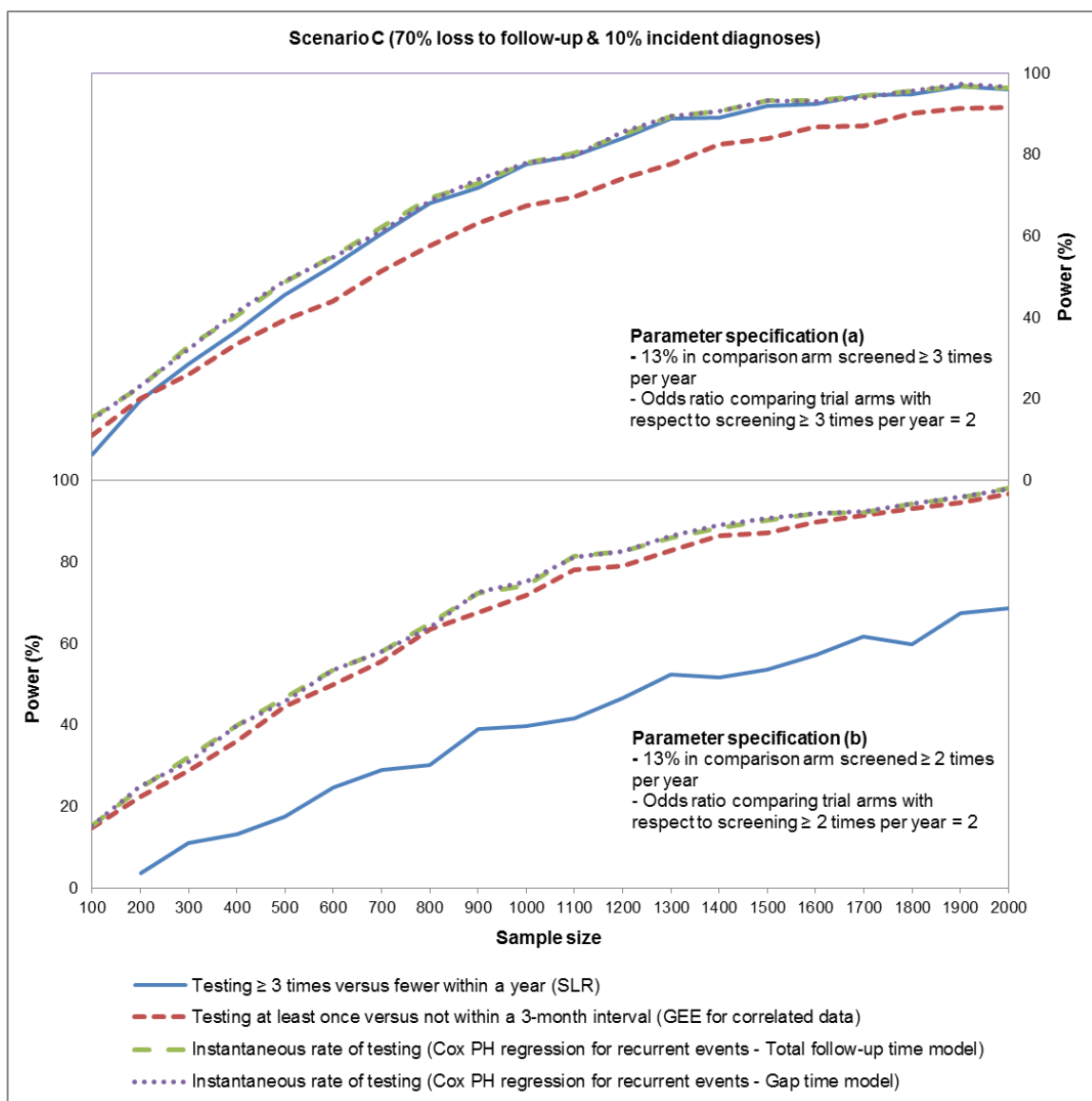


TABLE 4.1: Illustrative demographic characteristics of 1,200 simulated participants from a single replication of a randomized trial.

Characteristic	Intervention (N=570)		Comparison (N=630)	
	n	(%)	n	(%)
Race/Ethnicity:				
White, non-Hispanic	368	(65)	414	(66)
Black, non-Hispanic	88	(15)	93	(15)
Hispanic	114	(20)	123	(20)
Age group (years):				
18-24	229	(40)	246	(39)
25-34	199	(35)	219	(35)
35-44	79	(14)	106	(17)
≥ 45	63	(11)	59	(9)
Education:				
College or Post graduate	195	(34)	234	(37)
Some college or Associate's degree	271	(48)	290	(46)
High school or GED	94	(16)	91	(14)
Less than high school or unknown	10	(2)	15	(2)
Annual income:				
< \$14,999	183	(32)	200	(32)
\$15,000 – \$39,999	176	(31)	214	(34)
\$40,000 – \$74,999	119	(21)	117	(19)
> \$75,000	68	(12)	65	(10)
Unknown	24	(4)	34	(5)
Residential location:				
Urban	345	(61)	407	(65)
Rural	191	(34)	191	(30)
Unknown	34	(6)	32	(5)
Self-identified sexual orientation:				
Homosexual	485	(85)	529	(84)
Bisexual	72	(13)	84	(13)
Heterosexual or unknown	13	(2)	17	(3)

Abbreviations: GED, General educational development.

TABLE 4.2: Screening characteristics from a single replication of a randomized trial with 1,200 simulated participants assuming the first parameter specification under Scenario A.

Characteristic	Intervention		Comparison	
	(N=570)		(N=630)	
Categorical	n	(%)	n	(%)
Annual testing frequency:				
0	171	(30)	256	(41)
1	141	(25)	203	(32)
2	108	(19)	87	(14)
3	92	(16)	42	(7)
4	34	(6)	29	(5)
5	8	(1)	6	(1)
6	16	(3)	7	(1)
	(N=399 with a total of 905 screening tests)		(N=374 with a total of 691 screening tests)	
Continuous	Median	IQR	Median	IQR
Inter-test interval ^a (days)	79	(36-160)	97	(43-191)

Abbreviations: IQR, Inter-quartile range.

^a Defined as time between consecutive screenings within a year among participants who tested at least once, with time to the first test being recorded from study entry.

TABLE 4.3: Point estimates from different analytical approaches on 10 replications of a randomized trial each with 1,200 simulated participants assuming the first parameter specification under Scenario A.

Replication # (Intervention N, Comparison N)	SLR			GEE for correlated data ^a			Stratified Cox PH regression for recurrent events			
	Odds of testing ≥ 3 times within a year among:			Average 3-month interval-specific odds of testing at least once among:			Average rate of testing (per 1,000 person-days) among:		Total follow-up time model ^b	Gap time model ^c
	I	C	OR ^d	I	C	OR ^e	I	C	HR ^f	HR ^f
1 (614, 586)	0.36	0.14	2.62	0.43	0.29	1.49	4.19	2.91	1.36	1.36
2 (613, 587)	0.32	0.16	1.97	0.40	0.31	1.31	3.94	3.16	1.18	1.19
3 (582, 618)	0.31	0.20	1.55	0.37	0.35	1.06	3.90	3.55	1.08	1.09
4 (588, 612)	0.29	0.17	1.70	0.42	0.33	1.26	4.14	3.24	1.23	1.23
5 (611, 589)	0.29	0.17	1.69	0.38	0.34	1.13	3.86	3.30	1.13	1.14
6 (609, 591)	0.33	0.17	1.96	0.39	0.34	1.15	3.93	3.32	1.16	1.15
7 (597, 603)	0.41	0.19	2.19	0.48	0.33	1.44	4.59	3.34	1.29	1.30
8 (621, 579)	0.34	0.12	2.75	0.42	0.31	1.36	4.13	3.05	1.30	1.30
9 (596, 604)	0.32	0.14	2.28	0.40	0.31	1.27	3.96	3.10	1.23	1.24
10 (599, 601)	0.32	0.18	1.75	0.40	0.30	1.33	3.95	3.13	1.20	1.19

Note: Results from different approaches presented in a single table for convenience. Magnitudes of the ratio effect estimates are not directly comparable. Consult appendix for frequency distributions of point estimates from 1,000 replications (Figures 4.S5 and 4.S6).

Abbreviations: AR1, Autoregressive first order; C, Comparison; GEE, Generalized Estimating Equations; HR, Hazard Ratio; I, Intervention; OR, Odds Ratio; PH, Proportional Hazards; SLR, Standard Logistic Regression.

^a AR1 working correlation structure assumed for describing the relationship between 3-month interval-specific testing behaviors within each participant.

^b Modeling repeated screening test patterns over a year.

^c Modeling time between each of the recurrent screening test events.

^d Interpreted as the odds of testing ≥ 3 times within a year among intervention arm participants are x times the odds of testing ≥ 3 times within a year among comparison arm participants. Estimates in bold represent a statistically significant effect ($P < 0.05$).

^e Interpreted as the odds of testing at least once in a 3-month interval among intervention arm participants are x times the odds of testing at least once in a 3-month interval among comparison arm participants after accounting for correlation from repeated measures. Estimates in bold represent a statistically significant effect ($P < 0.05$).

^f Interpreted as the instantaneous rate of testing among intervention arm participants is x times the instantaneous rate of testing among comparison arm participants after accounting for recurrent screening events. Estimates in bold represent a statistically significant effect ($P < 0.05$).

SECTION 4.6: Supplementary content

[Material available as an online supplement to the manuscript in *Contemporary Clinical Trials* 2015, 41:152-159. Available at: [doi:10.1016/j.cct.2015.01.014](https://doi.org/10.1016/j.cct.2015.01.014). Permission for reuse of the full article in this dissertation (both print and electronic formats) obtained from Elsevier on April 3, 2015. (License number: 3601380699356)]

Simulation strategy details

Macro programming in SAS version 9.3[174] was used to simulate multiple iterations of a randomized controlled trial (RCT) of rapid human immunodeficiency virus (HIV) self-testing among men who have sex with men (MSM)[94] and to perform all subsequent analyses. The RAND function (which uses the Mersenne-Twister random number generator) was used to generate random numbers from different continuous and discrete assumed variable distributions. In this section, we explain how data were generated for hypothetical participants in a single replication of the trial, summarize our six parameter specifications, discuss the three censoring mechanisms, and outline how we obtained the final 360,000 datasets for power calculations.

Data generation steps for a single trial replication

First, men were randomly assigned certain baseline characteristics (race/ethnicity, age group, education, annual income, residential location, self-identified sexual orientation) using tabled distributions with probabilities based on demographic data obtained from an

actual year-long prospective study of HIV behavioral risks among internet-using MSM in the United States[76]. Men in that study were recruited through advertisements displayed on social networking websites in 2010, administered a baseline survey after which they were asked to voluntarily test using a single home dried blood spot (DBS) specimen collection kit, and then complete regular follow-up surveys at 2-month intervals.

Second, simulated participants were randomized to either the intervention or comparison arm using a Bernoulli distribution with a probability of success equal to 0.5.

Third, annual HIV testing frequencies were generated using different assumptions for men in each arm. In the actual prospective study referenced above, approximately 1% of participants had tested for HIV six times, 1% had tested five times, 3% had tested four times, 8% had tested three times, 17% had tested twice, 31% had tested once and 39% had not tested even once within a year. Data on annual testing frequencies for comparison arm participants were simulated assuming they had either similar or varying testing patterns as men in that study. Intervention arm participant data were simulated by specifying a positive intervention effect. To implement this, we assumed that their odds of testing for HIV certain number of times within the past year were greater than those for men in the comparison arm, i.e. the odds ratios (ORs) comparing trial arms were greater than the null value of 1. Participants in either arm could presumably test for a maximum of six times during the year.

Fourth, the days on which men tested were randomly generated from a uniform distribution on the interval 1 to 365, assuming that all days of the year were equally likely to be selected.

Finally, the days of testing were separated to obtain HIV testing behavior within each of the four 3-month time intervals.

Parameter specifications

Overall, six variations of a successful intervention were simulated assuming the following parameter specifications:

- (a) 13% of comparison arm tested ≥ 3 times annually with an OR comparing trial arms of 2
- (b) 13% of comparison arm tested ≥ 2 times annually with an OR comparing trial arms of 2
- (c) 13% of comparison arm tested ≥ 3 times annually with an OR comparing trial arms of 2.5
- (d) 13% of comparison arm tested ≥ 2 times annually with an OR comparing trial arms of 2.5
- (e) 5% of comparison arm tested ≥ 3 times annually with an OR comparing trial arms of 3
- (f) 5% of comparison arm tested ≥ 2 times annually with an OR comparing trial arms of 3

Censoring scenarios

Participants could be censored either because they were lost to follow-up during the year-long RCT or because they received a new HIV diagnosis. Data were generated under the following three retention and incidence scenarios:

Scenario A (Ideal retention) – No loss to follow-up and no incident HIV diagnoses by the end of the year-long trial

Scenario B (Modest retention) – 40% loss to follow-up and 5% incident HIV diagnoses by the end of the year-long trial

Scenario C (Poor retention) – 70% loss to follow-up and 10% incident HIV diagnoses by the end of the year-long trial

Censoring in Scenarios B and C was implemented by randomly flagging some testers from Scenario A as incident positives, followed by selecting men from the remaining pool to be lost over time. The days of censoring for both these sets of individuals were randomly generated from a uniform distribution on the interval 1 to 365.

Newly diagnosed positives were censored at the times of their last test in Scenario A which preceded the simulated censoring times in Scenarios B and C. For example, if Participant Y was designated to be censored on day 250 in Scenario B, and the time of his last HIV test (earlier than day 250) was day 200 in Scenario A, we assumed this was his first positive result and censored him on day 200. However, in the few cases where simulated censoring times in Scenarios B and C were earlier than the times of individuals' first HIV tests in Scenario A, they were censored at these earlier time points. For example, if Participant Z was designated to be censored on day 75 in Scenario C, and the time of his first HIV test was day 100 in Scenario A, we assumed he was lost to follow-up before he would have received a positive result and censored him on day 75.

Figures 4.S1, 4.S2 and 4.S3 depict screening patterns and inter-test intervals for an illustrative sample of 20 simulated MSM from a single trial replication assuming the first

set of parameters under each of the three censoring scenarios respectively. Retention at the 3, 6, 9 and 12-month time points for 1,200 simulated participants by study arm from a single trial replication under the different scenarios is presented in Figure 4.S4.

Final datasets for analyses

360,000 datasets were generated by sequentially replicating “successful” trials of increasing starting sample sizes for each of the parameter specifications under different censoring scenarios (1,000 trial replications * 20 sample sizes * 6 parameter specifications * 3 censoring scenarios). Each of these were analyzed using standard logistic regression (SLR) for the first outcome definition (testing ≥ 3 times versus fewer within a year), generalized estimating equations (GEE) for correlated data for the second outcome definition (testing at least once versus not within a 3-month interval), and the total follow-up time and gap time models for stratified Cox proportional hazards (PH) regression for recurrent events for the third outcome definition (instantaneous rate of testing). Further details regarding each of these approaches are presented in the following sections. Figures 4.S5 and 4.S6 depict the frequency distributions of point estimates obtained from utilizing these outcome definitions and performing corresponding analyses on 1,000 trial replications each with 1,200 simulated participants assuming the first parameter specification under Scenario A.

Figures 4.S7, 4.S8 and 4.S9 depict two sets of power curves assuming the third and fourth parameter specifications under the three censoring scenarios respectively. Figures 4.S10, 4.S11 and 4.S12 depict two sets of power curves assuming the fifth and sixth parameter specifications under the three censoring scenarios respectively. Each panel was generated from 1,000 trial replications each of 20 different starting sample sizes.

Standard Logistic Regression (SLR) for the first analytical outcome definition

SLR is a popular statistical modeling approach from the class of generalized linear models (GLM) that can be used to describe the relationship of several independent variables to a dichotomous dependent variable[140]. Because the standard logistic model assumes that observations on all participants are independent, data on multiple events of interest during follow-up need to be aggregated at the end of the study and dichotomized at some meaningful cut-point. The logistic function on which the model is based is designed to describe a probability, and provides estimates ranging between 0 and 1. In the context of this study, such a probability gives the cumulative incidence of HIV testing among trial participants.

SLR was used to analyze simulated data using the first outcome definition wherein we compared trial arms using a dichotomous outcome of screening certain number of times during the entire study period versus not, e.g. testing ≥ 3 times versus fewer within a year.

Basic data layout for analysis

Consider six hypothetical study participants F, G, H, P, Q and R whose screening patterns are depicted in Figure 4.S3.

Participants F, G and H were randomized to the intervention arm: F and G tested ≥ 3 times within the past year; H tested ≤ 2 times within the past year. Because F was newly diagnosed as HIV-positive during the trial, his information would be excluded altogether and he would not be represented as a row of data. G and H would each be represented as a single row of data.

Participants P, Q and R were randomized to the comparison arm: Q tested ≥ 3 times within the past year; P and R tested ≤ 2 times within the past year. P, Q and R would each be represented as a single row of data.

TABLE 4.S1: Data layout example for the first analytical outcome definition

Participant identifier (ID)	Study arm (E) Coded 1 for “Intervention” and 0 for “Comparison”	Tested ≥ 3 times within a year (Y) Coded 1 for “Yes” and 0 for “No”
G	1	1
H	1	0
P	0	0
Q	0	1
R	0	0

Model specification

$$\text{logit } P(Y = 1 | E) = \beta_0 + \beta_1 E$$

The function $\text{logit } P(Y = 1 | E)$ represents the log-odds of testing ≥ 3 times within the past year, given the value of the independent variable E for study arm. β_0 and β_1 represent unknown parameters that need to be estimated using maximum likelihood methods.

Sample SAS code

```
proc logistic data = slrstudy descending;
model Y = E / rl;
run;
```

Generalized Estimating Equations (GEE) for correlated data for the second analytical outcome definition

GEE is an extension of the GLM framework and is often used for data in which the responses are correlated within but independent between participants[140]. This offers an alternative approach for performing logistic regression using more detailed interval-

specific data, but the underlying assumptions, including the presence of correlation, and the manner in which parameters are estimated are different. Unlike SLR in which uses maximum likelihood estimation, GEE uses a generalization of quasi-likelihood estimation and a working correlation structure needs to be specified to describe the relationship between responses within each participant.

GEE for correlated data was used to analyze simulated data using the second outcome definition wherein we compared trial arms using a binary outcome of screening at least once within a clinically meaningful time period versus not, e.g. testing at least once versus not within a 3-month interval.

Basic data layout for analysis

Again, consider the same six hypothetical study participants F, G, H, P, Q and R whose screening patterns are depicted in Figure 4.S3.

Participants F, G and H were randomized to the intervention arm: F tested at least once in the second and third time interval; G tested at least once in the second, third and fourth time interval; H tested at least once in the third time interval. Because F was censored on being newly diagnosed as HIV-positive during the third time interval, he would be represented by three rows of data. G was followed for the entire duration of the trial and would be represented by four rows of data. H was lost to follow-up in the fourth time interval and would be represented by four rows of data.

Participants P, Q and R were randomized to the comparison arm: P did not test at least once in any time interval; Q tested at least once in the first and third time interval; R did not test at least once in the first time interval. Because P and Q were followed for the

entire duration of the trial they would each be represented by four rows of data. R was lost to follow-up in the first time interval and would be represented by one row of data.

TABLE 4.S2: Data layout example for the second analytical outcome definition

Participant identifier (ID)	Time interval (I) Coded 1 for “0 to 3 months”, 2 for “3 to 6 months”, 3 for “6 to 9 months” and 4 for “9 to 12 months”	Study arm (E) Coded 1 for “Intervention” and 0 for “Comparison”	Tested at least once in a 3-month time interval (Y) Coded 1 for “Yes” and 0 for “No”
F	1	1	0
F	2	1	1
F	3	1	1
G	1	1	0
G	2	1	1
G	3	1	1
G	4	1	1
H	1	1	0
H	2	1	0
H	3	1	1
H	4	1	0
P	1	0	0
P	2	0	0
P	3	0	0
P	4	0	0
Q	1	0	1
Q	2	0	0
Q	3	0	1
Q	4	0	0
R	1	0	0

Model specification

$$\text{logit } P(Y = 1 | E) = \beta_0 + \beta_1 E$$

The function $\text{logit } P(Y = 1 | E)$ represents the log-odds of testing at least once in a 3-month interval, given the value of the independent variable E for study arm. β_0 and β_1 represent unknown parameters that need to be estimated using a generalization of quasi-likelihood methods wherein the full likelihood is not formulated. Instead, a variance function is used to specify a relationship between the variance and mean for

each response, and a working correlation structure is used to describe how subsequent testing outcomes within each trial participant are related.

Correlation structure illustrations

Suppose a very small dataset has information on only three study participants (X, Y and Z) in which there are four responses for each individual. Because GEE models assume that the responses between participants are independent, the between-subject correlation is 0. However, the within-subject correlation between the j^{th} and k^{th} responses from each participant is denoted as ρ_{jk} . GEE models also require that each participant has a common set of correlation parameters, i.e. the correlation between testing outcomes in the second and third time intervals of one trial participant for example, is assumed to be equal to the correlation between testing outcomes in the second and third time intervals of another trial participant.

A block diagonal matrix can be used to illustrate the form of the correlation structure in which participant-specific correlation matrices are blocks along the diagonal. Here is an example of an unstructured (UN) correlation structure which has a separate correlation parameter for each pair of observations (j, k) within each participant. For example, the correlation between testing outcomes in the first and second time intervals for a participant can differ from the correlation in the second and third time intervals.

1	ρ_{12}	ρ_{13}	ρ_{14}	0	0	0	0	0	0	0	0
ρ_{12}	1	ρ_{23}	ρ_{24}	0	0	0	0	0	0	0	0
ρ_{13}	ρ_{23}	1	ρ_{34}	0	0	0	0	0	0	0	0
ρ_{14}	ρ_{24}	ρ_{34}	1	0	0	0	0	0	0	0	0
0	0	0	0	1	ρ_{12}	ρ_{13}	ρ_{14}	0	0	0	0
0	0	0	0	ρ_{12}	1	ρ_{23}	ρ_{24}	0	0	0	0

0	0	0	0	1	ρ^1	ρ^2	ρ^3	0	0	0	0
0	0	0	0	ρ^1	1	ρ^1	ρ^2	0	0	0	0
0	0	0	0	ρ^2	ρ^1	1	ρ^1	0	0	0	0
0	0	0	0	ρ^3	ρ^2	ρ^1	1	0	0	0	0
0	0	0	0	0	0	0	0	1	ρ^1	ρ^2	ρ^3
0	0	0	0	0	0	0	0	ρ^1	1	ρ^1	ρ^2
0	0	0	0	0	0	0	0	ρ^2	ρ^1	1	ρ^1
0	0	0	0	0	0	0	0	ρ^3	ρ^2	ρ^1	1
Participant L				Participant M				Participant N			

Sample SAS code

```
proc genmod data = geestudy descending;
class ID I;
model Y = E / dist = bin link = logit;
repeated subject = ID / type = ar(1) within = I corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' E 1 / exp;
run;
```

Stratified Cox proportional hazards (PH) regression for recurrent events for the third analytical outcome definition

Stratified Cox PH regression for recurrent events is a survival analysis approach in which there are separate baseline hazard functions for different strata, and the inference adjusts for the fact that events may be correlated within participants[169]. We considered two different “no-interaction” conditional models, both of which focus on survival time between events but differ in the way the risk set is determined for strata corresponding to events after the first event. Both account for the order of events because they assume that it is not possible to be “at risk” for a subsequent event without having experienced the previous one. The total follow-up time model is useful for modeling the full time course of the recurrent event process[166, 175]. Here the time interval of a subsequent event

starts at the end of the time interval for the previous event. This model uses the actual times of two events from study entry and the time until the first event affects the composition of the risk set for later events. The gap time model is useful for modeling the interval length between subsequent events rather than the full time course of the recurrent event process[166, 176]. Here each time interval starts at 0 and ends at the length of time until the next event. The time until the first event does not influence the composition of the risk set for a second or later event.

Both models were used to analyze simulated data using the third outcome definition wherein we examined whether the intervention increased the “hazard” or instantaneous rate of screening. The resulting hazard ratio (HR) estimate from both “no-interaction” models could be interpreted as the instantaneous rate of testing among intervention arm participants is x times that among comparison arm participants after accounting for recurrent screening events. “Interaction” models can help evaluate the effect of the intervention on the rate of a subsequent event among those who experienced preceding events, and yield separate HRs for each stratum number corresponding to a screening test. Therefore, if the maximum number of times someone can test is six, “interaction” models will each provide six different HRs. The estimate derived from an “interaction” total follow-up time model for the n^{th} testing event would represent the event-specific HR for the n^{th} event from study start in the intervention arm relative to the comparison arm, conditional on previously testing $n - 1$ times. The event-specific HR for the m^{th} testing event from an “interaction” gap time model would represent the instantaneous rate of the m^{th} event from the time of the previous event in the intervention arm relative to the comparison arm. We only considered the “no-interaction” approach because in our study

it was more meaningful to evaluate whether the intervention increased the overall rate of HIV testing.

Basic data layout for analysis

Once again, consider the six hypothetical study participants F, G, H, P, Q and R whose screening patterns are depicted in Figure 4.S3.

Participants F, G and H were randomized to the intervention arm: F tested three times on days 125, 130 and 183; G tested three times on days 156, 252 and 323; H tested once on day 236. Because F was censored on being newly diagnosed as HIV-positive on day 183, he would be represented by three rows of data, each corresponding to his testing events. G was followed for the entire 365-day duration of the trial and would be represented by four rows of data, the first three corresponding to each of his testing events and the fourth corresponding to his remaining time in the study. H was lost to follow-up on day 301 and would be represented by two rows of data, the first corresponding to his testing event and the second corresponding to his remaining length of time under observation.

Participants P, Q and R were randomized to the comparison arm: P did not test even once during the trial; Q tested five times on days 5, 11, 38, 190 and 256; R did not test even once during the trial. Because P did not test and was followed for the entire 365-day duration of the trial, he would be represented by a single row of data corresponding to his length of time in the study. Q was not lost to follow-up and would be represented by six rows of data, the first five corresponding to each of his testing events and the sixth corresponding to his remaining time in the study. R was lost to follow-up on day 4

without ever having tested and would be represented by a single row of data corresponding to this length of time under observation.

TABLE 4.S3: Data layout example for the third analytical outcome definition

Participant identifier (ID)	Study arm (E) Coded 1 for “Intervention” and 0 for “Comparison”	Testing event (Y) Coded 1 for “Yes” and 0 for “No”	Stratum (G)	Total follow-up time model		Gap time model	
				START1 (days)	STOP1 (days)	START2 (days)	STOP2 (days)
F	1	1	1	0	125	0	125
F	1	1	2	125	130	0	5
F	1	1	3	130	183	0	53
G	1	1	1	0	156	0	156
G	1	1	2	156	252	0	96
G	1	1	3	252	323	0	71
G	1	0	4	323	365	0	42
H	1	1	1	0	236	0	236
H	1	0	2	236	301	0	65
P	0	0	1	0	365	0	365
Q	0	1	1	0	5	0	5
Q	0	1	2	5	11	0	6
Q	0	1	3	11	38	0	27
Q	0	1	4	38	190	0	152
Q	0	1	5	190	256	0	66
Q	0	0	6	256	365	0	109
R	0	0	1	0	4	0	4

Model specifications

“No-interaction” model:

$$h_G(t, E) = h_{0G}(t) * \exp(\beta E)$$

where $G = 1, 2, 3, 4, 5, 6$ assuming someone can test for a maximum of six times per year

“Interaction” model:

$$h_G(t, E) = h_{0G}(t) * \exp(\beta_G E)$$

where $G = 1, 2, 3, 4, 5, 6$ assuming someone can test for a maximum of six times per year

Stratified Cox PH models give an expression for the hazard at time t for an individual with a given specification of the independent variable E for study arm for each stratum denoted by a subscript G . This model is semi-parametric as the baseline hazard $h_{0G}(t)$, which is allowed to be different for each stratum, is an unspecified function. The interaction model separates the data into six models, one for each of the strata.

Sample SAS code

“No-interaction” total follow-up time model:

```
proc phreg data = phstudy covs(aggregate);
model (START1, STOP1)*Y(0) = E / rl;
id ID; strata G;
run;
```

“No-interaction” gap time model:

```
proc phreg data = phstudy covs(aggregate);
model (START2, STOP2)*Y(0) = E / rl;
id ID; strata G;
run;
```

“Interaction” total follow-up time model:

```
proc phreg data = phstudy covs(aggregate);
model (START1, STOP1)*Y(0) = G1-G6 / rl;
G1 = E*(G=1); G2 = E*(G=2); G3 = E*(G=3); G4 = E*(G=4); G5 = E*(G=5);
G6 = E*(G=6);
id ID; strata G;
run;
```

“Interaction” gap time model:

```
proc phreg data = phstudy covs(aggregate);
model (START2, STOP2)*Y(0) = G1-G6 / rl;
G1 = E*(G=1); G2 = E*(G=2); G3 = E*(G=3); G4 = E*(G=4); G5 = E*(G=5);
G6 = E*(G=6);
id ID; strata G;
run;
```

FIGURE 4.S1: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario A.

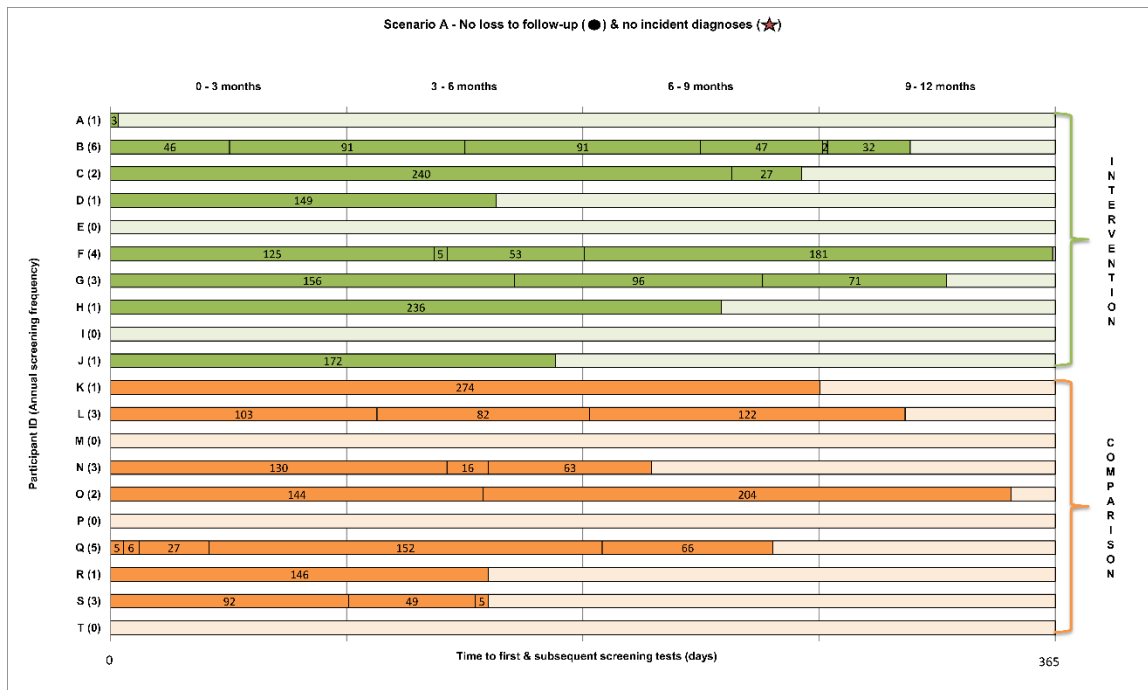


FIGURE 4.S2: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario B.

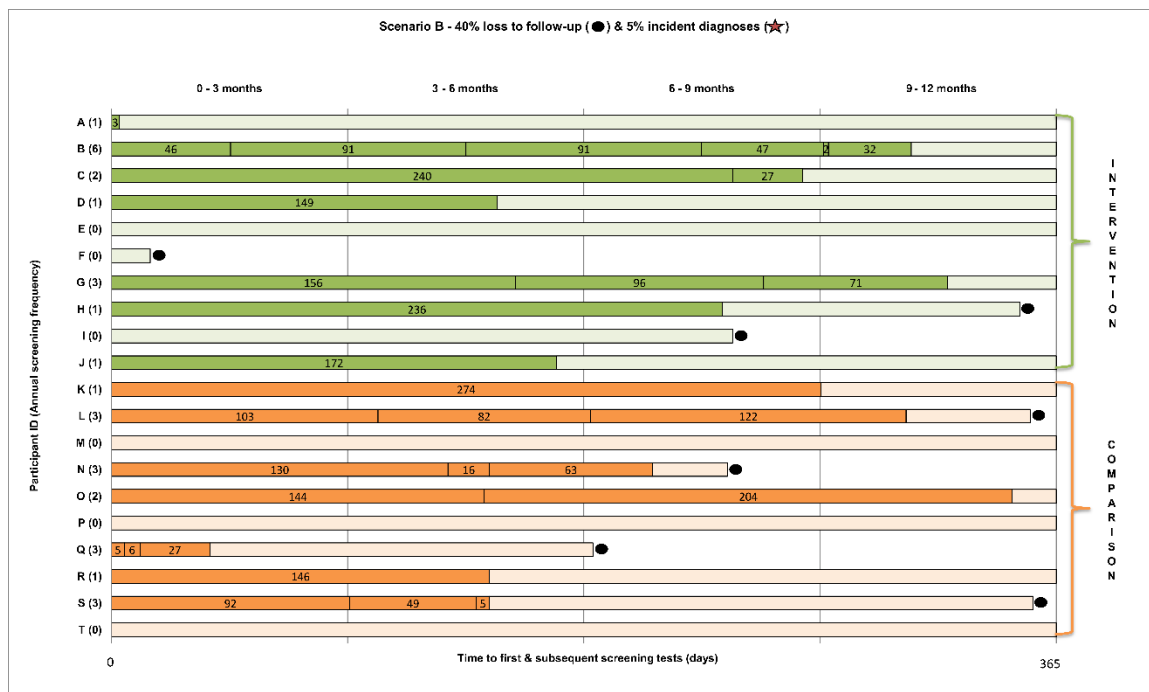


FIGURE 4.S3: Screening patterns and inter-test intervals for 20 simulated participants by study arm from a single trial replication assuming the first parameter specification under Scenario C.

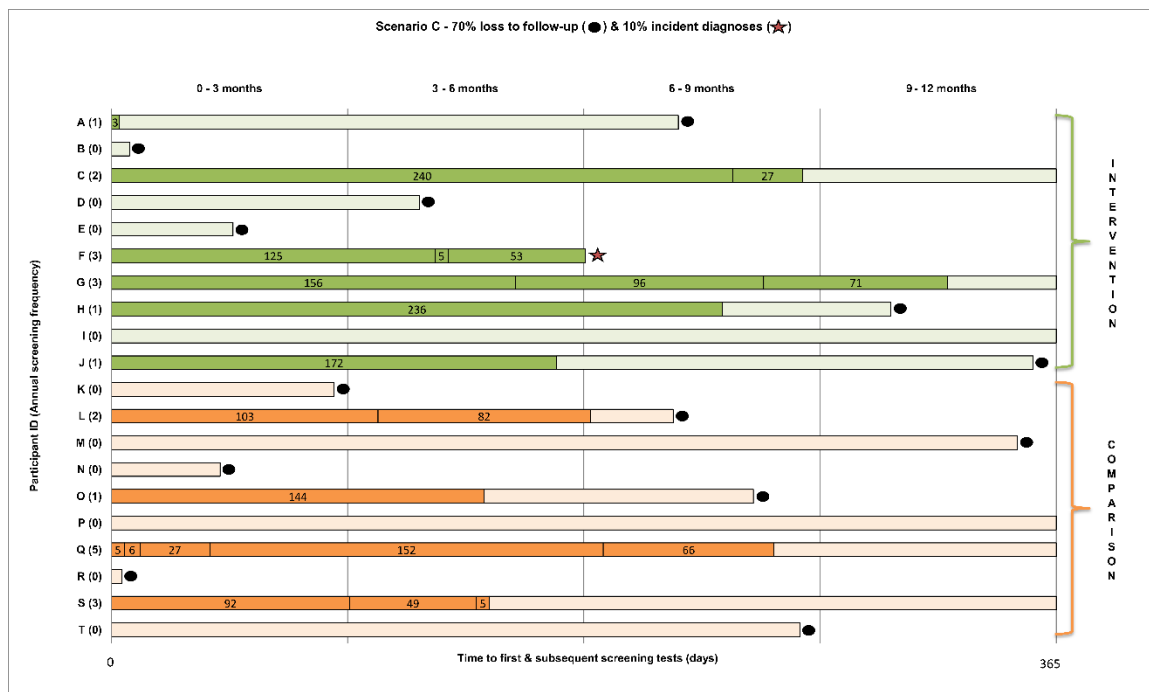


FIGURE 4.S4: Retention at 3, 6, 9 and 12 months for 1,200 simulated participants by study arm from a single trial replication under three censoring scenarios.

**Retention in Scenarios A (No loss to follow-up & no incident diagnoses),
B (40% loss to follow-up & 5% incident diagnoses) & C (70% loss to follow-up & 10% incident diagnoses)**

	0 - 3 months	3 - 6 months	6 - 9 months	9 - 12 months	
Scenario A	570	570	570	570	INTERVENTION
	(100%)	(100%)	(100%)	(100%)	
	522	468	398	334	
Scenario B	(92%)	(82%)	(70%)	(59%)	
	480	349	261	150	
Scenario C	(81%)	(61%)	(46%)	(26%)	
	630	630	630	630	COMPARISON
Scenario A	(100%)	(100%)	(100%)	(100%)	
	562	471	422	354	
Scenario B	(89%)	(75%)	(67%)	(56%)	
	514	399	266	169	
Scenario C	(82%)	(63%)	(45%)	(27%)	

N & (% of participants) retained at each follow-up time point

FIGURE 4.S5: Frequency distribution of point estimates generated from SLR and GEE for correlated data on 1,000 trial replications each with 1,200 simulated participants assuming the first parameter specification under Scenario A.

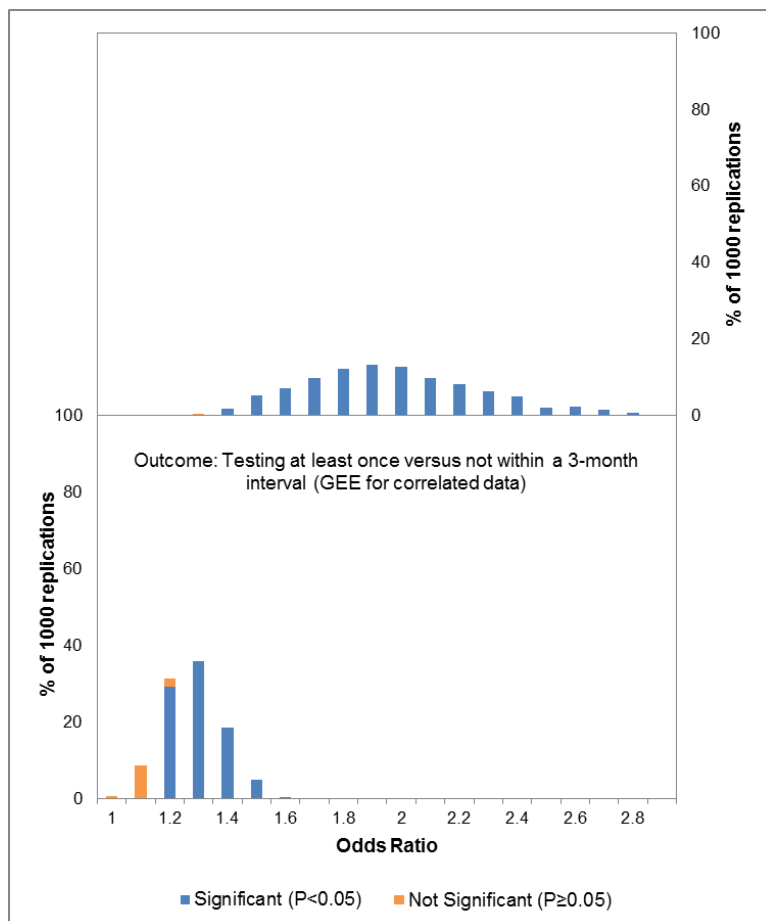


FIGURE 4.S6: Frequency distribution of point estimates generated from the total follow-up time and gap time models for stratified Cox PH regression on 1,000 trial replications each with 1,200 simulated participants assuming the first parameter specification under Scenario A.

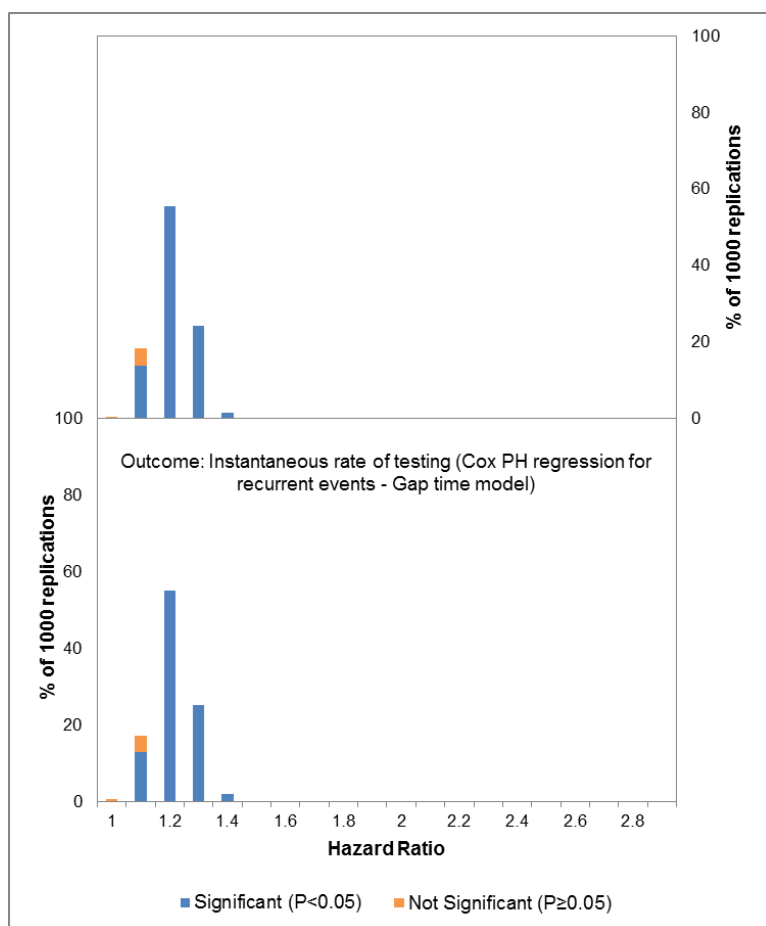


FIGURE 4.S7: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario A.

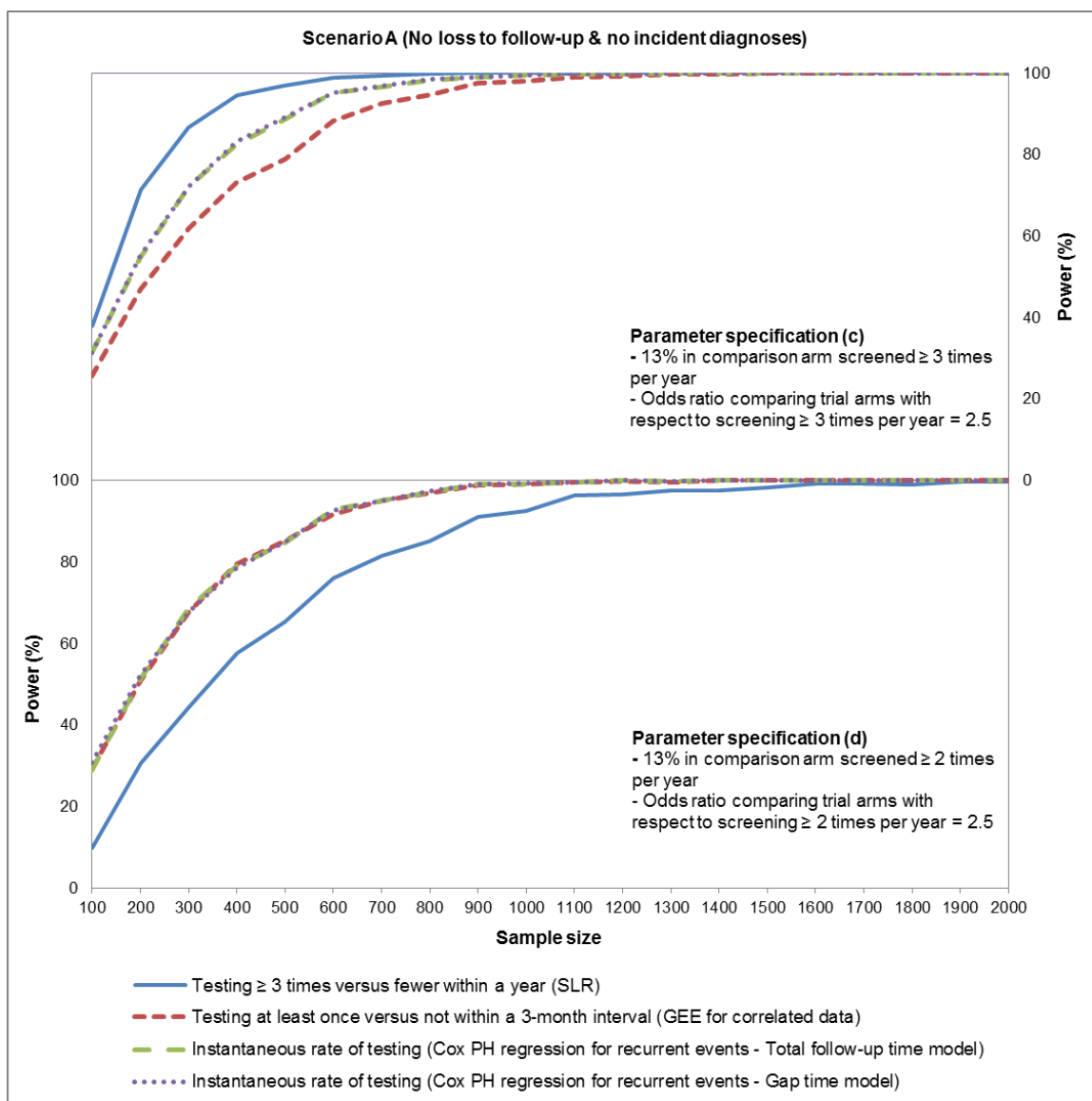


FIGURE 4.S8: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario B.

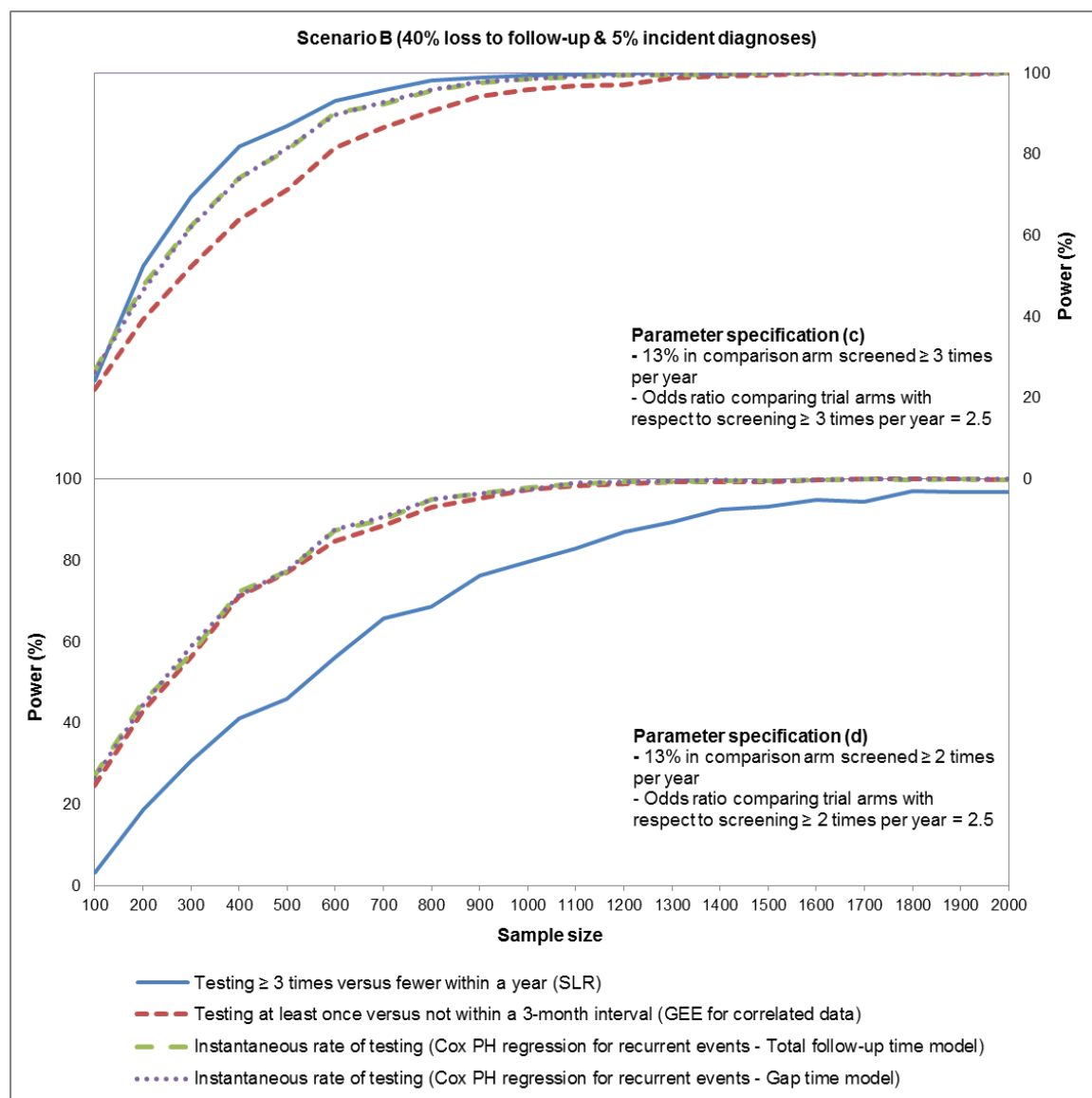


FIGURE 4.S9: Power curves from 1,000 trial replications assuming the third and fourth parameter specifications under Scenario C.

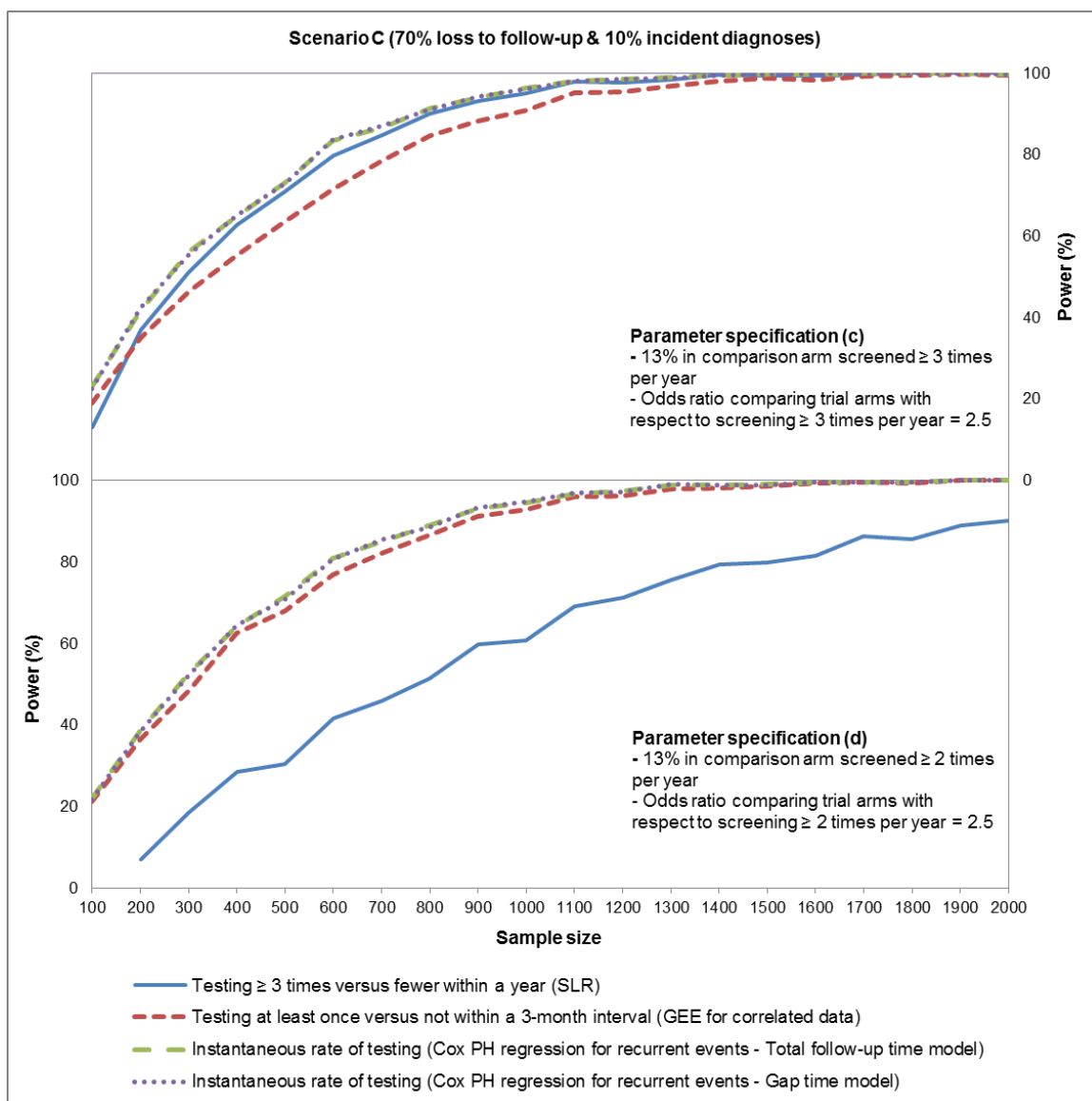


FIGURE 4.S10: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario A.

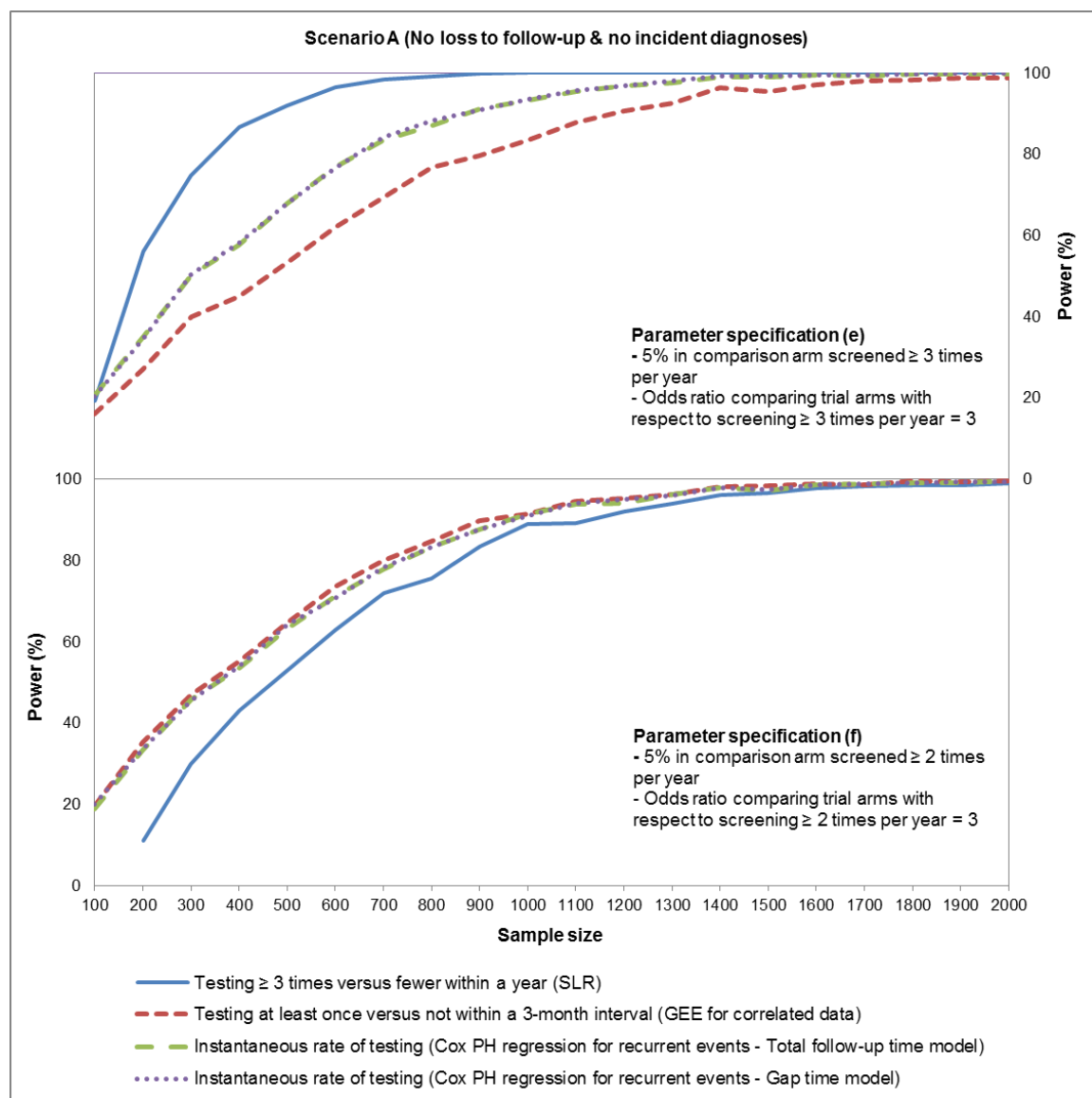


FIGURE 4.S11: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario B.

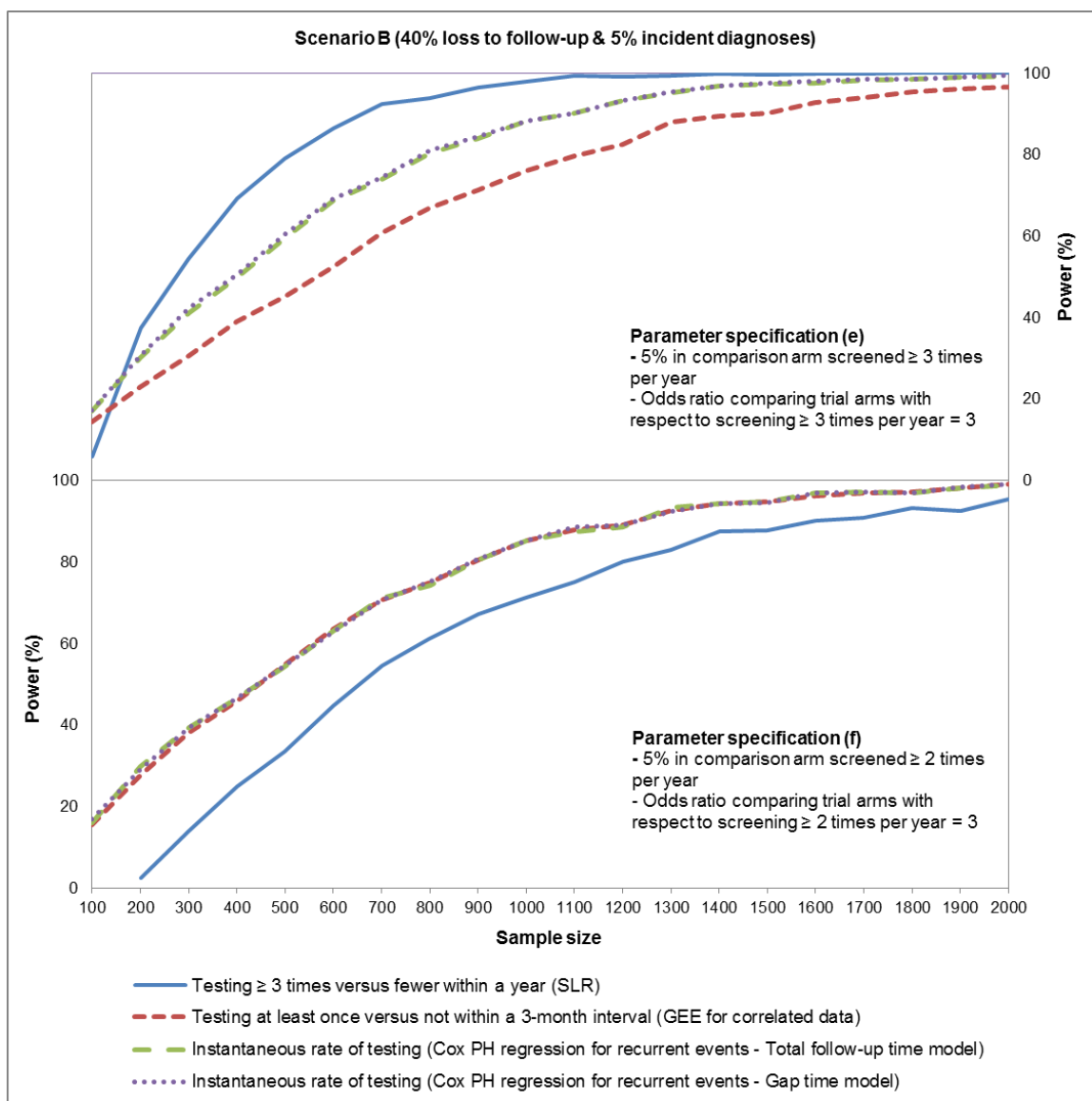
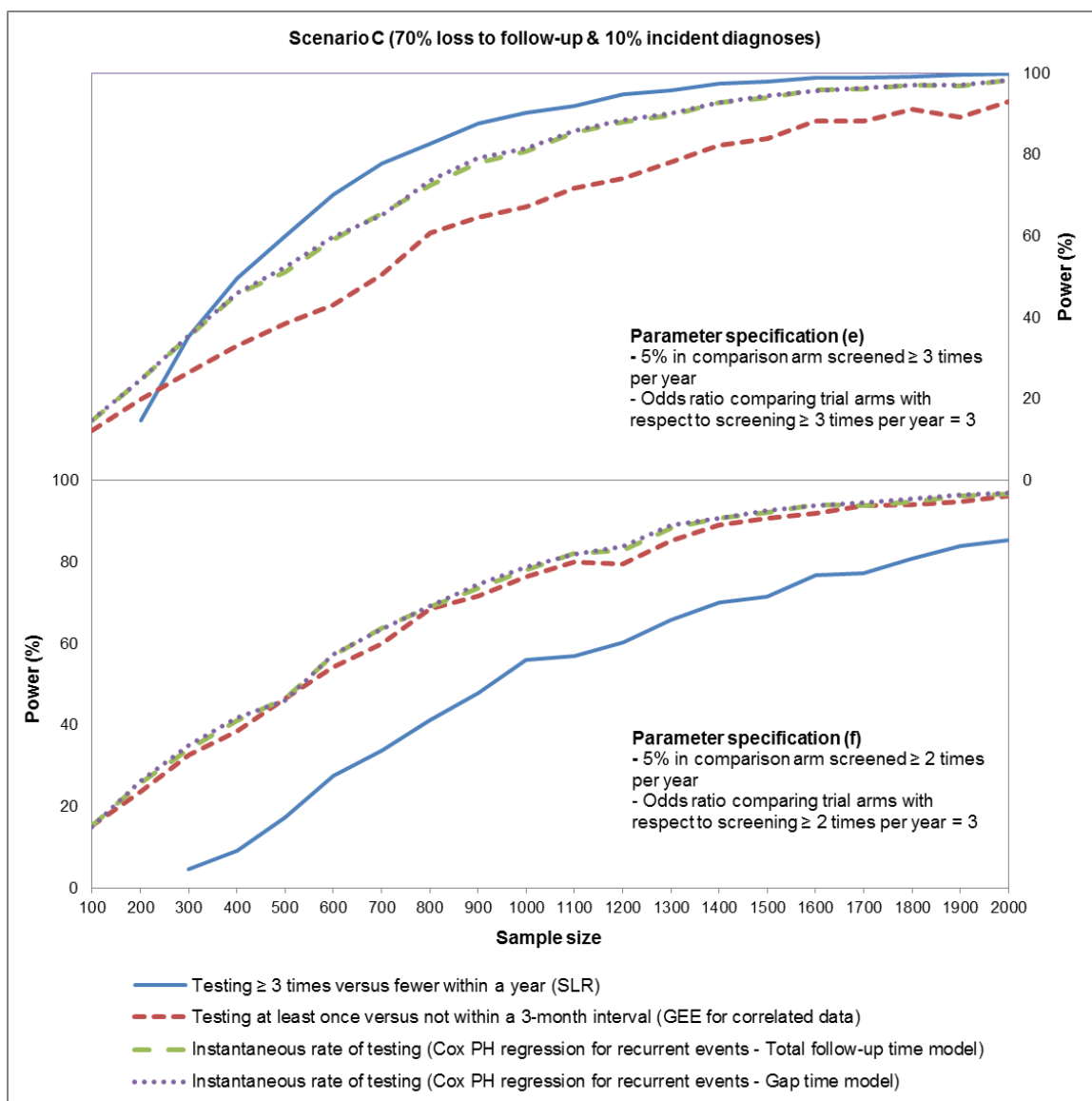


FIGURE 4.S12: Power curves from 1,000 trial replications assuming the fifth and sixth parameter specifications under Scenario C.



CHAPTER 5: Relevance and Contributions

SECTION 5.1: Review of major findings

Gay, bisexual and other men who have sex with men (collectively referred to as MSM) residing in the United States (US) are willing to adopt newer HIV testing technologies as part of future online prevention strategies, and appropriate analytical methods should be employed to evaluate the efficacy of such approaches in improving testing behaviors.

Study 1

Our first study sought to determine the acceptability and intended usage preferences for the following six HIV testing options hypothetically offered free of charge to internet-using MSM: Testing at a physician's office; Voluntary counseling and testing (VCT); Couples' HIV counseling and testing (CHCT); Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) specimen self-collection for laboratory testing. Majority of our 973 participants, recruited online between October and November 2012 through banner advertising on Facebook.com, reported being extremely likely or somewhat likely to use all HIV testing modalities except home DBS specimen self-collection for laboratory testing. Younger MSM reported greater acceptability for expedited/express testing, those with lower educational attainment indicated being more likely to use CHCT, and non-Hispanic black men indicated lower acceptability for VCT. Clear preferences emerged across demographic

and behavioral subgroups, with rapid home self-testing using an oral fluid test and testing at a physician's office being the two most desirable options for future testing. These results were published as an original scientific manuscript in SpringerPlus in February 2014. Our work received favorable comments from peer-reviewers in regard to its contribution to the field of HIV prevention among MSM. Since publication, our manuscript has been designated by the journal as '*Highly accessed*' (2,725 all-time accesses; 2,553 accesses in the last 365 days; 144 accesses in the last 30 days), and has been cited by at least four other articles researching HIV among MSM in Brazil[182], China[183], Spain[184] and the US[185] (as of April 1, 2015).

Study 2

In our second study, we sought to evaluate the awareness and previous use of over-the-counter (OTC) home HIV tests among internet-using MSM, identify factors associated with their stated likelihood of potentially distributing hypothetically offered free rapid home test kits and potentially testing with their social or sexual network associates, as well as describe self-perceived barriers to these alternative uses. Participants were recruited online between May and October 2014 in a formative phase of the KnowAtHome research project funded by the Centers for Disease Control and Prevention (CDC). Most of the 840 men in our sample were aware about home HIV tests, but less than a tenth reported using them within the past year due to concerns regarding cost and accuracy. However, almost everyone indicated being likely to self-test with an oral fluid rapid test and most men were willing to use a finger-stick blood rapid test being provided to them free of charge as part of KnowAtHome. A greater proportion of MSM reported

being likely to potentially distribute oral fluid tests compared to finger-stick blood tests. Men who self-identified as bisexual had reduced odds of being willing to distribute oral fluid tests, and those with lower educational levels had greater odds of being willing to distribute finger-stick blood tests in the future. Reported likelihood of self-testing with free rapid home HIV tests was positively associated with potential kit distribution among social or sexual networks. Almost three fourths of participants indicated being likely to potentially test with friends or sex partners in the same location, the odds being significantly higher for those who engaged in unprotected anal intercourse (UAI) with a single man within the past year, had previously tested for HIV, and who reported being likely to potentially distribute free rapid home HIV tests. Possible negative perceptions and reactions among kit recipients were the most common self-perceived barriers to potential distribution, and MSM unwilling to potentially test with others were most concerned about loss of privacy and confidentiality.

Study 3

Our third study sought to understand the efficiency of multiple analytical approaches for evaluating an intervention seeking to increase routine preventive health screenings. We simulated several replications of a “successful” randomized trial (i.e. one with a positive intervention effect) under various censoring scenarios, formulated three outcome definitions (testing a certain number of times during the entire study period versus not, testing at least once within a clinically meaningful time interval versus not, and the “hazard” or instantaneous rate of testing) to perform corresponding analyses, and finally compared them with respect to interpreting results and estimating power at different

sample sizes. Our results indicate that approaches which better utilize detailed prospective data, while also accounting for within-participant correlations (e.g. generalized estimating equations [GEE] for correlated data and survival analysis methods for recurrent events), are less likely to miss the actual underlying benefits conferred by a novel intervention compared to relying solely on a dichotomous measure derived from aggregating repeated events over the complete study duration. Furthermore, such analytical approaches are also more powerful in real world scenarios wherein some participants are expected to be lost to follow-up. These results were published as an original scientific manuscript in *Contemporary Clinical Trials* in January 2015. Our work received favorable comments from peer-reviewers who opined that this “topic is an important one to explore” and that our “results should be very useful for future studies”.

SECTION 5.2: Current relevance of dissertation studies

MSM are the most heavily impacted risk group for HIV in the US, despite comprising a small proportion of the general population[1]. Testing for HIV may be considered an important prevention activity as individuals aware of their positive serostatus take measures to protect their own health and reduce onward transmission[11]. HIV testing is also the cornerstone of almost all comprehensive risk reduction efforts, including long-established behavioral (e.g. risk reduction counseling) and newer biomedical (e.g. pre-exposure prophylaxis [PrEP]) strategies[10]. None of the currently available interventions are expected to be completely effective in eliminating the spread of HIV on their own, and there have been calls to develop combination packages which incorporate multiple prevention approaches[186, 187]. HIV testing may play a synergistic role with behavioral

and biomedical interventions in potentially averting incident infections among MSM. A recent agent-based modeling study evaluating the effectiveness of prevention packages among South African MSM demonstrated that increases in HIV testing could incrementally prevent the percentage of new infections, depending upon which other components were in the package (2.9% in conjunction with increases in antiretroviral treatment [ART] coverage, 3.1% in conjunction with increases in ART coverage and decreases in unprotected anal intercourse [UAI], and 4.9% in conjunction with increases in ART coverage, decreases in UAI and increases in PrEP acceptance among eligible individuals)[14].

Despite the high prevalence of lifetime HIV testing among US MSM[27, 97], a third of men interviewed in 2011 reported not having been tested within the past year[20], and a similar percentage of newly identified positive MSM did not know they were infected at the time of interview[21]. Increasing the proportion of seropositive individuals who are aware of their status is one of the goals of the National HIV/AIDS Strategy[13] and a Healthy People 2020 objective[99]. Although consensus regarding the optimum time intervals for HIV screening in different population subgroups has not yet been reached, the importance of regular testing in the MSM community has been emphasized in recommendations released by the CDC[17] and the United States Preventive Services Task Force (USPSTF)[19]. How recent healthcare policy changes will influence the demographics of HIV testing in the US is yet to be determined. According to January 2015 estimates from the Congressional Budget Office (CBO) and the Joint Committee on Taxation (JCT), about 42 million non-elderly residents were uninsured in 2014, which is approximately 12 million fewer than who would have been uninsured in the absence of

the Affordable Care Act (ACA)[188]. The annual numbers of uninsured persons are projected to decrease over the next several years, and the mandate that most new health insurance plans must now cover HIV testing for everyone aged 15 to 65 without additional cost-sharing such as copays or deductibles is encouraging[35].

Given the urgent need to stem the reemerging HIV epidemic among MSM in the US, coupled with expected improvements in preventive healthcare coverage, we believe this is an opportune time to be researching novel approaches aimed at increasing testing in this high risk population. Emerging technologies such as internet-based HIV prevention interventions have the advantage of potentially targeting geographically hard-to-reach communities, and effectively engaging MSM who are being overlooked in existing field efforts[77]. Online recruitment in our first two studies provided us diverse samples of men with varying demographic characteristics and risk behavior profiles from all four US census regions. Historically, research has often examined attitudes towards HIV testing services in high risk subgroups concentrated at specific venues, usually located in New York City[189, 190], Seattle[24, 191], and San Francisco[124, 192]. Although we note that important insights have emerged from these endeavors, MSM included in these samples may not be typical of those residing in rural and other urban areas. Our internet-based recruitment strategies and data collection methods enabled us gather information from large numbers of MSM across the country, and hold relevance in today's fast-paced dynamic environment. Another strength of this approach is that individuals tend to be more open and honest while reporting sensitive risk behavior information using computer-based technologies compared to traditional pen-and-paper questionnaires[118, 154], thereby improving data accuracy and reducing the potential for social desirability

bias[119, 153]. Additionally, using online technologies can offer significant benefits with respect to saving time and financial resources. Findings from our third study are relevant in light of the fact that many researchers still employ naive analytical approaches for recurrent events data which either ignore the existence of multiple events, their timing during follow-up, or the correlation between repeated measures[172, 173]. Simulation models offer more control than observational research as one can readily change input parameters for various assumed distributions. This approach allowed us to replicate multiple iterations of a “successful” RCT in realistic and extreme situations, and subsequently demonstrate how aggregating and dichotomizing information on repeated screening tests, without regard to when they were conducted, could miss potentially important benefits of a new prevention intervention to promote regular health monitoring.

SECTION 5.3: Public health contributions of this dissertation

Reducing the burden of HIV among US MSM requires coordinated efforts by entities in the field of research, implementation science and politics. The work undertaken in this dissertation directly contributes to at least two of these domains. Our studies add to the limited evidence base regarding nationwide MSM’s preferences for newer HIV testing modalities, and provide guidance on designing and analyzing future studies to evaluate their public health impact.

Previous research with MSM has focused on attitudes towards selected testing modalities in isolation[28, 56, 74]. However, we believe that our first study is the only one that has examined relative preferences for different options presented collectively. Identifying

variations in ranking orders within demographic and behavioral strata represents an initial step in conceptualizing packages of multiple HIV testing services targeting specific MSM subgroups. Such an intervention could potentially enable men in formulating annual personalized testing plans tailored to their specific needs, and promote more frequent screening, as has been recommended by the CDC[18]. Understanding circumstances wherein men would use particular HIV testing modalities, and which combinations would be preferred over others is the next step in researching this novel concept. Our work provided preliminary data to inform the development of HealthMindr, a smartphone HIV prevention application for MSM which provides individualized recommendations for the best manner in which men can manage their risk[193]. The application can help MSM create a personalized HIV testing strategy and get linked to prevention services including PrEP and non-occupational post-exposure prophylaxis (nPEP). Interested men can also order condoms and home test kits through HealthMindr.

Given that rapid home HIV testing with an oral fluid test is now a reality in the US, there is a need to prioritize the efficacy of this technology in improving testing behaviors among MSM. Our second study evaluated the awareness and adoption of this approach by men in real world settings, and identified factors associated with their reported willingness to distribute and test with social or sexual network associates using free kits in future research settings. We also described MSM's self-perceived barriers to multiple aspects of rapid home HIV testing. These are new findings which have not yet been reported in a large number of MSM from different regions of the US. The only published literature on related issues is from two very selective studies with 60[129] and 27[130] high risk MSM in New York City who never or rarely used condoms and had multiple

sex partners. Our work furthers the understanding of potential alternative uses of this recently introduced technology in a diverse sample of MSM. Peer-driven kit distribution strategies could facilitate the penetration of HIV testing into hidden high risk networks and should be investigated in future studies. The final phase of KnowAtHome will evaluate the distribution of free rapid home test kits by internet-using MSM among their friends or sex partners, as well as changes in their sexual behaviors following rapid HIV self-testing[94].

Finally, our third study provides empirical evidence about which analytical approaches have the potential to increase detection of any actual underlying positive impact of a specific prevention strategy. Using RCTs to evaluate new interventions, such as combination HIV testing packages or rapid home HIV tests, for encouraging frequent screenings in the MSM community can help inform evidence-based health promotion programs. However, researchers often face challenges with regard to specifying the most meaningful study outcomes, and frequently summarize repeated measures into less-informative dichotomous responses. We advocate for a careful consideration of analytical outcome choices and urge researchers to employ more efficient methods that model comprehensive event-specific information, including time of occurrence. This will help reduce the possibility of erroneously labeling a truly beneficial prevention strategy as a failure. Although we used the rationale of an RCT seeking to determine the effectiveness of rapid home HIV test kits in improving testing behaviors among MSM, our findings are applicable to trials evaluating new approaches aimed at promoting routine preventive screenings to manage chronic morbidities such as lipid disorders, type 2 diabetes, and hypertension. Further, we believe our work is a fine contribution to the emerging field of

implementation science, which according to the National Institutes of Health (NIH) intends to “test new approaches to improve health programming, as well as determine a causal relationship between the intervention and its impact”[194].

SECTION 5.4: Future directions emerging from this dissertation

This dissertation has furthered our understanding of new approaches to potentially increase HIV testing frequencies among MSM in the US. During this period of active debate regarding optimum screening intervals and whether a single HIV testing recommendation can successfully capture MSM’s risk diversity, modeling has suggested that individualized testing interventions could significantly help reduce incident infections[195]. Our findings relating to testing option preferences have set the stage for future research in this area by turning an initial concept into the HealthMindr smartphone application[193]. This project, currently underway in Atlanta and Seattle, will provide real world evidence of how MSM employ latest technologies for HIV prevention. During our course of planning KnowAtHome[94], an oral fluid rapid HIV test was approved by the Food and Drug Administration (FDA) for OTC sale in the US[60]. Next steps should involve determining whether this modality has the potential to improve self-testing behaviors among MSM, and further research into how rapid home HIV tests can best be incorporated into personalized testing plans is warranted. Alternative uses such as kit distribution and screening partners before sexual encounters need to be investigated in large national samples of high risk men. Evaluating the extent to which MSM’s high reported acceptability for different aspects of rapid home testing translates into actual use

with social or sexual network associates represents an important future direction of inquiry that we plan on pursuing.

In conclusion, multiple efficacious tools and approaches are needed to mitigate the disproportionate impact of HIV among MSM in the US. Exploring these issues and addressing common barriers to traditional and novel testing strategies can help in fully realizing the role of HIV testing as part of comprehensive prevention efforts.

REFERENCES

1. Purcell DW, Johnson CH, Lansky A, Prejean J, Stein R, Denning P, *et al.* Estimating the Population Size of Men Who Have Sex with Men in the United States to Obtain HIV and Syphilis Rates. *Open AIDS Journal* 2012,**6**:98-107.
2. CDC. Diagnoses of HIV Infection in the United States and Dependent Areas, 2013. http://www.cdc.gov/hiv/pdf/g-l/hiv_surveillance_report_vol_25.pdf. Accessed 03/03/2015.
3. Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, *et al.* Estimation of HIV incidence in the United States. *JAMA* 2008,**300**:520-529.
4. Prejean J, Song R, Hernandez A, Ziebell R, Green T, Walker F, *et al.* Estimated HIV incidence in the United States, 2006-2009. *PLOS ONE* 2011,**6**.
5. Balaji AB, Bowles KE, Le BC, Paz-Bailey G, Oster AM, Group ftNS. High HIV incidence and prevalence and associated factors among young MSM, 2008. *AIDS* 2013,**27**:269-278.
6. CDC. Highlights of CDC Activities Addressing HIV Prevention Among African American Gay, Bisexual, and Other Men Who Have Sex With Men. http://www.cdc.gov/hiv/pdf/library_factsheet_CDC_Highlights_aaGay_Bisexual_MSM.pdf. Accessed 03/03/2015.
7. CDC. Estimated HIV Incidence in the United States, 2007-2010. http://www.cdc.gov/hiv/pdf/statistics_hsr_vol_17_no_4.pdf. Accessed 03/03/2015.

8. CDC. Factsheet: HIV among Gay and Bisexual Men. <http://www.cdc.gov/hiv/pdf/HIV-MSM-english-508.pdf>. Accessed 03/03/2015.
9. CDC. Prevalence and Awareness of HIV Infection Among Men Who Have Sex With Men - 21 Cities, United States, 2008. *MMWR* 2010,**59**:1201-1207.
10. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, *et al.* Successes and challenges of HIV prevention in men who have sex with men. *Lancet* 2012,**380**:388-399.
11. Marks G, Crepaz N, Senterfitt JW, Janssen RS. Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: implications for HIV prevention programs. *JAIDS* 2005,**39**:446-453.
12. Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS* 2006,**20**:1447-1450.
13. National HIV/AIDS Strategy for the United States. <http://aids.gov/federal-resources/national-hiv-aids-strategy/nhas.pdf>. Accessed 03/03/2015.
14. Brookmeyer R, Boren D, Baral SD, Bekker L-G, Phaswana-Mafuya N, Beyrer C, *et al.* Combination HIV Prevention among MSM in South Africa: Results from Agent-based Modeling. *PLOS ONE* 2014,**9**.
15. Gardner EM, McLees MP, Steiner JF, del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clinical Infectious Diseases* 2011,**52**:793-800.
16. CDC. Revised Recommendations for HIV Testing of Adults, Adolescents, and Pregnant Women in Health-Care Settings. *MMWR* 2006,**55**:1-17.

17. CDC. Sexually Transmitted Diseases Treatment Guidelines, 2010. *MMWR* 2010,**59**:12-13.
18. CDC. HIV Testing Among Men Who Have Sex with Men - 21 Cities, United States, 2008. *MMWR* 2011,**60**:694-699.
19. USPSTF. Final Recommendation Statement: Human Immunodeficiency Virus (HIV) Infection: Screening.
<http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/human-immunodeficiency-virus-hiv-infection-screening>. Accessed 03/03/2015.
20. Cooley LA, Oster AM, Rose CE, Wejnert C, Le BC, Paz-Bailey G, *et al.* Increases in HIV Testing among Men Who Have Sex with Men — National HIV Behavioral Surveillance System, 20 U.S. Metropolitan Statistical Areas, 2008 and 2011. *PLOS ONE* 2014,**9**.
21. Wejnert C, Le B, Rose CE, Oster AM, Smith AJ, Zhu J. HIV infection and awareness among men who have sex with men—20 cities, United States, 2008 and 2011. *PLOS ONE* 2013,**8**.
22. Spielberg F, Kurth A, Gorbach PM, Goldbaum G. Moving from apprehension to action: HIV counseling and testing preferences in three at-risk populations. *AIDS Education and Prevention* 2001,**13**:524-540.
23. Kellerman SE, Lehman SJ, Lansky A, Stevens MR, Hecht FM, Bindman AB, *et al.* HIV testing within at-risk populations in the United States and the reasons for seeking or avoiding HIV testing. *JAIDS* 2002,**31**:202-210.

24. Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, Celum CL, *et al.* Overcoming barriers to HIV testing: preferences for new strategies among clients of a needle exchange, a sexually transmitted disease clinic, and sex venues for men who have sex with men. *JAIDS* 2003,**32**:318-327.
25. CDC. HIV Prevalence, Unrecognized Infection, and HIV Testing Among Men Who Have Sex with Men - Five U.S. Cities, June 2004 - April 2005. *MMWR* 2005,**54**:597-601.
26. CDC. Human Immunodeficiency Virus (HIV) Risk, Prevention, and Testing Behaviors - United States, National HIV Behavioral Surveillance System: Men Who Have Sex with Men, November 2003-April 2005. *MMWR* 2006,**55**:1-16.
27. CDC. HIV Risk, Prevention, and Testing Behaviors Among Men Who Have Sex With Men - National HIV Behavioral Surveillance System, 21 U.S. Cities, United States, 2008. *MMWR* 2011,**60**:1-34.
28. MacKellar DA, Hou S-I, Whalen CC, Samuelsen K, Sanchez T, Smith A, *et al.* Reasons for Not HIV Testing, Testing Intentions, and Potential Use of an Over-the-Counter Rapid HIV Test in an Internet Sample of Men Who Have Sex With Men Who Have Never Tested for HIV. *Sexually Transmitted Diseases* 2011,**38**:419-428.
29. Margolis AD, Joseph H, Belcher L, Hirshfield S, Chiasson MA. 'Never testing for HIV' among men who have sex with men recruited from a sexual networking website, United States. *AIDS and Behavior* 2012,**16**:23-29.
30. Qaseem A, Snow V, Shekelle P, Hopkins JR, Owens DK. Screening for HIV in Health Care Settings: A Guidance Statement From the American College of

- Physicians and HIV Medicine Association. *Annals of Internal Medicine* 2009,**150**:125-131.
31. Hutchinson AB, Corbie-Smith G, Thomas SB, Mohanan S, del Rio C. Understanding the patient's perspective on rapid and routine HIV testing in an inner-city urgent care center. *AIDS Education and Prevention* 2004,**16**:101-114.
 32. Copenhaver MM, Fisher JD. Experts outline ways to decrease the decade-long yearly rate of 40,000 new HIV infections in the US. *AIDS and Behavior* 2006,**10**:105-114.
 33. Medicare.gov. HIV screening. <http://www.medicare.gov/coverage/hiv-screening.html>. Accessed 03/03/2015.
 34. Kaiser Family Foundation: State Medicaid Coverage of Routine HIV Screening. <http://kff.org/hivaids/fact-sheet/state-medicaid-coverage-of-routine-hiv-screening/>. Accessed 03/03/2015.
 35. AIDS.gov. The Affordable Care Act and HIV/AIDS. <https://www.aids.gov/federal-resources/policies/health-care-reform/>. Accessed 03/03/2015.
 36. Burke RC, Sepkowitz KA, Bernstein KT, Karpati AM, Myers JE, Tsoi BW, *et al.* Why don't physicians test for HIV? A review of the US literature. *AIDS* 2007,**21**:1617-1624.
 37. Wortley PM, Chu SY, Diaz T, Ward JW, Doyle B, Davidson AJ, *et al.* HIV testing patterns: where, why, and when were persons with AIDS tested for HIV? *AIDS* 1995,**9**:487-492.

38. Coates TJ. Efficacy of voluntary HIV-1 counselling and testing in individuals and couples in Kenya, Tanzania, and Trinidad: a randomised trial. *Lancet* 2000,**356**:103-112.
39. Sweat M, Gregorich S, Sangiwa G, Furlonge C, Balmer D, Kamenga C, *et al.* Cost-effectiveness of voluntary HIV-1 counselling and testing in reducing sexual transmission of HIV-1 in Kenya and Tanzania. *Lancet* 2000,**356**:113-121.
40. Denison J, O'Reilly K, Schmid G, Kennedy C, Sweat M. HIV Voluntary Counseling and Testing and Behavioral Risk Reduction in Developing Countries: A Meta-analysis, 1990–2005. *AIDS and Behavior* 2008,**12**:363-373.
41. Weinhardt LS, Carey MP, Johnson BT, Bickham NL. Effects of HIV counseling and testing on sexual risk behavior: a meta-analytic review of published research, 1985-1997. *American Journal of Public Health* 1999,**89**:1397-1405.
42. UNAIDS. The impact of Voluntary Counselling and Testing: A global review of the benefits and challenges. http://data.unaids.org/Publications/IRC-pub02/jc580-vct_en.pdf. Accessed 03/03/2015.
43. Kamb ML, Fishbein M, Douglas Jr JM, Rhodes F, Rogers J, Bolan G, *et al.* Efficacy of risk-reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases. *JAMA* 1998,**280**:1161-1167.
44. Koblin BA. Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 2004,**364**:41-50.

45. Painter TM. Voluntary counseling and testing for couples: a high-leverage intervention for HIV/AIDS prevention in sub-Saharan Africa. *Social Science & Medicine* 2001,**53**:1397-1411.
46. Kamenga M, Ryder RW, Jingu M, Mbuyi N, Mbu L, Behets F, *et al.* Evidence of marked sexual behavior change associated with low HIV-1 seroconversion in 149 married couples with discordant HIV-1 serostatus: experience at an HIV counselling center in Zaire. *AIDS* 1991,**5**:61-68.
47. Allen S, Tice J, Van de Perre P, Serufilira A, Hudes E, Nsengumuremyi F, *et al.* Effect of serotesting with counselling on condom use and seroconversion among HIV discordant couples in Africa. *BMJ* 1992,**304**:1605-1609.
48. Allen S, Meinzen-Derr J, Kautzman M, Zulu I, Trask S, Fideli U, *et al.* Sexual behavior of HIV discordant couples after HIV counseling and testing. *AIDS* 2003,**17**:733-740.
49. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika C, *et al.* New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. *Lancet* 2008,**371**:2183-2191.
50. Gates GJ. Same-Sex Couples and the Gay, Lesbian, Bisexual Population: New Estimates from the American Community Survey. In: The Williams Institute; 2006.
51. Sullivan PS, Salazar L, Buchbinder S, Sanchez TH. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. *AIDS* 2009,**23**:1153-1162.

52. Goodreau SM, Carnegie NB, Vittinghoff E, Lama JR, Sanchez J, Grinsztejn B, *et al.* What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? *PLOS ONE* 2012,**7**.
53. Lewis MA, Gladstone E, Schmal S, Darbes LA. Health-related social control and relationship interdependence among gay couples. *Health Education Research* 2006,**21**:488-500.
54. EffectiveInterventions.org. Couples HIV Testing and Counseling.
<https://www.effectiveinterventions.org/en/HighImpactPrevention/PublicHealthStrategies/CHTC.aspx>. Accessed 03/03/2015.
55. Wagenaar BH, Christiansen-Lindquist L, Khosropour C, Salazar LF, Benbow N, Prachand N, *et al.* Willingness of US Men Who Have Sex with Men (MSM) to Participate in Couples HIV Voluntary Counseling and Testing (CVCT). *PLOS ONE* 2012,**7**.
56. Stephenson R, Sullivan PS, Salazar LF, Gratzner B, Allen S, Seelbach E. Attitudes towards couples-based HIV testing among MSM in three US cities. *AIDS and Behavior* 2011,**15**:80-87.
57. FDA. Approval Letter - INSTI™ HIV-1 Antibody Test Kit.
<http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm235239.htm>. Accessed 03/03/2015.
58. Bowles KE, Clark HA, Tai E, Sullivan PS, Song B, Tsang J, *et al.* Implementing rapid HIV testing in outreach and community settings: results from an advancing HIV prevention demonstration project conducted in seven US cities. *Public Health Reports* 2008,**123**:78-85.

59. Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, Rossini A, *et al.* Choosing HIV Counseling and Testing Strategies for Outreach Settings: A Randomized Trial. *JAIDS* 2005,**38**:348-355.
60. FDA. Approval Letter - OraQuick In-Home HIV Test.
<http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm310592.htm>. Accessed 03/03/2015.
61. Wright AA, Katz IT. Home Testing for HIV. *New England Journal of Medicine* 2006,**354**:437-440.
62. OraQuick: How-To Video. <http://www.oraquick.com/Taking-the-Test/How-To-Video>. Accessed 03/03/2015.
63. OraQuick: Where to Buy. <http://www.oraquick.com/where-to-buy>. Accessed 03/03/2015.
64. Phillips KA, Chen JL. Willingness to use instant home HIV tests: Data from the California Behavioral Risk Factor Surveillance Survey. *American Journal of Preventive Medicine* 2003,**24**:340-348.
65. Skolnik HdS, Phillips KA, Binson D, Dilley JW. Deciding Where and How to Be Tested for HIV: What Matters Most? *JAIDS* 2001,**27**:292-300.
66. Pai NP, Sharma J, Shivkumar S, Pillay S, Vadnais C, Joseph L, *et al.* Supervised and Unsupervised Self-Testing for HIV in High-and Low-Risk Populations: A Systematic Review. *PLOS Medicine* 2013,**10**.
67. Kimbrough LW, Fisher HE, Jones KT, Johnson W, Thadiparthi S, Dooley S. Accessing Social Networks With High Rates of Undiagnosed HIV Infection: The

- Social Networks Demonstration Project. *American Journal of Public Health* 2009,**99**:1093-1099.
68. Ventuneac A, Carballo-Diéguez A, Leu C-S, Levin B, Bauermeister J, Woodman-Maynard E, *et al.* Use of a Rapid HIV Home Test to Screen Sexual Partners: An Evaluation of its Possible Use and Relative Risk. *AIDS and Behavior* 2009,**13**:731-737.
69. Walensky RP, Paltiel AD. Rapid HIV Testing at Home: Does It Solve a Problem or Create One? *Annals of Internal Medicine* 2006,**145**:459-462.
70. FDA. Letter - Home Access HIV-1 Test System.
<http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm091848.htm>. Accessed 03/03/2015.
71. Home Access Health Corporation. <http://www.homeaccess.com/products.asp>. Accessed 03/03/2015.
72. Branson BM. Home sample collection tests for HIV infection. *JAMA* 1998,**280**:1699-1701.
73. Greensides DR, Berkelman R, Lansky A, Sullivan PS. Alternative HIV testing methods among populations at high risk for HIV infection. *Public Health Reports* 2003,**118**:531-539.
74. Spielberg F, Critchlow C, Vittinghoff E, Coletti AS, Sheppard H, Mayer KH, *et al.* Home collection for frequent HIV testing: acceptability of oral fluids, dried blood spots and telephone results. *AIDS* 2000,**14**:1819-1828.

75. Sharma A, Sullivan PS, Khosropour CM. Willingness to Take a Free Home HIV Test and Associated Factors among Internet-Using Men Who Have Sex with Men. *JIAPAC* 2011,**10**:357-364.
76. Khosropour CM, Johnson BA, Ricca AV, Sullivan PS. Enhancing retention of an internet-based cohort study of men who have sex with men (MSM) via text messaging: randomized controlled trial. *Journal of Medical Internet Research* 2013,**15**.
77. Sullivan PS, Grey JA, Rosser BRS. Emerging technologies for HIV prevention for MSM: What we've learned, and ways forward. *JAIDS* 2013,**63**:S102-S107.
78. Bowen A, Williams M, Horvath K. Using the Internet to Recruit Rural MSM for HIV Risk Assessment: Sampling Issues. *AIDS and Behavior* 2004,**8**:311-319.
79. Swendeman D, Rotheram-Borus MJ. Innovation in sexually transmitted disease and HIV prevention: Internet and mobile phone delivery vehicles for global diffusion. *Current Opinion in Psychiatry* 2010,**23**:139-144.
80. Pequegnat W, Rosser BS, Bowen AM, Bull SS, DiClemente RJ, Bockting WO, *et al.* Conducting Internet-based HIV/STD prevention survey research: Considerations in design and evaluation. *AIDS and Behavior* 2007,**11**:505-521.
81. Bowen AM, Horvath K, Williams ML. A randomized control trial of Internet-delivered HIV prevention targeting rural MSM. *Health Education Research* 2007,**22**:120-127.
82. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS* 2009,**23**:107-115.

83. Weatherburn P, Schmidt A, Hickson F, Reid D, Berg R, Hospers H, *et al.* The European Men-Who-Have-Sex-With-Men Internet Survey (EMIS): Design and Methods. *Sexuality Research and Social Policy* 2013,**10**:243-257.
84. File T, Ryan C. Computer and Internet Use in the United States: 2013. *American Community Survey Reports* 2014.
85. Benotsch EG, Kalichman S, Cage M. Men who have met sex partners via the Internet: Prevalence, predictors, and implications for HIV prevention. *Archives of Sexual Behavior* 2002,**31**:177-183.
86. Liao A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sexually Transmitted Diseases* 2006,**33**:576-584.
87. Rosser BS, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL. HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's INternet Sex Study-II (MINTS-II). *AIDS and Behavior* 2009,**13**:488-498.
88. Horvath KJ, Rosser BS, Remafedi G. Sexual risk taking among young internet-using men who have sex with men. *American Journal of Public Health* 2008,**98**:1059-1067.
89. Garofalo R, Herrick A, Mustanski BS, Donenberg GR. Tip of the iceberg: Young men who have sex with men, the Internet, and HIV risk. *American Journal of Public Health* 2007,**97**:1113-1117.

90. Horvath K, Nygaard K, Simon Rosser BR. Ascertaining Partner HIV Status and its Association with Sexual Risk Behavior Among Internet-Using Men Who Have Sex with Men. *AIDS and Behavior* 2010,**14**:1376-1383.
91. Bilardi JE, Walker S, Read T, Prestage G, Chen MY, Guy R, *et al.* Gay and Bisexual Men's Views on Rapid Self-Testing for HIV. *AIDS and Behavior* 2013,**17**:2093-2099.
92. Colfax G, Lehman J, Bindman A, Vittinghoff E, Vranizan K, Fleming P, *et al.* What happened to home HIV test collection kits? Intent to use kits, actual use, and barriers to use among persons at risk for HIV infection. *AIDS Care* 2002,**14**:675-682.
93. FDA. Approval Letter - Chembio SURE CHECK HIV 1/2 Assay. <http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm091414.htm>. Accessed 03/03/2015.
94. ClinicalTrials.gov. Evaluation of Rapid HIV Self-testing Among MSM (eSTAMP). <http://clinicaltrials.gov/show/NCT02067039>. Accessed 03/03/2015.
95. Crepaz N, Marks G, Liau A, Mullins MM, Aupont LW, Marshall KJ, *et al.* Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis. *AIDS* 2009,**23**:1617-1629.
96. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, *et al.* Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *New England Journal of Medicine* 2011,**365**:493-505.

97. Mimiaga MJ, Landers SJ, Conron KJ. Prevalence and correlates of lifetime HIV testing in a population-based sample of men who have sex with men in massachusetts. *AIDS Patient Care and STDs* 2011,**25**:323-326.
98. MacKellar DA, Valleroy LA, Secura GM, Behel S, Bingham T, Celentano DD, *et al.* Unrecognized HIV infection, risk behaviors, and perceptions of risk among young men who have sex with men: opportunities for advancing HIV prevention in the third decade of HIV/AIDS. *JAIDS* 2005,**38**:603-614.
99. Healthy People 2020 Summary of Objectives.
<http://www.healthypeople.gov/2020/topics-objectives/topic/hiv/objectives>.
Accessed 03/03/2015.
100. Sullivan PS, White D, Rosenberg ES, Barnes J, Jones J, Dasgupta S, *et al.* Safety and Acceptability of Couples HIV Testing and Counseling for US Men Who Have Sex with Men A Randomized Prevention Study. *JIAPAC* 2013,**13**:135-144.
101. Koo DJ, Begier EM, Henn MH, Sepkowitz KA, Kellerman SE. HIV counseling and testing: less targeting, more testing. *American Journal of Public Health* 2006,**96**:962-964.
102. Bavinton BR, Brown G, Hurley M, Bradley J, Keen P, Conway DP, *et al.* Which gay men would increase their frequency of HIV testing with home self-testing? *AIDS and Behavior* 2013,**17**:2084-2092.
103. SAS Version 9.3 [computer program]. Cary, NC: SAS Institute Inc.; 2011.
104. Conover W, Iman RL. Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician* 1981,**35**:124-129.

105. Abdi H. The Bonferonni and Šidák Corrections for Multiple Comparisons. In: *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007.
106. Borda J-C. Memoire sur les Elections au Scrutin. *Histoire de V Academie Royale des Sciences* 1781:657-664.
107. Katz DA, Swanson F, Stekler JD. Why do men who have sex with men test for HIV infection? Results from a community-based testing program in Seattle. *Sexually Transmitted Diseases* 2013,**40**:724-728.
108. Smith A. Americans and Text Messaging. <http://www.pewinternet.org/2011/09/19/americans-and-text-messaging/>. Accessed 03/03/2015.
109. Malebranche DJ, Peterson JL, Fullilove RE, Stackhouse RW. Race and sexual identity: perceptions about medical culture and healthcare among Black men who have sex with men. *Journal of the National Medical Association* 2004,**96**:97-107.
110. Mimiaga MJ, Reisner SL, Bland S, Skeer M, Cranston K, Isenberg D, *et al.* Health system and personal barriers resulting in decreased utilization of HIV and STD testing services among at-risk black men who have sex with men in Massachusetts. *AIDS Patient Care and STDs* 2009,**23**:825-835.
111. Nanín J, Osubu T, Walker JN, Powell B, Powell D, Parsons J. “HIV is still real”: perceptions of HIV testing and HIV prevention among black men who have sex with men in New York City. *American Journal of Men's Health* 2009,**3**:150-164.
112. Stephenson R, Chard A, Finneran C, Sullivan P. Willingness to use couples voluntary counseling and testing services among men who have sex with men in seven countries. *AIDS Care* 2013,**26**:191-198.

113. Stephenson R, Rentsch C, Sullivan P. High levels of acceptability of couples-based HIV testing among MSM in South Africa. *AIDS Care* 2012,**24**:529-535.
114. Cutler DM, Lleras-Muney A. Education and health: evaluating theories and evidence. In. Cambridge, MA: National Bureau of Economic Research; 2006.
115. Jansen IA, Geskus RB, Davidovich U, Jurriaans S, Coutinho RA, Prins M, *et al.* Ongoing HIV-1 transmission among men who have sex with men in Amsterdam: a 25-year prospective cohort study. *AIDS* 2011,**25**:493-501.
116. Chen M, Bilardi J, Lee D, Cummings R, Bush M, Fairley C. Australian men who have sex with men prefer rapid oral HIV testing over conventional blood testing for HIV. *International Journal of STD & AIDS* 2010,**21**:428-430.
117. Mimiaga MJ, Goldhammer H, Belanoff C, Tetu AM, Mayer KH. Men who have sex with men: perceptions about sexual risk, HIV and sexually transmitted disease testing, and provider communication. *Sexually Transmitted Diseases* 2007,**34**:113-119.
118. Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 1998,**280**:867-873.
119. Ellingson JE, Sackett PR, Hough LM. Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology* 1999,**84**:155-165.
120. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, *et al.* Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. *Journal of Medical Internet Research* 2011,**13**.

121. Smith A. Home Broadband 2010. <http://www.pewinternet.org/2010/08/11/home-broadband-2010/>. Accessed 03/03/2015.
122. Jaffe HW, Valdiserri RO, De Cock KM. The reemerging HIV/AIDS epidemic in men who have sex with men. *JAMA* 2007,**298**:2412-2414.
123. CDC. HIV Testing and Risk Behaviors Among Gay, Bisexual, and Other Men Who Have Sex with Men - United States. *MMWR* 2013,**62**:958-962.
124. Phillips KA, Maddala T, Johnson FR. Measuring preferences for health care interventions using conjoint analysis: an application to HIV testing. *Health Services Research* 2002,**37**:1681-1705.
125. Bavinton BR, Brown G, Hurley M, Bradley J, Keen P, Conway DP, *et al.* Which gay men would increase their frequency of HIV testing with home self-testing? *AIDS and Behavior* 2013,**17**:2084-2092.
126. Greacen T, Friboulet D, Fugon L, Hefez S, Lorente N, Spire B. Access to and use of unauthorised online HIV self-tests by internet-using French-speaking men who have sex with men. *Sexually Transmitted Infections* 2012,**88**:368-374.
127. Belza MJ, Figueroa C, Rosales-Statkus ME, Ruiz M, Vallejo F, de la Fuente L. Low knowledge and anecdotal use of unauthorized online HIV self-test kits among attendees at a street-based HIV rapid testing programme in Spain. *International Journal of Infectious Diseases* 2014,**25**:196-200.
128. CDC. Use of Social Networks to Identify Persons with Undiagnosed HIV Infection - Seven U.S. Cities, October 2003-September 2004. *MMWR* 2005,**54**:601-605.

129. Carballo-Diéguez A, Frasca T, Dolezal C, Balan I. Will Gay and Bisexually Active Men at High Risk of Infection Use Over-the-Counter Rapid HIV Tests to Screen Sexual Partners? *Journal of Sex Research* 2012,**49**:379-387.
130. Carballo-Diéguez A, Frasca T, Balan I, Ibitoye M, Dolezal C. Use of a rapid HIV home test prevents HIV exposure in a high risk sample of men who have sex with men. *AIDS and Behavior* 2012,**16**:1753-1760.
131. Fisher RA. On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society* 1922,**85**:87-94.
132. Osmond DH, Catania J, Pollack L, Canchola J, Jaffe D, MacKellar D, *et al.* Obtaining HIV test results with a home collection test kit in a community telephone sample. *JAIDS* 2000,**24**:363-368.
133. Hooper S, Rosser BS, Horvath KJ, Oakes JM, Danilenko G. An online needs assessment of a virtual community: what men who use the internet to seek sex with men want in Internet-based HIV prevention. *AIDS and Behavior* 2008,**12**:867-875.
134. CDC. HIV Testing Trends in the United States, 2000-2011. http://www.cdc.gov/hiv/pdf/testing_trends.pdf. Accessed 03/03/2015.
135. Schwarcz S, Richards TA, Frank H, Wenzel C, Chin Hsu L, Chin C-SJ, *et al.* Identifying barriers to HIV testing: personal and contextual factors associated with late HIV testing. *AIDS Care* 2011,**23**:892-900.
136. Kates J, Levi J. Insurance Coverage and Access to HIV Testing and Treatment: Considerations for Individuals at Risk for Infection and for Those with Undiagnosed Infection. *Clinical Infectious Diseases* 2007,**45**:S255-S260.

137. Lyons A, Pitts M, Grierson J, Smith A, McNally S, Couch M. Sexual behavior and HIV testing among bisexual men: a nationwide comparison of Australian bisexual-identifying and gay-identifying men. *AIDS and Behavior* 2012,**16**:1934-1943.
138. Horvath KJ, Oakes JM, Rosser BS. Sexual negotiation and HIV serodisclosure among men who have sex with men with their online and offline partners. *Journal of Urban Health* 2008,**85**:744-758.
139. Wald A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical society* 1943,**54**:426-482.
140. Kleinbaum D, Klein M. *Logistic Regression: A Self-Learning Text*. 3rd ed. New York: Springer Science + Business Media, LLC; 2010.
141. Hosmer DW, Jr., Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York: John Wiley & Sons; 2000.
142. FDA. Information regarding the OraQuick In-Home HIV Test.
<http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm311895.htm>. Accessed 03/03/2015.
143. FDA. Information regarding the Home Access HIV-1 Test System.
<http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/ucm311903.htm>. Accessed 02/24/2015.
144. Curren MT, Folkes VS. Attributional Influences on Consumers' Desires To Communicate About Products. *Psychology & Marketing* 1987,**4**:31-45.

145. Jacqueline Johnson B, Reingen PH. Social Ties and Word-of-Mouth Referral Behavior. *Journal of Consumer Research* 1987,**14**:350-362.
146. Duhan D, Johnson S, Wilcox J, Harrell G. Influences on consumer use of word-of-mouth recommendation sources. *Journal of the Academy of Marketing Science* 1997,**25**:283-295.
147. Halkitis P, Kupprat S, McCree D, Simons S, Jabouin R, Hampton M, *et al.* Evaluation of the Relative Effectiveness of Three HIV Testing Strategies Targeting African American Men Who Have Sex with Men (MSM) in New York City. *Annals of Behavioral Medicine* 2011,**42**:361-369.
148. Kennedy M, Doll LS. Male Bisexuality and HIV Risk. *Journal of Bisexuality* 2000,**1**:109-135.
149. Goldbaum G, Perdue T, Wolitski R, Rietmeijer C, Hedrich A, Wood R, *et al.* Differences in Risk Behavior and Sources of AIDS Information Among Gay, Bisexual, and Straight-Identified Men Who Have Sex with Men. *AIDS and Behavior* 1998,**2**:13-21.
150. Chakravarty D, Hoff CC, Neilands TB, Darbes LA. Rates of testing for HIV in the presence of serodiscordant UAI among HIV-negative gay men in committed relationships. *AIDS and Behavior* 2012,**16**:1944-1948.
151. Balan I, Carballo-Diéguez A, Frasca T, Dolezal C, Ibitoye M. The Impact of Rapid HIV Home Test Use with Sexual Partners on Subsequent Sexual Behavior Among Men Who have Sex with Men. *AIDS and Behavior* 2014,**18**:254-262.

152. Rosenberger JG, Reece M, Novak DS, Mayer KH. The Internet as a valuable tool for promoting a new framework for sexual health among gay men and other men who have sex with men. *AIDS and Behavior* 2011,**15**:88-90.
153. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity* 2013,**47**:2025-2047.
154. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. *JAIDS* 2004,**35**:421-426.
155. NPHPPHC. National Prevention Strategy.
<http://www.surgeongeneral.gov/initiatives/prevention/strategy/report.pdf>.
Accessed 03/03/2015.
156. USPSTF. Final Recommendation Statement: Lipid Disorders in Adults (Cholesterol, Dyslipidemia): Screening.
<http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/lipid-disorders-in-adults-cholesterol-dyslipidemia-screening>.
Accessed 03/03/2015.
157. USPSTF. Final Recommendation Statement: Diabetes Mellitus (Type 2) in Adults: Screening.
<http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/diabetes-mellitus-type-2-in-adults-screening>. Accessed 03/03/2015.
158. USPSTF. Final Recommendation Statement: Hepatitis B in Pregnant Women: Screening.

<http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/hepatitis-b-in-pregnant-women-screening>. Accessed 03/03/2015.

159. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine* 2010,**8**.
160. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012,**345**.
161. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012,**308**:2594-2604.
162. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* 2008,**148**:295-309.
163. Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Annals of Internal Medicine* 2006,**144**:364-367.
164. Eysenbach G, Group C-E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *Journal of Medical Internet Research* 2011,**13**.
165. Cook RJ. The design and analysis of randomized trials with recurrent events. *Statistics in Medicine* 1995,**14**:2081-2098.
166. Kuramoto L, Sobolev BG, Donaldson MG. On reporting results from randomized controlled trials with recurrent events. *BMC Medical Research Methodology* 2008,**8**.

167. Cook RJ, Bergeron P-J, Boher J-M, Liu Y. Two-stage design of clinical trials involving recurrent events. *Statistics in Medicine* 2009,**28**:2617-2638.
168. Delucchi KL. Sample size estimation in research with dependent measures and dichotomous outcomes. *American Journal of Public Health* 2004,**94**:372-377.
169. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981,**68**:373-379.
170. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Annals of Statistics* 1982:1100-1120.
171. Wei L-J, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* 1989,**84**:1065-1073.
172. Donaldson MG, Sobolev B, Cook WL, Janssen PA, Khan KM. Analysis of recurrent events: a systematic review of randomised controlled trials of interventions to prevent falls. *Age and Ageing* 2009,**38**:151-155.
173. Metcalfe C, Thompson S, Cowie M, Sharples L. The use of hospital admission data as a measure of outcome in clinical studies of heart failure. *European Heart Journal* 2003,**24**:105-112.
174. *SAS 9.3 Macro Language: Reference*. Cary, NC: SAS Institute; 2011.
175. Twisk JW, Smidt N, de Vente W. Applied analysis of recurrent events: a practical overview. *Journal of Epidemiology and Community Health* 2005,**59**:706-710.
176. Kleinbaum D, Klein M. *Survival Analysis: A Self-Learning Text*. 3rd ed. New York: Springer Science + Business Media, LLC; 2012.

177. CDC. PRS Efficacy Criteria for Best-Evidence Risk Reduction (RR) Individual-level, Group-level, and Couple-level Interventions (ILIs/GLIs/CPLs).
<http://www.cdc.gov/hiv/pdf/dhap/prb/prs/efficacy/rr/criteria/hiv-rr-efficacy-best-ili-gli-cpls.pdf>. Accessed 03/03/2015.
178. Katz DA, Dombrowski JC, Swanson F, Buskin SE, Golden MR, Stekler JD. HIV intertest interval among MSM in King County, Washington. *Sexually Transmitted Infections* 2013,**89**:32-37.
179. Lin D, Wei L, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 2000,**62**:711-730.
180. Cook RJ, Lawless JF. Analysis of repeated events. *Statistical Methods in Medical Research* 2002,**11**:141-166.
181. Dean CB, Balshaw R. Efficiency lost by analyzing counts rather than event times in Poisson and overdispersed Poisson regression models. *Journal of the American Statistical Association* 1997,**92**:1387-1398.
182. Lippman SA, Veloso VG, Buchbinder S, Fernandes NM, Terto V, Sullivan PS, *et al.* Over-the-counter human immunodeficiency virus self-test kits: time to explore their use for men who have sex with men in Brazil. *Brazilian Journal of Infectious Diseases* 2014,**18**:239-244.
183. Zhao Y, Zhang L, Zhang H, Xia D, Pan SW, Yue H, *et al.* HIV Testing and Preventive Services Accessibility Among Men Who Have Sex With Men at High Risk of HIV Infection in Beijing, China. *Medicine* 2015,**94**.

184. Rosales-Statkus ME, de la Fuente L, Fernández-Balbuena S, Figueroa C, Fernández-López L, Hoyos J, *et al.* Approval and Potential Use of Over-the-Counter HIV Self-Tests: The Opinion of Participants in a Street Based HIV Rapid Testing Program in Spain. *AIDS and Behavior* 2015,**19**:472-484.
185. Marlin RW, Young SD, Bristow CC, Wilson G, Rodriguez J, Ortiz J, *et al.* Piloting an HIV self-test kit voucher program to raise serostatus awareness of high-risk African Americans, Los Angeles. *BMC Public Health* 2014,**14**.
186. Bekker L-G, Beyrer C, Quinn TC. Behavioral and Biomedical Combination Strategies for HIV Prevention. *Cold Spring Harbor Perspectives in Medicine* 2012,**2**.
187. Cohen J. Prevention Cocktails: Combining Tools to Stop HIV's Spread. *Science* 2005,**309**:1002-1005.
188. Updated Estimates of the Insurance Coverage Provisions of the Affordable Care Act. <http://www.cbo.gov/sites/default/files/cbofiles/attachments/49892-breakout-AppendixB.pdf>. Accessed 03/03/2015.
189. Antonio-Gaddy MS, Richardson-Moore A, Burstein GR, Newman DR, Branson BM, Birkhead GS. Rapid HIV Antibody Testing in the New York State Anonymous HIV Counseling and Testing Program: Experience From the Field. *JAIDS* 2006,**43**:446-450.
190. Nanín J, Osubu T, Walker JN, Powell B, Powell D, Parsons J. "HIV Is Still Real": Perceptions of HIV Testing and HIV Prevention Among Black Men Who Have Sex With Men in New York City. *American Journal of Men's Health* 2008.

191. Helms DJ, Weinstock HS, Mahle KC, Bernstein KT, Furness BW, Kent CK, *et al.*
HIV Testing Frequency Among Men Who Have Sex With Men Attending Sexually Transmitted Disease Clinics: Implications for HIV Prevention and Surveillance. *JAIDS* 2009,**50**:320-326.
192. Skolnik HdS, Phillips KA, Binson D, Dilley JW. Deciding Where and How to Be Tested for HIV: What Matters Most? *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2001,**27**:292-300.
193. The HealthMindr.
<https://play.google.com/store/apps/details?id=com.thehealthmindr.mac&hl=en>.
Accessed 03/03/2015.
194. NIH. Implementation Science Information and Resources.
<http://www.fic.nih.gov/researchtopics/pages/implementationsscience.aspx>.
Accessed 03/03/2015.
195. Khanna A, Goodreau SM, Wohlfeiler D, Daar E, Little S, Gorbach PM.
Individualized diagnosis interventions can add significant effectiveness in reducing HIV incidence among men who have sex with men: insights from Southern California. *Annals of Epidemiology* 2015,**25**:1-6.

APPENDICES

A1: Survey questions on the acceptability and ranking of six HIV testing options

Acceptability

A wide range of services and tools have been developed to make HIV testing more accessible and acceptable. In the sections below, please read the description of each testing option and indicate how likely you would be to use it. For these questions, assume that all of the testing options described are given to you free of charge.

1. Testing with your physician: At your doctor's office, a sample of your blood or saliva is taken by a clinician. Your results are usually ready within approximately a week and provided either over the phone or in person.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

2. Individual voluntary counseling and testing (VCT): In a one-on-one session with a trained counselor, you receive education and prevention counseling, are tested for HIV, provided with your results, and you work with the counselor to develop a plan to for the future. VCT usually takes place at a testing center or community-based organization and you receive your HIV test results from the counselor in 30 to 45 minutes.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

3. Couples' HIV counseling and testing (CHCT): Together, couples are provided with education and prevention counseling, tested, and presented with their results. The testing is usually provided at a testing center or community-based organization. Based on these

results, the counselor then helps the couple develop a plan for the future. The process takes 45 minutes to 1 hour.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

4. Expedited/express testing: Express testing cuts out the education and prevention counseling portion of other modes of testing. You enter your contact information in a computer, after which an HIV tester collects a drop of blood from your fingertip. Your results can be sent to you by text message or email, or you can retrieve them by checking with a confidential PIN on the internet. If your test for HIV turned out to be positive, your result would be returned by phone or in person by an HIV counselor.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

5. Rapid home self-testing: In the privacy of your own home or other location, you collect a sample of saliva by swabbing the inside of your mouth along your gums and putting the swab inside the provided test tube. Results are ready in 20 minutes. You read the results of the test yourself, like women read a home pregnancy test.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

6. Home DBS self-collection: You prick your finger, collect a drop of blood on a card, and mail the card to a laboratory for testing using a pre-paid envelope enclosed in the testing kit. This testing option allows you to complete the test in the privacy of your own home or other location. Your test results will be available over the phone approximately seven business days from when you mail the card back.

On a scale from one to five, how likely would you be to use this service?

1	2	3	4	5
Extremely unlikely				Extremely likely

Ranking

The six testing options you just read about are listed below. Please rank them in order from 1 (most likely to use) to 6 (least likely to use).

Testing with your physician

Individual voluntary counseling and testing (VCT)

Couples' HIV counseling and testing (CHCT)

Expedited/express testing

Rapid home self-testing

Home DBS self-collection

A2: Press copy of Chapter 2 (*SpringerPlus* 2014, 3:109)

Sharma et al. *SpringerPlus* 2014, 3:109
<http://www.springerplus.com/content/3/1/109>

 SpringerPlus
 a SpringerOpen Journal

RESEARCH

Open Access

Acceptability and intended usage preferences for six HIV testing options among internet-using men who have sex with men

Akshay Sharma^{1,2*}, Rob B Stephenson³, Darcy White³ and Patrick S Sullivan^{1,2}

Abstract

Background: Men who have sex with men (MSM) continue to be disproportionately impacted by the Human Immunodeficiency Virus (HIV) epidemic in the United States (US). Testing for HIV is the cornerstone of comprehensive prevention efforts and the gateway to early engagement of infected individuals in medical care. We sought to determine attitudes towards six different HIV testing modalities presented collectively to internet-using MSM and identify which options rank higher than others in terms of intended usage preference.

Methods: Between October and November 2012, we surveyed 973 HIV-negative or -unknown status MSM and assessed their acceptability of each of the following services hypothetically offered free of charge: Testing at a physician's office; Individual voluntary counseling and testing (VCT); Couples' HIV counseling and testing (CHCT); Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) specimen self-collection for laboratory testing. Kruskal-Wallis tests were used to determine whether the stated likelihood of using each of these modalities differed by selected respondent characteristics. Men were also asked to rank these options in order of intended usage preference, and consensual rankings were determined using the modified Borda count (MBC) method.

Results: Most participants reported being extremely likely or somewhat likely to use all HIV testing modalities except DBS self-collection for laboratory testing. Younger MSM indicated greater acceptability for expedited/express testing ($P < 0.001$), and MSM with lower educational levels reported being more likely to use CHCT ($P < 0.001$). Non-Hispanic black MSM indicated lower acceptability for VCT ($P < 0.001$). Rapid home self-testing using an oral fluid test and testing at a physician's office were the two most preferred options across all demographic and behavioral strata.

Conclusions: Novel approaches to increase the frequency of HIV testing among US MSM are urgently needed. Combination testing packages could enable high risk MSM in putting together annual testing strategies personalized to their circumstances, and warrant due consideration as an element of combination HIV prevention packages.

Keywords: HIV testing preferences; Internet-using men who have sex with men; Combination prevention approaches; Rapid home HIV self-testing

Background

Men who have sex with men (MSM) comprise approximately 4% of the adult male United States (US) population (Purcell et al. 2012), but are the most heavily impacted risk group for Human Immunodeficiency Virus (HIV) infection. Since 2000, incident infections among

MSM have been increasing annually (Hall et al. 2008), with the rate of new diagnoses in this group being at least 44 times that of other men (Purcell et al. 2012). In 2010, MSM accounted for more than three fourths (78%) of new HIV infections among males, and almost two thirds (63%) of all new infections in the US (CDC 2012). Most incident diagnoses occurred among young (ages 13–24), black MSM relative to any other age or racial category (CDC 2012). Better prevention strategies are needed to help reverse current trends.

* Correspondence: ashar24@emory.edu

¹Department of Epidemiology, Emory University Laney Graduate School, 1518 Clifton Road NE, Atlanta, GA 30322, USA

²Department of Epidemiology, Emory University Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322, USA

Full list of author information is available at the end of the article



Behavioral interventions, such as risk reduction counseling, and biomedical approaches, such as condoms and pre-exposure prophylaxis, have complementary roles in HIV prevention. Modeling experiments have shown that offering packages of currently available interventions can avert at least 25% of new infections among MSM over a decade (Sullivan et al. 2012). Testing for HIV is not just a critical first step in developing client-specific recommendations regarding the adoption of these approaches, but can be considered an important prevention activity in itself. Meta-analytic evidence suggests that seropositive individuals aware of their status are motivated to interrupt onward transmission and reduce risky behaviors including unprotected anal intercourse (UAI) (Crepez et al. 2009; Marks et al. 2005). HIV testing is also the gateway to early engagement of infected individuals in treatment and care (Gardner et al. 2011), wherein resulting viral load reductions are known to offer substantial prevention benefits (Cohen et al. 2011).

The US Centers for Disease Control and Prevention (CDC) currently recommends that sexually active MSM should be tested for HIV annually, and that higher risk MSM who have multiple partners or use illicit drugs concurrent with sexual activity should be screened for sexually transmitted infections (STI) at 3-6 month intervals (CDC 2010). Although the nationwide prevalence of lifetime testing among MSM is high (CDC 2011; Mimiaga et al. 2011), many men report not being tested within the past year (CDC 2011) and a high proportion of seropositive MSM are unaware that they are infected (CDC 2005; MacKellar et al. 2005). The estimated HIV transmission rate from persons who are unaware of their infection is 3.5 times that from serostatus-aware individuals (Marks et al. 2006). MSM therefore remain a key risk group for expanded testing efforts. Increasing the percentage of infected individuals who know their serostatus is one of the goals of the National HIV/AIDS Strategy (National HIV/AIDS Strategy for the United States 2010) and a Healthy People 2020 objective (Healthy People 2020 Summary of Objectives 2010).

Depending upon their preferences or circumstances, MSM in the US can choose from several HIV testing approaches ranging from the traditional to the contemporary. Physician's offices, frequently offering screening as part of routine general physical exams, have remained one of the most common testing venues (CDC 2006a, 2011). Individual voluntary counseling and testing (VCT) is usually provided at community-based organizations, and involves one-on-one sessions comprising of pre-test risk assessments and post-test counseling. VCT has proven efficacious in promoting behavior change among high risk persons who learn they are living with HIV, and constitutes an opportunity for both primary and secondary prevention (UNAIDS 2001; Koblin 2004). In 2012,

the CDC initiated a national diffusion plan for couples' HIV counseling and testing (CHCT) targeting same-sex couples in 12 US jurisdictions with the highest HIV burden (EffectiveInterventions 2012). Here, partners participate in the whole cycle of counseling and testing together, and receive risk reduction messages tailored to their couple serostatus (sero-concordant negative, sero-concordant positive, or sero-discordant) (Sullivan et al. 2013).

Although prevention counseling is desirable for high risk individuals, the CDC recognizes that such counseling might not be appropriate or feasible in all settings (CDC 2006b), and it could pose a barrier to testing. States such as New York that have streamlined regulations regarding pre-test counseling have seen increases in HIV testing (Koo et al. 2006). Given the recent licensure of rapid tests with processing times as little as 60 seconds (FDA 2010), an expedited/express testing approach that excludes prevention counseling sessions could be provided through street outreach programs at large events such as gay pride. Individuals could choose to receive their results by text message or email, or retrieve them online using a confidential personal identification number (PIN) whenever ready. Preliminary positive persons would receive their results either by phone or in person by a trained counselor. In addition to saving time, this approach can help reduce stigma associated with HIV testing options requiring an assessment of risk behaviors (Copenhaver and Fisher 2006; Hutchinson et al. 2004). Rapid home self-testing with a recently US Food and Drug Administration (FDA) approved oral fluid test (FDA 2012), is another testing modality offering privacy and convenience (Pai et al. 2013). Individuals can read their own test results within 20 minutes, and have the option of calling a 24x7 support center toll free if they have questions or receive a preliminary positive result. This non-invasive approach differs from home dried blood spot (DBS) self-collection wherein specimens need to be returned for laboratory HIV testing, and results are available by phone within 7 days (FDA 1996).

Considering this menu of available options, we believe that analogous to combination prevention approaches, combination testing packages need due consideration as an element in continuing efforts to increase HIV testing frequencies among high risk populations. Such an intervention could enable individuals in putting together annual personalized testing strategies tailored to their needs and risk perceptions. Previous online and in-person research studies among MSM, each focusing on selected testing modalities in isolation, have found generally favorable attitudes towards their adoption (Bilardi et al. 2013; Bavinton et al. 2013; MacKellar et al. 2011; Phillips and Chen 2003; Wagenaar et al. 2012; Stephenson et al. 2011; Sharma et al. 2011; Spielberg et al. 2000). We sought to explore the acceptability of six HIV testing approaches

presented collectively to internet-using MSM in the US when hypothetically offered free of charge. Another objective of our study was determining which testing options rank higher than others in terms of intended usage preference. Identifying variations in ranking orders within demographic and behavioral strata of MSM represents an initial step in developing comprehensive packages of HIV testing services targeting specific subgroups.

Methods

MSM were recruited online through selective placement of banner advertisements displayed on a social networking website (Facebook.com) from October to November 2012. Recruitment was targeted only towards internet users in the US who indicated in their Facebook profile that they were male, 18 years of age or older and interested in men. Individuals who clicked through the banner advertisements were directed to an online informed consent module, and those who consented were screened to determine eligibility before being administered an internet-based survey. Eligibility criteria included being reportedly male, 18 years of age or older, currently residing within the US, and having at least one male sex partner in the past 6 months. This study was approved by the Emory University Institutional Review Board.

Demographic information collected from participants included age, race/ethnicity, state of residence, education, employment, self-identified sexual orientation and whether they had a main partner. Questions pertaining to the participants' behaviors included whether they had engaged in UAI with male sex partners in the past 6 months, and HIV testing characteristics detailing the timing, location and type of their most recent test. Men who reported being previously tested were asked to indicate one or more motivations for their decision to test from a list of pre-specified options based on subject area expertise, and provided with the choice of typing in an open-ended response.

Participants who reported not being infected with HIV were provided brief descriptions about different testing approaches, and then asked about their likelihood of using each option hypothetically offered free of charge. Acceptability was assessed by the question: "On a scale from one to five, how likely would you be to use this service?" Six questions of this form were asked, one for each of the following approaches: Testing at a physician's office; VCT; CHCT; Expedited/express testing; Rapid home self-testing using an oral fluid test; Home dried blood spot (DBS) specimen self-collection for laboratory testing. Responses were collected in the following Likert item format: 1 = Extremely unlikely; 2 = Somewhat unlikely; 3 = Neutral; 4 = Somewhat likely; 5 = Extremely likely. Finally, men were asked to rank these options in order of intended usage preference from the one they were

most likely to use (assigned Rank 1) to the one they were least likely to use (assigned Rank 6).

Statistical analyses were performed using SAS version 9.3 (SAS 2011). The analytic sample only included self-reported HIV negative or unknown status MSM who answered at least one of six questions on the acceptability of various testing approaches. Fisher's exact tests were used to compare these respondents with men who were excluded. Demographic and behavioral characteristics of all study participants and HIV testing characteristics of men who reported being previously tested were tabulated. Responses for their decisions to test were summarized, and open-ended comments were manually reviewed and reassigned to appropriate pre-specified options.

The acceptability of various HIV testing approaches stratified by selected demographic and behavioral characteristics was summarized by finding the median and mean of participants' five-point Likert item responses. Given the ordinal nature of these data, non-parametric tests are preferable for statistical inferences (Conover and Iman 1981). The Kruskal-Wallis ANOVA, a generalized form of the Wilcoxon-Mann-Whitney test, was used to determine whether the intended usage likelihood of a particular testing option differed by the following respondent characteristics: age; race/ethnicity; education; whether they had a main partner; whether they had engaged in UAI with male sex partners in the past 6 months; HIV testing history. Because of our *a priori* decision to conduct 36 independent tests (6 testing approaches times 6 participant characteristics), the alpha level was corrected using the Sidak equation to limit the overall risk of making at least one Type I error to 0.05 (Abdi 2007). Each Kruskal-Wallis test was considered statistically significant only if its associated probability was smaller than 0.001. Additional analyses were performed to examine whether participants' stated likelihood of using any testing option differed by geographic region of residence.

The modified Borda count (MBC) method was used to identify the relative orders of preferences for the various testing modalities overall, as well as stratified by selected participant characteristics. The original system invented by Jean-Charles de Borda in 1770 was intended for use in elections with a single winner (Borda 1781). Each testing approach was assigned a certain number of points corresponding to the position in which it was ranked by individual respondents. The number of points given for a participant's first and subsequent choices was determined by the total number of options he actually ranked, rather than the total number of options available. Points for each approach were summed to determine ranking orders representing the collective best compromise within each stratum. This method effectively penalized respondents who did not rank all six testing approaches,

Table 1 Demographic and behavioral characteristics of 973 HIV negative or unknown status MSM^a respondents to a national online health survey, United States, 2012

Characteristic	n ^b	%
Age group (years)^c:		
18-24	410	42
25-34	269	28
35-44	123	13
≥ 45	171	18
Race/Ethnicity:		
White, non-Hispanic	751	77
Black, non-Hispanic	14	1
Hispanic	117	12
Other ^d	91	9
Census region:		
West	279	29
Midwest	218	22
Northeast	190	20
South	285	29
Unknown	1	0
Education:		
College, Post graduate, or Professional school	384	39
Some college, Associate's degree, and/or Technical school	391	40
High school, GED ^e or less	193	20
Unknown	5	1
Employment:		
Employed full-time	515	53
Employed part-time	221	23
Unemployed	191	20
Retired	39	4
Unknown	7	1
Self-identified sexual orientation:		
Homosexual/Gay	945	97
Bisexual	19	2
Other ^f	9	1
Had a main partner^g:		
Yes, for ≥ 1 year	366	38
Yes, for < 1 year	175	18
No	430	44
Unknown	2	0
Had UAI^h with a male sex partner in the past 6 months:		
Yes, with ≥ 2 men	196	20
Yes, with 1 man	409	42
No	333	34
Unknown	35	4

Table 1 Demographic and behavioral characteristics of 973 HIV negative or unknown status MSM^a respondents to a national online health survey, United States, 2012 (Continued)

HIV testing history:	
Never tested	160 16
Tested at least once	795 82
Unknown	18 2
HIV status (Result of most recent HIV test):	
Negative	773 79
Unknown ⁱ	200 21

^aMSM: Men who have sex with men.^bSample size (N) = 973.^cAge: Mean = 31, Median = 26, Range = 18-77.^dIncludes 31 Asian/Pacific Islander, 12 Native American/Alaska Native, 36 multiracial, 9 other and 3 unknown.^eGED: General educational development.^fIncludes 2 who indicated they were "Heterosexual/Straight", 2 who indicated they were "Questioning/Unsure", 4 who indicated "Other" as their response and 1 unknown.^gDefined as "Someone you feel committed to above all others. You might call this person your boyfriend, partner, significant other, spouse, or husband".^hUAI: Unprotected anal intercourse. Neither the respondent nor his partner used a condom.ⁱIncludes 160 who never tested, 9 who tested but did not receive a result, 1 who tested and received an indeterminate result, and 30 who declined to answer.

by diminishing the number of points their ranking distributed among these options, thereby favoring approaches supported by a broad consensus.

Results

Overall, 432,632 advertising impressions on Facebook resulted in 4,638 click-throughs to the survey over a 10-day period; 1,739 (38% of click-throughs) consented and were asked questions used to determine eligibility. Of these, 15 identified their gender other than male, 37 were less than 18 years of age, 15 did not reside within the US, 335 did not self-report sex with a man in the past 6 months, and 86 did not respond to one or more of the eligibility questions, yielding a sample of 1,285 (74% of respondents to eligibility questions) who could potentially complete the survey. We restricted our analyses to 973 (81% of 1,204 HIV negative or unknown status participants) who answered at least one of the six acceptability questions. Compared to these participants, excluded men were more likely to be non-Hispanic black, but similar with respect to all other characteristics (data not shown in table).

Table 1 summarizes the demographic and behavioral characteristics of respondents included in our analyses. Majority of the participants were young (mean age in years = 31; median = 26) non-Hispanic white men with some college education or higher. More than one third had a main partner for ≥ 1 year, one fifth reported having UAI with ≥ 2 men within the past 6 months, and almost one fifth had never been tested for HIV.

The HIV testing characteristics of 795 participants who reported being previously tested are described in Table 2. More than two fifths of the men had their most recent test > 1 year prior to the survey. Among the 56% who tested within the past year, almost a quarter indicated testing routinely every 12 months, approximately one third tested routinely every 6 months and almost one fifth tested routinely every 3 months. Private doctors' offices and public health clinics were the most commonly reported testing locations, followed by HIV counseling and testing sites.

Regarding participants' decisions to test for HIV, 55% of the 795 indicated they got tested routinely, 25% before they started having sex with a new partner, 21% whenever they had the opportunity, 15% after they had UAI with someone whose HIV status they did not know,

10% whenever someone they had sex with told them they had an STI, 10% if they started to notice or feel symptoms of an STI, 3% after they had UAI with someone they knew to be HIV positive, and 2% whenever they felt the need to test. Respondents could have indicated more than one motivation for their decision to previously test for HIV: 23% chose multiple reasons, 72% chose a single reason and 5% did not specify a reason (data not shown in table).

Figure 1 depicts MSM's stated likelihood of using each of the six testing approaches hypothetically offered free of charge. Overall, majority of the men reported being extremely likely or somewhat likely to use different options. DBS self-collection for laboratory testing was the only approach that appealed to less than half the participants.

The intended usage likelihood for each testing approach stratified by selected demographic and behavioral characteristics is summarized in Table 3. On adjusting for multiple comparisons, younger participants were significantly more likely to use expedited/express testing ($P < 0.001$), non-Hispanic black participants reported lower acceptability for VCT ($P < 0.001$), and participants with lower educational levels were more likely to use CHCT ($P < 0.001$). The stated likelihood of using any particular option did not significantly differ by the behavioral characteristics of respondents or by their region of residence (data not shown in table).

The MBC ranking orders for all six HIV testing approaches are presented in Figure 2. Overall, rapid home self-testing using an oral fluid test and testing at a physician's office were the two most preferred options. Expedited/express testing and VCT were next, followed by DBS self-collection for laboratory testing and CHCT. Similar patterns were observed on stratifying by HIV testing history, relationship status, and history of UAI with a male sex partner within the past 6 months.

Discussion

Our study sought to explore attitudes towards using long established and newer HIV testing modalities available in the US. Specifically, we were interested in determining the acceptability and intended usage preferences for six testing options hypothetically offered free of charge to internet-using MSM. Based on self-reported likelihood of using each approach, our results indicate high overall acceptability, demonstrating the potential for combining multiple options as part of comprehensive packages to promote regular testing in this disproportionately impacted population. Motivations for our participants' decisions to test are comparable to MSM undergoing HIV testing at a community-based program in Seattle (Katz et al. 2013). Clear preferences for test types emerged across subgroups, revealing which approaches

Table 2 HIV testing characteristics of 795 HIV negative or unknown status MSM^a respondents to a national online health survey who had previously tested, United States, 2012

Characteristic	n ^b	%
Time of most recent HIV test:		
More than 2 years ago	151	19
Between the past 1 - 2 years	187	24
Within the past 1 year		
Test routinely every year	99	12
Test routinely every 6 months	141	18
Test routinely every 3 months	71	9
Test routinely at other intervals	10	1
Do not test routinely	128	16
Unknown	8	1
Location of most recent HIV test:		
Private doctor's office (including HMO ^c) ^d	325	41
Public health clinic/Community health center/STD ^e clinic	242	30
HIV counseling and testing site	106	13
Street outreach program/Mobile unit	41	5
Home or other private location	21	3
Other ^f	60	8
Most recent HIV test type:		
Test that required drawing blood with a syringe	427	54
Finger-stick blood rapid test	171	22
Oral fluid rapid test	154	19
Unknown	43	5

^aMSM: Men who have sex with men.

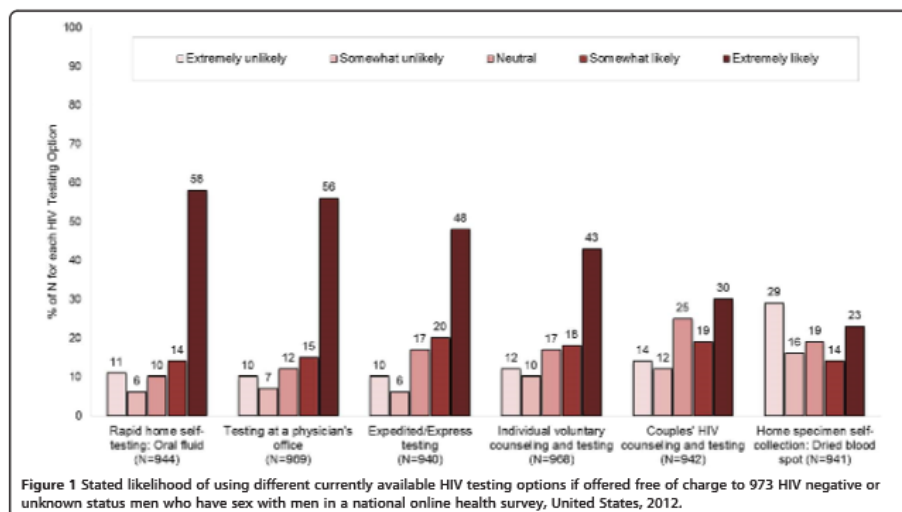
^bSample size (N) = 795.

^cHMO: Health maintenance organization.

^dIncludes 12 who tested in the Emergency Room and 19 who tested as an inpatient.

^eSexually transmitted disease.

^fIncludes 8 who tested in the military, 3 who tested at a correctional facility (jail or prison), 41 other and 8 unknown.



men would potentially employ in developing personalized testing strategies.

Across demographic and behavioral participant characteristics, MSM in our study generally reported being extremely or somewhat likely to use various testing modalities. Younger men significantly preferred expedited/express testing compared to older men. Possible explanations include the reduced time associated with this approach, not having to receive counseling, and the likely appeal of receiving results through text messages if so desired. Young US adults are avid users of text messaging, current statistics indicating that cell phone owners between 18–24 years exchange an average of 109.5 messages per day (Smith 2011). Previous research with MSM has found higher acceptability for rapid versus traditional testing in outreach settings, and an increase in testing when counseling was made optional (Spielberg et al. 2001, 2005). Compared to men of other racial/ethnic categories, non-Hispanic black men reported being less likely to use VCT. Black MSM's experiences with societal and institutional racism, coupled with a general distrust of the medical community and heightened perceptions of stigma, have posed personal and systemic barriers to them accessing HIV prevention resources (Malebranche et al. 2004). Although we agree with advocates of providing culturally competent counseling and testing services (Mimiaga et al. 2009; Nanin et al. 2009; CDC 2010), this result needs to be interpreted with caution due to the underrepresentation of black MSM in our study. Finally, lower educational levels were significantly

associated with a higher stated likelihood of using CHCT, the direction of this result being consistent with a recent study among internet-using MSM in seven countries (Stephenson et al. 2013). Another online study found that South African MSM with more schooling were significantly less likely to express willingness to utilize CHCT services (Stephenson et al. 2012). Possible explanations could include greater financial resources enabling more access to health care providers (Cutler and Lleras-Muney 2006), and a lower perceived or actual risk of acquiring HIV among better educated individuals (Jansen et al. 2011).

Given that rapid home self-testing and testing at a physician's office consensually emerged as the top ranked choices, these warrant consideration as key components of potential future combination HIV testing packages for MSM. Privacy, convenience, ease of specimen collection, almost instantaneous results, and not having to visit a testing facility have been reported as favorable attributes of rapid home oral fluid self-testing by high risk populations globally (Phillips and Chen 2003; Spielberg et al. 2003; Chen et al. 2010; Bilardi et al. 2013; Bavinton et al. 2013; Skolnik et al. 2001). Depending upon the kind of relationship and levels of trust MSM share with their physicians, some men may feel more comfortable getting tested at their doctor's office. Favorable attributes of this option that distinguish it from rapid home self-testing include the availability of in-person post-test counseling for newly diagnosed positives as well as the potential for early initiation of treatment. Expedited/express testing was consistently ranked as the third choice across risk groups in our study,

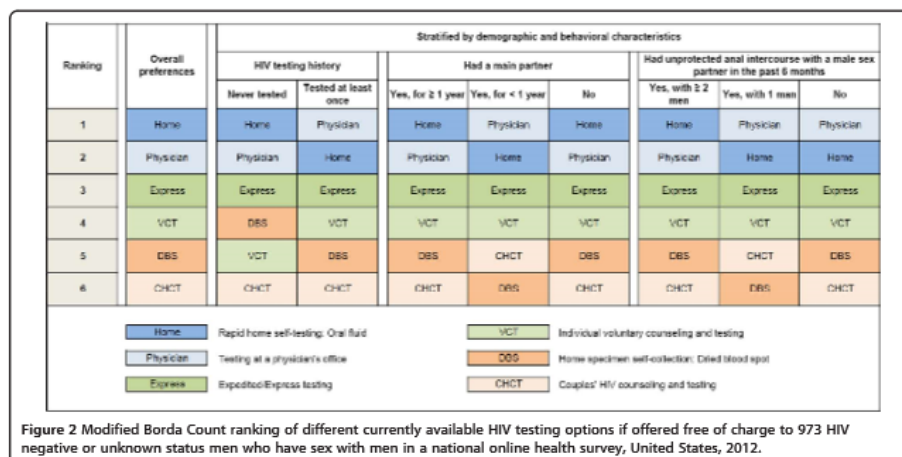
Table 3 Stated likelihood of using different HIV testing options if offered free of charge by selected demographic and behavioral characteristics of 973 HIV negative or unknown status MSM^a respondents to a national online health survey, United States, 2012

Characteristic	HIV testing option					
	Testing at a physician's office	Individual voluntary counseling and testing (VCT)	Couples' HIV counseling and testing (CHCT)	Expedited/express testing	Rapid home self-testing: Oral fluid	Home specimen self-collection: DBS ^b
	Median ^c (Mean ^d)					
Age group (years) ^e :						
18-24	5 (4.2)	4 (3.8)	4 (3.5)	4 (4.0)	5 (4.0)	3 (2.8)
25-34	5 (4.0)	4 (3.8)	3 (3.4)	5 (4.1)	5 (4.1)	3 (3.1)
35-44	5 (3.8)	4 (3.6)	3 (3.3)	4 (3.8)	5 (4.0)	3 (2.8)
≥ 45	5 (3.7)	4 (3.3)	3 (3.1)	4 (3.5)	5 (3.9)	3 (2.7)
Race/Ethnicity ^f :						
White, non-Hispanic	5 (4.0)	4 (3.6)	3 (3.3)	4 (3.9)	5 (4.0)	3 (2.9)
Black, non-Hispanic	5 (4.1)	3 (3.2)	4 (3.7)	5 (4.2)	2 (2.6)	1 (2.6)
Hispanic	5 (4.0)	5 (3.9)	4 (3.6)	4 (3.9)	5 (4.1)	3 (2.7)
Other ^g	5 (3.8)	5 (4.1)	4 (3.5)	5 (4.1)	5 (3.9)	3 (2.9)
Education ^h :						
College, Post graduate, or Professional school	5 (3.8)	4 (3.6)	3 (3.2)	4 (4.0)	5 (4.0)	3 (2.9)
Some college, Associate's degree, and/or Technical school	5 (4.1)	4 (3.8)	4 (3.5)	4 (3.9)	5 (4.0)	3 (2.8)
High school, GED ^h or less	5 (4.2)	4 (3.7)	4 (3.6)	4 (3.9)	5 (4.1)	3 (3.2)
Had a main partner ⁱ :						
Yes, for ≥ 1 year	5 (3.9)	4 (3.5)	3 (3.2)	4 (3.8)	5 (3.9)	3 (2.9)
Yes, for < 1 year	5 (4.1)	4 (3.7)	4 (3.6)	4 (3.9)	5 (4.0)	3 (2.7)
No	5 (4.0)	4 (3.9)	3 (3.4)	4 (4.0)	5 (4.1)	3 (3.0)
Had UAI ^j with a male sex partner in the past 6 months:						
Yes, with ≥ 2 men	5 (4.0)	4 (3.7)	3 (3.3)	5 (4.1)	5 (4.2)	3 (3.0)
Yes, with 1 man	5 (4.0)	4 (3.6)	3 (3.4)	4 (3.8)	5 (3.9)	3 (2.8)
No	5 (3.9)	4 (3.8)	4 (3.4)	4 (3.9)	5 (4.0)	3 (2.8)
HIV testing history:						
Never tested	5 (4.1)	4 (3.5)	3 (3.2)	4 (3.9)	5 (4.2)	3 (3.2)
Tested at least once	5 (4.0)	4 (3.7)	4 (3.4)	4 (3.9)	5 (4.0)	3 (2.8)

^aMSM: Men who have sex with men.^bDBS: Dried blood spot.^cFive-point Likert item format: 1 = Extremely unlikely, 2 = Somewhat unlikely, 3 = Neutral, 4 = Somewhat likely, 5 = Extremely likely.^dKruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using expedited/express testing differed by age group was significant ($P < 0.001$).^eKruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using VCT differed by race/ethnicity was significant ($P < 0.001$).^fIncludes 31 Asian/Pacific Islander, 12 Native American/Alaska Native, 36 multiracial, 9 other and 3 unknown.^gKruskal-Wallis nonparametric ANOVA testing whether stated likelihood of using CHCT differed by education was significant ($P < 0.001$).^hGED: General educational development.ⁱDefined as "Someone you feel committed to above all others. You might call this person your boyfriend, partner, significant other, spouse, or husband".^jUAI: Unprotected anal intercourse. Neither the respondent nor his partner used a condom.

higher than both VCT and CHCT, suggesting a niche role for this modality in frequent testing strategies. Previous studies among MSM have reported mixed reactions towards pre-test counseling, ranging from generally positive attitudes (Mimiaga et al. 2007) to considering it 'repetitive' and 'unnecessary' (Spielberg et al. 2001).

Despite only a quarter of our participants' negative perceptions about CHCT, comparable to online research with MSM in the US (Wagenaar et al. 2012), Australia and the United Kingdom (Stephenson et al. 2013), this approach ranked low in terms of intended usage preference when presented in conjunction with other testing



alternatives. Limited awareness about the intentions and content of this relatively new intervention for MSM (EffectiveInterventions 2012) could explain why even men in main partner relationships for longer than a year preferred other alternatives. DBS specimen self-collection kits for laboratory HIV testing have been unevenly adopted in the US. Although their acceptability and use in the context of research studies have been high (Sharma et al. 2011; Spielberg et al. 2000), this approach has had minimal impact on the testing behavior of high risk individuals due to concerns regarding privacy and accuracy (Colfax et al. 2002).

Strengths of our study include the evaluation of attitudes towards using six different HIV testing approaches presented collectively to a group of MSM recruited through the internet in a time, cost and resource efficient manner. Considering that online negotiations of both high-risk and safe sex have become increasingly prevalent among members of this community (Rosser et al. 2009; Horvath et al. 2008; Garofalo et al. 2007; Liao et al. 2006), we believe that understanding their testing preferences is critical in advancing internet-based HIV prevention efforts. Participants could only enter the online survey by clicking on banner advertisements displayed on Facebook, and because multiple surveys could not be completed from the same browser, it is unlikely that the same individual completed the survey more than once. People tend to be more open and honest while reporting sensitive risk behavior information using computer-based technologies compared to traditional questionnaires (Turner et al. 1998), thereby improving data accuracy and reducing the possibility of social desirability bias (Ellingson et al. 1999).

Limitations of our study include not being able to generalize to all MSM users of Facebook, users of other online social networks, or MSM in the general population. Because our banner advertisements were displayed only to men who had reported being interested in men in their Facebook profile, MSM who did not disclose their interest in men in their online profile were systematically underrepresented. One limitation of collecting data online is the inability to verify participants' self-reported demographic characteristics. Non-Hispanic black men comprised a smaller proportion of our sample relative to the general US population prevalence, an unfortunate reality that has plagued online research studies (Sullivan et al. 2011). Reduced access to and use of both basic and high-speed internet among black Americans compared to white or Hispanic individuals may explain this disparity (Smith 2010). Because questions in our survey involving a 6-month recall period were answered based on memory, our results could be subjected to recall bias. Although we collected data on participants' geographic region of residence, the lack of information regarding whether they lived in urban versus rural areas limited our ability to explore preference patterns within these strata. Additionally, usage intentions do not always translate into actions (Colfax et al. 2002), and the extent to which newer modalities will be adopted by MSM in research as well as real world settings is yet to be determined.

Despite these concerns, we believe that our results have important implications for future HIV prevention research. In this time of great challenge and opportunity, we envision an approach of combination testing packages to enable individuals form personalized HIV testing

strategies. The fact that MSM belonging to all demographic and behavioral risk strata in our study were willing to use most testing approaches is encouraging. Moreover, their order of intended usage preferences suggest that newer options such as rapid home self-testing could be incorporated as key components of comprehensive interventions to promote testing and increase serostatus awareness. Further research, especially among black MSM, is needed to explain the relative ranks assigned to these options and explore how different modalities can be packaged together. Given the challenges with recruiting racial and ethnic MSM online (Sullivan et al. 2011), additional in-person surveys or qualitative work with black MSM may be required to fully capture the perspectives of this critical population. Understanding circumstances in which men would use particular approaches, and how they would combine multiple options to test in a year is imperative. To this end, we are conducting qualitative research with MSM using online focus group discussions, the results of which we hope will provide an in-depth understanding of these issues. The efficacy of each approach in increasing HIV testing frequencies should be a high priority as part of developing comprehensive prevention strategies for MSM in the US.

Abbreviations

CDC: Centers for Disease Control and Prevention; CHCT: Couples' HIV Counseling and Testing; DBS: Dried Blood Spot; FDA: Food and Drug Administration; GED: General Educational Development; HIV: Human Immunodeficiency Virus; MBC: Modified Borda Count; MSM: Men who have Sex with Men; STI: Sexually Transmitted Infection; UAI: Unprotected Anal Intercourse; US: United States; VCT: Individual Voluntary Counseling and Testing.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AS participated in designing the study and collecting data, performed the statistical analyses, and drafted the manuscript. RBS participated in conceiving and designing the study, contributed to data interpretation, and helped draft the manuscript. DW participated in coordinating the study and collecting data, and helped draft the manuscript. PSS participated in conceiving and designing the study, contributed to data interpretation, and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was facilitated by the Center for AIDS Research at Emory University (P30 AI050409).

Author details

¹Department of Epidemiology, Emory University Laney Graduate School, 1518 Clifton Road NE, Atlanta, GA 30322, USA. ²Department of Epidemiology, Emory University Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322, USA. ³Hubert Department of Global Health, Emory University Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322, USA.

Received: 13 September 2013 Accepted: 20 February 2014
Published: 24 February 2014

References

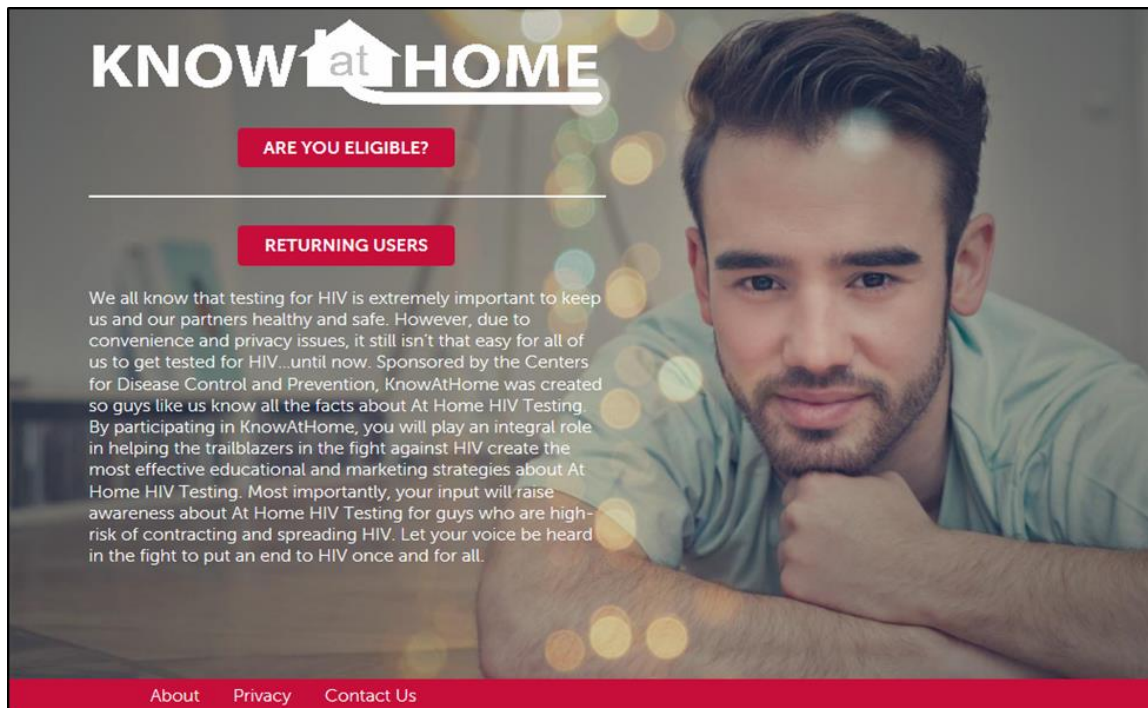
- Abdi H (2007) The Bonferroni and Sidák corrections for multiple comparisons. In: Encyclopedia of measurement and statistics. Sage, Thousand Oaks, CA
- Bavinton BR, Brown G, Hurley M, Bradley J, Keen P, Conway DP, Guy R, Grulich AE, Prestage G (2013) Which gay men would increase their frequency of HIV testing with home self-testing? *AIDS Behav* 17(6):2084–2092
- Billardi JE, Walker S, Read T, Prestage G, Chen MY, Guy R, Bradshaw C, Fairley CK (2013) Gay and Bisexual Men's Views on Rapid Self-Testing for HIV. *AIDS Behav* 17(6):2093–2099
- Borda J-C (1781) *Memoire sur les Elections au Scrutin*. Histoire de l'Academie Royale des Sciences, pp 657–664
- CDC (2005) HIV prevalence, unrecognized infection, and HIV testing among men who have sex with men—five U.S. cities, June 2004–April 2005. *MMWR Morb Mortal Wkly Rep* 54(24):597–601
- CDC (2006a) Human Immunodeficiency Virus (HIV) risk, prevention, and testing behaviors - United States, National HIV Behavioral Surveillance System: men who have sex with men, November 2003–April 2005. *MMWR Morb Mortal Wkly Rep* 55(SS-06):1–16
- CDC (2006b) Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Morb Mortal Wkly Rep* 55(RR-14):1–17
- CDC (2010) Sexually transmitted diseases treatment guidelines, 2010. *MMWR Morb Mortal Wkly Rep* 59(RR-12):12–13
- CDC (2011) HIV risk, prevention, and testing behaviors among men who have sex with men - National HIV Behavioral Surveillance System, 21 U.S. cities, United States, 2008. *MMWR Morb Mortal Wkly Rep* 60(SS-14):1–34
- CDC (2012) Estimated HIV incidence in the United States, 2007–2010. http://www.cdc.gov/hiv/pdf/statistics_hsr_vo1_17_no_4.pdf. Accessed 02/26/2014
- Chen M, Billardi J, Lee D, Cummings R, Bush M, Fairley C (2010) Australian men who have sex with men prefer rapid oral HIV testing over conventional blood testing for HIV. *Int J STD AIDS* 21(6):428–430
- Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, Hakim JG, Kumwenda J, Grinsztajn B, Pilloto JHS, Godbole SV, Miehendale S, Charlyalertsak S, Santos BR, Mayer KH, Hoffman IF, Eshleman SH, Piwovar-Manning E, Wang L, Makhema J, Mills LA, de Bruyn G, Sanne I, Eron J, Gallant J, Havir D, Swindells S, Ribaudo H, Elharrar V, Burns D, et al (2011) Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 365(6):498–505. doi:10.1056/NEJMoa1105243
- Colfax G, Lehman J, Bindman A, Vittinghoff E, Vranizan K, Fleming P, Chesney M, Osmond D, Hecht F (2002) What happened to home HIV test collection kits? Intent to use kits, actual use, and barriers to use among persons at risk for HIV infection. *AIDS Care* 14(5):673–682
- Conover W, Iman RL (1981) Rank transformations as a bridge between parametric and nonparametric statistics. *Am Stat* 35(3):124–129
- Copenhaver MM, Fisher JD (2006) Experts outline ways to decrease the decade-long yearly rate of 40,000 new HIV infections in the US. *AIDS Behav* 10(1):103–114
- Crepaz N, Marks G, Liaw A, Mullins MM, Aupont LW, Marshall KJ, Jacobs ED, Wolitski RJ (2009) Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis. *AIDS* 23(13):1617–1629
- Cutler DM, Uleras-Muney A (2006) Education and health: evaluating theories and evidence. National Bureau of Economic Research, Cambridge, MA
- Effectivelnterventions (2012) Couples HIV testing and counseling. <https://www.effectivelnterventions.org/en/HighImpactPrevention/PublicHealthStrategies/CHTC.aspx>. Accessed 07/10/2013
- Ellingson JE, Sackett PR, Hough LM (1999) Social desirability corrections in personality measurement: issues of applicant comparison and construct validity. *J Appl Psychol* 84:153–165
- FDA (1996) Letter - Home Access HIV-1 Test System. <http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMA/ucm091848.htm>. Accessed 04/12/2013
- FDA (2010) Approval letter - INSTI HIV-1 Antibody Test Kit. <http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMA/ucm235239.htm>. Accessed 08/14/2013
- FDA (2012) Approval letter - OraQuick In-Home HIV Test. <http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMA/ucm310592.htm>. Accessed 07/10/2013
- Gardner EM, McLees MP, Steiner JF, del Rio C, Burman WJ (2011) The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis* 52(6):799–800

- Garofalo R, Herrick A, Mustanski BS, Donenberg GR (2007) Tip of the iceberg: Young men who have sex with men, the Internet, and HIV risk. *Am J Public Health* 97(6):1113–1117
- Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, Karon J, Brookmeyer R, Kaplan EH, McKenna MT (2008) Estimation of HIV Incidence in the United States. *JAMA* 300(5):520–529
- Healthy People 2020 Summary of Objectives (2010) <http://healthypeople.gov/2020/topics/objectives/2020/pdfs/HIV.pdf>. Accessed 07/10/2013
- Horvath KJ, Rosser BS, Remafedi G (2008) Sexual risk taking among young Internet-using men who have sex with men. *Am J Public Health* 98(6):1059–1067
- Hutchinson AB, Corble-Smith G, Thomas SB, Mohanan S, del Rio C (2004) Understanding the patient's perspective on rapid and routine HIV testing in an inner-city urgent care center. *AIDS Educ Prev* 16(2):101–114
- Jansen IA, Gaskus RB, Davidovich U, Jurrians S, Coutinho RA, Prins M, Stolte IG (2011) Ongoing HIV-1 transmission among men who have sex with men in Amsterdam: a 25-year prospective cohort study. *AIDS* 25(4):493–501
- Katz DA, Swanson F, Stekler JD (2013) Why do men who have sex with men test for HIV infection? Results from a community-based testing program in Seattle. *Sex Transm Dis*. Published Ahead of Print
- Koblin BA (2004) Effects of a behavioural intervention to reduce acquisition of HIV infection among men who have sex with men: the EXPLORE randomised controlled study. *Lancet* 364(9428):41–50
- Koo DJ, Begler EM, Henn MH, Sepkowitz KA, Kelleman SE (2006) HIV counseling and testing: less targeting, more testing. *Am J Public Health* 96(6):962
- Liau A, Millett G, Marks G (2006) Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sex Transm Dis* 33(9):576–584
- Mackellar DA, Valeriy LA, Secura GM, Behel S, Bingham T, Celentano DD, Koblin BA, LaLota M, McFarland W, Shehan D (2005) Unrecognized HIV infection, risk behaviors, and perceptions of risk among young men who have sex with men: opportunities for advancing HIV prevention in the third decade of HIV/AIDS. *J Acquir Immune Defic Syndr* 38(5):603–614
- Mackellar DA, Hou S-I, Whalen CC, Samuelsen K, Sanchez T, Smith A, Denson D, Lansky A, Sullivan P (2011) Reasons for not HIV testing, testing intentions, and potential use of an over-the-counter rapid HIV test in an Internet sample of men who have sex with men who have never tested for HIV. *Sex Transm Dis* 38(5):419–428
- Malebranche DJ, Peterson JL, Fullilove RE, Stackhouse RW (2004) Race and sexual identity: perceptions about medical culture and healthcare among Black men who have sex with men. *J Natl Med Assoc* 96(1):97–107
- Marks G, Crepaz N, Senterfitt JW, Janssen RS (2005) Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: Implications for HIV prevention programs. *J Acquir Immune Defic Syndr* 39(4):446–453
- Marks G, Crepaz N, Janssen RS (2006) Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS* 20(10):1447–1450
- Mimlaga MJ, Goldhammer H, Belanoff C, Tetu AM, Mayer KH (2007) Men who have sex with men: perceptions about sexual risk, HIV and sexually transmitted disease testing, and provider communication. *Sex Transm Dis* 34(2):113–119
- Mimlaga MJ, Reinsner SL, Bland S, Skeer M, Cranston K, Isenberg D, Vega BA, Mayer KH (2009) Health system and personal barriers resulting in decreased utilization of HIV and STD testing services among at-risk black men who have sex with men in Massachusetts. *AIDS Patient Care STDS* 23(10):825–833
- Mimlaga MJ, Landers SJ, Connor KJ (2011) Prevalence and correlates of lifetime HIV testing in a population-based sample of men who have sex with men in Massachusetts. *AIDS Patient Care STDS* 25(6):323–326
- Nanin J, Osubu T, Walker JN, Powell B, Powell D, Parsons J (2009) "HIV is still real": perceptions of HIV testing and HIV prevention among black men who have sex with men in New York City. *Am J Mens Health* 3(2):150–164
- National HIV/AIDS Strategy for the United States (2010) <http://aids.gov/federal-resources/national-hiv-aids-strategy/nhas.pdf>. Accessed 07/10/2013
- Pai NP, Sharma J, Shivkumar S, Pillay S, Vadhani C, Joseph L, Dheda K, Peeling RW (2013) Supervised and unsupervised self-testing for HIV in high- and low-risk populations: a systematic review. *PLoS Med* 10(4):e1001414
- Phillips KA, Chen JL (2003) Willingness to use instant home HIV tests: data from the California Behavioral Risk Factor Surveillance Survey. *American Journal of Preventive Medicine* 24(4):340–348
- Purcell DW, Johnson CH, Lansky A, Prejean J, Stein R, Denning P, Gau Z, Weinstock H, Su J, Crepaz N (2012) Estimating the population size of men who have sex with men in the United States to obtain HIV and syphilis rates. *Open AIDS J* 6(1):98–107
- Rosser BS, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL (2009) HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's Internet Sex Study-II (MINTS-II). *AIDS Behav* 13(3):488–498
- SAS (2011) Version 9.3. SAS Institute Inc, Cary, NC
- Sharma A, Sullivan PS, Khosropour CM (2011) Willingness to take a free home HIV test and associated factors among Internet-using men who have sex with men. *J Int Assoc Physicians AIDS Care* 10(6):357–364
- Skolnik HS, Phillips KA, Binson D, Dilley JW (2001) Deciding where and how to be tested for HIV: what matters most? *J Acquir Immune Defic Syndr* 27(3):292–300
- Smith A (2010) Home broadband 2010. Pew Research Center. <http://www.pewinternet.org/Reports/2010/Home-Broadband-2010.aspx>. Accessed 07/15/2013
- Smith A (2011) Americans and text messaging. Pew Research Center. <http://www.pewinternet.org/Reports/2011/Cell-Phone-Texting-2011.aspx>. Accessed 07/12/2013
- Spielberg F, Crichtlow C, Vittinghoff E, Coletti AS, Sheppard H, Mayer KH, Metzger D, Judson FN, Buchbinder S, Chesney M (2000) Home collection for frequent HIV testing: acceptability of oral fluids, dried blood spots and telephone results. *AIDS* 14(12):1819–1828
- Spielberg F, Kurth A, Gorbach PM, Goldbaum G (2001) Moving from apprehension to action: HIV counseling and testing preferences in three at-risk populations. *AIDS Educ Prev* 13(6):524–540
- Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, Celum CL, Rossini A, Crichtlow CW, Wood RW (2003) Overcoming barriers to HIV testing: preferences for new strategies among clients of a needle exchange, a sexually transmitted disease clinic, and sex venues for men who have sex with men. *J Acquir Immune Defic Syndr* 32(3):318–327
- Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, Rossini A, Wood RW (2005) Choosing HIV counseling and testing strategies for outreach settings: a randomized trial. *J Acquir Immune Defic Syndr* 38(3):348–355
- Stephenson R, Sullivan PS, Salazar LF, Gratzler B, Allen S, Seelbach E (2011) Attitudes towards couples-based HIV testing among MSM in three US cities. *AIDS Behav* 15(1):80–87
- Stephenson R, Rentsch C, Sullivan P (2012) High levels of acceptability of couples-based HIV testing among MSM in South Africa. *AIDS Care* 24(4):529–535
- Stephenson R, Chard A, Finneran C, Sullivan P (2013) Willingness to use couples voluntary counseling and testing services among men who have sex with men in seven countries. *AIDS Care*:1–8 (ahead of print)
- Sullivan PS, Khosropour CM, Lutsi N, Amsden M, Coggia T, Wingood GM, DiClemente RJ (2011) Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. *J Med Internet Res* 13(2):8
- Sullivan PS, Carballo-Dieguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, Smith A, Goswami P, Sanchez J (2012) Successes and challenges of HIV prevention in men who have sex with men. *Lancet* 380(9839):388–399
- Sullivan PS, White D, Rosenberg E, Barnes J, Jones J, Dasgupta S, O'Hara B, Scales L, Salazar LF, Wingood G, DiClemente RJ, Wall K, Hoff C, Gratzler B, Allen S, Stephenson R (2013) Safety and acceptability of couples HIV testing and counseling for US men who have sex with men: a randomized prevention study. *J Int Assoc Provid AIDS Care*. In Press
- Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL (1998) Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science* 280(5365):867–873
- UNAIDS (2001) The impact of voluntary counselling and testing: a global review of the benefits and challenges. http://data.unaids.org/Publications/IRC-pub02/jc580-vct_en.pdf. Accessed 04/09/2013
- Wagenaar BH, Christiansen-Lindquist L, Khosropour C, Salazar LF, Benbow N, Prachand N, Sineath RC, Stephenson R, Sullivan PS (2012) Willingness of US men who have sex with men (MSM) to participate in couples HIV voluntary counseling and testing (CVCT). *PLoS One* 7(8):e42953

doi:10.1186/2193-1801-3-109

Cite this article as: Sharma et al.: Acceptability and intended usage preferences for six HIV testing options among Internet-using men who have sex with men. *SpringerPlus* 2014 **3**:109.

A3: KnowAtHome.org landing page

The landing page features a background image of a man with a beard and short dark hair, wearing a light blue shirt, resting his chin on his hand. The page has a dark grey background with a red footer bar. The logo 'KNOW at HOME' is at the top left, with 'at' in a white house icon. Below the logo is a red button with white text 'ARE YOU ELIGIBLE?'. A horizontal line follows. Below that is another red button with white text 'RETURNING USERS'. The main text is a paragraph explaining the importance of HIV testing and the role of the user. The footer bar contains three links: 'About', 'Privacy', and 'Contact Us' in white text.

KNOW at HOME

ARE YOU ELIGIBLE?

RETURNING USERS

We all know that testing for HIV is extremely important to keep us and our partners healthy and safe. However, due to convenience and privacy issues, it still isn't that easy for all of us to get tested for HIV...until now. Sponsored by the Centers for Disease Control and Prevention, KnowAtHome was created so guys like us know all the facts about At Home HIV Testing. By participating in KnowAtHome, you will play an integral role in helping the trailblazers in the fight against HIV create the most effective educational and marketing strategies about At Home HIV Testing. Most importantly, your input will raise awareness about At Home HIV Testing for guys who are high-risk of contracting and spreading HIV. Let your voice be heard in the fight to put an end to HIV once and for all.

[About](#) [Privacy](#) [Contact Us](#)

A4: Survey questions on the awareness and previous use of home HIV tests, and willingness to potentially self-test, distribute and test with friends or sex partners using rapid home test kits

Awareness & previous use of home HIV tests

A home diagnostic test is one you can buy at a store or online, and use to test yourself for cholesterol, blood glucose, etc.

OT-1. Have you ever used any home diagnostic test before?

No

Yes

I prefer not to answer

OT-2. Have you ever heard about home HIV tests? A home HIV test is one that you can use to test yourself for HIV at home or another private location.

No

Yes

I prefer not to answer

If OT-2= "No" skip to next section.

OT-3. In the past 12 months, have you used a home HIV test?

No

Yes

I prefer not to answer

If OT-3 = "Yes", skip to OT-6.

OT-4. What are some of your reasons for not using a home HIV test? Check all that apply.

- I'm concerned about the cost of a home HIV test
- I'm concerned about the accuracy of such a test
- I'm concerned I would not be able to perform the test correctly
- I'm concerned I would not be able to read the result properly
- I do not know where to get a home HIV test kit
- I would rather talk to a counselor when I get an HIV test
- I do not want to stick my finger to get a drop of blood
- I do not want to mail my blood sample to a lab
- I got tested at a different location such as my doctor's office
- I'm afraid of finding out that I have HIV
- Other reason (Specify _____)
- I prefer not to answer

OT-5. Among the reasons you indicated, what is the main reason you have not used a home HIV test? Choose only one.

Display response options based on selections in OT-4.

If OT-3 ≠ "Yes", skip to next section.

OT-6. In the past 12 months, which of these home HIV tests have you used? Check all that apply.

- Home Access® HIV-1 Test System (where you prick your finger, collect a blood sample on a card and mail that card to a lab for testing)
- OraQuick® In-Home HIV Test (where you collect your own oral fluid sample by swabbing your mouth, use the testing device yourself and read test results within 20 minutes)
- Other rapid HIV test (Specify _____)

OT-7. In the past 12 months, how many times have you used a home HIV test?

___ (1 to N)

I prefer not to answer

OT-8. You mentioned you used [insert option from OT-6] to test yourself for HIV. Where did you get the home HIV test(s)? Check all that apply.

Bought Online

From a pharmacy

From a friend

From a sex partner

From my doctor's office

From an HIV counseling and testing site

From a Health Department

From another research study

Other location or person (Specify_____)

I prefer not to answer

If participant chose more than one option in OT-6, loop back to ask OT-8 for each option.

OT-9. People use home HIV tests for many different reasons. In the past 12 months, did you use the home HIV test(s) for any of the following reasons? Check all that apply.

It was more convenient than getting tested by a doctor or at an HIV testing site

It was more private than getting tested by a doctor or at an HIV testing site

I didn't want other people to know I am testing

To test together with someone, before having sex

To test myself, before having sex

To test myself, after having sex

A sex partner asked me to take a home HIV test

Other reason (Specify_____)

I prefer not to answer

Willingness to potentially self-test, distribute and test with others

We will be mailing you free, anonymous rapid HIV tests as part of this online study. You can use these to test yourself at home or another private location and read your results within a few minutes. The package will contain two types of tests:

Option 1: Here you could collect your own oral fluid sample by swabbing your mouth, use the testing device yourself and read your HIV test results within 20 minutes.

Option2: Here you could prick your finger to collect a blood sample, use the testing device yourself and read your HIV test results within 15 minutes.

If any test were to come back positive, you would be able to call a toll-free, 24-hour hotline and get immediate counseling.

IQ-1. How likely would you be to test yourself using the free, anonymous oral fluid rapid HIV test (Option 1)?

- Very likely
- Somewhat likely
- Not sure
- Somewhat unlikely
- Very unlikely
- I prefer not to answer

IQ-2. How likely would you be to test yourself using the free, anonymous finger-stick blood rapid HIV test (Option 2)?

- Very likely
- Somewhat likely
- Not sure
- Somewhat unlikely
- Very unlikely
- I prefer not to answer

If IQ-1 = "Very likely" or "Somewhat likely", and IQ-2 = "Very likely" or "Somewhat likely", skip to IQ-5.

IQ-3. What are some of your reasons for not being willing to use a free, anonymous rapid HIV test? Check all that apply.

I'm afraid of finding out that I have HIV

I live with people who might see the test kits arrive by mail

I live with people who might see me performing an HIV test

I do not want to swab my mouth to collect an oral fluid sample

I do not want to prick my finger to collect a blood sample

I'm concerned about the accuracy of an oral fluid rapid HIV test

I'm concerned about the accuracy of a finger-stick blood rapid HIV test

I'm concerned I would not be able to perform the test correctly

I'm concerned I would not be able to read the result properly

I would rather talk to a counselor when I get an HIV test

I would rather be tested by someone who is trained to conduct the test

Other reason (Specify _____)

I prefer not to answer

IQ-4. Among the reasons you indicated, what is the main reason you are not willing to use a free, anonymous rapid HIV test? Choose only one.

Display response options based on selections in IQ-3.

Now imagine you could give away free, anonymous rapid HIV tests to your friends or sex partners. They can use these to test themselves at home or another private location, and read their results within a few minutes. Suppose the same two types of tests were available:

Option 1: Here they could collect their own oral fluid sample by swabbing their mouth, use the testing device themselves and read their HIV test results within 20 minutes.

Option2: Here they could prick their finger to collect a blood sample, use the testing device themselves and read their HIV test results within 15 minutes.

If any test were to come back positive, they would be able to call a toll-free, 24-hour hotline and get immediate counseling.

IQ-5. How likely would you be to give away a free, anonymous oral fluid rapid HIV test (Option 1) to your friends or sex partners?

- Very likely
- Somewhat likely
- Not sure
- Somewhat unlikely
- Very unlikely
- I prefer not to answer

IQ-6. How likely would you be to give away a free, anonymous finger-stick blood rapid HIV test (Option 2) to your friends or sex partners?

- Very likely
- Somewhat likely
- Not sure
- Somewhat unlikely
- Very unlikely
- I prefer not to answer

If IQ-5 = "Very likely" or "Somewhat likely", and IQ-6 = "Very likely" or "Somewhat likely", skip to IQ-9.

IQ-7. What are some of your reasons for not being willing to give away free, anonymous rapid HIV tests to your friends or sex partners? Check all that apply.

- I'm concerned this might affect our friendship
- I'm concerned this might affect our sexual relationship
- I think they would get upset or angry
- I'm afraid they would think I have HIV
- I'm concerned about the accuracy of an oral fluid rapid HIV test
- I'm concerned about the accuracy of a finger-stick blood rapid HIV test
- I'm concerned they would not be able to perform the test correctly
- I'm concerned they would not be able to read the result properly
- I would rather they talk to a counselor when they get an HIV test
- I would rather they get tested by someone who is trained to conduct the test

Other reason (Specify_____)

I prefer not to answer

IQ-8. Among the reasons you indicated, what is the main reason you are not willing to give away free, anonymous rapid HIV tests to your friends or sex partners? Choose only one.

Display response options based on selections in IQ-7.

If IQ-5 ≠ “Very likely” or “Somewhat likely”, and IQ-6 ≠ “Very likely” or “Somewhat likely”, skip to IQ-10.

IQ-9. To whom would you most likely give away free, anonymous rapid HIV tests? Check all that apply.

Main sex partner (someone you feel committed to above all others)

Casual sex partner (someone you do not feel committed to above all others)

One-time sex partner (someone you have sex with only once and probably never again)

Exchange sex partner (someone you have sex with in exchange for money, food or drugs)

Family member (who is not a sex partner)

Friend (who is not a sex partner)

Stranger (who is not a sex partner)

Acquaintance (who is not a sex partner)

Someone else (Specify_____)

I prefer not to answer

IQ-10. How likely would you be to test yourself together with your friends or sex partners using free, anonymous rapid HIV tests? This means that both of you take the rapid test together in the same place.

Very likely

Somewhat likely

Not sure

Somewhat unlikely

Very unlikely

I prefer not to answer

If IQ-10 = “Very likely” or “Somewhat likely”, skip to IQ-13.

IQ-11. What are some of the reasons you would not test yourself together with your friends or sex partners using free, anonymous rapid HIV tests? Check all that apply.

I do not want them to find out my test result

I do not want to find out their test result

I’m concerned they would tell someone my test result

I think they would get upset or angry

I’m concerned this might affect our friendship

I’m concerned this might affect our sexual relationship

I’m concerned we might be drunk or high on drugs

Other reason (Specify_____)

I prefer not to answer

IQ-12. Among the reasons you indicated, what is the main reason you are not willing to test yourself together with your friends or sex partners using free, anonymous rapid HIV tests? Choose only one.

Display response options based on selections in IQ-11.

If IQ-10 ≠ “Very likely” or “Somewhat likely”, skip to IQ-15.

IQ-13. With whom would you most likely test together using free, anonymous rapid HIV tests? Check all that apply.

Main sex partner (someone you feel committed to above all others)

Casual sex partner (someone you do not feel committed to above all others)

One-time sex partner (someone you have sex with only once and probably never again)

Exchange sex partner (someone you have sex with in exchange for money, food or drugs)

Family member (who is not a sex partner)

Friend (who is not a sex partner)

Stranger (who is not a sex partner)

Acquaintance (who is not a sex partner)

Other person (Specify_____)

I prefer not to answer

IQ-14. Where would you be most likely to test yourself together with your friends or sex partners using free, anonymous rapid HIV tests? Check all that apply.

At my house

At their house

At a bar or dance club

At a bathhouse or sex club

At a private sex party

At a circuit party

In a car

Other location (Specify_____)

I prefer not to answer

A5: SAS code for simulations and subsequent analyses

ScenarioA_Specifications135.sas

```

*Suppress listing of log;
options nonotes nosource nosource2;

*Macro - Program to run simulations multiple times using different
parameters
Macro name: simtrialmacro
Macro variables: seed (seed)
                    nsim (number of simulations)
                    m (number of participants in the trial)
                    beta0 (log odds of testing >= 3 times in a
year for Comparison)
                    betalor (beta1 depending upon the odds ratio)
                    timeseed (seed)
                    corrstruc (working correlation structure)
Syntax: Use ampersand (&) before variable names within a macro;

%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseed,
corrstruc);

*Creating dataset Simtrial;

data simtrial;

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

do z=1 to &nsim; *Number of simulations;

    do i=1 to &m; *Person ID;

        group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

        *Control variables from CheckingIn data;
        race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
        age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
        educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
        income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;

```



```

        residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
        orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

        subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

        *Assuming different Odds Ratios (Odds of testing >= 3 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
        lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 3 times in a year;
        odds=round(exp(lnodds),0.00001); *Odds of testing >= 3
times in a year;
        probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 3 times in a year;

        *Outcome for Standard Logistic Regression;
        test3ormore=rand('bernoulli',probtest); *Tested >= 3 times
in a year (based on probability of testing) - Yes(1)/No(0);

        if test3ormore=0 then
howmany=rand('table',0.44855,0.35209,0.19936);
        else howmany=rand('table',0.58511,0.23404,0.09574,0.08511);

        *Annual testing frequency;
        if test3ormore=0 and howmany=1 then annualtests=0;
        if test3ormore=0 and howmany=2 then annualtests=1;
        if test3ormore=0 and howmany=3 then annualtests=2;
        if test3ormore=1 and howmany=1 then annualtests=3;
        if test3ormore=1 and howmany=2 then annualtests=4;
        if test3ormore=1 and howmany=3 then annualtests=5;
        if test3ormore=1 and howmany=4 then annualtests=6;

output;

        end;

end;

run;

*Creating new dataset SimtrialLogistic;

data simtriallogistic;
set simtrial;

keep z i group race age educ income residence orientation test3ormore
annualtests;

run;

*Creating new dataset SimtrialConditional;

```

```

data simtrialcox0;
set simtrial;

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

    if j=0 then timeoftest=0;
    else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
    if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

    *Calendar time between each test;
    lagtime=lag(timeoftest);
    if first.i then timebetweentests=timeoftest;
    if not first.i then timebetweentests=timeoftest-lagtime;
    if timeoftest=0 then timebetweentests=0;

run;

proc sort data=simtrialcox1;
by z i timeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i timeoftest;

    if timeoftest=0 and timebetweentests=0 then delete;

    *Variables for Conditional model 1;
    start1=lagtime;
    stop1=timeoftest;

    *Variables for Conditional model 2;
    start2=0;
    stop2=timebetweentests;

run;

```

```

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    *Outcome for Stratified Cox PH Regression for Recurrent Events;
    testingevent=1;
    if annualtests ne intervalph and stop1=365 then testingevent=0;

keep z i group race age educ income residence orientation test3ormore
annualtests start1 stop1 start2 stop2 testingevent intervalph;

run;

*Creating new dataset SimtrialRepeated;

data simtrialgee0;
set simtrialconditional;

    *Interval in which person tested;
    if 1<=start1<=90 then testinterval3=1;
    else testinterval3=0;
    if 91<=start1<=181 then testinterval6=1;
    else testinterval6=0;
    if 182<=start1<=273 then testinterval9=1;
    else testinterval9=0;
    if 274<=start1<=365 then testinterval12=1;
    else testinterval12=0;

run;

proc sort data=simtrialgee0;
by z i;
run;

data simtrialgeel;
set simtrialgee0;
by z i;

    if first.i then testnum0to3=0;
    testnum0to3+testinterval3;

    if first.i then testnum3to6=0;
    testnum3to6+testinterval6;

    if first.i then testnum6to9=0;
    testnum6to9+testinterval9;

```

```

        if first.i then testnum9to12=0;
        testnum9to12+testintervall12;

run;

proc sort data=simtrialgee1;
by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

        if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialrepeated;
set simtrialgee2;
by z i;

        do intervalgee=1 to 4;

                *Outcome for GEE for Correlated Data;
                testedp3mo=0;
                if intervalgee=1 and testnum0to3>0 then testedp3mo=1;
                if intervalgee=2 and testnum3to6>0 then testedp3mo=1;
                if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
                if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

keep z i group race age educ income residence orientation test3ormore
annualtests testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox0 simtrialcox1 simtrialcox2 simtrialgee0 simtrialgee1
simtrialgee2;
run;

ods select all;

```

```

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

    if group=0 then group=2;
    if test3ormore=0 then test3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))
  Creating output dataset Chi with values for estimates of Relative risk
and their CIs;

ods select none;

proc freq data=test2by2;
table group*test3ormore / all;
by z;
output out = test2by2res relrisk;
run;

ods select all;

*Power - Proportion of those CIs that DO NOT include 1 where Relative
risk > 1;

data power2by2;
set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;

run;

ods select none;

proc freq data=power2by2;
tables powerpr / out=a;
run;

ods select all;

*Power and sample size - Standard Logistic Regression;

ods select none;

proc logistic data=simtriallogistic descending;
model test3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;

```

```

run;

ods select all;

data powerlogisticcrude;
set testlogisticcrude;

    if effect ne "group" then delete;

    if lowercl gt 1 then powerslrcrude=1;
    else powerslrcrude=2;

run;

ods select none;

proc freq data=powerlogisticcrude;
tables powerslrcrude / out=b;
run;

ods select all;

*Power and sample size - GEE for Correlated Data;

ods select none;

proc genmod data=simtrialrepeated descending;
class i intervalgee;
model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;
ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

    if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

    if lbetalowercl gt 1 then powergeerepcrude=1;
    else powergeerepcrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepcrude / out=d;
run;

ods select all;

```

```
*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;
```

```
*Conditional model 1 - Models the full time course of the recurrent
event process
```

```
(Time interval of a subsequent event
starts at the end of the time interval for the previous event);
```

```
ods select none;
```

```
proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional1crude;
run;
```

```
ods select all;
```

```
data powerconditional1crude;
set testconditional1crude;
```

```
if parameter ne "group" then delete;
```

```
if hrlowercl gt 1 then powerphreg1crude=1;
else powerphreg1crude=2;
```

```
run;
```

```
ods select none;
```

```
proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;
```

```
ods select all;
```

```
*Conditional model 2 - Models the time between each of the recurring
events
```

```
(Time interval starts at 0 and ends at
the length of time until the next event);
```

```
ods select none;
```

```
proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;
```

```
ods select all;
```

```
data powerconditional2crude;
```

```

set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepcrude=1 then powerlabel="Power GEE Crude";
    if powerphreg1crude=1 then powerlabel="Power Cox PH Conditional I
Crude";
    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

    power=percent;

run;

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

proc print data=crudepowerresults noobs;
title "Crude Analyses: &m Participants, Beta0 = &beta0, Beta1 =
&beta1or ";
run;

title;

%mend;

*****;

*Assuming OR=2 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 1;

```



```
%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 1, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\100checker2.csv';  
run;  
  
%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 2, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\200checker2.csv';  
run;  
  
%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 3, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\300checker2.csv';  
run;  
  
%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 4, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\400checker2.csv';  
run;  
  
%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 5, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\500checker2.csv';  
run;  
  
%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 6, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\600checker2.csv';  
run;  
  
%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 7, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\700checker2.csv';  
run;  
  
%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 8, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\800checker2.csv';  
run;  
  
%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 9, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR  
2\900checker2.csv';  
run;  
  
%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 10, ar(1));
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1000checkor2.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 11, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1100checkor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 12, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1200checkor2.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 13, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1300checkor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1400checkor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1500checkor2.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1600checkor2.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1700checkor2.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1800checkor2.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\1900checkor2.csv';
```

```

run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 20, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
2\2000checkor2.csv';
run;

*****;

*Assuming OR=2.5 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 3;

%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 1, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\100checkor25.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 2, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\200checkor25.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 3, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\300checkor25.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 4, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\400checkor25.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 5, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\500checkor25.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 6, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\600checkor25.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 7, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\700checkor25.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 8, ar(1));

```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\800checkor25.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 9, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\900checkor25.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 10, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1000checkor25.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 11, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1100checkor25.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 12, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1200checkor25.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 13, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1300checkor25.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1400checkor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1500checkor25.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1600checkor25.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1700checkor25.csv';
```

```

run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1800checker25.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\1900checker25.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 20, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CheckingIn OR
25\2000checker25.csv';
run;

*****;

*Assuming OR=3 and 5% of Comparison arm participants test >= 3 times in
a year - Specification 5;

%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 1, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\100cdc3.csv';
run;

%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 2, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\200cdc3.csv';
run;

%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 3, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\300cdc3.csv';
run;

%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 4, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\400cdc3.csv';
run;

%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 5, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\500cdc3.csv';
run;

%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 6, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\600cdc3.csv';
run;

%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 7, ar(1));

```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\700cdcor3.csv';
run;

%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 8, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\800cdcor3.csv';
run;

%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 9, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\900cdcor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 10, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1000cdcor3.csv';
run;

%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 11, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1100cdcor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 12, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1200cdcor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 13, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1300cdcor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1400cdcor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1500cdcor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1600cdcor3.csv';
run;

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1700cdcor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1800cdcor3.csv';
```

```

run;

%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\1900cdcor3.csv';
run;

%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 20, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Crude CDC OR 3\2000cdcor3.csv';
run;

*****;

*Start listing log;
options notes source source2;

```

ScenarioA_Specifications246.sas

```

*Suppress listing of log;
options nonotes nosource nosource2;

*Macro - Program to run simulations multiple times using different
parameters
Macro name: simtrialmacro
Macro variables: seed (seed)
                    nsim (number of simulations)
                    m (number of participants in the trial)
                    beta0 (log odds of testing >= 2 times in a
year for Comparison)
                    betalor (beta1 depending upon the odds ratio)
                    timeseed (seed)
                    corrstruc (working correlation structure)
Syntax: Use ampersand (&) before variable names within a macro;

%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseed,
corrstruc);

*Creating dataset Simtrial;

data simtrial;

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

do z=1 to &nsim; *Number of simulations;

```

```

do i=1 to &m; *Person ID;

    group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

    *Control variables from CheckingIn data;
    race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
    age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
    educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
    income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;
    residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
    orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

    subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

    *Assuming different Odds Ratios (Odds of testing >= 2 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
    lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 2 times in a year;
    odds=round(exp(lnodds),0.00001); *Odds of testing >= 2
times in a year;
    probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 2 times in a year;

    test2ormore=rand('bernoulli',probtest); *Tested >= 2 times
in a year (based on probability of testing) - Yes(1)/No(0);

    if test2ormore=0 then
howmany=rand('table',0.56024,0.43976);
    else
howmany=rand('table',0.56881,0.25229,0.10092,0.04128,0.03669);

    *Annual testing frequency;
    if test2ormore=0 and howmany=1 then annualtests=0;
    if test2ormore=0 and howmany=2 then annualtests=1;
    if test2ormore=1 and howmany=1 then annualtests=2;
    if test2ormore=1 and howmany=2 then annualtests=3;
    if test2ormore=1 and howmany=3 then annualtests=4;
    if test2ormore=1 and howmany=4 then annualtests=5;
    if test2ormore=1 and howmany=5 then annualtests=6;

    *Outcome for Standard Logistic Regression - Still keep at 3
or more times;
    if annualtests=0 then test3ormore=0;
    if annualtests=1 then test3ormore=0;
    if annualtests=2 then test3ormore=0;
    if annualtests=3 then test3ormore=1;
    if annualtests=4 then test3ormore=1;
    if annualtests=5 then test3ormore=1;

```



```

        if annualtests=6 then test3ormore=1;

output;

        end;

end;

run;

*Creating new dataset SimtrialLogistic;

data simtriallogistic;
set simtrial;

keep z i group race age educ income residence orientation test3ormore
annualtests;

run;

*Creating new dataset SimtrialConditional;

data simtrialcox0;
set simtrial;

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

        if j=0 then timeoftest=0;
        else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
        if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

        *Calendar time between each test;
        lagtime=lag(timeoftest);
        if first.i then timebetweentests=timeoftest;
        if not first.i then timebetweentests=timeoftest-lagtime;
        if timeoftest=0 then timebetweentests=0;

```

```

run;

proc sort data=simtrialcox1;
by z i timeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i timeoftest;

    if timeoftest=0 and timebetweentests=0 then delete;

    *Variables for Conditional model 1;
    start1=lagtime;
    stop1=timeoftest;

    *Variables for Conditional model 2;
    start2=0;
    stop2=timebetweentests;

run;

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    *Outcome for Stratified Cox PH Regression for Recurrent Events;
    testingevent=1;
    if annualtests ne intervalph and stop1=365 then testingevent=0;

keep z i group race age educ income residence orientation test3ormore
annualtests start1 stop1 start2 stop2 testingevent intervalph;

run;

*Creating new dataset SimtrialRepeated;

data simtrialgee0;
set simtrialconditional;

    *Interval in which person tested;
    if 1<=start1<=90 then testinterval3=1;
    else testinterval3=0;
    if 91<=start1<=181 then testinterval6=1;
    else testinterval6=0;
    if 182<=start1<=273 then testinterval9=1;

```

```

        else testinterval9=0;
        if 274<=start1<=365 then testinterval12=1;
        else testinterval12=0;

run;

proc sort data=simtrialgee0;
by z i;
run;

data simtrialgee1;
set simtrialgee0;
by z i;

        if first.i then testnum0to3=0;
        testnum0to3+testinterval3;

        if first.i then testnum3to6=0;
        testnum3to6+testinterval6;

        if first.i then testnum6to9=0;
        testnum6to9+testinterval9;

        if first.i then testnum9to12=0;
        testnum9to12+testinterval12;

run;

proc sort data=simtrialgee1;
by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

        if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialrepeated;
set simtrialgee2;
by z i;

        do intervalgee=1 to 4;

                *Outcome for GEE for Correlated Data;
                testedp3mo=0;
                if intervalgee=1 and testnum0to3>0 then testedp3mo=1;
                if intervalgee=2 and testnum3to6>0 then testedp3mo=1;

```

```

        if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
        if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

keep z i group race age educ income residence orientation test3ormore
annualtests testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox0 simtrialcox1 simtrialcox2 simtrialgee0 simtrialgee1
simtrialgee2;
run;

ods select all;

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

    if group=0 then group=2;
    if test3ormore=0 then test3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))
Creating output dataset Chi with values for estimates of Relative risk
and their CIs;

ods select none;

proc freq data=test2by2;
table group*test3ormore / all;
by z;
output out = test2by2res relrisk;
run;

ods select all;

*Power - Proportion of those CIs that DO NOT include 1 where Relative
risk > 1;

data power2by2;

```

```

set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;

run;

ods select none;

proc freq data=power2by2;
tables powerpr / out=a;
run;

ods select all;

*Power and sample size - Standard Logistic Regression;

ods select none;

proc logistic data=simtriallogistic descending;
model test3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;
run;

ods select all;

data powerlogisticcrude;
set testlogisticcrude;

    if effect ne "group" then delete;

    if lowercl gt 1 then powerslr crude=1;
    else powerslr crude=2;

run;

ods select none;

proc freq data=powerlogisticcrude;
tables powerslr crude / out=b;
run;

ods select all;

*Power and sample size - GEE for Correlated Data;

ods select none;

proc genmod data=simtrialrepeated descending;
class i intervalgee;
model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;

```

```

ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

    if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

    if lbetalowercl gt 1 then powergeerepcrude=1;
    else powergeerepcrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepcrude / out=d;
run;

ods select all;

*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;

*Conditional model 1 - Models the full time course of the recurrent
event process
                                (Time interval of a subsequent event
starts at the end of the time interval for the previous event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditionalcrude;
run;

ods select all;

data powerconditionalcrude;
set testconditionalcrude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreglcrude=1;
    else powerphreglcrude=2;

run;

ods select none;

```

```

proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;

ods select all;

*Conditional model 2 - Models the time between each of the recurring
events
                                (Time interval starts at 0 and ends at
the length of time until the next event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;

ods select all;

data powerconditional2crude;
set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepcrude=1 then powerlabel="Power GEE Crude";
    if powerphreg1crude=1 then powerlabel="Power Cox PH Conditional I
Crude";
    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

    power=percent;

```

```

run;

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

proc print data=crudepowerresults noobs;
title "Crude Analyses: &m Participants, Beta0 = &beta0, Beta1 =
&betalor ";
run;

title;

%mend;

*****;

*Assuming OR=2 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 2;

%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 1, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\100m2checker2.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 2, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\200m2checker2.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 3, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\300m2checker2.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 4, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\400m2checker2.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 5, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\500m2checker2.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 6, ar(1));

```



```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\600m2checkor2.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 7, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\700m2checkor2.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 8, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\800m2checkor2.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 9, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\900m2checkor2.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 10, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1000m2checkor2.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 11, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1100m2checkor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 12, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1200m2checkor2.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 13, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1300m2checkor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1400m2checkor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1500m2checkor2.csv';
```

```

run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1600m2checkor2.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1700m2checkor2.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1800m2checkor2.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\1900m2checkor2.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 20, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
2\2000m2checkor2.csv';
run;

*****;

*Assuming OR=2.5 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 4;

%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 1, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\100m2checkor25.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 2, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\200m2checkor25.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 3, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\300m2checkor25.csv';
run;

```

```
%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 4, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\400m2checkor25.csv';  
run;  
  
%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 5, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\500m2checkor25.csv';  
run;  
  
%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 6, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\600m2checkor25.csv';  
run;  
  
%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 7, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\700m2checkor25.csv';  
run;  
  
%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 8, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\800m2checkor25.csv';  
run;  
  
%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 9, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\900m2checkor25.csv';  
run;  
  
%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 10, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\1000m2checkor25.csv';  
run;  
  
%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 11, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\1100m2checkor25.csv';  
run;  
  
%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 12, ar(1));  
proc export data=crudepowerresults dbms = csv  
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR  
25\1200m2checkor25.csv';  
run;  
  
%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 13, ar(1));
```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1300m2checkor25.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1400m2checkor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1500m2checkor25.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1600m2checkor25.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1700m2checkor25.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1800m2checkor25.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\1900m2checkor25.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 20, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal Modified2 OR
25\2000m2checkor25.csv';
run;

*****;

*Assuming OR=3 (odds of testing >= 2 times comparing study arms) and 5%
of Comparison arm participants test >= 2 times in a year -
Specification 6;

%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 1, ar(1));

```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\100m2cdcor3.csv';
run;

%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 2, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\200m2cdcor3.csv';
run;

%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 3, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\300m2cdcor3.csv';
run;

%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 4, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\400m2cdcor3.csv';
run;

%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 5, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\500m2cdcor3.csv';
run;

%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 6, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\600m2cdcor3.csv';
run;

%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 7, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\700m2cdcor3.csv';
run;

%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 8, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\800m2cdcor3.csv';
run;

%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 9, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\900m2cdcor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 10, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1000m2cdcor3.csv';
```

```
run;

%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 11, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1100m2cdcor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 12, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1200m2cdcor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 13, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1300m2cdcor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 14, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1400m2cdcor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 15, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1500m2cdcor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 16, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1600m2cdcor3.csv';
run;

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 17, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1700m2cdcor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 18, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1800m2cdcor3.csv';
run;

%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 19, ar(1));
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\1900m2cdcor3.csv';
run;

%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 20, ar(1));
```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Ideal CDC Modified2 OR
3\2000m2cdcor3.csv';
run;

```

```

*****;

```

```

*Start listing log;
options notes source source2;

```

ScenarioB_Specifications135.sas

```

*Suppress listing of log;
options nonotes nosource nosource2;

```

```

*Macro - Program to run simulations multiple times using different
parameters

```

```

Macro name: simtrialmacro

```

```

Macro variables: seed (seed)

```

```

nsim (number of simulations)

```

```

m (number of participants in the trial)

```

```

beta0 (log odds of testing >= 3 times in a

```

```

year for Comparison)

```

```

betalor (beta1 depending upon the odds ratio)

```

```

timeseedloss (seed)

```

```

timeseed (seed)

```

```

corrstruc (working correlation structure)

```

```

pos (proportion newly diagnosed as positive)

```

```

neg (proportion not positive)

```

```

loss (proportion lost to follow-up)

```

```

keep (proportion staying until end)

```

```

Syntax: Use ampersand (&) before variable names within a macro;

```

```

%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseedloss,
timeseed, corrstruc, pos, neg, loss, keep);

```

```

*Creating dataset Simtrial;

```

```

data simtrial;

```

```

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

```

```

do z=1 to &nsim; *Number of simulations;

```

```

do i=1 to &m; *Person ID;

```

```

        group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

        *Control variables from CheckingIn data;
        race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
        age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
        educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
        income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;
        residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
        orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

        subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

        *Assuming different Odds Ratios (Odds of testing >= 3 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
        lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 3 times in a year;
        odds=round(exp(lnodds),0.00001); *Odds of testing >= 3
times in a year;
        probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 3 times in a year;

        *Outcome for Standard Logistic Regression;
        test3ormore=rand('bernoulli',probtest); *Tested >= 3 times
in a year (based on probability of testing) - Yes(1)/No(0);

        if test3ormore=0 then
howmany=rand('table',0.44855,0.35209,0.19936);
        else howmany=rand('table',0.58511,0.23404,0.09574,0.08511);

        *Annual testing frequency;
        if test3ormore=0 and howmany=1 then annualtests=0;
        if test3ormore=0 and howmany=2 then annualtests=1;
        if test3ormore=0 and howmany=3 then annualtests=2;
        if test3ormore=1 and howmany=1 then annualtests=3;
        if test3ormore=1 and howmany=2 then annualtests=4;
        if test3ormore=1 and howmany=3 then annualtests=5;
        if test3ormore=1 and howmany=4 then annualtests=6;

output;

        end;

end;

run;

*Creating new dataset SimtrialConditional;

```



```

data simtrialcox00;
set simtrial;

call streaminit(&seed);

    if annualtests ne 0 then positive=rand('table',&pos,&neg);
    else positive=2;

    if positive=1 then lostovertime=1;
    else lostovertime=rand('table',&loss,&keep);

    if lostovertime=1 then
timeofloss=round(364*ranuni(&timeseedloss)+1,1);
    else timeofloss=0;

run;

data simtrialcox0;
set simtrialcox00;

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

    if j=0 then timeoftest=0;
    else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
    if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

    ntimeoftest=timeoftest;
    if lostovertime=1 and timeoftest > timeofloss then
ntimeoftest=timeofloss;

    *Calendar time between each test;
    lagtime=lag(ntimeoftest);
    if first.i then timebetweentests=ntimeoftest;
    if not first.i then timebetweentests=ntimeoftest-lagtime;
    if ntimeoftest=0 then timebetweentests=0;

```

```

run;

proc sort data=simtrialcox1;
by z i ntimeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i ntimeoftest;

    if ntimeoftest=0 and timebetweentests=0 then delete;

    *Variables for Conditional model 1;
    start1=lagtime;
    stop1=ntimeoftest;

    *Variables for Conditional model 2;
    start2=0;
    stop2=timebetweentests;

run;

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional0;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

run;

data simtrialconditional;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    *Outcome for Stratified Cox PH Regression for Recurrent Events;
    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;

```

```

        if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

        if stop2=0 then delete;

        if positive=1 and annualtests ne 0 and intervalph > annualtests
then delete;
        if positive=1 and testingevent=0 and intervalph ne 1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime timeofloss positive start1 stop1 start2 stop2
testingevent intervalph;

run;

*Creating new dataset SimtrialLogistic;

data simtriallogistic;
set simtrialconditional;
by z i;

        if first.i then nannualtests=0;
        nannualtests+testingevent;

        if not last.i then delete;

        if nannualtests ge 3 then ntest3ormore=1;
        else ntest3ormore=0;

        if positive=1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime positive ntest3ormore nannualtests;

run;

*Creating new dataset SimtrialRepeated;

data simtrialgee0;
set simtrialconditional0;

        *Interval in which person tested;
        if 1<=start1<=90 then testinterval3=1;
        else testinterval3=0;
        if 91<=start1<=181 then testinterval6=1;
        else testinterval6=0;
        if 182<=start1<=273 then testinterval9=1;
        else testinterval9=0;
        if 274<=start1<=365 then testinterval12=1;
        else testinterval12=0;

run;

proc sort data=simtrialgee0;
by z i;

```

```

run;

data simtrialgee1;
set simtrialgee0;
by z i;

    if first.i then testnum0to3=0;
    testnum0to3+testinterval3;

    if first.i then testnum3to6=0;
    testnum3to6+testinterval6;

    if first.i then testnum6to9=0;
    testnum6to9+testinterval9;

    if first.i then testnum9to12=0;
    testnum9to12+testinterval12;

run;

proc sort data=simtrialgee1;
by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

    if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialgee3;
set simtrialgee2;
by z i;

    do intervalgee=1 to 4;

        *Outcome for GEE for Correlated Data;
        testedp3mo=0;
        if intervalgee=1 and testnum0to3>0 then testedp3mo=1;
        if intervalgee=2 and testnum3to6>0 then testedp3mo=1;
        if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
        if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

run;

```

```

data simtrialrepeated;
set simtrialgee3;
by z i;

    if positive=1 and 1<=stop1<=90 and intervalgee in (2,3,4) then
delete;
    if positive=1 and 91<=stop1<=181 and intervalgee in (3,4) then
delete;
    if positive=1 and 182<=stop1<=273 and intervalgee in (4) then
delete;

    if positive=2 and lostovertime=1 and 1<=timeofloss<=90 and
intervalgee in (2,3,4) then delete;
    if positive=2 and lostovertime=1 and 91<=timeofloss<=181 and
intervalgee in (3,4) then delete;
    if positive=2 and lostovertime=1 and 182<=timeofloss<=273 and
intervalgee in (4) then delete;

    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum3to6=0
and testnum6to9=0 and testnum9to12=0 and intervalgee=2 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum6to9=0
and testnum9to12=0 and intervalgee=3 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and
testnum9to12=0 and intervalgee=4 then delete;

keep z i group race age educ income residence orientation lostovertime
timeofloss positive testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox00 simtrialcox0 simtrialcox1 simtrialcox2
simtrialconditional0 simtrialgee0 simtrialgee1 simtrialgee2
simtrialgee3;
run;

ods select all;

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

```

```

    if group=0 then group=2;
    if ntest3ormore=0 then ntest3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))
  Creating output dataset Chi with values for estimates of Relative risk
and their CIs;

ods select none;

proc freq data=test2by2;
table group*ntest3ormore / all;
by z;
output out = test2by2res relrisk;
run;

ods select all;

*Power - Proportion of those CIs that DO NOT include 1 where Relative
risk > 1;

data power2by2;
set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;

run;

ods select none;

proc freq data=power2by2;
tables powerpr / out=a;
run;

ods select all;

*Power and sample size - Standard Logistic Regression;

ods select none;

proc logistic data=simtriallogistic descending;
model ntest3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;
run;

ods select all;

data powerlogisticcrude;
set testlogisticcrude;

    if effect ne "group" then delete;

```

```

        if lowercl gt 1 then powerslr crude=1;
        else powerslr crude=2;

run;

ods select none;

proc freq data=powerlogistic crude;
tables powerslr crude / out=b;
run;

ods select all;

*Power and sample size - GEE for Correlated Data;

ods select none;

proc genmod data=simtrialrepeated descending;
class i intervalgee;
model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;
ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

        if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

        if lbetalowercl gt 1 then powergeerepcrude=1;
        else powergeerepcrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepcrude / out=d;
run;

ods select all;

*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;

*Conditional model 1 - Models the full time course of the recurrent
event process
                                (Time interval of a subsequent event
starts at the end of the time interval for the previous event);

```

```

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional1crude;
run;

ods select all;

data powerconditional1crude;
set testconditional1crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg1crude=1;
    else powerphreg1crude=2;

run;

ods select none;

proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;

ods select all;

*Conditional model 2 - Models the time between each of the recurring
events
                                (Time interval starts at 0 and ends at
the length of time until the next event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;

ods select all;

data powerconditional2crude;
set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

```



```

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepcrude=1 then powerlabel="Power GEE Crude";
    if powerphreg1crude=1 then powerlabel="Power Cox PH Conditional I
Crude";
    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

    power=percent;

run;

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

proc print data=crudepowerresults noobs;
title "Crude Analyses (&loss Loss to follow-up, &pos Positive): &m
Participants, Beta0 = &beta0, Beta1 = &betalor ";
run;

title;

%mend;

*****;

*Assuming OR=2 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 1;

%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\100checklosspor2.csv';
run;

```

```
%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\200checklosspor2.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\300checklosspor2.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\400checklosspor2.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\500checklosspor2.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\600checklosspor2.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\700checklosspor2.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\800checklosspor2.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\900checklosspor2.csv';
run;
```

```
%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1000checklosspor2.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1100checklosspor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1200checklosspor2.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1300checklosspor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1400checklosspor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1500checklosspor2.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1600checklosspor2.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1700checklosspor2.csv';
run;
```

```

%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1800checklosspor2.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\1900checklosspor2.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
2\2000checklosspor2.csv';
run;

*****;

*Assuming OR=2.5 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 3;

%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\100checklosspor25.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\200checklosspor25.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\300checklosspor25.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\400checklosspor25.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\500checklosspor25.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\600checklosspor25.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\700checklosspor25.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\800checklosspor25.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\900checklosspor25.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1000checklosspor25.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1100checklosspor25.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1200checklosspor25.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1300checklosspor25.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1400checklosspor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1500checklosspor25.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1600checklosspor25.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1700checklosspor25.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1800checklosspor25.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\1900checklosspor25.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CheckingIn OR
25\2000checklosspor25.csv';
run;

*****;
```

*Assuming OR=3 and 5% of Comparison arm participants test ≥ 3 times in a year - Specification 5;

```
%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\100cdclosspor3.csv';
run;
```

```
%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\200cdclosspor3.csv';
run;
```

```
%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\300cdclosspor3.csv';
run;
```

```
%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\400cdclosspor3.csv';
run;
```

```
%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\500cdclosspor3.csv';
run;
```

```
%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\600cdclosspor3.csv';
run;
```

```
%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\700cdclosspor3.csv';
run;
```

```
%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\800cdclosspor3.csv';
```

```
run;

%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\900cdclosspor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1000cdclosspor3.csv';
run;

%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1100cdclosspor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1200cdclosspor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1300cdclosspor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1400cdclosspor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1500cdclosspor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1600cdclosspor3.csv';
run;
```



```

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1700cdclosspor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1800cdclosspor3.csv';
run;

%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\1900cdclosspor3.csv';
run;

%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC OR
3\2000cdclosspor3.csv';
run;

*****;

*Start listing log;
options notes source source2;

```

ScenarioB_Specifications246.sas

```

*Suppress listing of log;
options nonotes nosource nosource2;

*Macro - Program to run simulations multiple times using different
parameters
Macro name: simtrialmacro
Macro variables: seed (seed)
                    nsim (number of simulations)
                    m (number of participants in the trial)
                    beta0 (log odds of testing >= 2 times in a
year for Comparison)
                    betalor (beta1 depending upon the odds ratio)
                    timeseedloss (seed)
                    timeseed (seed)

```

```

        corrstruc (working correlation structure)
        pos (proportion newly diagnosed as positive)
        neg (proportion not positive)
        loss (proportion lost to follow-up)
        keep (proportion staying until end)
Syntax: Use ampersand (&) before variable names within a macro;

%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseedloss,
timeseed, corrstruc, pos, neg, loss, keep);

*Creating dataset Simtrial;

data simtrial;

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

do z=1 to &nsim; *Number of simulations;

    do i=1 to &m; *Person ID;

        group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

        *Control variables from CheckingIn data;
        race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
        age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
        educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
        income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;
        residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
        orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

        subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

        *Assuming different Odds Ratios (Odds of testing >= 2 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
        lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 2 times in a year;
        odds=round(exp(lnodds),0.00001); *Odds of testing >= 2
times in a year;
        probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 2 times in a year;

        test2ormore=rand('bernoulli',probtest); *Tested >= 2 times
in a year (based on probability of testing) - Yes(1)/No(0);

```

```

        if test2ormore=0 then
howmany=rand('table',0.56024,0.43976);
        else
howmany=rand('table',0.56881,0.25229,0.10092,0.04128,0.03669);

        *Annual testing frequency;
if test2ormore=0 and howmany=1 then annualtests=0;
if test2ormore=0 and howmany=2 then annualtests=1;
if test2ormore=1 and howmany=1 then annualtests=2;
if test2ormore=1 and howmany=2 then annualtests=3;
if test2ormore=1 and howmany=3 then annualtests=4;
if test2ormore=1 and howmany=4 then annualtests=5;
if test2ormore=1 and howmany=5 then annualtests=6;

        *Outcome for Standard Logistic Regression - Still keep at 3
or more times;
if annualtests=0 then test3ormore=0;
if annualtests=1 then test3ormore=0;
if annualtests=2 then test3ormore=0;
if annualtests=3 then test3ormore=1;
if annualtests=4 then test3ormore=1;
if annualtests=5 then test3ormore=1;
if annualtests=6 then test3ormore=1;

output;

        end;

end;

run;

*Creating new dataset SimtrialConditional;

data simtrialcox0;
set simtrial;

call streaminit(&seed);

        if annualtests ne 0 then positive=rand('table',&pos,&neg);
        else positive=2;

        if positive=1 then lostovertime=1;
        else lostovertime=rand('table',&loss,&keep);

        if lostovertime=1 then
timeofloss=round(364*ranuni(&timeseedloss)+1,1);
        else timeofloss=0;

run;

data simtrialcox0;
set simtrialcox0;

```

```

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

    if j=0 then timeoftest=0;
    else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
    if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

    ntimeoftest=timeoftest;
    if lostovertime=1 and timeoftest > timeofloss then
ntimeoftest=timeofloss;

    *Calendar time between each test;
    lagtime=lag(ntimeoftest);
    if first.i then timebetweentests=ntimeoftest;
    if not first.i then timebetweentests=ntimeoftest-lagtime;
    if ntimeoftest=0 then timebetweentests=0;

run;

proc sort data=simtrialcox1;
by z i ntimeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i ntimeoftest;

    if ntimeoftest=0 and timebetweentests=0 then delete;

    *Variables for Conditional model 1;
    start1=lagtime;
    stop1=ntimeoftest;

    *Variables for Conditional model 2;
    start2=0;
    stop2=timebetweentests;

run;

```

```

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional0;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

run;

data simtrialconditional;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    *Outcome for Stratified Cox PH Regression for Recurrent Events;
    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

    if positive=1 and annualtests ne 0 and intervalph > annualtests
then delete;
    if positive=1 and testingevent=0 and intervalph ne 1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime timeofloss positive start1 stop1 start2 stop2
testingevent intervalph;

run;

*Creating new dataset SimtrialLogistic;

data simtriallogistic;
set simtrialconditional;
by z i;

```

```

    if first.i then nannualtests=0;
    nannualtests+testingevent;

    if not last.i then delete;

    if nannualtests ge 3 then ntest3ormore=1;
    else ntest3ormore=0;

    if positive=1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime positive ntest3ormore nannualtests;

run;

*Creating new dataset SimtrialRepeated;

data simtrialgee0;
set simtrialconditional0;

    *Interval in which person tested;
    if 1<=start1<=90 then testinterval3=1;
    else testinterval3=0;
    if 91<=start1<=181 then testinterval6=1;
    else testinterval6=0;
    if 182<=start1<=273 then testinterval9=1;
    else testinterval9=0;
    if 274<=start1<=365 then testinterval12=1;
    else testinterval12=0;

run;

proc sort data=simtrialgee0;
by z i;
run;

data simtrialgeel;
set simtrialgee0;
by z i;

    if first.i then testnum0to3=0;
    testnum0to3+testinterval3;

    if first.i then testnum3to6=0;
    testnum3to6+testinterval6;

    if first.i then testnum6to9=0;
    testnum6to9+testinterval9;

    if first.i then testnum9to12=0;
    testnum9to12+testinterval12;

run;

proc sort data=simtrialgeel;

```

```

by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

    if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialgee3;
set simtrialgee2;
by z i;

    do intervalgee=1 to 4;

        *Outcome for GEE for Correlated Data;
        testedp3mo=0;
        if intervalgee=1 and testnum0to3>0 then testedp3mo=1;
        if intervalgee=2 and testnum3to6>0 then testedp3mo=1;
        if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
        if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

run;

data simtrialrepeated;
set simtrialgee3;
by z i;

    if positive=1 and 1<=stop1<=90 and intervalgee in (2,3,4) then
delete;
    if positive=1 and 91<=stop1<=181 and intervalgee in (3,4) then
delete;
    if positive=1 and 182<=stop1<=273 and intervalgee in (4) then
delete;

    if positive=2 and lostovertime=1 and 1<=timeofloss<=90 and
intervalgee in (2,3,4) then delete;
    if positive=2 and lostovertime=1 and 91<=timeofloss<=181 and
intervalgee in (3,4) then delete;
    if positive=2 and lostovertime=1 and 182<=timeofloss<=273 and
intervalgee in (4) then delete;

```

```

        if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum3to6=0
and testnum6to9=0 and testnum9to12=0 and intervalgee=2 then delete;
        if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum6to9=0
and testnum9to12=0 and intervalgee=3 then delete;
        if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and
testnum9to12=0 and intervalgee=4 then delete;

keep z i group race age educ income residence orientation lostovertime
timeofloss positive testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox0 simtrialcox0 simtrialcox1 simtrialcox2
simtrialconditional0 simtrialgee0 simtrialgee1 simtrialgee2
simtrialgee3;
run;

ods select all;

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

        if group=0 then group=2;
        if ntest3ormore=0 then ntest3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))
Creating output dataset Chi with values for estimates of Relative risk
and their CIs;

ods select none;

proc freq data=test2by2;
table group*ntest3ormore / all;
by z;
output out = test2by2res relrisk;
run;

ods select all;

```



```
*Power - Proportion of those CIs that DO NOT include 1 where Relative
risk > 1;
```

```
data power2by2;
set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;
```

```
run;
```

```
ods select none;
```

```
proc freq data=power2by2;
tables powerpr / out=a;
run;
```

```
ods select all;
```

```
*Power and sample size - Standard Logistic Regression;
```

```
ods select none;
```

```
proc logistic data=simtriallogistic descending;
model ntest3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;
run;
```

```
ods select all;
```

```
data powerlogisticcrude;
set testlogisticcrude;
```

```
    if effect ne "group" then delete;
```

```
    if lowercl gt 1 then powerslr crude=1;
    else powerslr crude=2;
```

```
run;
```

```
ods select none;
```

```
proc freq data=powerlogisticcrude;
tables powerslr crude / out=b;
run;
```

```
ods select all;
```

```
*Power and sample size - GEE for Correlated Data;
```

```
ods select none;
```

```
proc genmod data=simtrialrepeated descending;
class i intervalgee;
```

```

model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;
ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

    if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

    if lbetalowercl gt 1 then powergeerepccrude=1;
    else powergeerepccrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepccrude / out=d;
run;

ods select all;

*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;

*Conditional model 1 - Models the full time course of the recurrent
event process
                                (Time interval of a subsequent event
starts at the end of the time interval for the previous event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditionalcrude;
run;

ods select all;

data powerconditionalcrude;
set testconditionalcrude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreglcrude=1;
    else powerphreglcrude=2;

```

```

run;

ods select none;

proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;

ods select all;

*Conditional model 2 - Models the time between each of the recurring
events
                                (Time interval starts at 0 and ends at
the length of time until the next event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;

ods select all;

data powerconditional2crude;
set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepcrude=1 then powerlabel="Power GEE Crude";
    if powerphreg1crude=1 then powerlabel="Power Cox PH Conditional I
Crude";

```

```

    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

```

```

    power=percent;

```

```

run;

```

```

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

```

```

proc print data=crudepowerresults noobs;
title "Crude Analyses (&loss Loss to follow-up, &pos Positive): &m
Participants, Beta0 = &beta0, Beta1 = &betalor ";
run;

```

```

title;

```

```

&mend;

```

```

*****;

```

```

*Assuming OR=2 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 2;

```

```

%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\100m2checklosspor2.csv';
run;

```

```

%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\200m2checklosspor2.csv';
run;

```

```

%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\300m2checklosspor2.csv';
run;

```

```

%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\400m2checklosspor2.csv';

```

```
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\500m2checklosspor2.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\600m2checklosspor2.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\700m2checklosspor2.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\800m2checklosspor2.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\900m2checklosspor2.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1000m2checklosspor2.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1100m2checklosspor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1200m2checklosspor2.csv';
run;
```

```
%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1300m2checklosspor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1400m2checklosspor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1500m2checklosspor2.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1600m2checklosspor2.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1700m2checklosspor2.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1800m2checklosspor2.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\1900m2checklosspor2.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
2\2000m2checklosspor2.csv';
run;
```

```

*****;

*Assuming OR=2.5 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 4;

%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\100m2checklosspor25.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\200m2checklosspor25.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\300m2checklosspor25.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\400m2checklosspor25.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\500m2checklosspor25.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\600m2checklosspor25.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\700m2checklosspor25.csv';
run;

```

```
%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\800m2checklosspor25.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\900m2checklosspor25.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1000m2checklosspor25.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1100m2checklosspor25.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1200m2checklosspor25.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1300m2checklosspor25.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1400m2checklosspor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1500m2checklosspor25.csv';
run;
```



```
%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1600m2checklosspor25.csv';
run;
```

```
%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1700m2checklosspor25.csv';
run;
```

```
%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1800m2checklosspor25.csv';
run;
```

```
%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\1900m2checklosspor25.csv';
run;
```

```
%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive Modified2 OR
25\2000m2checklosspor25.csv';
run;
```

```
*****;
```

```
*Assuming OR=3 (odds of testing >= 2 times comparing study arms) and 5%
of Comparison arm participants test >= 2 times in a year -
Specification 6;
```

```
%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 51, 1, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\100m2cdclosspor3.csv';
run;
```

```
%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 52, 2, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\200m2cdclosspor3.csv';
run;
```

```
%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 53, 3, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\300m2cdclosspor3.csv';
run;

%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 54, 4, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\400m2cdclosspor3.csv';
run;

%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\500m2cdclosspor3.csv';
run;

%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 56, 6, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\600m2cdclosspor3.csv';
run;

%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 57, 7, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\700m2cdclosspor3.csv';
run;

%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 58, 8, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\800m2cdclosspor3.csv';
run;

%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 59, 9, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\900m2cdclosspor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 60, 10, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1000m2cdclosspor3.csv';
run;
```

```
%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 61, 11, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1100m2cdclosspor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1200m2cdclosspor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 63, 13, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1300m2cdclosspor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 64, 14, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1400m2cdclosspor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 65, 15, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1500m2cdclosspor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 66, 16, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1600m2cdclosspor3.csv';
run;

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 67, 17, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1700m2cdclosspor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 68, 18, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1800m2cdclosspor3.csv';
run;
```

```
%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 69, 19, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\1900m2cdclosspor3.csv';
run;
```

```
%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 70, 20, ar(1), 0.05,
0.95, 0.4, 0.6);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Loss Positive CDC Modified2 OR
3\2000m2cdclosspor3.csv';
run;
```

```
*****;
```

```
*Start listing log;
options notes source source2;
```

ScenarioC_Specifications135.sas

```
*Suppress listing of log;
options nonotes nosource nosource2;
```

```
*Macro - Program to run simulations multiple times using different
parameters
```

```
Macro name: simtrialmacro
```

```
Macro variables: seed (seed)
```

```
nsim (number of simulations)
```

```
m (number of participants in the trial)
```

```
beta0 (log odds of testing  $\geq 3$  times in a
```

```
year for Comparison)
```

```
betalor (beta1 depending upon the odds ratio)
```

```
timeseedloss (seed)
```

```
timeseed (seed)
```

```
corrstruc (working correlation structure)
```

```
pos (proportion newly diagnosed as positive)
```

```
neg (proportion not positive)
```

```
loss (proportion lost to follow-up)
```

```
keep (proportion staying until end)
```

```
Syntax: Use ampersand (&) before variable names within a macro;
```

```
%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseedloss,
timeseed, corrstruc, pos, neg, loss, keep);
```

```
*Creating dataset Simtrial;
```

```
data simtrial;
```

```

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

do z=1 to &nsim; *Number of simulations;

    do i=1 to &m; *Person ID;

        group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

        *Control variables from CheckingIn data;
        race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
        age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
        educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
        income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;
        residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
        orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

        subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

        *Assuming different Odds Ratios (Odds of testing >= 3 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
        lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 3 times in a year;
        odds=round(exp(lnodds),0.00001); *Odds of testing >= 3
times in a year;
        probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 3 times in a year;

        *Outcome for Standard Logistic Regression;
        test3ormore=rand('bernoulli',probtest); *Tested >= 3 times
in a year (based on probability of testing) - Yes(1)/No(0);

        if test3ormore=0 then
howmany=rand('table',0.44855,0.35209,0.19936);
        else howmany=rand('table',0.58511,0.23404,0.09574,0.08511);

        *Annual testing frequency;
        if test3ormore=0 and howmany=1 then annualtests=0;
        if test3ormore=0 and howmany=2 then annualtests=1;
        if test3ormore=0 and howmany=3 then annualtests=2;
        if test3ormore=1 and howmany=1 then annualtests=3;
        if test3ormore=1 and howmany=2 then annualtests=4;
        if test3ormore=1 and howmany=3 then annualtests=5;
        if test3ormore=1 and howmany=4 then annualtests=6;

    output;

```

```

        end;

end;

run;

*Creating new dataset SimtrialConditional;

data simtrialcox00;
set simtrial;

call streaminit(&seed);

    if annualtests ne 0 then positive=rand('table',&pos,&neg);
    else positive=2;

    if positive=1 then lostovertime=1;
    else lostovertime=rand('table',&loss,&keep);

    if lostovertime=1 then
timeofloss=round(364*ranuni(&timeseedloss)+1,1);
    else timeofloss=0;

run;

data simtrialcox0;
set simtrialcox00;

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

    if j=0 then timeoftest=0;
    else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
    if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

    ntimeoftest=timeoftest;

```

```

        if lostovertime=1 and timeoftest > timeofloss then
ntimeoftest=timeofloss;

        *Calendar time between each test;
        lagtime=lag(ntimeoftest);
        if first.i then timebetweentests=ntimeoftest;
        if not first.i then timebetweentests=ntimeoftest-lagtime;
        if ntimeoftest=0 then timebetweentests=0;

run;

proc sort data=simtrialcox1;
by z i ntimeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i ntimeoftest;

        if ntimeoftest=0 and timebetweentests=0 then delete;

        *Variables for Conditional model 1;
        start1=lagtime;
        stop1=ntimeoftest;

        *Variables for Conditional model 2;
        start2=0;
        stop2=timebetweentests;

run;

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional0;
set simtrialcox2;
by z i;

        if first.i then intervalph=0;
        intervalph+1;

        testingevent=1;
        if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
        if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

        if stop2=0 then delete;

run;

data simtrialconditional;
set simtrialcox2;

```

```

by z i;

    if first.i then intervalph=0;
    intervalph+1;

    *Outcome for Stratified Cox PH Regression for Recurrent Events;
    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

    if positive=1 and annualtests ne 0 and intervalph > annualtests
then delete;
    if positive=1 and testingevent=0 and intervalph ne 1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime timeofloss positive start1 stop1 start2 stop2
testingevent intervalph;

run;

```

```

*Creating new dataset SimtrialLogistic;

```

```

data simtriallogistic;
set simtrialconditional;
by z i;

    if first.i then nannualtests=0;
    nannualtests+testingevent;

    if not last.i then delete;

    if nannualtests ge 3 then ntest3ormore=1;
    else ntest3ormore=0;

    if positive=1 then delete;

keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime positive ntest3ormore nannualtests;

run;

```

```

*Creating new dataset SimtrialRepeated;

```

```

data simtrialgee0;
set simtrialconditional0;

    *Interval in which person tested;
    if 1<=start1<=90 then testinterval3=1;
    else testinterval3=0;
    if 91<=start1<=181 then testinterval6=1;
    else testinterval6=0;

```



```

        if 182<=start1<=273 then testinterval9=1;
        else testinterval9=0;
        if 274<=start1<=365 then testinterval12=1;
        else testinterval12=0;

run;

proc sort data=simtrialgee0;
by z i;
run;

data simtrialgee1;
set simtrialgee0;
by z i;

        if first.i then testnum0to3=0;
        testnum0to3+testinterval3;

        if first.i then testnum3to6=0;
        testnum3to6+testinterval6;

        if first.i then testnum6to9=0;
        testnum6to9+testinterval9;

        if first.i then testnum9to12=0;
        testnum9to12+testinterval12;

run;

proc sort data=simtrialgee1;
by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

        if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialgee3;
set simtrialgee2;
by z i;

do intervalgee=1 to 4;

*Outcome for GEE for Correlated Data;
testedp3mo=0;
if intervalgee=1 and testnum0to3>0 then testedp3mo=1;

```

```

    if intervalgee=2 and testnum3to6>0 then testedp3mo=1;
    if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
    if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

run;

data simtrialrepeated;
set simtrialgee3;
by z i;

    if positive=1 and 1<=stop1<=90 and intervalgee in (2,3,4) then
delete;
    if positive=1 and 91<=stop1<=181 and intervalgee in (3,4) then
delete;
    if positive=1 and 182<=stop1<=273 and intervalgee in (4) then
delete;

    if positive=2 and lostovertime=1 and 1<=timeofloss<=90 and
intervalgee in (2,3,4) then delete;
    if positive=2 and lostovertime=1 and 91<=timeofloss<=181 and
intervalgee in (3,4) then delete;
    if positive=2 and lostovertime=1 and 182<=timeofloss<=273 and
intervalgee in (4) then delete;

    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum3to6=0
and testnum6to9=0 and testnum9to12=0 and intervalgee=2 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum6to9=0
and testnum9to12=0 and intervalgee=3 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and
testnum9to12=0 and intervalgee=4 then delete;

keep z i group race age educ income residence orientation lostovertime
timeofloss positive testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox00 simtrialcox0 simtrialcox1 simtrialcox2
simtrialconditional0 simtrialgee0 simtrialgee1 simtrialgee2
simtrialgee3;
run;

ods select all;

```

```

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

    if group=0 then group=2;
    if ntest3ormore=0 then ntest3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))
  Creating output dataset Chi with values for estimates of Relative risk
and their CIs;

ods select none;

proc freq data=test2by2;
table group*ntest3ormore / all;
by z;
output out = test2by2res relrisk;
run;

ods select all;

*Power - Proportion of those CIs that DO NOT include 1 where Relative
risk > 1;

data power2by2;
set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;

run;

ods select none;

proc freq data=power2by2;
tables powerpr / out=a;
run;

ods select all;

*Power and sample size - Standard Logistic Regression;

ods select none;

proc logistic data=simtriallogistic descending;
model ntest3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;

```

```

run;

ods select all;

data powerlogisticcrude;
set testlogisticcrude;

    if effect ne "group" then delete;

    if lowercl gt 1 then powerslrcrude=1;
    else powerslrcrude=2;

run;

ods select none;

proc freq data=powerlogisticcrude;
tables powerslrcrude / out=b;
run;

ods select all;

*Power and sample size - GEE for Correlated Data;

ods select none;

proc genmod data=simtrialrepeated descending;
class i intervalgee;
model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;
ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

    if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

    if lbetalowercl gt 1 then powergeerepcrude=1;
    else powergeerepcrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepcrude / out=d;
run;

ods select all;

```

```
*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;
```

```
*Conditional model 1 - Models the full time course of the recurrent
event process
```

```
(Time interval of a subsequent event
starts at the end of the time interval for the previous event);
```

```
ods select none;
```

```
proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional1crude;
run;
```

```
ods select all;
```

```
data powerconditional1crude;
set testconditional1crude;
```

```
if parameter ne "group" then delete;
```

```
if hrlowercl gt 1 then powerphreg1crude=1;
else powerphreg1crude=2;
```

```
run;
```

```
ods select none;
```

```
proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;
```

```
ods select all;
```

```
*Conditional model 2 - Models the time between each of the recurring
events
```

```
(Time interval starts at 0 and ends at
the length of time until the next event);
```

```
ods select none;
```

```
proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;
```

```
ods select all;
```

```
data powerconditional2crude;
```

```

set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepcrude=1 then powerlabel="Power GEE Crude";
    if powerphreg1crude=1 then powerlabel="Power Cox PH Conditional I
Crude";
    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

    power=percent;

run;

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

proc print data=crudepowerresults noobs;
title "Crude Analyses (&loss Loss to follow-up, &pos Positive): &m
Participants, Beta0 = &beta0, Beta1 = &betalor ";
run;

title;

%mend;

*****;

*Assuming OR=2 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 1;

```

```
%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\100checkelosspor2.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\200checkelosspor2.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\300checkelosspor2.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\400checkelosspor2.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\500checkelosspor2.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\600checkelosspor2.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\700checkelosspor2.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\800checkelosspor2.csv';
run;
```

```
%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\900checkelosspor2.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1000checkelosspor2.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1100checkelosspor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1200checkelosspor2.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1300checkelosspor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1400checkelosspor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1500checkelosspor2.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1600checkelosspor2.csv';
run;
```



```
%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1700checkelosspor2.csv';
run;
```

```
%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1800checkelosspor2.csv';
run;
```

```
%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\1900checkelosspor2.csv';
run;
```

```
%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
2\2000checkelosspor2.csv';
run;
```

```
*****;
```

```
*Assuming OR=2.5 and 13.128% of Comparison arm participants test >= 3
times in a year - Specification 3;
```

```
%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\100checkelosspor25.csv';
run;
```

```
%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\200checkelosspor25.csv';
run;
```

```
%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\300checkelosspor25.csv';
run;
```

```
%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\400checkelosspor25.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\500checkelosspor25.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\600checkelosspor25.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\700checkelosspor25.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\800checkelosspor25.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\900checkelosspor25.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1000checkelosspor25.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1100checkelosspor25.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
```

```
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1200checkelosspor25.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1300checkelosspor25.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1400checkelosspor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1500checkelosspor25.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1600checkelosspor25.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1700checkelosspor25.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1800checkelosspor25.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\1900checkelosspor25.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CheckingIn OR
25\2000checkelosspor25.csv';
run;

*****;

*Assuming OR=3 and 5% of Comparison arm participants test >= 3 times in
a year - Specification 5;

%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\100cdcelosspor3.csv';
run;

%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\200cdcelosspor3.csv';
run;

%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\300cdcelosspor3.csv';
run;

%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\400cdcelosspor3.csv';
run;

%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\500cdcelosspor3.csv';
run;

%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\600cdcelosspor3.csv';
run;

%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\700cdcellosspor3.csv';
run;

%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\800cdcellosspor3.csv';
run;

%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\900cdcellosspor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1000cdcellosspor3.csv';
run;

%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1100cdcellosspor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1200cdcellosspor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1300cdcellosspor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1400cdcellosspor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);

```

```

proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1500cdcellosspor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1600cdcellosspor3.csv';
run;

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1700cdcellosspor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1800cdcellosspor3.csv';
run;

%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\1900cdcellosspor3.csv';
run;

%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC OR
3\2000cdcellosspor3.csv';
run;

*****;

*Start listing log;
options notes source source2;

ScenarioC_Specifications246.sas

*Suppress listing of log;
options nonotes nosource nosource2;

```

```

*Macro - Program to run simulations multiple times using different
parameters
Macro name: simtrialmacro
Macro variables: seed (seed)
                    nsim (number of simulations)
                    m (number of participants in the trial)
                    beta0 (log odds of testing >= 2 times in a
year for Comparison)
                    betalor (beta1 depending upon the odds ratio)
                    timeseedloss (seed)
                    timeseed (seed)
                    corrstruc (working correlation structure)
                    pos (proportion newly diagnosed as positive)
                    neg (proportion not positive)
                    loss (proportion lost to follow-up)
                    keep (proportion staying until end)
Syntax: Use ampersand (&) before variable names within a macro;

%macro simtrialmacro (seed, nsim, m, beta0, betalor, timeseedloss,
timeseed, corrstruc, pos, neg, loss, keep);

*Creating dataset Simtrial;

data simtrial;

call streaminit(&seed); *Syntax: call streaminit(seed) - Specifies a
seed value to create a reproducible stream of random numbers with the
RAND function;

do z=1 to &nsim; *Number of simulations;

    do i=1 to &m; *Person ID;

        group=rand('bernoulli',0.5); *Randomization to
Intervention(1)/Comparison(0);

        *Control variables from CheckingIn data;
        race=rand('table',0.63,0.17,0.20); *White, Black, Hispanic;
        age=rand('table',0.38,0.35,0.16,0.11); *18-24, 25-34, 35-
44, >=45;
        educ=rand('table',0.36,0.44,0.17,0.03); *College/Post-
graduate, Some college/Associate degree, HS or GED, Less than HS or
Unknown;
        income=rand('table',0.33,0.32,0.20,0.11,0.04); *< $14,999,
$15,000 - $39,999, $40,000 - $74,999, > $75,000, Unknown;
        residence=rand('table',0.63,0.33,0.04); *Urban, Rural,
Unknown;
        orientation=rand('table',0.84,0.13,0.03); *Homosexual,
Bisexual, Other;

        subjectre=round(rand('normal',0,0.5),0.00001); *Subject
specific random error;

```

```

*Assuming different Odds Ratios (Odds of testing >= 2 times
in a year in the Intervention are X times that of the Comparison)-
Change betalor;
    lnodds=round(&beta0+(&betalor*group)+subjectre,0.00001);
*Log odds of testing >= 2 times in a year;
    odds=round(exp(lnodds),0.00001); *Odds of testing >= 2
times in a year;
    probtest=round(odds/(1+odds),0.00001); *Probability of
testing >= 2 times in a year;

    test2ormore=rand('bernoulli',probtest); *Tested >= 2 times
in a year (based on probability of testing) - Yes(1)/No(0);

    if test2ormore=0 then
howmany=rand('table',0.56024,0.43976);
    else
howmany=rand('table',0.56881,0.25229,0.10092,0.04128,0.03669);

*Annual testing frequency;
if test2ormore=0 and howmany=1 then annualtests=0;
if test2ormore=0 and howmany=2 then annualtests=1;
if test2ormore=1 and howmany=1 then annualtests=2;
if test2ormore=1 and howmany=2 then annualtests=3;
if test2ormore=1 and howmany=3 then annualtests=4;
if test2ormore=1 and howmany=4 then annualtests=5;
if test2ormore=1 and howmany=5 then annualtests=6;

*Outcome for Standard Logistic Regression - Still keep at 3
or more times;
if annualtests=0 then test3ormore=0;
if annualtests=1 then test3ormore=0;
if annualtests=2 then test3ormore=0;
if annualtests=3 then test3ormore=1;
if annualtests=4 then test3ormore=1;
if annualtests=5 then test3ormore=1;
if annualtests=6 then test3ormore=1;

output;

    end;

end;

run;

*Creating new dataset SimtrialConditional;

data simtrialcox00;
set simtrial;

call streaminit(&seed);

if annualtests ne 0 then positive=rand('table',&pos,&neg);
else positive=2;

if positive=1 then lostovertime=1;

```



```

else lostovertime=rand('table',&loss,&keep);

if lostovertime=1 then
timeofloss=round(364*ranuni(&timeseedloss)+1,1);
else timeofloss=0;

run;

data simtrialcox0;
set simtrialcox00;

do j=0 to annualtests+1; *Generate rows depending upon the annual
testing frequency - Two extra rows for each person created for data
manipulation;

if j=0 then timeoftest=0;
else timeoftest=round(364*ranuni(&timeseed)+1,1); *Syntax:
a*ranuni(seed)+b - Returns a random number generated from the uniform
distribution on the interval (b=1,a+b=365);
if j=annualtests+1 then timeoftest=365;

output;

end;

run;

proc sort data=simtrialcox0;
by z i timeoftest;
run;

data simtrialcox1;
set simtrialcox0;
by z i timeoftest;

ntimeoftest=timeoftest;
if lostovertime=1 and timeoftest > timeofloss then
ntimeofloss=timeofloss;

*Calendar time between each test;
lagtime=lag(ntimeoftest);
if first.i then timebetweentests=ntimeoftest;
if not first.i then timebetweentests=ntimeoftest-lagtime;
if ntimeoftest=0 then timebetweentests=0;

run;

proc sort data=simtrialcox1;
by z i ntimeoftest;
run;

data simtrialcox2;
set simtrialcox1;
by z i ntimeoftest;

```

```

if ntimeoftest=0 and timebetweentests=0 then delete;

*Variables for Conditional model 1;
start1=lagtime;
stop1=ntimeoftest;

*Variables for Conditional model 2;
start2=0;
stop2=timebetweentests;

run;

proc sort data=simtrialcox2;
by z i;
run;

data simtrialconditional0;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

    testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

run;

data simtrialconditional;
set simtrialcox2;
by z i;

    if first.i then intervalph=0;
    intervalph+1;

*Outcome for Stratified Cox PH Regression for Recurrent Events;
testingevent=1;
    if lostovertime=2 and annualtests ne intervalph and stop1=365
then testingevent=0;
    if lostovertime=1 and timeofloss < timeoftest then
testingevent=0;

    if stop2=0 then delete;

    if positive=1 and annualtests ne 0 and intervalph > annualtests
then delete;
    if positive=1 and testingevent=0 and intervalph ne 1 then delete;

```

```
keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime timeofloss positive start1 stop1 start2 stop2
testingevent intervalph;
```

```
run;
```

```
*Creating new dataset SimtrialLogistic;
```

```
data simtriallogistic;
set simtrialconditional;
by z i;
```

```
    if first.i then nannualtests=0;
    nannualtests+testingevent;
```

```
    if not last.i then delete;
```

```
    if nannualtests ge 3 then ntest3ormore=1;
    else ntest3ormore=0;
```

```
    if positive=1 then delete;
```

```
keep z i group race age educ income residence orientation test3ormore
annualtests lostovertime positive ntest3ormore nannualtests;
```

```
run;
```

```
*Creating new dataset SimtrialRepeated;
```

```
data simtrialgee0;
set simtrialconditional0;
```

```
    *Interval in which person tested;
    if 1<=start1<=90 then testinterval3=1;
    else testinterval3=0;
    if 91<=start1<=181 then testinterval6=1;
    else testinterval6=0;
    if 182<=start1<=273 then testinterval9=1;
    else testinterval9=0;
    if 274<=start1<=365 then testinterval12=1;
    else testinterval12=0;
```

```
run;
```

```
proc sort data=simtrialgee0;
by z i;
run;
```

```
data simtrialgee1;
set simtrialgee0;
by z i;
```

```
    if first.i then testnum0to3=0;
    testnum0to3+testinterval3;
```

```

    if first.i then testnum3to6=0;
    testnum3to6+testinterval6;

    if first.i then testnum6to9=0;
    testnum6to9+testinterval9;

    if first.i then testnum9to12=0;
    testnum9to12+testinterval12;

run;

proc sort data=simtrialgee1;
by z i;
run;

data simtrialgee2;
set simtrialgee1;
by z i;

    if not last.i then delete;

run;

proc sort data=simtrialgee2;
by z i;
run;

data simtrialgee3;
set simtrialgee2;
by z i;

    do intervalgee=1 to 4;

        *Outcome for GEE for Correlated Data;
        testedp3mo=0;
        if intervalgee=1 and testnum0to3>0 then testedp3mo=1;
        if intervalgee=2 and testnum3to6>0 then testedp3mo=1;
        if intervalgee=3 and testnum6to9>0 then testedp3mo=1;
        if intervalgee=4 and testnum9to12>0 then testedp3mo=1;

output;

end;

run;

data simtrialrepeated;
set simtrialgee3;
by z i;

    if positive=1 and 1<=stop1<=90 and intervalgee in (2,3,4) then
delete;

```

```

    if positive=1 and 91<=stop1<=181 and intervalgee in (3,4) then
delete;
    if positive=1 and 182<=stop1<=273 and intervalgee in (4) then
delete;

    if positive=2 and lostovertime=1 and 1<=timeofloss<=90 and
intervalgee in (2,3,4) then delete;
    if positive=2 and lostovertime=1 and 91<=timeofloss<=181 and
intervalgee in (3,4) then delete;
    if positive=2 and lostovertime=1 and 182<=timeofloss<=273 and
intervalgee in (4) then delete;

    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum3to6=0
and testnum6to9=0 and testnum9to12=0 and intervalgee=2 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and testnum6to9=0
and testnum9to12=0 and intervalgee=3 then delete;
    if positive=1 and
testnum0to3+testnum3to6+testnum6to9+testnum9to12 ne 0 and
testnum9to12=0 and intervalgee=4 then delete;

keep z i group race age educ income residence orientation lostovertime
timeofloss positive testnum0to3 testnum3to6 testnum6to9 testnum9to12
intervalgee testedp3mo;

run;

*Deleting intermediate datasets;
ods select none;

proc datasets library=work;
delete simtrialcox00 simtrialcox0 simtrialcox1 simtrialcox2
simtrialconditional0 simtrialgee0 simtrialgee1 simtrialgee2
simtrialgee3;
run;

ods select all;

*Power and sample size - 2 x 2;

*Creating new dataset Test2by2 - Renaming dichotomous outcome (Tested 3
or more times vs. less than 3 times);

data test2by2;
set simtriallogistic;

    if group=0 then group=2;
    if ntest3ormore=0 then ntest3ormore=2;

run;

*Obtaining estimates of Relative risk (Risk ratio (Coll RR - When
Exposure on left and Outcome on top))

```

Creating output dataset Chi with values for estimates of Relative risk and their CIs;

```
ods select none;
```

```
proc freq data=test2by2;
table group*ntest3ormore / all;
by z;
output out = test2by2res relrisk;
run;
```

```
ods select all;
```

*Power - Proportion of those CIs that DO NOT include 1 where Relative risk > 1;

```
data power2by2;
set test2by2res;

    if l_rrc1 gt 1 then powerpr=1;
    else powerpr=2;

run;
```

```
ods select none;
```

```
proc freq data=power2by2;
tables powerpr / out=a;
run;
```

```
ods select all;
```

*Power and sample size - Standard Logistic Regression;

```
ods select none;
```

```
proc logistic data=simtriallogistic descending;
model ntest3ormore = group;
by z;
ods output oddsratios = testlogisticcrude;
run;
```

```
ods select all;
```

```
data powerlogisticcrude;
set testlogisticcrude;

    if effect ne "group" then delete;

    if lowercl gt 1 then powerslrcrude=1;
    else powerslrcrude=2;
```

```
run;
```

```
ods select none;
```

```

proc freq data=powerlogisticcrude;
tables powerslrcrude / out=b;
run;

ods select all;

*Power and sample size - GEE for Correlated Data;

ods select none;

proc genmod data=simtrialrepeated descending;
class i intervalgee;
model testedp3mo = group / dist=bin link=logit;
repeated subject=i / type=&corrstruc within=intervalgee corrw;
estimate 'Log odds ratio (Intervention vs. Comparison)' group 1 / exp;
by z;
ods output estimates = testgeecrude;
run;

ods select all;

data powergeecrude;
set testgeecrude;

    if label="Log odds ratio (Intervention vs. Comparison)" then
delete;

    if lbetalowercl gt 1 then powergeerepcrude=1;
    else powergeerepcrude=2;

run;

ods select none;

proc freq data=powergeecrude;
tables powergeerepcrude / out=d;
run;

ods select all;

*Power and sample size - Stratified Cox PH Regression for Recurrent
Events;

*Conditional model 1 - Models the full time course of the recurrent
event process
                                (Time interval of a subsequent event
starts at the end of the time interval for the previous event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start1,stop1)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;

```

```

ods output parameterestimates = testconditional1crude;
run;

ods select all;

data powerconditional1crude;
set testconditional1crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg1crude=1;
    else powerphreg1crude=2;

run;

ods select none;

proc freq data=powerconditional1crude;
tables powerphreg1crude / out=f;
run;

ods select all;

*Conditional model 2 - Models the time between each of the recurring
events
                                (Time interval starts at 0 and ends at
the length of time until the next event);

ods select none;

proc phreg data=simtrialconditional covs(aggregate);
model (start2,stop2)*testingevent(0) = group / rl;
id i;
strata intervalph;
by z;
ods output parameterestimates = testconditional2crude;
run;

ods select all;

data powerconditional2crude;
set testconditional2crude;

    if parameter ne "group" then delete;

    if hrlowercl gt 1 then powerphreg2crude=1;
    else powerphreg2crude=2;

run;

ods select none;

proc freq data=powerconditional2crude;
tables powerphreg2crude / out=h;
run;

ods select all;

```



```

*Combined dataset for values of power from different analytic methods;

data finalpowercrude;
set a b d f h;

    powerlabel="";
    if powerpr=1 then powerlabel="Power RR";
    if powerslrcrude=1 then powerlabel="Power SLR Crude";
    if powergeerepocrude=1 then powerlabel="Power GEE Crude";
    if powerphreglcrude=1 then powerlabel="Power Cox PH Conditional I
Crude";
    if powerphreg2crude=1 then powerlabel="Power Cox PH Conditional
II Crude";

    power=percent;

run;

data crudepowerresults;
set finalpowercrude;
keep powerlabel power;
where powerlabel in ("Power SLR Crude", "Power GEE Crude", "Power Cox
PH Conditional I Crude", "Power Cox PH Conditional II Crude");
run;

proc print data=crudepowerresults noobs;
title "Crude Analyses (&loss Loss to follow-up, &pos Positive): &m
Participants, Beta0 = &beta0, Beta1 = &betalor ";
run;

title;

%mend;

*****;

*Assuming OR=2 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 2;

%simtrialmacro (1, 1000, 100, -1.88968, 0.69315, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\100m2checklosspor2.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.69315, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\200m2checklosspor2.csv';
run;

```

```
%simtrialmacro (3, 1000, 300, -1.88968, 0.69315, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\300m2checkelosspor2.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.69315, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\400m2checkelosspor2.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\500m2checkelosspor2.csv';
run;

%simtrialmacro (6, 1000, 600, -1.88968, 0.69315, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\600m2checkelosspor2.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.69315, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\700m2checkelosspor2.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.69315, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\800m2checkelosspor2.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.69315, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\900m2checkelosspor2.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.69315, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1000m2checkelosspor2.csv';
run;
```

```
%simtrialmacro (11, 1000, 1100, -1.88968, 0.69315, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1100m2checkelosspor2.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1200m2checkelosspor2.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.69315, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1300m2checkelosspor2.csv';
run;

%simtrialmacro (14, 1000, 1400, -1.88968, 0.69315, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1400m2checkelosspor2.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.69315, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1500m2checkelosspor2.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.69315, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1600m2checkelosspor2.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.69315, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1700m2checkelosspor2.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.69315, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1800m2checkelosspor2.csv';
run;
```

```

%simtrialmacro (19, 1000, 1900, -1.88968, 0.69315, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\1900m2checkelosspor2.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.69315, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
2\2000m2checkelosspor2.csv';
run;

*****;

*Assuming OR=2.5 (odds of testing >= 2 times comparing study arms) and
13.128% of Comparison arm participants test >= 2 times in a year -
Specification 4;

%simtrialmacro (1, 1000, 100, -1.88968, 0.91629, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\100m2checkelosspor25.csv';
run;

%simtrialmacro (2, 1000, 200, -1.88968, 0.91629, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\200m2checkelosspor25.csv';
run;

%simtrialmacro (3, 1000, 300, -1.88968, 0.91629, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\300m2checkelosspor25.csv';
run;

%simtrialmacro (4, 1000, 400, -1.88968, 0.91629, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\400m2checkelosspor25.csv';
run;

%simtrialmacro (5, 1000, 500, -1.88968, 0.91629, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\500m2checkelosspor25.csv';
run;

```

```
%simtrialmacro (6, 1000, 600, -1.88968, 0.91629, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\600m2checkelosspor25.csv';
run;

%simtrialmacro (7, 1000, 700, -1.88968, 0.91629, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\700m2checkelosspor25.csv';
run;

%simtrialmacro (8, 1000, 800, -1.88968, 0.91629, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\800m2checkelosspor25.csv';
run;

%simtrialmacro (9, 1000, 900, -1.88968, 0.91629, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\900m2checkelosspor25.csv';
run;

%simtrialmacro (10, 1000, 1000, -1.88968, 0.91629, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1000m2checkelosspor25.csv';
run;

%simtrialmacro (11, 1000, 1100, -1.88968, 0.91629, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1100m2checkelosspor25.csv';
run;

%simtrialmacro (12, 1000, 1200, -1.88968, 0.91629, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1200m2checkelosspor25.csv';
run;

%simtrialmacro (13, 1000, 1300, -1.88968, 0.91629, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1300m2checkelosspor25.csv';
run;
```

```

%simtrialmacro (14, 1000, 1400, -1.88968, 0.91629, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1400m2checkelosspor25.csv';
run;

%simtrialmacro (15, 1000, 1500, -1.88968, 0.91629, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1500m2checkelosspor25.csv';
run;

%simtrialmacro (16, 1000, 1600, -1.88968, 0.91629, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1600m2checkelosspor25.csv';
run;

%simtrialmacro (17, 1000, 1700, -1.88968, 0.91629, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1700m2checkelosspor25.csv';
run;

%simtrialmacro (18, 1000, 1800, -1.88968, 0.91629, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1800m2checkelosspor25.csv';
run;

%simtrialmacro (19, 1000, 1900, -1.88968, 0.91629, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\1900m2checkelosspor25.csv';
run;

%simtrialmacro (20, 1000, 2000, -1.88968, 0.91629, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive Modified2 OR
25\2000m2checkelosspor25.csv';
run;

*****;

*Assuming OR=3 (odds of testing >= 2 times comparing study arms) and 5%
of Comparison arm participants test >= 2 times in a year -
Specification 6;

```

```
%simtrialmacro (1, 1000, 100, -2.94447, 1.09861, 51, 1, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\100m2cdcellosspor3.csv';
run;

%simtrialmacro (2, 1000, 200, -2.94447, 1.09861, 52, 2, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\200m2cdcellosspor3.csv';
run;

%simtrialmacro (3, 1000, 300, -2.94447, 1.09861, 53, 3, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\300m2cdcellosspor3.csv';
run;

%simtrialmacro (4, 1000, 400, -2.94447, 1.09861, 54, 4, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\400m2cdcellosspor3.csv';
run;

%simtrialmacro (5, 1000, 500, -2.94447, 1.09861, 55, 5, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\500m2cdcellosspor3.csv';
run;

%simtrialmacro (6, 1000, 600, -2.94447, 1.09861, 56, 6, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\600m2cdcellosspor3.csv';
run;

%simtrialmacro (7, 1000, 700, -2.94447, 1.09861, 57, 7, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\700m2cdcellosspor3.csv';
run;

%simtrialmacro (8, 1000, 800, -2.94447, 1.09861, 58, 8, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\800m2cdcellosspor3.csv';
run;
```

```
%simtrialmacro (9, 1000, 900, -2.94447, 1.09861, 59, 9, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\900m2cdcelosspor3.csv';
run;

%simtrialmacro (10, 1000, 1000, -2.94447, 1.09861, 60, 10, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1000m2cdcelosspor3.csv';
run;

%simtrialmacro (11, 1000, 1100, -2.94447, 1.09861, 61, 11, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1100m2cdcelosspor3.csv';
run;

%simtrialmacro (12, 1000, 1200, -2.94447, 1.09861, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1200m2cdcelosspor3.csv';
run;

%simtrialmacro (13, 1000, 1300, -2.94447, 1.09861, 63, 13, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1300m2cdcelosspor3.csv';
run;

%simtrialmacro (14, 1000, 1400, -2.94447, 1.09861, 64, 14, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1400m2cdcelosspor3.csv';
run;

%simtrialmacro (15, 1000, 1500, -2.94447, 1.09861, 65, 15, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1500m2cdcelosspor3.csv';
run;

%simtrialmacro (16, 1000, 1600, -2.94447, 1.09861, 66, 16, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1600m2cdcelosspor3.csv';
run;
```



```

%simtrialmacro (17, 1000, 1700, -2.94447, 1.09861, 67, 17, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1700m2cdcelosspor3.csv';
run;

%simtrialmacro (18, 1000, 1800, -2.94447, 1.09861, 68, 18, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1800m2cdcelosspor3.csv';
run;

%simtrialmacro (19, 1000, 1900, -2.94447, 1.09861, 69, 19, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\1900m2cdcelosspor3.csv';
run;

%simtrialmacro (20, 1000, 2000, -2.94447, 1.09861, 70, 20, ar(1), 0.1,
0.9, 0.7, 0.3);
proc export data=crudepowerresults dbms = csv
outfile='H:\Simulations\SAS Output\Extreme Loss Positive CDC Modified2
OR 3\2000m2cdcelosspor3.csv';
run;

*****;

*Start listing log;
options notes source source2;

```

SimulationTablesFigures.sas

```

*Sample statistics for tables;

*Run SimulationMacro_Ideal: 25 replications of 1200 subjects assuming
OR=2 and 13.128% of Comparison arm participants test >= 3 times in a
year - CheckingIn;
%simtrialmacro (12, 25, 1200, -1.88968, 0.69315, 12, ar(1));

*Descriptive statistics tables;

*Table 1 - Demographic characteristics;

proc freq data=work.simtrial;
tables (race age educ income residence orientation)*group / norow
nopercent;
where z=5;
run;

```

```

*Table 2 - Annual HIV testing characteristics;

proc freq data=work.simtriallogistic;
tables (test3ormore annualtests)*group / norow nopercnt;
where z=5;
run;

proc freq data=work.simtrialrepeated;
tables (testnum0to3 testnum3to6 testnum6to9 testnum9to12)*group / norow
nopercnt;
where z=5 and intervalgee=1;
run;

proc means data=work.simtrialconditional median q1 q3 qrange mean min
max range maxdec=2;
class group;
var stop2;
where z=5 and testingevent=1;
run;

*Effect estimates table - ar(1) corelation structure;

*Table 3 - Point estimates from different analytic techniques;

proc print data=work.powerlogisticcrude;
where z in (11,12,13,14,15,16,17,18,19,20);
var z oddsratioest powerslrcrude;
run;

proc print data=work.powergeecrude;
where z in (11,12,13,14,15,16,17,18,19,20);
var z lbetaestimate powergeerepcrude;
run;

proc print data=work.powerconditionalcrude;
where z in (11,12,13,14,15,16,17,18,19,20);
var z hazardratio powerphreglcrude;
run;

proc print data=work.powerconditional2crude;
where z in (11,12,13,14,15,16,17,18,19,20);
var z hazardratio powerphreg2crude;
run;

*Num and Den for Standard Logistic Regression;

*Odds of testing >= 3 times per year among each arm;

proc freq data=work.simtriallogistic;
tables test3ormore*group / noprint out=logisticodds0;
where z in (11,12,13,14,15,16,17,18,19,20);
by z;
run;

data logisticodds1;

```

```

set logisticodds0;
if test3ormore=0 and group=0 then a1=count;
if test3ormore=0 and group=1 then a2=count;
if test3ormore=1 and group=0 then a3=count;
if test3ormore=1 and group=1 then a4=count;
if a1=. then a1=0;
if a2=. then a2=0;
if a3=. then a3=0;
if a4=. then a4=0;
run;

proc sort data=logisticodds1;
by z;
run;

proc means data=logisticodds1 noprint;
var a1 a2 a3 a4;
output out=logisticodds2 (drop=_type_ _freq_)
sum(a1-a4)=asum1-asum4;
by z;
run;

data logisticodds;
set logisticodds2;
oddsamongintervention=round(asum4/asum2,0.01);
oddsamongcomparison=round(asum3/asum1,0.01);
keep z oddsamongintervention oddsamongcomparison;
run;

proc print data=work.logisticodds;
var z oddsamongintervention oddsamongcomparison;
run;

*Num and Den for Stratified Cox PH Regression for Recurrent Events;

*Average rates of testing among each arm;

data conditionalavg0;
set simtrialconditional;
where stop1=365 and stop2 ne 0 and z in
(11,12,13,14,15,16,17,18,19,20);
run;

data conditionalavg1;
set conditionalavg0;
a1=annualtests;
a2=stop1;
run;

proc sort data=conditionalavg1;
by z group;
run;

proc means data=conditionalavg1 noprint;
var a1 a2;
output out=conditionalavg2 (drop=_type_ _freq_)
sum(a1-a2)=asum1-asum2;

```

```

by z group;
run;

data conditionalavg;
set conditionalavg2;
avgrateper1000pd=(round(asum1/asum2,0.00001))*1000;
keep z group avgrateper1000pd;
run;

proc print data=work.conditionalavg;
var z avgrateper1000pd;
where group=1;
run;

proc print data=work.conditionalavg;
var z avgrateper1000pd;
where group=0;
run;

*Num and Den for GEE for Correlated Data;

*Average 3-month interval specific odds of testing at least once among
each arm;

proc freq data=work.simtrial;
table group*z / norow nocol nopercnt;
where z in (11,12,13,14,15,16,17,18,19,20);
run;

*Intervention;
data repeatedavg0;
set simtrialrepeated;
where z in (11,12,13,14,15,16,17,18,19,20) and group=1;
run;

data repeatedavg1;
set repeatedavg0;
if intervalgee=1 then a1=testedp3mo;
if intervalgee=2 then a2=testedp3mo;
if intervalgee=3 then a3=testedp3mo;
if intervalgee=4 then a4=testedp3mo;
if a1=. then a1=0;
if a2=. then a2=0;
if a3=. then a3=0;
if a4=. then a4=0;
run;

proc sort data=repeatedavg1;
by z i;
run;

proc means data=repeatedavg1 noprint;
var a1 a2 a3 a4;
output out=repeatedavg2 (drop=_type_ _freq_)
sum(a1-a4)=asum1-asum4;
by z i;
run;

```

```

proc means data=repeatedavg2 noprint;
var asum1 asum2 asum3 asum4;
output out=repeatedavg3 (drop=_type_ _freq_)
sum(asum1-asum4)=oddsnum1-oddsnum4;
by z;
run;

data repeatedavg1;
set repeatedavg3;

if z=11 then oddsden1=614-oddsnum1;
if z=11 then oddsden2=614-oddsnum2;
if z=11 then oddsden3=614-oddsnum3;
if z=11 then oddsden4=614-oddsnum4;

if z=12 then oddsden1=613-oddsnum1;
if z=12 then oddsden2=613-oddsnum2;
if z=12 then oddsden3=613-oddsnum3;
if z=12 then oddsden4=613-oddsnum4;

if z=13 then oddsden1=582-oddsnum1;
if z=13 then oddsden2=582-oddsnum2;
if z=13 then oddsden3=582-oddsnum3;
if z=13 then oddsden4=582-oddsnum4;

if z=14 then oddsden1=588-oddsnum1;
if z=14 then oddsden2=588-oddsnum2;
if z=14 then oddsden3=588-oddsnum3;
if z=14 then oddsden4=588-oddsnum4;

if z=15 then oddsden1=611-oddsnum1;
if z=15 then oddsden2=611-oddsnum2;
if z=15 then oddsden3=611-oddsnum3;
if z=15 then oddsden4=611-oddsnum4;

if z=16 then oddsden1=609-oddsnum1;
if z=16 then oddsden2=609-oddsnum2;
if z=16 then oddsden3=609-oddsnum3;
if z=16 then oddsden4=609-oddsnum4;

if z=17 then oddsden1=597-oddsnum1;
if z=17 then oddsden2=597-oddsnum2;
if z=17 then oddsden3=597-oddsnum3;
if z=17 then oddsden4=597-oddsnum4;

if z=18 then oddsden1=621-oddsnum1;
if z=18 then oddsden2=621-oddsnum2;
if z=18 then oddsden3=621-oddsnum3;
if z=18 then oddsden4=621-oddsnum4;

if z=19 then oddsden1=596-oddsnum1;
if z=19 then oddsden2=596-oddsnum2;
if z=19 then oddsden3=596-oddsnum3;
if z=19 then oddsden4=596-oddsnum4;

if z=20 then oddsden1=599-oddsnum1;

```

```

if z=20 then oddsden2=599-oddsnum2;
if z=20 then oddsden3=599-oddsnum3;
if z=20 then oddsden4=599-oddsnum4;

ratio1=oddsnum1/oddsden1;
ratio2=oddsnum2/oddsden2;
ratio3=oddsnum3/oddsden3;
ratio4=oddsnum4/oddsden4;

avgoddsintervention=(ratio1+ratio2+ratio3+ratio4)/4;

run;

proc print data=work.repeatedavg1;
var z avgoddsintervention;
run;

*Comparison;
data repeatedavg4;
set simtrialrepeated;
where z in (11,12,13,14,15,16,17,18,19,20) and group=0;
run;

data repeatedavg5;
set repeatedavg4;
if intervalgee=1 then a1=testedp3mo;
if intervalgee=2 then a2=testedp3mo;
if intervalgee=3 then a3=testedp3mo;
if intervalgee=4 then a4=testedp3mo;
if a1=. then a1=0;
if a2=. then a2=0;
if a3=. then a3=0;
if a4=. then a4=0;
run;

proc sort data=repeatedavg5;
by z i;
run;

proc means data=repeatedavg5 noprint;
var a1 a2 a3 a4;
output out=repeatedavg6 (drop=_type_ _freq_)
sum(a1-a4)=asum1-asum4;
by z i;
run;

proc means data=repeatedavg6 noprint;
var asum1 asum2 asum3 asum4;
output out=repeatedavg7 (drop=_type_ _freq_)
sum(asum1-asum4)=oddsnum1-oddsnum4;
by z;
run;

data repeatedavgc;
set repeatedavg7;

if z=11 then oddsden1=586-oddsnum1;

```

```
if z=11 then oddsden2=586-oddsnum2;
if z=11 then oddsden3=586-oddsnum3;
if z=11 then oddsden4=586-oddsnum4;

if z=12 then oddsden1=587-oddsnum1;
if z=12 then oddsden2=587-oddsnum2;
if z=12 then oddsden3=587-oddsnum3;
if z=12 then oddsden4=587-oddsnum4;

if z=13 then oddsden1=618-oddsnum1;
if z=13 then oddsden2=618-oddsnum2;
if z=13 then oddsden3=618-oddsnum3;
if z=13 then oddsden4=618-oddsnum4;

if z=14 then oddsden1=612-oddsnum1;
if z=14 then oddsden2=612-oddsnum2;
if z=14 then oddsden3=612-oddsnum3;
if z=14 then oddsden4=612-oddsnum4;

if z=15 then oddsden1=589-oddsnum1;
if z=15 then oddsden2=589-oddsnum2;
if z=15 then oddsden3=589-oddsnum3;
if z=15 then oddsden4=589-oddsnum4;

if z=16 then oddsden1=591-oddsnum1;
if z=16 then oddsden2=591-oddsnum2;
if z=16 then oddsden3=591-oddsnum3;
if z=16 then oddsden4=591-oddsnum4;

if z=17 then oddsden1=603-oddsnum1;
if z=17 then oddsden2=603-oddsnum2;
if z=17 then oddsden3=603-oddsnum3;
if z=17 then oddsden4=603-oddsnum4;

if z=18 then oddsden1=579-oddsnum1;
if z=18 then oddsden2=579-oddsnum2;
if z=18 then oddsden3=579-oddsnum3;
if z=18 then oddsden4=579-oddsnum4;

if z=19 then oddsden1=604-oddsnum1;
if z=19 then oddsden2=604-oddsnum2;
if z=19 then oddsden3=604-oddsnum3;
if z=19 then oddsden4=604-oddsnum4;

if z=20 then oddsden1=601-oddsnum1;
if z=20 then oddsden2=601-oddsnum2;
if z=20 then oddsden3=601-oddsnum3;
if z=20 then oddsden4=601-oddsnum4;

ratio1=oddsnum1/oddsden1;
ratio2=oddsnum2/oddsden2;
ratio3=oddsnum3/oddsden3;
ratio4=oddsnum4/oddsden4;

avgoddscomparison=(ratio1+ratio2+ratio3+ratio4)/4;

run;
```

```

proc print data=work.repeatedavgc;
var z avgoddscomparison;
run;

*****;

*ITI figures;

*Run SimulationMacro_Ideal (no loss)
10 replications of 500 subjects assuming OR=2 and 13.128% of Comparison
arm participants test >= 3 times in a year - CheckingIn;
%simtrialmacro (5, 10, 500, -1.88968, 0.69315, 5, ar(1));

*Run SimulationMacro_Loss_Positive (40% loss and 5% incident diagnoses)
10 replications of 500 subjects assuming OR=2 and 13.128% of Comparison
arm participants test >= 3 times in a year - CheckingIn;
%simtrialmacro (5, 10, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.05,
0.95, 0.4, 0.6);

*Run SimulationMacro_ExtremeLoss_Positive (70% loss and 10% incident
diagnoses)
10 replications of 500 subjects assuming OR=2 and 13.128% of Comparison
arm participants test >= 3 times in a year - CheckingIn;
%simtrialmacro (5, 10, 500, -1.88968, 0.69315, 55, 5, ar(1), 0.1, 0.9,
0.7, 0.3);

proc print data=work.simtriallogistic;
where z=7 and group=1 and i in
(245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,26
2,263,264);
var i group annualtests;
run;

proc print data=work.simtriallogistic;
where z=7 and group=0 and i in
(196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,21
3,214,215);
var i group annualtests;
run;

proc print data=work.simtrialconditional;
where z=7 and group=1 and i in
(245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,26
2,263,264);
var i group testingevent stop2 positive lostovertime;
run;

proc print data=work.simtrialconditional;
where z=7 and group=0 and i in
(196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,21
3,214,215);
var i group testingevent stop2 positive lostovertime;
run;

```



```

*****;

*Frequency figures;

*Run SimulationMacro_Ideal: 1000 replications of 1200 subjects assuming
OR=2 and 13.128% of Comparison arm participants test >= 3 times in a
year - CheckingIn;
%simtrialmacro (12, 1000, 1200, -1.88968, 0.69315, 12, ar(1));

*SLR;
data logisticfreq;
set powerlogisticcrude;
slroddsratio=round(oddsratioest,0.1);
run;

proc freq data=logisticfreq;
table slroddsratio*powerslr crude / nofreq norow nocol;
run;

*GEE for correlated data;
data geefreq;
set powergeecrude;
geeoddsratio=round(lbetaestimate,0.1);
run;

proc freq data=geefreq;
table geeoddsratio*powergeerep crude / nofreq norow nocol;
run;

*Stratified Cox PH regression for recurrent events - Conditional 1;
data conditionallfreq;
set powerconditionallcrude;
hazardratio1=round(hazardratio,0.1);
run;

proc freq data=conditionallfreq;
table hazardratio1*powerphreg1 crude / nofreq norow nocol;
run;

*Stratified Cox PH regression for recurrent events - Conditional 2;
data conditional2freq;
set powerconditional2crude;
hazardratio2=round(hazardratio,0.1);
run;

proc freq data=conditional2freq;
table hazardratio2*powerphreg2 crude / nofreq norow nocol;
run;

*****;

*Retention figure;

```

```

*Scenario B - 40% loss and 5% incident diagnoses;

*Run SimulationMacro_Loss_Positive: 10 replications of 1200 subjects
assuming OR=2 and 13.128% of Comparison arm participants test >= 3
times in a year - CheckingIn;
%simtrialmacro (12, 10, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.05,
0.95, 0.4, 0.6);

data testb;
set simtrialconditional;
by z i;

if not last.i then delete;
if z ne 5 then delete;

run;

proc freq data=testb;
table group;
run;

proc freq data=testb;
table i;
where stop1 < 90 and group=1;
run;

proc freq data=testb;
table i;
where 90 <= stop1 < 181 and group=1;
run;

proc freq data=testb;
table i;
where 181 <= stop1 < 273 and group=1;
run;

proc freq data=testb;
table i;
where 273 <= stop1 < 365 and group=1;
run;

proc freq data=testb;
table i;
where stop1 < 90 and group=0;
run;

proc freq data=testb;
table i;
where 90 <= stop1 < 181 and group=0;
run;

proc freq data=testb;
table i;

```

```

where 181 <= stop1 < 273 and group=0;
run;

proc freq data=testb;
table i;
where 273 <= stop1 < 365 and group=0;
run;

*Scenario C - 70% loss and 10% incident diagnoses;

*Run SimulationMacro_ExtremeLoss_Positive: 10 replications of 1200
subjects assuming OR=2 and 13.128% of Comparison arm participants test
>= 3 times in a year - CheckingIn;
%simtrialmacro (12, 10, 1200, -1.88968, 0.69315, 62, 12, ar(1), 0.1,
0.9, 0.7, 0.3);

data testc;
set simtrialconditional;
by z i;

if not last.i then delete;
if z ne 5 then delete;

run;

proc freq data=testc;
table group;
run;

proc freq data=testc;
table i;
where stop1 < 90 and group=1;
run;

proc freq data=testc;
table i;
where 90 <= stop1 < 181 and group=1;
run;

proc freq data=testc;
table i;
where 181 <= stop1 < 273 and group=1;
run;

proc freq data=testc;
table i;
where 273 <= stop1 < 365 and group=1;
run;

proc freq data=testc;
table i;
where stop1 < 90 and group=0;
run;

```

```
proc freq data=testc;  
table i;  
where 90 <= stop1 < 181 and group=0;  
run;
```

```
proc freq data=testc;  
table i;  
where 181 <= stop1 < 273 and group=0;  
run;
```

```
proc freq data=testc;  
table i;  
where 273 <= stop1 < 365 and group=0;  
run;
```

```
*****;
```

A6: Press copy of Chapter 4 (*Contemporary Clinical Trials* 2015, 41:152-159)

Contemporary Clinical Trials 41 (2015) 152–159



Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Evaluating interventions to promote routine preventive screenings: A comparison of analytical outcomes

Akshay Sharma^{a,*}, Brent A. Johnson^b, Patrick S. Sullivan^a

^a Department of Epidemiology, Emory University Laney Graduate School and Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322, United States
^b Department of Biostatistics and Computational Biology, University of Rochester, 601 Elmwood Avenue, Rochester, NY 14642, United States

ARTICLE INFO

Article history:

Received 19 November 2014
 Received in revised form 20 January 2015
 Accepted 21 January 2015
 Available online 29 January 2015

Keywords:

Preventive screenings
 Analytical outcome choices
 Randomized trials with recurrent events
 Rapid home HIV self-testing
 Men who have sex with men

ABSTRACT

Background: Often in public health, we are interested in promoting routine preventive screenings (e.g., blood glucose monitoring, hypertension screening, or mammography). Evaluating novel interventions to encourage frequent screenings using randomized controlled trials can help inform evidence-based health promotion programs. When the desired behavior change is a recurrent event, specifying the most meaningful study outcomes may prove challenging.

Methods: To understand the efficiency of multiple approaches for evaluating an intervention seeking to increase regular health screenings we (a) simulated several replications of a trial with a positive intervention effect under various censoring scenarios, (b) formulated three different analytical outcome definitions (screening a certain number of times during the entire study period versus not, screening at least once within a clinically meaningful time period versus not, “hazard” or instantaneous rate of screening), and (c) compared them with regard to interpreting results and estimating power at different sample sizes.

Results: Approaches which better utilize detailed prospective data, while also accounting for within-participant correlations, are less likely to miss the actual underlying benefits conferred by a new prevention strategy compared to relying on a dichotomous measure derived from aggregating events over the study duration. Such approaches are also more powerful in realistic scenarios wherein some participants are lost to follow-up over time.

Conclusions: Researchers should carefully consider the choice of analytical outcomes and strive to employ more efficient approaches that model comprehensive event-specific information, rather than summarizing repeated measures into less-informative dichotomous responses, while designing and conducting trials with recurrent preventive screenings.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Preventive screenings are an important component of health promotion efforts which can potentially reduce the significant economic burden of diseases [1]. Subgroup-specific recommendations have been developed to guide the timely identification and treatment for numerous morbidities. For example, the United States Preventive Services Task Force (USPSTF) recommends screening for lipid disorders in all men aged ≥ 35 years and women at an increased risk for coronary heart disease aged ≥ 45 years every five years [2], screening for

Abbreviations: AR1, autoregressive first order; CDC, Centers for Disease Control and Prevention; CED, General Educational Development; GEE, Generalized Estimating Equations; HIV, human immunodeficiency virus; HR, hazard ratio; IQR, inter-quartile range; MSM, men who have sex with men; OR, odds ratio; PH, proportional hazards; RCT, randomized controlled trial; SLR, standard logistic regression; USPSTF, United States Preventive Services Task Force.

* Corresponding author at: Department of Epidemiology, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, United States. Tel.: +1 404 857 8460. E-mail address: ashar24@emory.edu (A. Sharma).

<http://dx.doi.org/10.1016/j.cct.2015.01.014>
 1551-7144/© 2015 Elsevier Inc. All rights reserved.

type 2 diabetes in adults with hypertension or hyperlipidemia at 3-year intervals [3], and screening for Hepatitis B infection in pregnant women at the time of their first prenatal visit [4].

Given the emergence of new screening technologies, it is important to evaluate public health strategies to promote regular health exams. Parallel group randomized controlled trials (RCTs), reported according to established standards [5], are considered to be the most rigorous scientific tool for testing new interventions. Despite available guidance for variations in trial design [6,7], intervention content [8,9] and mode of delivery [10], limited discussion exists regarding design and methodological aspects unique to RCTs with recurrent events during follow-up [11–13]. Examples of such events include episodes of healthcare utilization, screening mammography, self-monitoring blood glucose and cholesterol levels, and screening for human immunodeficiency virus (HIV) infection.

Researchers conducting trials to promote preventive screening behaviors need to determine a priori what constitutes a meaningful outcome. Imagine a situation where we would like all adults get examined for hypertension annually, but only 20% actually follow this recommendation. Suppose we are studying a new intervention aimed at increasing the frequency of screening. How should we define our outcome? One option is a dichotomous measure, such as checking for high blood pressure ≥ 5 times versus not over a 5-year period. However, this definition might misclassify meaningful behavioral changes (e.g. increasing from one to four screenings) as failures, and a promising intervention could be wrongly described as being ineffective. Further, someone could screen five times in a short period (e.g. within a year), but not again for the remaining 5 years, and still be counted as a success. Potential alternative outcomes are screening at least once within a 1-year interval or the rate of screening. Depending upon the choice of our outcome, different analytical approaches are needed to answer the primary question “Does the intervention work?” directly impacting adequate sample size estimation [14].

Despite the availability of different techniques to analyze recurrent events data [15–17], researchers often adopt naive approaches which either ignore the existence of multiple events, their timing during follow-up, or the correlation between repeated measures. For example, a recent systematic review of 83 RCTs evaluating interventions to prevent falls among the elderly indicated that more than half the studies inappropriately employed proportions/odds-ratio (OR) based approaches [18]. Another review of 105 longitudinal studies examining hospitalization data among heart failure patients found that 70% based their analyses on outcomes incorporating only the first admission, and almost one-third compared proportions of individuals experiencing one or more hospitalizations using either a chi-squared test or standard logistic regression (SLR) [19].

Given that researchers continue to summarize repeated measures into less-informative dichotomous responses, we sought to demonstrate how different choices of analytical outcomes impact the ability to detect true underlying intervention effects. To understand the efficiency of multiple approaches for evaluating an intervention seeking to increase routine preventive screenings we (a) simulated several replications of a “successful” RCT (i.e. one with a positive intervention effect) under various censoring scenarios, (b) formulated three outcome definitions (screening a certain number of times during

the entire study period versus not, screening at least once within a clinically meaningful time period versus not, “hazard” or instantaneous rate of screening) and performed corresponding analyses, and (c) compared them with regard to interpreting results and estimating power at different sample sizes. For demonstration purposes and our own scientific research interests, we are using the rationale of a randomized trial which seeks to determine the effectiveness of rapid HIV self-test kits in increasing testing among men who have sex with men (MSM) in the United States [20].

2. Methods

2.1. Simulation strategy

Consider an RCT among HIV-negative or unknown status MSM prospectively followed for one year. The intervention to be evaluated is one to increase the frequency of HIV screening by distributing rapid HIV self-test kits that can be used at home. Intervention arm participants are given self-test kits and comparison arm participants are provided resources for identifying local HIV testing services. Men can report their test results online at the time of screening or during quarterly surveys. Participants are censored either because they are newly diagnosed as HIV-positive or because they are lost to follow-up.

SAS version 9.3 [21] was used to simulate 360,000 iterations of such a trial under different assumptions and perform all subsequent analyses. Hypothetical participants were assigned demographic characteristics based on a previous study of behavioral risks involving voluntary HIV testing with a home specimen collection kit [22] and randomized to either the intervention or comparison arm. In that prospective study, 1% of participants had tested for HIV six times within a year, 1% had tested five times, 3% had tested four times, 8% had tested three times, 17% had tested twice, 31% had tested once and 39% had not tested even once within a year. Screening frequencies for simulated participants were generated using different assumptions for men in either trial arm. Annual HIV testing frequencies for men in the intervention arm were simulated such that the intervention was effective and that participants could screen for a maximum of six times. Testing days were uniformly generated on the interval 1 to 365, assuming that all days of the year were equally likely to be selected, and the specific days of screening were separated to obtain HIV testing behavior within four 3-month time intervals.

First, two variations of a “successful” RCT (i.e. one with a positive intervention effect) were simulated assuming that 13% of comparison arm men screened (a) ≥ 3 times annually (as previously observed [22]) and (b) ≥ 2 times annually, with the intervention truly doubling these odds (i.e., assuming an OR of 2). Essentially, in the first parameter specification the odds of testing ≥ 3 times per year among intervention arm participants were twice the odds of testing ≥ 3 times per year among men in the comparison arm, and in the second parameter specification the odds of testing ≥ 2 times per year among intervention arm participants were twice the odds of testing ≥ 2 times per year among men in the comparison arm. For sensitivity analyses, we considered four additional parameter specifications such that 13% of the comparison arm screened either ≥ 3 or ≥ 2 times annually with an OR of 2.5, and 5% of the comparison arm screened either ≥ 3 or ≥ 2 times annually with an OR of 3.

Datasets were simulated by increasing starting sample sizes from 100 to 2000 in increments of 100 for each specification under the following retention and incidence scenarios:

Scenario A (Ideal retention) – No loss to follow-up and no incident HIV diagnoses.

Scenario B (Modest retention) – 40% loss to follow-up and 5% incident HIV diagnoses.

Scenario C (Poor retention) – 70% loss to follow-up and 10% incident HIV diagnoses.

Censoring in Scenarios B and C was implemented by randomly designating some testers from Scenario A as incident positives, followed by selecting men from the remaining pool to be lost over time, and uniformly generating their last day of follow-up on the interval 1 to 365. Consult appendix for further details regarding the simulation strategy.

2.2. Analytical outcome definitions

Outcome definition 1 – Trial arms could naively be compared using a dichotomous outcome of screening a certain number of times during the entire study period versus not, e.g. testing ≥ 3 times versus fewer within a year, and performing SLR [23].

Outcome definition 2 – Alternatively, one could use a binary outcome of screening at least once within a clinically meaningful time period versus not, e.g. testing at least once versus not within a 3-month interval. Because screening behaviors for each individual may be related across consecutive time intervals, Generalized Estimating Equations (GEE) with an autoregressive first order (AR1) correlation structure could be employed for analyses [23].

Outcome definition 3 – Finally, one could examine whether the intervention increases the “hazard” or instantaneous rate of screening. Stratified Cox proportional hazards (PH) regression for recurrent events could be used to model survival times between screening tests. The total follow-up time approach uses “event times”, i.e. the actual time of two tests from baseline [12,24], whereas the gap time approach uses “gap times”, i.e. the interval length between subsequent tests [12,25]. Both models adjust for the fact that screening events may be correlated within individuals.

2.3. Power and sample size estimation

SLR, GEE for correlated data, and stratified Cox PH regression for recurrent events were used to analyze simulated datasets. For SLR, we excluded all those participants who were newly diagnosed but retained available screening data for those lost to follow-up. For GEE and stratified Cox PH regression, we used available information from all participants (including incident positives) until they were censored. Power for different approaches was estimated from 1000 trial replications for a fixed sample size by determining the proportion for which the intervention effect magnitudes were significantly greater than the null. This process was sequentially repeated with 20 starting sample sizes under each censoring scenario for all six parameter specifications. Consult appendix for data layouts, model specifications and sample SAS code.

3. Results

3.1. Preliminaries

Table 1 illustrates our randomization process with sample data on 1200 simulated participants from a single trial replication. Study arms are balanced with respect to race/ethnicity, age, education, income, urban or rural residence, and self-identified sexual orientation. Annual screening characteristics under the first set of parameters in Scenario A are presented in Table 2. As specified, men in the intervention arm had higher testing frequencies versus those in the comparison arm, resulting in shorter inter-test intervals. Consult appendix for an illustration of screening patterns under different censoring scenarios (Figs. A1, A2 and A3).

Table 3 presents point estimates from different analytical approaches on 10 trial replications under the first set of parameters in Scenario A, each with 1200 participants. By design, the SLR ORs comparing trial arms with respect to testing ≥ 3 times annually are close to 2. Although we present results from different methods in a single table for convenience, their interpretations differ and their magnitudes are not directly comparable. We also performed a subset of analyses wherein we adjusted for demographic covariates, and our findings remained unchanged (data not shown in table). Frequency distributions of estimates from 1000 replications are presented in the appendix (Figs. A5 and A6).

Table 1
Illustrative demographic characteristics of 1200 simulated participants from a single replication of a randomized trial.

Characteristic	Intervention (N = 570)		Comparison (N = 630)	
	n	(%)	n	(%)
Race/ethnicity				
White, non-Hispanic	368	(65)	414	(66)
Black, non-Hispanic	88	(15)	93	(15)
Hispanic	114	(20)	123	(20)
Age group (years)				
18–24	229	(40)	246	(39)
25–34	199	(35)	219	(35)
35–44	79	(14)	106	(17)
≥ 45	63	(11)	59	(9)
Education				
College or post graduate	195	(34)	234	(37)
Some college or associate's degree	271	(48)	290	(46)
High school or GED	94	(16)	91	(14)
Less than high school or unknown	10	(2)	15	(2)
Annual income				
<\$14,999	183	(32)	200	(32)
\$15,000–\$39,999	176	(31)	214	(34)
\$40,000–\$74,999	119	(21)	117	(19)
>\$75,000	68	(12)	65	(10)
Unknown	24	(4)	34	(5)
Residential location				
Urban	345	(61)	407	(65)
Rural	191	(34)	191	(30)
Unknown	34	(6)	32	(5)
Self-identified sexual orientation				
Homosexual	485	(85)	529	(84)
Bisexual	72	(13)	84	(13)
Heterosexual or unknown	13	(2)	17	(3)

Abbreviations: GED, General Educational Development.

Table 2
Screening characteristics from a single replication of a randomized trial with 1200 simulated participants assuming the first parameter specification under Scenario A.

Characteristic	Intervention (N = 570)		Comparison (N = 630)	
	n	(%)	n	(%)
Categorical				
Annual testing frequency:				
0	171	(30)	256	(41)
1	141	(25)	203	(32)
2	108	(19)	87	(14)
3	92	(16)	42	(7)
4	34	(6)	29	(5)
5	8	(1)	6	(1)
6	16	(3)	7	(1)
	(N = 399 with a total of 905 screening tests)		(N = 374 with a total of 691 screening tests)	
Continuous				
Inter-test interval* (days)	Median	IQR	Median	IQR
	79	(36–160)	97	(43–191)

Abbreviations: IQR, inter-quartile range.

* Defined as time between consecutive screenings within a year among participants who tested at least once, with time to the first test being recorded from study entry.

3.2. Graphical comparison of analytical outcomes

Fig. 1 depicts two sets of power curves generated under the first and second parameter specifications in Scenario A. Each

panel includes multiple analytical approaches corresponding to different outcome choices on 20,000 trial replications. Compared to other methods, SLR (using the first outcome definition of testing ≥ 3 times versus fewer within a year) was more sensitive to variations in the underlying benefit conferred by the intervention. Despite being powerful when the intervention actually increased the odds of screening ≥ 3 times annually (top panel), this approach performed poorly when the intervention increased the odds of screening ≥ 2 times annually (bottom panel). Expressed alternatively, sample size estimation in order to attain a desired level of statistical power was subject to more uncertainty when using the first outcome definition compared to both other choices. For example, depending upon the actual success of the intervention in improving screening behaviors, recruiting either 500 participants would suffice (top panel) or 1200 would be needed (bottom panel) to detect a difference between trial arms with the conventional 80% power.

Figs. 2 and 3 demonstrate that compared to SLR, models for GEE for correlated data (using the second outcome definition of testing at least once versus not within a 3-month interval) and stratified Cox PH regression for recurrent events (using the third outcome definition of the instantaneous rate of testing) performed increasingly better in detecting the intervention effect as greater proportions of participants were censored over time (Scenarios B and C respectively). Similar patterns were observed on comparing results from sensitivity analyses with four additional parameter specifications (Figs. A7, A8, A9, A10, A11 and A12 presented in the appendix).

Table 3
Point estimates from different analytical approaches on 10 replications of a randomized trial each with 1200 simulated participants assuming the first parameter specification under Scenario A.

Replication # (Intervention N, Comparison N)	SLR			GEE for correlated data ^a			Stratified Cox PH regression for recurrent events				
	Odds of testing ≥ 3 times within a year among:			Average 3-month interval-specific odds of testing at least once among:			Average rate of testing (per 1000 person-days) among:		Total follow-up time model ^b	Gap time model ^c	
	Intervention	Comparison	OR ^d	Intervention	Comparison	OR ^e	Intervention	Comparison	HR ^f	HR ^g	HR ^h
1 (614, 586)	0.36	0.14	2.62	0.43	0.29	1.49	4.19	2.91	1.36	1.36	1.36
2 (613, 587)	0.32	0.16	1.97	0.40	0.31	1.31	3.94	3.16	1.18	1.19	1.19
3 (582, 618)	0.31	0.20	1.55	0.37	0.35	1.06	3.90	3.55	1.08	1.09	1.09
4 (588, 612)	0.29	0.17	1.70	0.42	0.33	1.26	4.14	3.24	1.23	1.23	1.23
5 (611, 589)	0.29	0.17	1.69	0.38	0.34	1.13	3.86	3.30	1.13	1.14	1.14
6 (609, 591)	0.33	0.17	1.96	0.39	0.34	1.15	3.93	3.32	1.16	1.15	1.15
7 (597, 603)	0.41	0.19	2.19	0.48	0.33	1.44	4.59	3.34	1.29	1.30	1.30
8 (621, 579)	0.34	0.12	2.75	0.42	0.31	1.36	4.13	3.05	1.30	1.30	1.30
9 (596, 604)	0.32	0.14	2.28	0.40	0.31	1.27	3.96	3.10	1.23	1.24	1.24
10 (599, 601)	0.32	0.18	1.75	0.40	0.30	1.33	3.95	3.13	1.20	1.19	1.19

Note: Results from different approaches presented in a single table for convenience. Magnitudes of the ratio effect estimates are not directly comparable. Consult appendix for frequency distributions of point estimates from 1000 replications (Figs. A5 and A6).

Note: Results from different approaches presented in a single table for convenience. Magnitudes of the ratio effect estimates are not directly comparable. Consult appendix for frequency distributions of point estimates from 1000 replications (Figs. A5 and A6).

Abbreviations: AR1, autoregressive first order; GEE, Generalized Estimating Equations; HR, hazard ratio; OR, odds ratio; PH, proportional hazards; SLR, standard logistic regression.

^a AR1 working correlation structure assumed for describing the relationship between 3-month interval-specific testing behaviors within each participant.

^b Modeling repeated screening test patterns over a year.

^c Modeling time between each of the recurrent screening test events.

^d Interpreted as the odds of testing ≥ 3 times within a year among intervention arm participants are x times the odds of testing ≥ 3 times within a year among comparison arm participants. Estimates in bold represent a statistically significant effect ($P < 0.05$).

^e Interpreted as the odds of testing at least once in a 3-month interval among intervention arm participants are x times the odds of testing at least once in a 3-month interval among comparison arm participants after accounting for correlation from repeated measures. Estimates in bold represent a statistically significant effect ($P < 0.05$).

^f Interpreted as the instantaneous rate of testing among intervention arm participants is x times the instantaneous rate of testing among comparison arm participants after accounting for recurrent screening events. Estimates in bold represent a statistically significant effect ($P < 0.05$).

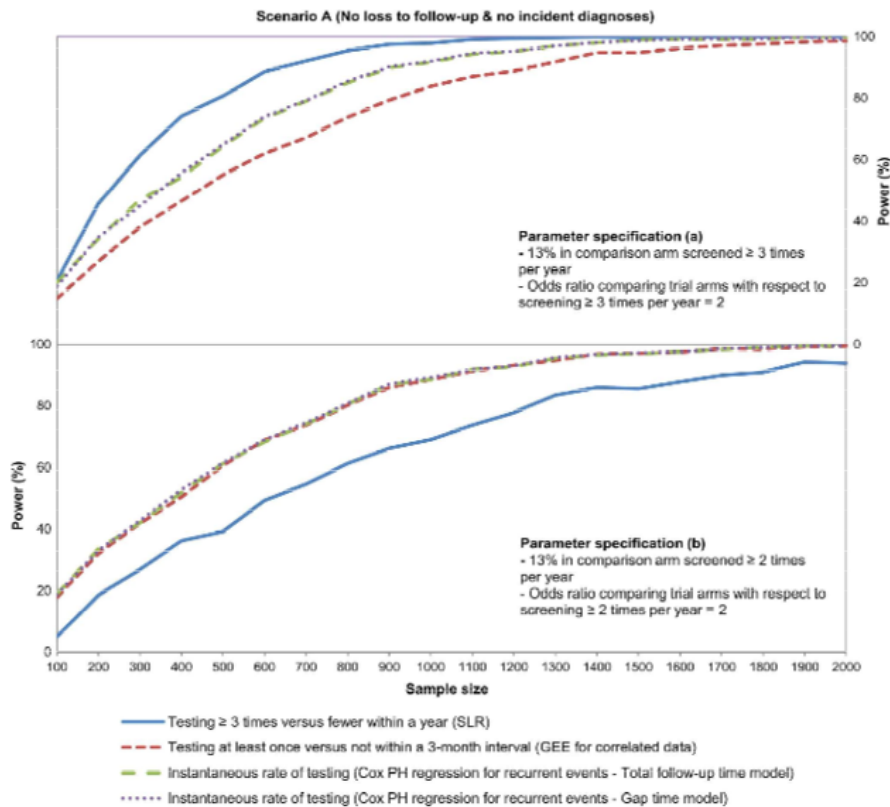


Fig. 1. Power curves from 1000 trial replications assuming the first and second parameter specifications under Scenario A.

4. Discussion

Researchers designing trials with recurrent screening events are often challenged with defining appropriate study endpoints. Our goal was to demonstrate how different analytical outcome choices can impact the ability to detect true underlying intervention effects. Therefore, we used simulations of varyingly “successful” RCTs to examine the efficiency of three outcome definition alternatives and their corresponding analyses. We conclude that approaches that model detailed prospective data while accounting for within-participant correlations are better than those which depend upon appropriately specifying a dichotomous outcome measure. In particular, when the dichotomous outcome is defined as number of tests exceeding a certain cut-point, SLR may fail in correctly assessing the effect of the intervention unless the theoretical and user-defined cut-point align closely; GEE for

correlated data and survival analysis methods for recurrent events do not suffer this limitation. Besides being less likely to miss the actual benefits of a new prevention strategy, they are also more powerful in realistic scenarios encompassing attrition over time. Although we based our simulations on a trial of rapid HIV self-testing among MSM [20], our findings are more widely applicable to longitudinal studies with repeated health screenings. Depending upon the condition for which someone is getting evaluated, a positive result could lead them to either test more often or to cease screening altogether. For example, pregnant women might increase their frequency of screening for gestational diabetes if their glucose levels are not under control, but someone who receives a positive HIV test result would most likely stop screening, seek confirmatory testing and initiate treatment. This highlights the importance of employing statistical methods which utilize all available information to the fullest extent. Next, we focus on practical

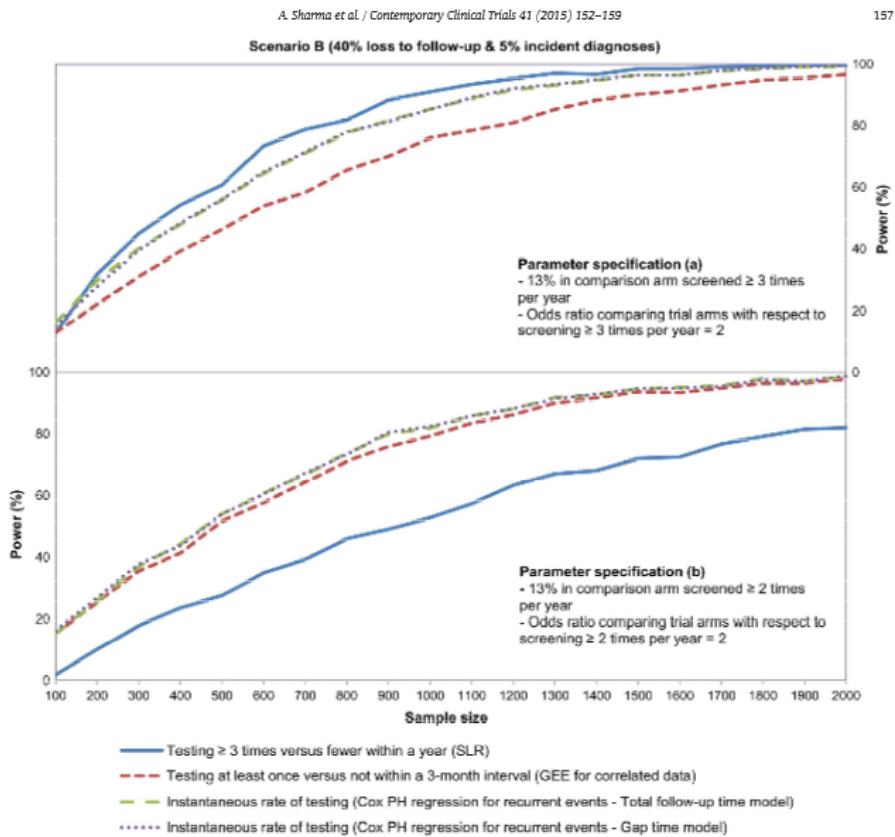


Fig. 2. Power curves from 1000 trial replications assuming the first and second parameter specifications under Scenario B.

issues while choosing relevant outcome measures and some methodological caveats of corresponding analytical approaches.

Suppose an intervention arm participant from our example self-tested three times on the first day of follow-up, and then did not test during the remainder of the study. Using the first outcome definition he would be considered a success at the end of a year, even though the practical value may only have been that of a single screening test. Contrast this with a comparison arm participant who tested twice annually at regular 6-month intervals (say first with his physician and later at a community based organization). Despite having a more reasonable interest interval, he would not be considered a success. Simply aggregating the number of screenings per individual at study conclusion and dichotomizing for SLR while ignoring the key aspect of when they were performed can therefore be problematic. Further, our simulations demonstrated that if the choice of cut-point is not reflective of the actual prevention

implications of a novel screening approach, researchers may fail to identify a genuinely superior intervention. Depending upon the disease for which one is screening (e.g. HIV, certain cancers), analyzing data in this manner could also necessitate excluding all new diagnoses, even if they had tested a few times before identification. From our example, someone who was identified as HIV-positive before his third test should not be considered a failure with regard to screening ≥ 3 times per year because he might have continued to screen for HIV, and eventually be counted as a success had his previous test result been negative. Yet another statistical decision involves handling missing data due to incomplete follow-up (i.e., should such observations be retained or removed from the dataset?)

Choosing an analytical outcome of any screening within shorter and perhaps more clinically meaningful intervals can help resolve some of these issues. Consider someone from the intervention arm in our example who decided to test because

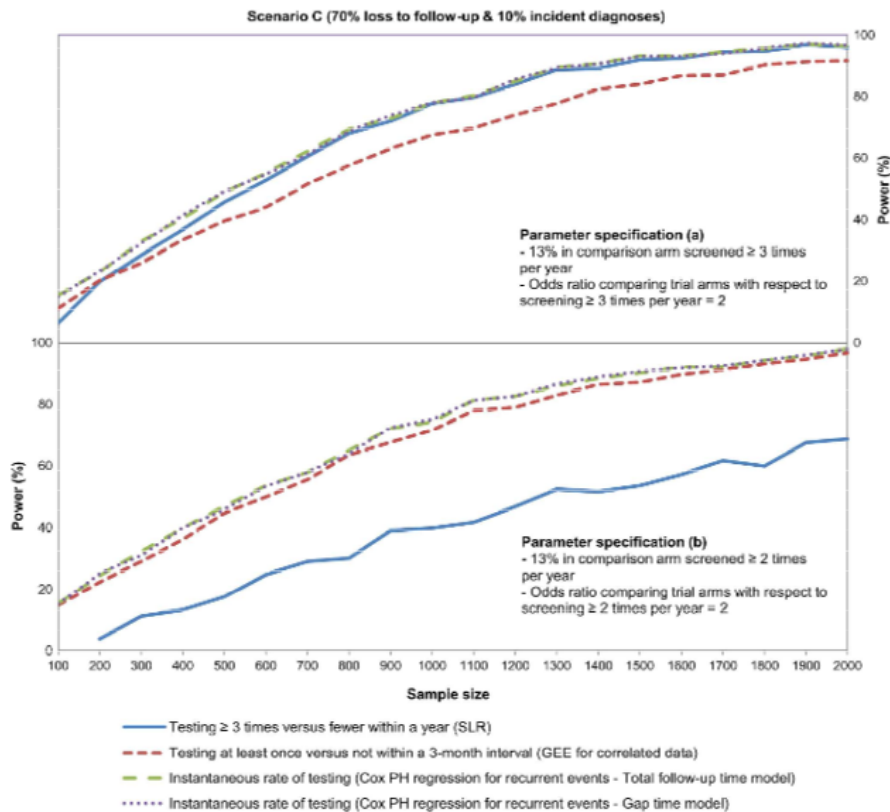


Fig. 3. Power curves from 1000 trial replications assuming the first and second parameter specifications under Scenario C.

he recently had unprotected sex. Suppose his first test result was negative but he self-tested again on the same day “just to be sure”. He later learned about the approximately 3-month long window period (interval between the time of infection and the production of measurable antibodies to HIV), and decided to test again after 3 months. Using the second outcome definition, his first two tests would be appropriately counted as a single success in one interval and his third test would be counted as a single success in the next interval. Although GEE for correlated data allows for an assessment of the intervention effect in shorter time periods, even this approach does not take into account the actual timing of screening tests. However unlike SLR, this option is less sensitive to variations in the true underlying intervention effects, newly diagnosed individuals need not be excluded and one can better utilize information from individuals who are lost to follow-up.

Researchers who plan on collecting data about the days on which each screening test was performed should consider

using the third outcome definition. Although several analytical methods exist to compare instantaneous rates between trial arms [12], we employed the total follow-up time and gap time models for stratified Cox PH regression for recurrent events. Given the level of detail in our data, we consider these choices to be more appropriate than count models. Both utilize all available information including the order and actual timing of subsequent screening tests. Unlike the previous options considered, one does not have to pre-specify a binary outcome such as testing certain number of times versus not within a particular time period. Power calculations assuming different parameter specifications also showed how both models for this approach were least affected by the manner and extent to which the intervention truly increased preventive screenings in the study population.

Our simulations also demonstrated that methods which more fully utilized available data (stratified Cox PH regression and GEE) versus aggregating information (SLR) performed

better in detecting the intervention effect as greater proportions of participants were censored over time. These results were in the expected direction because men who dropped out during later months before their third screening test were coded as failures when using the first outcome definition, and all new diagnoses had to be excluded. Although 40% (Scenario B) and 70% (Scenario C) losses to follow-up by year-end seem unusually high, more than 80% of the sample was retained in each arm at the first follow-up assessment in each scenario (Fig. A4 presented in the appendix). This exceeds the 70% retention threshold set by the Centers for Disease Control and Prevention (CDC) to support a designation of “best-evidence” HIV prevention interventions [26]. Loss to follow-up might not be problematic in short-term studies, but trials of a longer duration (e.g. an RCT evaluating an intervention to increase routine screenings where the recommended interval is a year) may benefit from analytical approaches that utilize maximum information.

Limitations of our study include making relatively simplistic assumptions while choosing probability distributions and parameters for the simulation process. HIV testing frequencies for comparison arm participants in three of the six parameter specifications were based on an actual behavioral study among MSM [22], but these might not be representative of current trends. However, similar results were observed on performing sensitivity analyses with different specifications in each of the three censoring scenarios and we are confident regarding the validity of our findings. Although the assumption that nobody could test more than six times annually is arbitrary, we believe that counts higher than this would be extremely rare. Because we simulated screening patterns within a year-long trial, the median inter-test intervals of our hypothetical participants were shorter than what have been reported in a recent study following MSM over eight years [27]. We also acknowledge we have not considered all analytical options for the third outcome definition (e.g. marginal approach for stratified Cox PH regression [17], independent-increment model for the rate of recurrent events [28]). The well-established counting process approach was not employed as this assumes independence between different screening tests on the same individual [16,25]. Negative binomial and zero-inflated models were not considered as these necessitate summarizing repeated outcomes on each participant with a single, one-dimensional measure of the total count of screening tests [29]. However, if the study design is such that only aggregated data rather than actual event times are collected, these alternatives could be used to estimate intervention effects without losing much efficiency [30].

Ultimately, the research question of interest and goals of the RCT should dictate the manner in which recurrent events data are measured and analyzed. Simple and familiar approaches are often appealing, but aggregating and dichotomizing information on repeated screening tests without regard to their timing could miss what are otherwise significant implications of a novel prevention strategy to increase routine health monitoring.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.cct.2015.01.014>.

Funding acknowledgement

This work was facilitated by the Center for AIDS Research at Emory University (P30 AI050409).

References

- [1] NPHPPHC. National prevention strategy. Rockville, MD: National Prevention, Health Promotion and Public Health Council; 2011.
- [2] USPSTF. Final recommendation statement: lipid disorders in adults (cholesterol, dyslipidemia): screening; 2013.
- [3] USPSTF. Final recommendation statement: diabetes mellitus (Type 2) in adults: screening; 2008.
- [4] USPSTF. Final recommendation statement: hepatitis B in pregnant women: screening; 2014.
- [5] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010; 8:18.
- [6] Campbell MK, Piaggio G, Elbourne DR, Altman DG. CONSORT 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345.
- [7] Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 2012;308:2594–604.
- [8] Boutron I, Moher D, Altman DG, Schulz KF, Ravau P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295–309.
- [9] Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med* 2006;144:364–7.
- [10] Eysenbach G, Group C-E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13.
- [11] Cook RJ. The design and analysis of randomized trials with recurrent events. *Stat Med* 1995;14:2081–98.
- [12] Kuramoto L, Sobolev BC, Donaldson MG. On reporting results from randomized controlled trials with recurrent events. *BMC Med Res Methodol* 2008;8:35.
- [13] Cook RJ, Bergeron P-J, Boher J-M, Liu Y. Two-stage design of clinical trials involving recurrent events. *Stat Med* 2009;28:2617–38.
- [14] Delucchi KL. Sample size estimation in research with dependent measures and dichotomous outcomes. *Am J Public Health* 2004;94:372–7.
- [15] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981;68:373–9.
- [16] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982;1100–20.
- [17] Wei L-J, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Stat Assoc* 1989;84:1065–75.
- [18] Donaldson MG, Sobolev B, Cook WL, Janssen PA, Khan KM. Analysis of recurrent events: a systematic review of randomised controlled trials of interventions to prevent falls. *Age Ageing* 2009;38:151–5.
- [19] Metcalfe C, Thompson S, Cowie M, Sharples L. The use of hospital admission data as a measure of outcome in clinical studies of heart failure. *Eur Heart J* 2003;24:105–12.
- [20] ClinicalTrials.gov. Evaluation of rapid HIV self-testing among MSM (eSTAMP); 2014.
- [21] SAS 9.3 macro language: reference. Cary, NC: SAS Institute; 2011.
- [22] Khosropour CM, Johnson BA, Ricca AV, Sullivan PS. Enhancing retention of an internet-based cohort study of men who have sex with men (MSM) via text messaging: randomized controlled trial. *J Med Internet Res* 2013;15.
- [23] Kleinbaum D, Klein M. Logistic regression: a self-learning text. 3rd ed. New York: Springer Science + Business Media, LLC; 2010.
- [24] Twisk JW, Smit D, de Vente W. Applied analysis of recurrent events: a practical overview. *J Epidemiol Community Health* 2005;59:706–10.
- [25] Kleinbaum D, Klein M. Survival analysis: a self-learning text. 3rd ed. New York: Springer Science + Business Media, LLC; 2012.
- [26] CDC. FRS efficacy criteria for best-evidence (tier 1) risk reduction (RR) individual-level, group-level, and couple-level interventions (IIs/CLIs/CPLs); 2014.
- [27] Katz DA, Dombrowski JC, Swanson F, Buskin SE, Golden MR, Stedler JD. HIV intertest interval among MSM in King County, Washington. *Sex Transm Infect* 2013;89:32–7.
- [28] Lin D, Wei L, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *J R Stat Soc Ser B Stat Methodol* 2000; 62:711–30.
- [29] Cook RJ, Lawless JF. Analysis of repeated events. *Stat Methods Med Res* 2002;11:141–66.
- [30] Dean CB, Balshaw R. Efficiency lost by analyzing counts rather than event times in Poisson and overdispersed Poisson regression models. *J Am Stat Assoc* 1997;92:1387–98.