



Evaluating the Impact of OSHA's Enforcement Interventions

Evaluation Design Report

December 20, 2010

Authors

Daniel Weissbein

Marios Michaelides

Jacob Benus

Submitted to:

Kristin Lantz
Program Specialist
Center for Program Planning and Results
US Department of Labor
200 Constitution Avenue, NW Rm. S3317
Washington, DC 20210

Submitted by:

Daniel Weissbein – Project Director
Labor and Social Policy Research Division
IMPAQ International, LLC
10420 Little Patuxent Parkway # 300
Columbia, MD 21044
443-539-9763

Evaluation Design Report

Table of Contents

EXECUTIVE SUMMARY	i
INTRODUCTION	1
1. REVIEW OF THE SST PROGRAM.....	3
2. EVALUATION DESIGN OPTIONS	7
2.1. Evaluation Design 1 – Primary Inspection List Sites.....	9
2.1.1. Implementation Steps.....	9
2.1.2. Required Sample Sizes.....	12
2.1.3. Evaluation Analyses.....	13
2.1.3. Summary of Evaluation Design 1	18
2.2. Evaluation Design 2 – Primary and Secondary Inspection List Sites.....	19
2.2.1. Implementation Steps.....	20
2.2.2. Required Sample Sizes.....	23
2.2.3. Evaluation Analyses.....	23
2.2.3. Summary of Evaluation Design 2	30
2.3. Evaluation Design 3 – All High Rate Letter Sites	31
2.3.1. Implementation Steps.....	32
2.3.2. Required Sample Sizes.....	35
2.3.3. Evaluation Analyses.....	35
2.3.3. Summary of Evaluation Design 3	38
2.4. Comparison of Evaluation Designs and Implications for the SST Program.....	39
3. EVALUATION DESIGN IMPLEMENTATION	43
3.1. Random Assignment	44
3.2. Sample Attrition.....	44
3.3. Process for Minimizing the Effects of Attrition.....	46
3.4. Timing of Inspections	48
4. EMPLOYER BEHAVIOR	51
5. CONCLUSION AND RECOMMENDATIONS.....	54

EXECUTIVE SUMMARY

The Occupational Safety and Health Administration (OSHA) is the U.S. Department of Labor agency responsible for promoting safe and healthful working conditions. Under the 1970 Occupational Safety and Health Act, OSHA is authorized to conduct inspections of worksites to determine whether employers are compliant with workplace safety and health standards. The majority of OSHA inspections are programmed inspections, where the agency selects and targets establishments based on a number of criteria, including injury/illness rates, citation history, or random selection. OSHA also conducts inspections initiated as a result of imminent danger, fatalities, referrals, and complaints. One of the most important OSHA-programmed inspection initiatives is the Site-Specific Targeting (SST) program. The SST program focuses on inspections of non-construction employers with 40 or more workers. In a typical year, the SST program conducts approximately 2,900 inspections of employers with high injury/illness rates.

The purpose of the current project, *Evaluating the Impact of OSHA's Enforcement Interventions*, is to design an evaluation to assess the impact of OSHA's SST program on employer compliance. Specifically, the scope of the project is to design a rigorous evaluation of the impact of SST program which can provide answers to the following research questions:

- What is the impact of SST program inspections on future employer?
- Which employer characteristics are strong indicators of future compliance?
- What are the best practices and measures to reduce future occupational injuries and illnesses among employers, investigated or otherwise?

This report provides a detailed description of three random assignment evaluation design options that can be used to assess the impact of the SST program's on employer compliance. We also discuss some important implementation issues that OSHA should take into account. Finally, we make recommendations regarding which evaluation design option to adopt and discuss the implementation process for ensuring the evaluation is rigorously implemented.

The SST Program

The SST program is one of OSHA's programmed inspection programs, which targets non-construction workplaces with at least 40 workers in high-hazard industries. The program targets establishments based on their reported injury/illness rates, using two types of interventions:

- *High Rate Letter* – This letter informs employers they are among workplaces with high injury/illness rates, urges them to take corrective action, and informs them of available OSHA resources.
- *Inspections* – OSHA Area Offices conduct inspections of employers with the highest injury/illness rates to determine whether they are in compliance with OSHA workplace safety and health standards.

Each year, OSHA uses data from the OSHA Data Initiative (ODI) survey to calculate site-specific injury/illness rates for employers that are under OSHA's jurisdiction. OSHA then determines an injury/illness rate threshold which is set to identify approximately 14,000 sites to receive the high rate letter. The SST program categorizes these employers into three groups:

- *Primary inspection list* – Includes sites with the highest injury/illness rates among those receiving the high rate letter. In a typical year, there are 3,500 primary inspection list sites. These sites have the highest priority for receiving an SST inspection.
- *Secondary inspection list* – Includes sites with injury/illness rates that are lower than the rates of primary inspection list sites but still well above the high rate letter threshold. In a typical year, there are 7,000 secondary inspection list sites. These sites are inspected only after all primary inspection list sites in an Area Office's jurisdiction are inspected.
- *High rate letter only* – Includes sites with the lowest injury/illness rates among those receiving the high rate letter. In a typical year, there are 3,500 high rate letter only sites. These sites are generally not inspected under the SST program.

Evaluation Design Options

The evaluation design options presented in this report will allow for the assessment of the impact of the high rate letter, the incremental impact of an SST inspection, and the total impact of the

SST program (high rate letter and inspection). This report presents three rigorous random assignment evaluation options which can be implemented by OSHA to estimate the impact of the SST program:

- ***Evaluation Design 1*** – This design enables OSHA to estimate the impact of the SST program on compliance for primary inspection list sites. This design involves the random selection of primary list sites in fiscal year (FY) 2011 into three groups: 1) sites that receive high rate letter and an inspection in FY 2011, 2) sites that receive the high rate letter but no inspection in FY 2011, and 3) control group sites that receive neither a letter nor an inspection in FY 2011. All sites in these three groups will receive a follow-up inspection to determine their level of compliance in FY 2012. This design enables OSHA to assess whether the SST program’s interventions are effective in improving compliance for primary inspection list sites which are the main focus of the SST program.
- ***Evaluation Design 2*** – This design expands the scope of Evaluation Design 1 as it enables OSHA to estimate the impact of the SST program on compliance for primary inspection list *and* for secondary inspection list sites. This design involves a random selection process similar to Evaluation Design 1, but including both primary inspection list and secondary inspection list sites. This design enables OSHA to assess whether the SST program is effective for primary and secondary inspection list sites and whether there is value in expanding the program to conduct a higher number of inspections of secondary inspection list sites.
- ***Evaluation Design 3*** – This design expands the scope of Evaluation Design 2 as it enables OSHA to estimate the impact of the SST program for all employers exceeding the high rate letter threshold (primary inspection list, secondary inspection list, and high rate letter only). This design involves the same random selection process as the first two designs. Evaluation Design 3 will enable OSHA to assess whether the SST program is effective for all sites receiving the high rate letter and whether there is value in expanding the program to conduct a higher number of inspections of secondary inspection list and high rate letter only sites.

All three evaluation design options require some alterations to the regular operations of the SST program in FY 2011, including: 1) avoid sending high rate letters to some eligible sites and 2)

avoid inspecting a portion of sites that are eligible for inspection. Additionally, all sites in the evaluation sample must receive a follow-up inspection in FY 2012 regardless of their reported injury/illness rates. As Evaluation Design 1 only involves primary list sites, this design will enable OSHA to estimate the SST program's impact while causing a minimal disruption to regular program operations. Evaluation Design 2 involves both primary and secondary list sites, so it may require a small increase in the regular number of inspections of secondary list sites. Evaluation Design 3 is the most comprehensive design as it allows OSHA to assess the program for all sites exceeding the high rate letter threshold. However, this design would require OSHA to inspect sites that are not on the inspection lists.

Evaluation Design Implementation

In implementing the selected evaluation design, there are a number of implementation issues that OSHA needs to take into account to maintain the validity of the design.

Random Assignment – The rigorous implementation of random assignment is the single most important element for ensuring the validity of the selected evaluation design. The evaluation contractor must use appropriate software to randomly assign sites into the evaluation sample.

Attrition – In performing random assignment, the evaluation contractor must account for sample attrition to ensure that a sufficiently high number of evaluation sites are available for the follow-up inspections. This will provide sufficient statistical power to detect the impact of the SST program. Sample attrition may take two forms: 1) *initial attrition* involving sites that may be out of business or ineligible for the SST program in FY 2011, and 2) *attrition over time* involving sites that may go out of business before the follow-up inspections. We propose a process that would help eliminate the impact of attrition on the validity of the design. This process requires the implementation contractor to randomly select a sufficient number of sites to offset expected attrition.

Timing of Inspections – If OSHA implements Evaluation Design 3, the agency would need to conduct a high number of follow-up inspections of high rate letter only sites in the year following random assignment. As such sites are not typically inspected by the SST program, this

design would require reducing the total number of inspections of primary list sites to accommodate the evaluation. The report provides three implementation options for the timing of inspections under Evaluation Design 3:

- *Option 1* – Conduct all initial inspections in FY 2011 and all follow-up inspections in FY 2012. This option is the simplest to implement and minimizes the effects of attrition or other external conditions on the validity of the evaluation.
- *Option 2* – Conduct all initial inspections in FY 2011 and split follow-up inspections between FY 2012 and FY 2013. This option reduces the number of follow-up inspections of high rate letter only sites that must be conducted in a given year, but may result in higher sample attrition over time.
- *Option 3* – Conduct initial inspections in FY 2011 and in FY 2012 and conduct follow-up inspections from FY 2012 through FY 2014. This is the most complex of the three options. This option substantially reduces the number of inspections of high rate letter only sites that must be conducted in a given year, but may result in greater sample attrition due to the length of the evaluation.

Our recommendation is that OSHA adopts option 1 for implementing the selected design. If Evaluation Design 3 is selected, OSHA should consider option 2 and 3 only if inspecting a high number of high rate letter only sites in a given year is a primary concern.

Employer Behavior – We also developed an approach that can be used to identify how employers change their practices to improve workplace safety and health as a result of SST interventions. Specifically, we recommend that the evaluation contractor work with OSHA to develop a set of questions for employers regarding specific behaviors undertaken over the prior year, such as changes to the safety and health program, safety and health training, equipment additions or changes, personnel changes, policies and procedures, etc. These questions could be asked as part of follow-up inspections. A better approach would be to develop a survey that could be sent to establishments participating in the evaluation. The data could then be linked to the evaluation data, including the injury/illness rates and the violation outcome data. This survey data could be used to assess whether particular behaviors were related to reductions in injury/illness rates or violations.

Conclusion and Recommendations

This document provides OSHA with three options for implementing a rigorous experimental design evaluation of the SST program. Our recommendation is that OSHA consider adopting Evaluation Design 2 to rigorously evaluate the impact of the SST program on employer compliance for primary inspection list and secondary inspection list sites. We believe that this design strikes a balance between providing a rigorous assessment of the effectiveness of the SST program for both primary and secondary list sites, while requiring only small changes to the regular operations of the SST program. Evaluation Design 2 would provide important insight on the effectiveness of the program for primary list sites and would allow OSHA to assess the value of adjusting the SST program's focus in the future.

To complement the implementation of the selected evaluation design, we also recommend that OSHA consider implementing a survey of employers in the evaluation sample to gather information on behavior regarding workplace safety and health. This survey would enable OSHA to examine the impact of the SST program on employer behavior and to identify best practices that may be related to improved employer compliance and injury/illness rates. Our recommendation is that this survey be implemented near the end of the follow-up inspection period.

INTRODUCTION

The Occupational Safety and Health Administration (OSHA) is charged under the 1970 Occupational Safety and Health Act to assure, so far as possible, safe and healthful working conditions for every working individual in the nation. Under the Act, OSHA is authorized to conduct workplace inspections and investigations to determine whether employers are complying with standards issued by OSHA for safe and healthful workplaces. OSHA also enforces Section 5(a)(1) of the Act, known as the “General Duty Clause,” which requires that every working man and woman must be provided with a safe and healthful workplace.

The majority of OSHA inspections are programmed inspections, where the agency selects and targets establishments based on a number of criteria, including injury incidence rates, previous citation history, worker exposure to toxic substances, or random selection. OSHA also conducts un-programmed inspections, which are initiated as a result of imminent danger, fatalities, catastrophe, complaints, referrals, or follow-up.

There are two categories of programmed inspections. The Local and National Emphasis program inspection programs focus on a particular safety or health hazard (e.g., amputations, silica) or the hazards of a specific industry (e.g., logging, trenching). The Site-Specific Targeting (SST) program focuses on non-construction workplaces that have 40 or more workers. The SST program targets establishments based on high injury/illness rates. Typically, of the approximately 40,000 inspections OSHA conducts annually, about 2,900 are SST inspections.

The purpose of the current project, *Evaluating the Impact of OSHA’s Enforcement Interventions*, is to design an evaluation to assess the impact of OSHA’s SST program on employer compliance. Specifically, the scope of the project is to design a rigorous evaluation of the impact of SST program, and provide answers to the following research questions:

- What is the impact of SST program inspections on future employer compliance?
- Which employer characteristics are strong indicators of future compliance?
- What are the best practices and measures to reduce future occupational injuries and illnesses among employers, investigated or otherwise?

In addition, the project has been expanded to incorporate a design for evaluating the impact of the *high rate* letter on employer behavior. High rate letters are sent to roughly 14,000 employers annually whose injury/illness rates place them among the highest group of the 80,000 businesses in the targeted industries surveyed in the OSHA Data Initiative.

In this report, we present evaluation design options that can be used to evaluate the impact of the SST program on employer compliance. Specifically, the options presented are designed to evaluate the following:

- Impact of the high rate letter.
- Incremental impact of an inspection.
- Total impact of the SST program (high rate letter and inspection).

To estimate the impact of the SST program on employer compliance, the IMPAQ team proposes the use of a random assignment experimental design, which is widely regarded as the gold standard for evaluating program impacts. Such a design would enable OSHA to obtain accurate estimates of the impact of the SST program on employer compliance and gain critical information for implementing the program in the future.

This report is organized as follows. Section 1 provides an overview of the SST program and the selection procedures used by OSHA under this program. Section 2 presents a discussion of evaluation design options for evaluating the impact of the SST program on employer compliance. This section provides a detailed description of each design and its objectives and the impact analyses that can be produced based on each design. It also provides a comparison of the design options and the implications for the SST program for each. Section 3 provides a discussion of evaluation implementation issues including random assignment, attrition, and options for implementing the evaluation designs over time. Section 4 provides an overview of an option for collecting information on employer behavior related to safety and health as part of the evaluation designs. Section 5 provides a summary and conclusions.

1. REVIEW OF THE SST PROGRAM

The Site-Specific Targeting (SST) program is one of OSHA's programmed inspection programs, which targets private industry non-construction sites with 40 or more workers.¹ The SST process of selecting sites for inspection is based on data collected in the OSHA Data Initiative (ODI) survey. ODI collects data on injuries and acute illnesses attributable to work-related activities in private-sector industries from approximately 80,000 establishments in selected high-hazard industries. Since 1999, the ODI surveys have been sent to establishments with 40 or more employees, requesting the following from each employer: 1) the average number of workers who worked for the employer during the prior calendar year, 2) the total hours the workers worked during the prior year, and 3) the summary injury/illness data from the employer's OSHA Log form. OSHA uses these data to calculate establishment-specific injury/illness rates and to target enforcement and compliance assistance activities.

High Rate Letters. As shown in Exhibit 1, about 80,000 sites are sent the ODI survey, of which 48,000 are under Federal jurisdiction, thus eligible for the SST program. Based on the injury/illness rate information provided in the prior year's ODI, specific injury/illness rate thresholds are set for the SST program to select approximately 14,000 employer sites annually with injury/illness rates substantially higher than the national average. These sites receive an initial intervention known as the *high rate letter*. The letter informs the employer they are among the workplaces with the highest injury/illness rates, urges them to take corrective action, and informs them of OSHA resources that may be of assistance. In fiscal year (FY) 2008, OSHA sent approximately 14,000 high rate letters to employers with a lost workday injury/illness rate of 8.0 or higher (i.e., for every 100 employees there were eight or more injuries or illnesses which resulted in lost work days).²

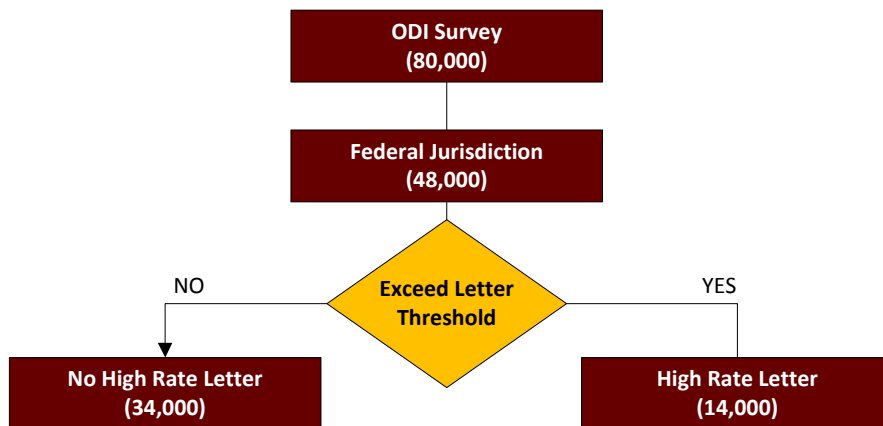
Prior to FY 2004, the high rate letter threshold was set by the Lost Workday Injury and Illness (LWDII) rate, which is calculated as the ratio of the number of lost work day injuries and illnesses combined divided by the total number of hours worked by all workers during the

¹ A site is what is referred to in the directives as an establishment, a single physical location where business is conducted or where services or industrial operations are performed.

² The national average injury/illness rate at the time was three.

calendar year per 100 full time equivalent workers. In FY 2002, a revised recordkeeping system went into place and was reflected in FY 2004, when the injury/illness rate used for the threshold was changed to the *Days Away from work, Restricted work, or job Transfer due to injury and illness* (DART) rate, which uses the number of cases involving days away and/or restricted work activity, and/or job transfer as the numerator. Whichever rate was used, the specific threshold has varied over time to keep number of letters sent roughly consistent across the years. For example, between FY 2004 and FY 2008, the DART rate threshold to receive a high rate letter has dropped from 7.0 to 5.0 as the national DART average declined from 2.8 to 2.1. SST program staff also considered other information such as employer size or the raw number of injuries/illnesses in making the final selection of letter recipients. As shown in Exhibit 1, of the approximately 48,000 sites under federal jurisdiction each fiscal year, about 34,000 do not exceed the high rate letter threshold, while the remaining 14,000 exceed the threshold and receive the high rate letter.

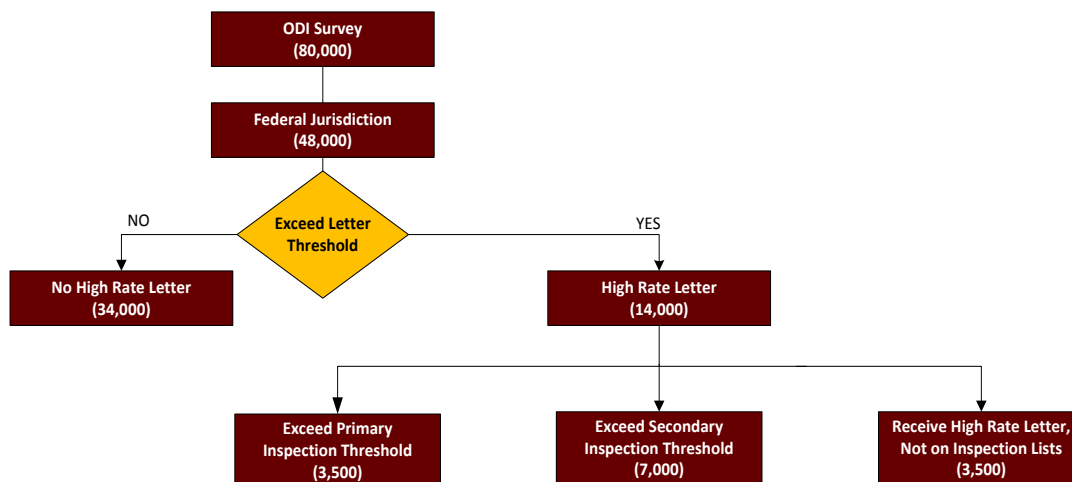
Exhibit 1: Overview of SST Program, High Rate Letter



SST Inspections. Subsequent to the high rate letters, each year OSHA sets injury/illness rate thresholds to determine priority sites for inspections under the SST program. There are generally two thresholds set for selection of sites. A higher injury/illness rate threshold is set for a *primary inspection list*, considered higher priority sites. A second, lower threshold is set for a *secondary inspection list* of sites to be inspected – resources permitting – when the primary inspection list is exhausted by an OSHA Area Office. The inspection thresholds are higher for the primary list, so this list includes sites with the highest injury/illness rates.

The specific injury/illness rates used for the two inspection lists have varied over the years. Initially, the LWDII was used to set selection list thresholds. Then starting in FY 2003, a second rate was taken into consideration, the Days Away from Work Injuries and Illness (DAFWII) case rate.³ Separate thresholds were set such that if the ODI data indicated that the site exceeded *either* the LWDII *or* DAFWII thresholds, the site would be eligible for inspection. From 2004 and in subsequent years, the thresholds were set using the DAFWII and DART rates. Similar to the high rate letters, the thresholds for the inspection lists have been lowered somewhat over time. As shown in Exhibit 2, based on the thresholds set every year, there are approximately 3,500 sites on the primary inspection list; 7,000 on the secondary inspection list; and 3,500 sites are not on these inspection lists as they do not exceed the injury/illness rate thresholds for either the primary or secondary inspection lists.⁴

Exhibit 2: Overview of SST Program, SST Inspection Lists



To schedule SST inspections, the OSHA National Office provides each OSHA Area Office with access to the secure SST website which contains the database of establishments on the primary inspection list for its coverage area. Using the website, OSHA Area Offices generate *inspection cycles* – the size of the inspection cycle can vary from 5 to 50 sites and is determined by each

³ DAFWII is a component of the LWDII comprised of injury and illness cases that involve at least one day away from work.

⁴ It should be noted that, through FY 2008, the SST program had separate thresholds for nursing home sites and non-nursing home sites. Additionally, in FY 2009, the SST program began setting separate thresholds for manufacturing and non-manufacturing sites. Based on interviews with OSHA personnel, our understanding is that nursing homes will be the focus of a National Emphasis program and removed from the SST program for the immediate future. Therefore, this report only focuses on evaluation design options for manufacturing and non-manufacturing sites.

Area Office based on the its available resources and geographic range. Once an inspection cycle is created, the Area Office begins inspecting sites in the cycle until all sites are inspected. Sites within an inspection cycle may be scheduled and inspected in any order that makes efficient use of available resources. All establishments in a cycle that have not been inspected within the past 36 months (based on the most recent directive) and with no consultation visit in progress or other disqualifying activity based on the directive must be inspected before any establishments in a new cycle may be inspected. When a cycle is completed, the OSHA Area Office can generate a new cycle using the SST website.

Once an OSHA Area Office inspects all sites on its primary inspection list, it may request additional sites for inspection from the National Office. The National Office then provides the secondary inspection list which the OSHA Area Office uses to inspect additional sites. From FY 2001 through FY 2008, OSHA conducted an average of approximately 2,900 site inspections under the SST program – nearly three quarters of these (about 2,150 sites) were inspections of primary inspection list sites while the remaining (about 750 sites) were inspections of secondary inspection list sites.

Other Inspections. In FY 2004 and subsequent years, a very small number of non-respondents to the ODI survey were added to the primary inspection list. Additionally, in FY 2007, to verify the reliability of claims by establishments that they have achieved low DART rates, analysts in OSHA's Office of Statistical Analysis selected approximately 100 low-rate establishments in high-rate industries by applying a random number table to all establishments within the jurisdiction of each OSHA Area Office that have reported both a DART rate from 0.0 to 2.4 and 100 workers or more. In FY 2008, this number was increased to approximately 175 low-rate establishments (DART rate from 0.0 through 2.4).

2. EVALUATION DESIGN OPTIONS

In this section, we present evaluation design options for assessing the impact of the SST program on employer compliance with OSHA regulations. The SST program focuses on the following groups of eligible establishments:

- *Primary inspection list* – These sites have the highest injury/illness rates. As a result, they receive the high rate letter and are the highest priority for receiving an inspection.
- *Secondary inspection list* – These sites have injury/illness rates lower than those of the primary list, but still sufficiently high to receive the high rate letter. These sites may be inspected only after all eligible primary inspection list sites within an OSHA Area Office’s jurisdiction are inspected.
- *High rate letter, not on inspection lists* – These sites have injury/illness rates sufficiently high to receive the high rate letter, but not high enough to qualify for the primary or secondary inspection lists. These sites are unlikely to be inspected under the SST program.

All employers exceeding the annually established high rate letter threshold receive the high rate letter under the current SST program. At the same time, primary list establishments (which have the highest reported injury/illness rates) are the main focus of SST inspections, with secondary list establishments being inspected only after all primary list employers in the OSHA Area Offices’ jurisdiction are inspected.

As part of the present study, IMPAQ conducted an analysis of employer violation outcomes and the relationship of the violations to prior SST program interventions.⁵ Results of this analysis showed that non-nursing home employers that were previously inspected by OSHA were between 4 and 12 percentage points less likely to be in violation than employers that were not previously inspected. Although these results provide evidence that SST program interventions are associated with lower future violation rates, they cannot be viewed as reliable impact

⁵ See *Further Examination of OSHA Data* (May 31, 2010).

estimates of the program's effectiveness. At this point, there are no impact estimates on the effectiveness of SST interventions on the compliance behavior of employers.

The main objective of the evaluation design options presented in this section is to estimate the impact of the SST program on employer compliance. These design options will allow an assessment of the:

- Impact of the high rate letter.
- Incremental impact of an SST inspection.
- Total impact of the SST program (high rate letter and inspection).

Furthermore, the evaluation design options will enable OSHA to produce estimates of the overall compliance rate and to conduct analyses of employer characteristics and how they relate to employer compliance.

IMPAQ proposes the use of a rigorous random assignment evaluation design, which is widely regarded as the gold standard for estimation of program impacts. Below, we describe three random assignment evaluation design options that can be used to estimate the impact of the SST program on employers with different injury/illness rates:

- *Evaluation Design 1* – estimates the impact of the SST program on primary inspection list employers.
- *Evaluation Design 2* – extends the design to estimate the impact of the SST program on both primary and secondary inspection list employers.
- *Evaluation Design 3* – further extends the design to estimate the impact of the SST program on all employers exceeding the high rate letter threshold.

In Section 2.1 we describe Evaluation Design 1, including a detailed description of the design, the analyses that can be conducted using this design, and the implementation procedures OSHA must follow to implement it. Sections 2.2 and 2.3 provide a similar summary for Evaluation Design 2 and Evaluation Design 3, respectively. Section 2.4 provides a comparison of the three designs and discusses their advantages and disadvantages.

2.1. Evaluation Design 1 – Primary Inspection List Sites

One of the main objectives of the SST program is to conduct inspections of establishments with the highest reported injury/illness rates (i.e., primary inspection list sites). These inspections aim to assess employer compliance with OSHA regulations and help employers improve the safety of their workplace. In the past ten years, about 75 percent of all inspections conducted under the SST program targeted primary inspection list employers.

The focus of Evaluation Design 1 is to estimate the impact of the SST program on employers in the primary inspection list using a random assignment experimental design approach. Using Evaluation Design 1, OSHA will be able to obtain an accurate estimate of the impact of the high rate letter, the incremental impact of the inspection, and the total impact of the SST program on employers in the primary inspection list. A detailed description of the design, the analyses that can be produced based on the design, and the design implementation requirements are presented below.

2.1.1. Implementation Steps

The purpose of Evaluation Design 1 is to enable OSHA to produce a rigorous evaluation of the SST program for primary inspection list sites. This design involves the following steps:

Step 1 – Random Selection of Primary List Sites

As shown in Exhibit 3, Evaluation Design 1 requires OSHA and their evaluation contractor to use the list of sites exceeding the primary inspection list in FY 2011 to randomly select sites and assign them into one of three groups: 1) control group, 2) treatment group A, and 3) treatment group B. These sites are hereafter referred to as the evaluation sample. Each group will include one third of the randomly selected sites from the primary inspection list in FY 2011. Sites assigned to the control group will not receive any intervention, that is, they will not receive the high rate letter and will not be inspected regardless of their reported injury/illness rate. Sites in treatment group A will receive the high rate letter but will not be inspected in FY 2011, while sites in treatment group B will receive the high rate letter and an inspection in FY 2011.

Step 2 – Send High Rate Letters to Treatment Group A and B Sites in FY 2011

As per the regular SST program operations, OSHA will send high rate letters in FY 2011 to sites assigned in treatment groups A and B. However, OSHA must not send high rate letters to the sites assigned in the control group despite the fact that their injury/illness rates would ordinarily qualify them to receive this letter.

Step 3 – Inspect Treatment Group B Sites in FY 2011

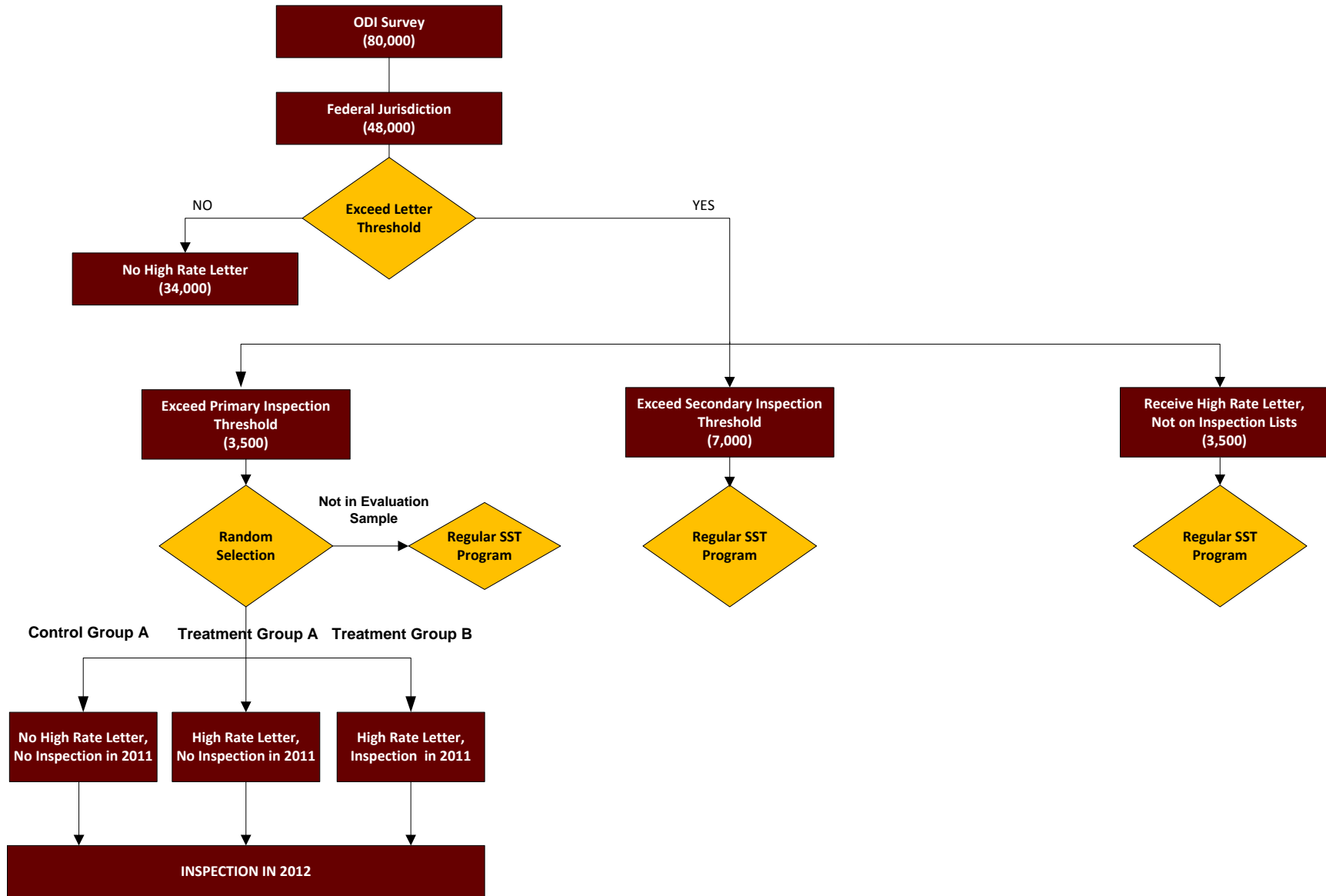
In FY 2011, OSHA must inspect sites in treatment group B, but must not inspect any of the sites in the control group and in treatment group A. Following the inspections, OSHA will document the OSHA inspection outcomes, including employer violations, types of violations, and financial penalties assessed and paid. This is the typical information OSHA collects and stores in the *Outcomes of OSHA Inspections* database.

Step 4 – Inspect All Control and Treatment Sites in FY 2012

Following the completion of the SST program in FY 2011, OSHA must inspect all sites in the evaluation sample, that is, all sites in the control group, treatment group A, and treatment group B. We propose that these follow-up inspections take place in FY 2012, which is one year following random assignment. This would minimize sample attrition and the possibility that the design is affected by changes in economic and other external conditions that occur over time.⁶ The FY 2012 follow-up inspections will be similar to the usual SST inspections, where OSHA inspectors gather information on employer compliance and types of violations. It is important to note that OSHA must inspect all sites in the control and two treatment groups, even if the sites' injury/illness rates in FY 2012 drop to a level that would ordinarily preclude them from an SST inspection. The objective of these inspections is to observe employer compliance for the three groups following the FY 2011 interventions.

⁶ Alternatively, the follow-up inspections can be distributed over multiple years following the end of the FY 2011 SST program. Section 3 discusses different options for conducting the follow-up inspections.

Exhibit 3: Evaluation Design 1 – Primary Inspection List Sites



2.1.2. Required Sample Sizes

One important consideration in implementing a random assignment design is to include a sufficiently high number of sites to ensure that the evaluation detects statistically significant estimates of the program's impact. Depending on the number of sites included in the evaluation, one can calculate the Minimum Detectable Effect (MDE). The MDE is the smallest program impact that the evaluation will likely be able to detect for a given outcome of interest (e.g., likelihood of violation). If the actual impact of the SST program on a given outcome is higher than the MDE, then one can be reasonably confident that the impact will be detected by the evaluation.

Let us consider the impact of the SST program on the likelihood of violation for primary list sites, which is the primary outcome of interest for this evaluation. The MDE for the likelihood of violation under Evaluation Design 1 is calculated based on: 1) the mean likelihood of violation for primary list sites,⁷ 2) the number of sites that will be included in the evaluation sample, and 3) the level of statistical significance and statistical power we want to achieve.⁸ Exhibit 4 presents the MDE for the likelihood of violation based on the total number of sites included in the evaluation. As shown in Exhibit 4, if Evaluation Design 1 includes 600 primary list sites (200 control, 200 treatment A, and 200 treatment B), the MDE for the likelihood of violation will be 10.9 percentage points. This means that if the impact of the SST program is at least 10.9 percentage points, an evaluation sample of 600 sites will enable us to detect the program's impact.

If OSHA includes a higher number of sites in the evaluation sample, the MDE would be lower. A sample of 900 sites would enable OSHA to detect if the SST program led to at least an 8.9 percentage-point impact on the violation rate of primary inspection list sites. By the same token, the MDE will be 6.3 percentage points if the evaluation sample includes 1,800 sites. Finally, an evaluation sample of 3,600 primary list sites will produce a MDE of 4.4 percentage points.

⁷ In the past few years, 71 percent of primary list sites inspected by the SST program were in violation.

⁸ The typical standard used in program evaluation is that the estimated impact is statistically significant at the 5 percent level. The typical standard for statistical power (i.e., the probability that a statistically significant impact is detected) is 80 percent. We also assume a .07 R-squared from the regression model estimating the likelihood of violation.

Exhibit 4: Minimum Detectable Effect on Likelihood of Violation, Evaluation Design 1

	Number of Sites in the Evaluation Sample				
	N = 600	N = 900	N = 1,200	N = 1,800	N = 3,600
Primary List Sites	10.9%	8.9%	7.7%	6.3%	4.4%

Note: Power calculations are based on a 0.71 average likelihood of violation.

There are two important considerations in determining the number of primary list sites included in Evaluation Design 1. First, a sufficiently high number of sites must be included in the evaluation to ensure that the Evaluation Design 1 will detect the program’s impact. Second, the number of sites included in the evaluation must be limited to a reasonable number so that the regular SST program is not burdened by the evaluation. Our recommended option is that OSHA includes 1,800 primary list sites in the evaluation (600 in the control, 600 in treatment A, and 600 in treatment B). This will produce a MDE of 6.3 percentage points using a relatively low number of sites. To facilitate the discussions for the remainder of this report, we assume OSHA will include 1,800 sites in the sample for Evaluation Design 1.

2.1.3. Evaluation Analyses

Random assignment of primary list sites to the control or to one of the two treatment groups ensures that, on average, employers in each group will have similar observed and unobserved characteristics at the time of random assignment. As a result, any subsequent differences in violation outcomes between treatment and control group sites would be attributed to the SST program. Below, we describe the types of analyses that can be conducted using the data gathered under Evaluation Design 1.

Descriptive Analyses

Prior to producing any impact analyses, it is useful to develop descriptive statistics for the control and the two treatment groups. These statistics will provide an overview of the baseline characteristics of treatment and control group sites, including employer size, industry, union status, and region. Descriptive analyses are necessary to provide an overall characterization of

the primary list sites included in the evaluation sample. In addition, these analyses are necessary to determine whether employers in the treatment and control groups were statistically similar in terms of their observable characteristics. Examining the degree to which the characteristics of the treatment and control groups are similar is important in assessing whether randomization was performed accurately and to establish the validity of comparing treatment with control group outcomes to estimate program impacts.

To formally test if randomization was done effectively, two linear regression models will be estimated where the dependent variables are the likelihood of being assigned to treatment group A and to treatment group B and control variables include all available site characteristics. Using t-tests, the evaluation contractor will assess whether any of the observed site characteristics have a statistically significant relationship to the likelihood of being assigned to treatment group A or to treatment group B. If the results indicate that baseline site characteristics do not have a significant relationship to the likelihood of treatment group assignment, this would suggest that randomization was done accurately and that a comparison of the treatment-control group outcomes would provide a consistent estimate of the program's impact.

Analyses Conducted Prior to Follow-up Inspections

The first set of analyses can be conducted after the FY 2011 inspections are concluded and before the completion of the follow-up inspections in FY 2012. Using the inspection data for treatment group B sites from FY 2011, the following analyses can be produced:

- *Estimate overall violation rate and other outcomes* – Using the inspection data for the 600 treatment group B sites, the evaluation contractor can estimate the violation rate (i.e., the total number of inspected sites with a violation over total inspected sites). As sites were randomly assigned to treatment group B, this violation rate is representative of the rate for all sites in the primary inspection list. Similar estimates can be obtained for other outcomes, including the number of violations, number of serious violations, and financial penalties assessed.
- *Characteristics of violators* – OSHA data typically report a number of site characteristics, including establishment size, union status, industry, and region. Using this information,

one can estimate the relationship between site characteristics and the likelihood of violation. This will help produce descriptive analyses to identify which types of employers are likely to be in violation based on their available characteristics.

In addition to the descriptive analyses above, it is feasible to obtain some preliminary estimates of the impact of the SST program using the ODI survey responses of sites in the control and the two treatment groups following FY 2011. For example, many of the 1,800 sites in the control and treatment groups will be required to report their injury/illness rates in 2012.⁹ Using this information, one can produce preliminary estimates of the impact of the SST program on employer injury/illness rates as follows:

- *Impact of high rate letter on injury/illness rates* – This is produced by a means comparison of the average injury/illness rate reported in 2012 between sites in treatment group A (high rate letter, no inspection) and sites in the control group (no high rate letter, no inspection). A simple t-test will be produced to assess whether the injury/illness rate for treatment group A sites is significantly lower than the rate for control group sites – i.e., if the receipt of the high rate letter led to a reduction in reported injury/illness rates for primary list sites.
- *Incremental impact of inspection on injury/illness rates* – This is produced by a means comparison of the average injury/illness rate reported in 2012 between sites in treatment group B (high rate letter, inspection) and sites in treatment group A (high rate letter, no inspection). This would estimate the additional impact of an SST inspection for primary list employers that received the high rate letter. A simple t-test will be produced to assess whether the injury/illness rates for treatment group B sites are significantly lower than the rate for control group sites, i.e., if the inspection resulted in a reduction in injury/illness rates for primary list sites, beyond the impact of the high rate letter.
- *Total impact of SST program on injury/illness rates* – This can be produced by a means comparison of the average injury/illness rate reported in 2012 between sites in treatment group B (high rate letter, inspection) and sites in control group A (no high rate letter, no inspection). A simple t-test will be produced to assess whether the injury/illness rate for

⁹ To ensure that analyses using the ODI data can be produced, OSHA will need to ensure all employers in the evaluation sample are selected for the ODI survey in subsequent years.

treatment group B sites is significantly lower than the rate for control group sites – i.e., if the SST program (high rate letter and inspection) led to a significant reduction in injury/illness rates for primary list sites.

These analyses will provide preliminary evidence for the effectiveness of the SST program. However, it should be noted that these results use employer reported injury/illness rates as their dependent variables, which are not validated by OSHA inspections and could be subject to reporting bias. The best estimates of the impact of the program would come from the evidence collected as a result of follow-up inspections to the establishments in the evaluation sample.

Analyses Conducted Once Follow-up Inspections are Completed

Once the inspections of all sites in the control and the two treatment groups are completed in FY 2012, the evaluator can use the available information to estimate the impact of the SST program on the likelihood that an employer is in violation. Using the outcomes of the FY 2012 inspections, a number of impact and other analyses can be produced.

Impact Analyses using Means Differences – Once it is verified that random assignment was successful, a means comparison of the violation rate between the control and the two treatment groups will produce unbiased estimates of the SST program’s impact on employer compliance:

- *Impact of high rate letter on employer violation* – A means comparison of the FY 2012 violation rate between treatment group A sites (high rate letter, no inspection in FY 2011) and control group sites (no high rate letter, no inspection in FY 2011) produces the impact of the high rate letter on the likelihood of violation.
- *Incremental impact of inspection on employer violation* – A means comparison of the FY 2012 violation rate between treatment group B sites (high rate letter, inspection in FY 2011) and treatment group A sites (high rate letter, no inspection in FY 2011) produces the incremental impact of the inspection over the high rate letter only on the likelihood of violation.
- *Total impact of SST program on employer violation* – A means comparison of the FY 2012 violation rate between treatment group B sites and control group sites produces the

total impact of the SST intervention of the high rate letter and inspection on the likelihood of violation.

Similar analyses can be conducted for other violation outcomes, including the likelihood of a serious violation, number of violations, number of serious violations, and financial penalties assessed.

Impact Analyses using Regression Models – To estimate the impact of the SST program on the likelihood of violation for primary list employers with increased statistical efficiency, multivariate regression models can be used to adjust for differences in site characteristics between the treatment and the control group members. Multivariate regression models also enable the evaluator to perform subgroup analyses that can identify impacts differences across employers with different characteristics. The impact analyses regression model can be expressed by the following equation:

$$Y = \alpha + X \cdot \beta + \gamma_A \cdot T_A + \gamma_B \cdot T_B + u \quad (1)$$

The dependent variable in the above model (Y) is the likelihood of violation (equals 1 if employer was found to be in violation in the follow-up inspection and 0 otherwise) following the FY 2012 inspection. Control variables include a constant term (α), and a vector (X) of available site characteristics (e.g., size, industry, union status, and region), and a zero mean disturbance term (u). The model also includes two treatment indicators:

- T_A , which equals 1 if the site was in treatment group A (high rate letter, no inspection in FY 2011) and 0 otherwise.
- T_B , which equals 1 if the site was in treatment group B (high rate letter and inspection in FY 2011) and 0 otherwise.

The first parameter of interest is γ_A , which is the regression-adjusted impact of the high-rate letter on the likelihood of violation. The second parameter of interest is γ_B , which is the regression-adjusted total impact of the SST program (high rate letter and inspection) on the

likelihood of violation. The difference between the two parameters (i.e., $\gamma_2 - \gamma_1$) is the regression-adjusted incremental impact of the inspection on the likelihood of violation. Once this model is estimated, t-tests will be produced to assess whether γ_1 (the impact of the high rate letter), $\gamma_2 - \gamma_1$ (the incremental impact of the inspection), and γ_2 (the total impact of the SST program) are statistically significant.

The above model can also be modified to estimate whether the SST program had a differential impact across employers. For example, the model can be modified to estimate the impact of the SST program for manufacturing and non-manufacturing sites. This will enable the evaluator to assess the degree to which the SST program has a different impact on manufacturing relative to non-manufacturing sites. Finally, note that all the aforementioned analyses can also be conducted for other available employer compliance outcomes, including the likelihood of a serious violation, number of violations, and amount of financial penalties assessed.

2.1.3. Summary of Evaluation Design 1

Evaluation Design 1 allows the assessment of the effectiveness of the SST program for sites on the primary inspection list, that is, those with the highest injury/illness rates. This design examines the program with the emphasis on the primary inspection list which is the focus of the SST program as currently implemented. Based on this design, OSHA will be able to:

- Estimate the likelihood of violation and other violation outcomes for primary inspection list sites.
- Identify characteristics of sites in the primary inspection list that are strong indicators of employer compliance.
- Produce a preliminary estimate of the impact of the SST program (high rate letter, inspection, high rate letter plus inspection) on the injury/illness rate of primary list sites.
- Estimate the impact of the SST program (high rate letter, inspection, and high letter plus inspection) on compliance for primary inspection list sites. Based on an evaluation sample of 1,800 sites, Evaluation Design 1 would enable us to detect if the SST program led to at least a 6.3 percentage-point impact on the violation rate of primary inspection list sites.

Evaluation Design 1 requires OSHA to make minimal changes in the FY 2011 SST program. OSHA typically inspects about 2,900 sites each year, three quarters of which are in the primary inspection list. So, assuming an evaluation sample of 1,800 primary list sites, Evaluation Design 1 will not require OSHA to increase the usual number of inspections under the SST program in FY 2011 or in FY 2012. To successfully implement Evaluation Design 1, OSHA would still need to make some changes to the regular operations of the SST program, including:

- Not sending high rate letters in FY 2011 to primary inspection list sites in the control group, despite the fact that they exceed the high rate letter threshold.
- Not conducting FY 2011 inspections for primary list sites in the control group and in treatment group A.
- Ensuring that the primary list sites in treatment group B are inspected in FY 2011.
- Inspecting all primary list sites in the evaluation sample in FY 2012, regardless of their reported injury/illness rates.

2.2. Evaluation Design 2 – Primary and Secondary Inspection List Sites

Under the SST program, once OSHA Area Offices complete the inspections of eligible primary list sites, they proceed to inspections of secondary list sites. Sites on the secondary inspection list have lower reported injury/illness than sites in the primary inspection list. In the past ten years, about 25 percent of SST program inspections targeted secondary inspection list sites.

There is no research indicating if the SST program is effective in improving compliance among employers in the secondary inspection list. For this reason, it is unknown whether the impact of the program is different for primary inspection list sites than it is for secondary list sites. Estimating the impact of the SST program for both primary and secondary list sites is important from a policy standpoint for determining the focus of the program. If the program's impact on compliance is higher for employers on the primary list than for employers on the secondary list, then the program's focus on the primary list is the appropriate approach for improving overall employer compliance. However, if the program's impact is equal or higher for employers on the secondary list, it may be valuable for OSHA to consider providing more emphasis on inspecting these employers.

Evaluation Design 1, outlined in Section 2.1, can be used to estimate the impact of the SST program only for primary list sites. In this section, we present Evaluation Design 2 which focuses on estimating the impact of the SST program for sites in the primary inspection list *and* for sites in the secondary inspection list. Using Evaluation Design 2, OSHA will be able to obtain an accurate estimate of the impact of the high rate letter, the incremental impact of the inspection, and the total impact of the SST program on primary and secondary inspection list sites. A detailed description of the design, the analyses that can be produced based on the design, and the design implementation requirements are presented below. Much of the discussion that follows is similar to the earlier discussion presented for Evaluation Design 1. The main difference is that in the discussion of Evaluation Design 2, we expand the proposed analysis to include both primary inspection list and secondary inspection list sites.

2.2.1. Implementation Steps

A visual representation of Evaluation Design 2 is presented in Exhibit 5. As shown, Evaluation Design 2 is similar to Evaluation Design 1 with the exception that Design 2 extends the design to include secondary list sites in the evaluation sample. In particular, Evaluation Design 2 involves the following steps:

Step 1 – Random Selection of Primary and Secondary List Sites

As Exhibit 5 shows, this design requires OSHA (or their evaluation contractor) to randomly select primary and secondary list sites for the evaluation. Similar to Evaluation Design 1, design 2 requires the random selection of primary list sites in FY 2011 and their assignment into three groups: 1) control group A, 2) treatment group A, and 3) treatment group B. In addition, Evaluation Design 2 requires the random selection of secondary list sites in FY 2011 that will be assigned into three groups: 1) control group B, 2) treatment group C, and 3) treatment group D. The primary and secondary list sites selected are the evaluation sample for Design 2.

Sites assigned in control groups A and B will not receive the high rate letter and will not be inspected in FY 2011. Sites in treatment group A and treatment group C will receive the high

rate letter but will not be inspected in FY 2011, while sites in treatment group B and treatment group D will receive the letter and will be inspected in FY 2011.

Step 2 – Send High Rate Letters to Treatment Groups Sites in FY 2011

OSHA will use their normal SST program operations to send high rate letters in FY 2011 to all primary inspection list sites assigned in treatment groups A and B and to all secondary inspection list sites assigned in treatment groups C and D. OSHA must not send high rate letters to primary list sites assigned in control group A and to secondary list sites assigned in control group B despite the fact that their injury/illness rates would ordinarily qualify them to receive this letter.

Step 3 – Inspect Sites in Treatment Groups B and D in FY 2011

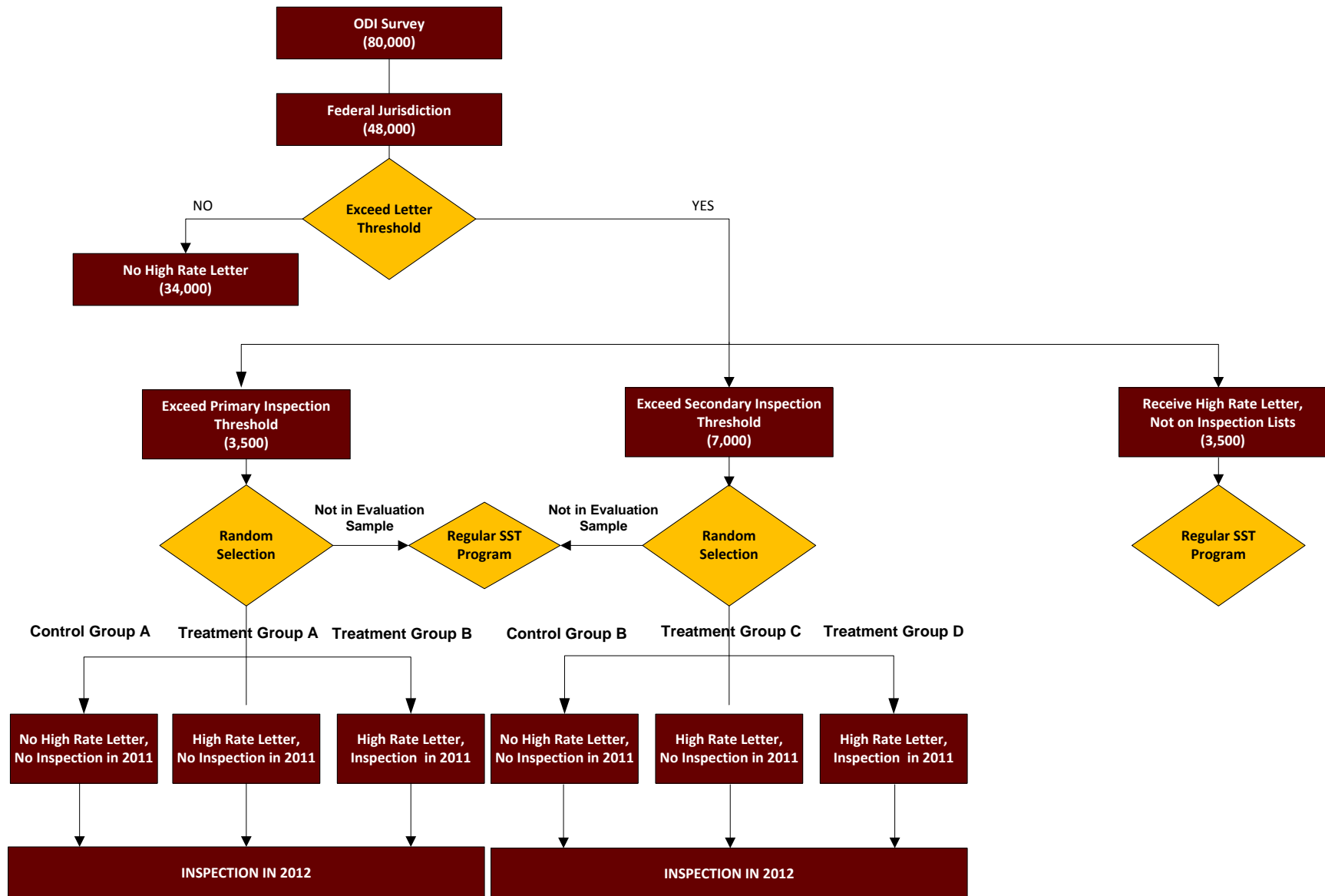
In FY 2011, OSHA must inspect all primary inspection list sites in treatment group B and all secondary inspection list sites in treatment group D. Note that OSHA will need to inspect the secondary inspection list sites in treatment group D regardless of whether the OSHA Area Offices assigned to inspect these sites have exhausted the primary inspection list. In contrast, OSHA must not inspect any of the sites in the remaining control and treatment groups.

Step 4 – Inspect All Control and Treatment Sites in FY 2012

Following the completion of the SST program in FY 2011, OSHA must inspect all sites in the evaluation sample in FY 2012.¹⁰ These follow-up inspections will be similar to the usual inspections conducted under the SST program. It is important to note that OSHA must inspect all sites in the control and treatment groups, even if the sites' injury/illness rates reported in 2012 have dropped to a level that would ordinarily preclude them from SST inspection. The objective of these inspections is to observe employer compliance for all employers in the evaluation sample following the FY 2011 interventions.

¹⁰ See Section 3.4 for a discussion of the different options for conducting the inspections.

Exhibit 5: Evaluation Design 2 – Primary and Secondary Inspection List Sites



2.2.2. Required Sample Sizes

As discussed in Section 2.1.2, one important consideration in implementing a random assignment design is to include a sufficiently high number of sites to ensure that the evaluation detects statistically significant estimates of the program's impact. In Section 2.1.2, we recommended that 1,800 primary list sites are included in Evaluation Design 1. The main difference between Evaluation Design 1 and Evaluation Design 2 is that the latter includes both primary and secondary inspection lists. Our recommendation is that if OSHA adopts Evaluation Design 2, it includes 1,800 sites in the evaluation sample: 900 primary list sites and 900 secondary list sites. This will enable OSHA to detect a minimum effect on the likelihood of violation of 8.9 percentage points for primary list and for secondary list sites.¹¹ Additionally, by combining the two samples, OSHA will be able to assess if there was at least a 6.3 percentage point impact on the likelihood of violation for both primary and secondary list sites.

2.2.3. Evaluation Analyses

Many of the analyses that can be conducted for Evaluation Design 2 are similar to the analyses for Evaluation Design 1, as described in Section 2.1.2. The key difference is that, under Evaluation Design 2, those analyses can be conducted for both primary and secondary inspection list and useful conclusions can be drawn regarding differences in the effectiveness of the SST program between the two groups. Below, we provide a summary of the types of analyses that can be conducted under this design.

Descriptive Analyses

Before any impact analyses are produced, it is useful to provide an overview of the baseline characteristics of treatment and control group sites. This overview, which will be produced separately for primary and for secondary inspection list sites, will provide descriptive statistics of the characteristics of employers in each group. Additionally, a comparison of the characteristics of primary and secondary inspection list employers would enable us to identify the degree to which there are important differences in characteristics between the two groups.

¹¹ For simplicity, we assume an average violation likelihood of 0.71 for both primary list and secondary list sites.

This information will then be used to assess whether primary inspection list employers and secondary inspection list employers were successfully randomized into treatment and control groups. Specifically, using a similar approach as the one described in Section 2.1.2, the evaluation contractor will verify: 1) there were no differences in primary inspection list employer characteristics between control group A, treatment group A, and treatment group B, and 2) there were no differences in secondary inspection list employer characteristics between control group B, treatment group C, and treatment group D. This step is necessary to ensure the validity of comparing treatment with control group outcomes to estimate SST program impacts for primary and secondary inspection list sites.

Analyses Conducted Prior to Follow-up Inspections

Using the outcomes of the FY 2011 inspections of primary inspection list sites in treatment group B and of secondary inspection list sites in treatment group D, the evaluation contractor can produce the following analyses:

- *Estimate overall violation rate and other outcomes* – Estimate the violation rate for primary inspection list sites based on FY 2011 inspections of sites in treatment group B. Similarly, one can estimate the violation rate for secondary inspection list sites based on FY 2011 inspections of sites in treatment group D. Then, one can compare the violation rate between primary and secondary inspection list sites to determine whether the gap in the reported injury/illness rates is reflected in the observed violation rate. Similar analyses can be produced for other violation outcomes, including the number of violations, serious violations, and financial penalties assessed.
- *Characteristics of violators* – Using the available site characteristics, one can estimate the relationship between site characteristics and the likelihood of violation. This analysis will be conducted separately for primary and for secondary inspection list employers and will help assess which types of employers are likely to be in violation. Additionally, the evaluation contractor will be able to assess the degree to which primary list violators share similar characteristics with secondary list violators.

In addition to these descriptive analyses, the ODI survey responses of evaluation sites in 2012 can be used to produce estimates of the impact of the SST program on employer reported injury/illness rates. Using the same approach as described in Section 2.2.1, the evaluation contractor can produce the following impact analyses:

- *Impact of high rate letter on injury/illness rates* – A means comparison of the average injury/illness rates reported in 2012 between treatment group A sites (letter, no inspection) and control group A sites (no letter, no inspection) represents the impact of the high rate letter on the injury/illness rate for primary list employers. The respective estimate for secondary list sites is produced by comparing the average injury/illness rate between treatment group C and control group B. T-tests can be used to: 1) assess if the impact of the high rate letter is statistically significant for primary and/or for secondary inspection sites, and 2) assess whether there is a difference in the impact of the high rate letter between primary and secondary list sites.
- *Incremental impact of inspection on injury/illness rates* – This impact is produced by comparing the injury/illness rates between treatment group B and treatment group A for primary list sites and by comparing the rates between treatment group D and treatment group C for secondary list sites. T-tests will be used to assess whether these impacts are statistically significant and if the impact for primary list sites differed from the impact for secondary list sites.
- *Total impact of SST program on injury/illness rates* – A comparison of the injury/illness rates between treatment group B and control group A produces the total impact of the SST program for primary inspection list sites, while a comparison of the injury/illness rates between treatment group D and control group B produces the total impact of the program for secondary inspection list sites. T-tests will be used to assess if these impacts are statistically significant and whether there is a significant difference between them.

These analyses will provide some preliminary evidence on the effectiveness of the SST program for primary and secondary inspection list sites, as well as an opportunity to assess whether the program may have a different impact across the two types of employers. As noted above, these results using employer reported injury/illness rates that have not been validated by OSHA inspection as their dependent variables would be preliminary evidence. The best estimates of the

impact of the program would come from the evidence collected as a result of follow-up inspections of the sites in the evaluation sample.

Analyses Conducted Once Follow-up Inspections are Completed

Once the FY 2012 follow-up inspections are completed, the available information will be used to estimate the impact of the SST program on employer violation outcomes for primary inspection list and for secondary inspection list sites. Below, we provide an overview of the impact analyses.

Impact Analyses using Means Differences – Assuming random assignment was implemented successfully, simple treatment-control comparisons of the mean violation rate will produce unbiased estimates of the SST program’s impact for primary and secondary list sites:

- *Impact of high rate letter on employer violation* – A means comparison of the violation rate between treatment group A (high rate letter, no inspection in FY 2011) and control group A (no high rate letter, no inspection in FY 2011) produces the impact of the high rate letter on the likelihood of violation for primary list sites. A similar comparison between treatment group C and control group B produces the impact of the high rate letter for secondary list sites.
- *Incremental impact of inspection on employer violation* – A means comparison of the violation rate between treatment group B (high rate letter, inspection in FY 2011) and treatment group A (high rate letter, no inspection in FY 2011) produces the incremental impact of the inspection on the likelihood of violation for primary inspection list sites. A means comparison of the violation rate between treatment group D and treatment group C produces the same impact estimate for secondary inspection list sites.
- *Total impact of SST program on employer violation* – A means comparison of the violation rate between treatment group B (high rate letter, inspection in FY 2011) and control group A (no high rate letter, no inspection in FY 2011) produces the incremental impact of the inspection on the likelihood of violation for primary inspection list sites. Similarly, the total impact of the SST program on secondary inspection list sites is produced by comparing the mean violation rate between treatment group D and control group B.

The same analyses can be conducted to obtain the impact of the SST program on other violation outcomes, including the likelihood of a serious violation, number of violations, number of serious violations, and financial penalties assessed.

Impact Analyses using Regression Models – Using regression models, the evaluation contractor can estimate the impact of the SST program for primary and for secondary inspection list sites, while controlling for differences in site characteristics. Based on these analyses, one will be able to assess if the impact of the SST program differs between primary and secondary inspection list sites. Specifically, the following regression model will be estimated for primary list sites:

Primary List:
$$Y^{PL} = \alpha + X \cdot \beta^{PL} + \gamma_A^{PL} \cdot T_A + \gamma_B^{PL} \cdot T_B + u \quad (2)$$

The dependent variable in model 2 (Y^{PL}) is the likelihood of violation (equals 1 if employer was found to be in violation in the subsequent inspection and 0 otherwise); the control variables include a constant term (α), a vector (X) of available site characteristics (e.g., size, industry, union status, and region), and a zero mean disturbance term (u). This model also includes two treatment indicators:

- T_A , which equals 1 if the site was in treatment group A (high rate letter, no inspection in FY 2011) and 0 otherwise.
- T_B , which equals 1 if the site was in treatment group B (high rate letter and inspection in FY 2011) and 0 otherwise.

The parameters of interest from model 2 are: 1) γ_A^{PL} , the regression-adjusted impact of the high rate letter on the likelihood of violation for primary list sites, 2) γ_B^{PL} , the regression-adjusted total impact of the SST program on the likelihood of violation for primary list sites, and 3) $\gamma_B^{PL} - \gamma_A^{PL}$, the regression-adjusted incremental impact of the inspection on the likelihood of violation for primary list sites.

Additionally, a similar regression model can be estimated for secondary list sites:

Secondary List:
$$Y^{SL} = \alpha + X \cdot \beta^{SL} + \gamma_A^{SL} \cdot T_A + \gamma_B^{SL} \cdot T_B + u \quad (3)$$

The parameters of interest from this model are: 1) γ_A^{SL} , the regression-adjusted impact of the high rate letter on the likelihood of violation for secondary list sites, 2) γ_B^{SL} , the regression-adjusted total impact of the SST program on the likelihood of violation for secondary list sites, and 3) $\gamma_B^{SL} - \gamma_A^{SL}$, the regression-adjusted incremental impact of the inspection on the likelihood of violation for secondary list sites.

Once these two models are estimated, t-tests will be produced to assess whether the parameters of interest are statistically significant. Additionally, the evaluation contractor will be able to test whether the program impacts are different between primary and secondary inspection list sites. For example, to test whether the SST program had a higher total impact on primary inspection list sites than on secondary inspection list sites, the evaluation contractor will compare γ_B^{PL} with γ_B^{SL} . If γ_B^{PL} is significantly higher than γ_B^{SL} , then one will conclude that the SST program is more effective for primary list sites than for secondary list sites.

Using the two models presented above, the evaluation contractor can also estimate whether the SST program had differential impacts across available employer characteristics (e.g., size, industry). Finally, all the above analyses can be conducted for other employer compliance outcomes, including the likelihood of a serious violation, number of violations, and amount of financial penalties assessed.

To estimate the impact of the SST program with higher efficiency, the evaluation contractor can also estimate a model that uses both the primary and secondary list sites included in the evaluation sample. This model will take the following format:

$$Y = \alpha + X \cdot \beta + \gamma_A \cdot T_A + \gamma_B \cdot T_B + \gamma_C \cdot T_A \cdot D^{SL} + \gamma_D \cdot T_B \cdot D^{SL} + u \quad (4)$$

The dependent variable in model 4 is the likelihood of violation (equals 1 if employer was found to be in violation in the subsequent inspection and 0 otherwise); control variables include a constant term (α), a vector (X) of available site characteristics (e.g., size, industry, union status,

and region), and a zero mean disturbance term (u). This model also includes two treatment indicators and two treatment interactions:

- T_A , which equals 1 if the site was in treatment group A or treatment group C (i.e., high rate letter and no inspection in FY 2011) and 0 otherwise.
- T_B , which equals 1 if the site was in treatment group B or treatment group D (i.e., high rate letter and inspection in FY 2011) and 0 otherwise.
- $T_A \cdot D^{SL}$, which equals 1 if the site was a secondary list site in treatment group C (i.e., high rate letter and no inspection in FY 2011) and 0 otherwise.
- $T_B \cdot D^{SL}$, which equals 1 if the site was a secondary list site in treatment group D (i.e., high rate letter and inspection in FY 2011) and 0 otherwise.

The parameters of interest from model 4 are: 1) γ_A , the impact of the high rate letter on the likelihood of violation for primary list sites, 2) γ_B , the total impact of the SST program (letter and inspection) on the likelihood of violation for primary list sites, 3) $\gamma_A + \gamma_C$, the impact of the high rate letter on the likelihood of violation for secondary list sites, and 4) $\gamma_B + \gamma_D$, the total impact of the SST program (letter and inspection) on the likelihood of violation for secondary list sites. T-tests will be produced to assess if these impact estimates are statistically significant and whether the impacts for primary list sites differ from the impacts for secondary list sites.

The advantage of estimating program impacts using the whole evaluation sample is that one can detect program impacts with higher efficiency than estimating program impacts separately for primary and secondary list sites. For instance, using the whole sample one can detect at least a 6.3 percentage-point impact for both primary and secondary list sites. Separating the analyses for primary and secondary list sites one can detect at least an 8.9 percentage-point impact. It should be noted, however, that using the whole evaluation sample assumes that the relationship between employer characteristics and violation outcomes is identical between primary and secondary inspection lists. Although this assumption may not affect the impact estimates, it would affect the estimated relationship between employer characteristics and the violation outcome.

Our recommendation is that OSHA uses both approaches (i.e., separate analyses and combined analyses) in estimating program impacts. These would enable OSHA to identify program impacts with higher precision and test the sensitivity of the impact estimates to different estimation approaches.

2.2.3. Summary of Evaluation Design 2

Evaluation Design 2 allows the assessment of the effectiveness of the SST program for primary inspection list sites and for secondary inspection list sites. Based on this design, OSHA will be able to:

- Estimate the likelihood of violation and other violation outcomes for primary inspection list and secondary inspection list sites.
- Compare violation outcomes between primary and secondary inspection list sites.
- Identify characteristics of employers in the primary and secondary inspection list that are strong indicators of employer compliance.
- Produce a preliminary estimate of the impact of the SST program (high rate letter, inspection, and high letter plus inspection) on the injury/illness rate of primary and secondary inspection list employers.
- Produce a reliable estimate of the impact of the SST program (high rate letter, inspection, and high letter plus inspection) on compliance for primary and secondary inspection list sites. Evaluation design 2 would enable us to detect if the SST program led to at least a 6.3 percentage-point impact on the violation rate of primary and secondary inspection list sites.
- Assess whether there is a difference in the SST program impact between primary and secondary inspection list sites. If differential impacts are detected, additional analyses can be performed to test if these differentials are due to

Evaluation Design 2 requires OSHA to make some changes in the regular SST program. OSHA typically inspects about 2,900 sites each year, of which about 2,150 are from the primary inspection list and about 750 sites are from the secondary inspection list. So, assuming an evaluation sample that includes 900 primary list and 900 secondary list sites, Evaluation Design 2 will not require OSHA to increase the usual number of inspections under the SST program in

FY 2011. However, Evaluation Design 2 would require OSHA to inspect a slightly higher number of secondary list sites in FY 2012 (900 sites) than the current annual average number of sites inspected by the regular SST program (750 sites). Evaluation Design 2 also requires OSHA to make the following changes to the regular operations of the SST program:

- Not sending high rate letters in FY 2011 to primary list sites in control group A and to secondary list sites in control group B, despite the fact that they exceed the high rate letter threshold.
- Not conducting FY 2011 inspections for primary list sites in treatment group A and for secondary list sites in treatment group C.
- Ensuring that primary list sites in treatment group B and secondary list sites in treatment group D are inspected in FY 2011, regardless of whether Area Offices complete their primary inspection lists.
- Inspecting all sites in the evaluation sample in FY 2012, regardless of their reported injury/illness rates.

2.3. Evaluation Design 3 – All High Rate Letter Sites

The SST program focuses on inspecting employers with the highest reported injury/illness rates. For this reason, OSHA Area Offices use their resources to inspect primary inspection list sites in their geographic area. Once an OSHA Area Office inspects all primary inspection list sites, it proceeds to inspect sites in the secondary inspection list, resources permitting. The SST program does not focus on sites that have injury/illness rates that exceed the high rate letter threshold but are not high enough to exceed the secondary inspection list threshold. Area Offices that exhaust their secondary inspection list can request an additional, tertiary list if resources permit additional inspections, but in practice this is rare. As a result, very few sites that are not on the two inspection lists are inspected each year under the SST program.

It is possible, however, that the impact of the SST interventions on compliance is as high or higher for sites that exceed the high rate letter threshold but not the secondary inspection list threshold. If this is the case, then there may be value to OSHA's expanding the SST program to inspect some of these sites. This section presents Evaluation Design 3, a random assignment evaluation design which expands the prior designs to estimate the SST program's impact for all

sites exceeding the high rate letter threshold: 1) primary inspection list sites, 2) secondary inspection list sites, and 3) sites that receive the high rate letter but are not on the two inspection lists.

A detailed description of the design, the analyses that can be produced based on the design, and the design implementation requirements is presented below. Again, much of the discussion that follows is similar to the earlier discussion presented for Evaluation Design 1 and Evaluation Design 2. The main difference between the following discussion and the earlier discussions is that, below, we discuss expanding the proposed analysis to high rate letter list sites.

2.3.1. Implementation Steps

Evaluation design 3 is presented in Exhibit 6. This design is similar to Evaluation Design 2 in that it involves the random selection of primary and secondary inspection list sites for the evaluation sample. In addition, Evaluation Design 3 includes sites that exceed the high rate letter threshold but are not on the two inspection lists. Evaluation design 3 involves the following steps:

Step 1 – Random Selection of High Rate Letter Sites

As Exhibit 6 shows, this design requires OSHA (or their evaluation contractor) to randomly select sites that exceed the high rate letter threshold for the evaluation. These randomly selected sites will include:

- Primary inspection list sites, randomly assigned to control group A, treatment group A, and treatment group B.
- Secondary inspection list sites, randomly assigned to control group B, treatment group C, and treatment group D.
- Sites with injury/illness rates that exceed the high rate letter threshold but are not high enough to qualify them for the primary or secondary inspection lists, randomly assigned to control group C, treatment group E, and treatment group F.

These sites are the evaluation sample for Evaluation Design 3. Sites assigned in the three control groups will not receive the high rate letter and will not be inspected in FY 2011. Sites in

treatment groups A, C, and E will receive the letter but will not be inspected in FY 2011, while sites in treatment groups B, D, and F will receive both the letter and an inspection in FY 2011.

Step 2 – Send High Rate Letters to Treatment Group Sites in FY 2011

OSHA will send high rate letters in FY 2011 to all sites assigned in the six treatment groups. OSHA must not send high rate letters to sites in control groups A, B, and C, regardless of the fact that their injury/illness rates qualify them to receive the letter.

Step 3 – Inspect Sites in Treatment Groups B, D, and F in FY 2011

In FY 2011, OSHA must inspect: 1) primary list sites in treatment group B, 2) secondary list sites in treatment group D, and 3) high rate letter only sites in treatment group F. Note that OSHA will need to inspect secondary inspection list sites in treatment group D and the high rate letter only sites in treatment group F, regardless of whether the OSHA Area Offices assigned to inspect these sites have exhausted the primary inspection list. However, OSHA must not inspect any of the sites in the remaining control and treatment groups.

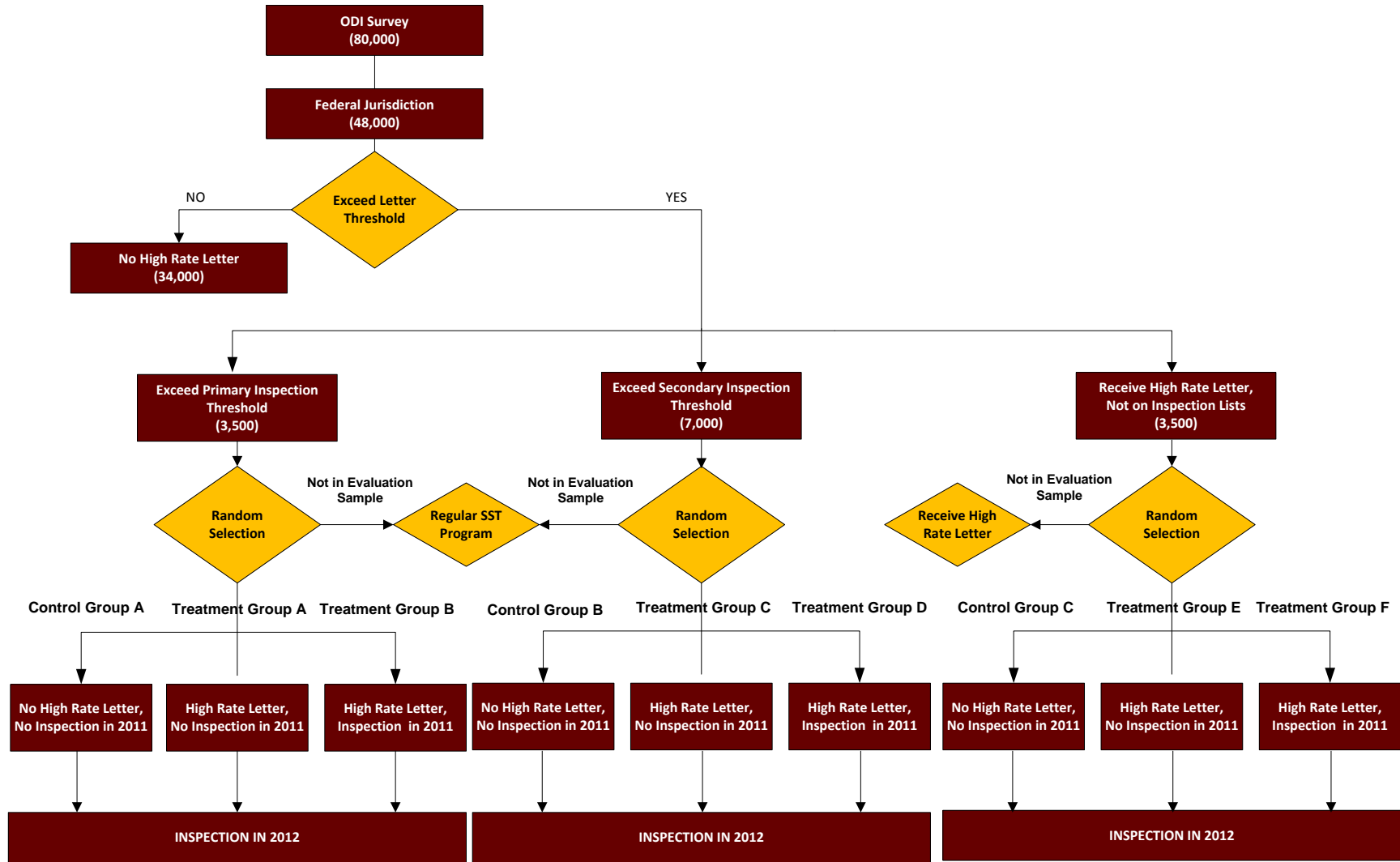
Step 4 – Inspect All Control and Treatment Sites in FY 2012

OSHA must inspect all sites in the evaluation sample in FY 2012.¹² In the follow-up inspections, OSHA will gather information on employer compliance.

This design requires OSHA to adopt one important change in the FY 2011 SST program, that is, inspect the high rate letter only sites in treatment group F. Furthermore, OSHA will need to conduct follow-up inspections to a number of high rate letter only sites, regardless of whether these sites are on the inspection lists in FY 2012. These are important changes as such sites are typically not targeted by the SST program. Otherwise, Evaluation Design 3 has similar implementation requirements as Evaluation Design 2.

¹² Section 3 provides different options for conducting the subsequent inspections.

Exhibit 6: Evaluation Design 3 – All High Rate Letter Sites



2.3.2. Required Sample Sizes

Following our discussion of required sample sizes in Section 2.1.2 (Evaluation Design 1) and in Section 2.2.2 (Evaluation Design 2), we recommend that OSHA includes 600 primary list, 600 secondary list, and 600 high rate letter only sites in Evaluation Design 3 for a total evaluation sample of 1,800 sites. This will enable OSHA to detect a 10.9 percentage point minimum effect on the likelihood of violation for primary list, secondary list, and high rate letter only sites.¹³ By combining the three samples, OSHA will be able to assess if there was at least a 6.3 percentage-point impact on the likelihood of violation for primary, secondary, and high rate letter only sites. This approach can also be used to assess differences in SST impact across the three groups.

2.3.3. Evaluation Analyses

The descriptive and impact analyses that can be conducted under Evaluation Design 3 are similar to those for the first two evaluation designs. The key difference is that, under Evaluation Design 3, those analyses can be conducted for all sites exceeding the high rate letter threshold, which will help draw conclusions regarding the effectiveness of the SST program for primary list, secondary list, and high rate letter only sites. Below, we summarize the analyses that can be conducted under this design.

Descriptive Analyses

Using the available information, the evaluation contractor will produce an overview of the characteristics of the evaluation sites – this overview will be produced separately for the three groups of employers (primary inspection list, secondary inspection list, and high rate letter only). This overview will enable us to compare site characteristics between primary inspection list, secondary inspection list, and high rate letter only sites. This information can then be used to assess whether randomization of sites to treatment and control groups was done effectively and ensure the validity of the evaluation design.

¹³ For simplicity, we assume an average likelihood of violation of 0.71 for primary list, secondary list, and high rate letter only sites.

Analyses Conducted Prior to Follow-up Inspections

Using the outcomes of FY 2011 inspections of primary inspection list sites in treatment group B, secondary inspection list sites in treatment group D, and high rate letter only sites in treatment group F, the evaluation contractor can produce the following analyses:

- *Estimate overall violation rate and other outcomes* – Estimate the violation rate and other violation outcomes for primary list, secondary list, and high rate letter only sites. Then, compare these outcomes across the three groups to assess whether group differences in the reported injury/illness rates are reflected in the violation outcomes.
- *Characteristics of violators* – Assess the relationship between site characteristics and violation outcomes. This analysis will be conducted separately for each group, enabling us to compare how violators in each group compare to violators in the remaining groups.

The responses of evaluation sample employers to the ODI survey in 2012 can be used to produce estimates of the impact of the SST program on reported injury/illness rates. Using the same approach as described in Sections 2.1.1 and 2.2.1, one can produce the following impact estimates for primary list, secondary list, and high rate letter only sites:

- *Impact of high rate letter on injury/illness rates.*
- *Incremental impact of inspection on injury/illness rates.*
- *Total impact of SST program on injury/illness rates.*

Using statistical tests, the evaluation contractor will assess if there were significant differences in the impact of the SST program across the three groups. These analyses will provide some preliminary evidence on the effectiveness of the SST program for all sites exceeding the high rate letter threshold and an opportunity to assess if the program has a different impact across the three types of sites.

Analyses Conducted Once Follow-up Inspections are Completed

The outcomes of the follow-up inspections can be used to estimate the impact of the SST program on employer compliance for all sites exceeding the high rate letter threshold. Below, we provide an overview of the impact analyses.

Impact Analyses using Means Differences – Treatment-control comparisons of the mean violation rate will produce consistent estimates of the SST program’s impact for primary list sites, secondary list sites, and high rate letter only sites. Using a similar approach as the one described in Sections 2.1.3 (Evaluation Design 1) and 2.2.3 (Evaluation Design 2), one can estimate the following impacts:

- *Impact of high rate letter on employer violation.*
- *Incremental impact of inspection on employer violation.*
- *Total impact of SST program on employer violation.*

These impacts will be produced separately for primary list, secondary list, and high rate letter only sites. Similar analyses can be produced for other violation outcomes such as serious violations or penalties.

Impact Analyses using Regression Models – Regression models can be used to estimate the impact of the SST program for employers exceeding the high rate letter threshold, while controlling for available site characteristics. For these analyses, the evaluation contractor will estimate a separate regression model for primary list, secondary list, and high rate letter only sites. Following a similar empirical approach as the one described for Evaluation Designs 1 and 2, the following estimates will be produced:

- *Regression-adjusted impact of high rate letter on employer violation.*
- *Regression-adjusted incremental impact of inspection on employer violation.*
- *Regression-adjusted total impact of SST program on employer violation.*

Once these impacts are estimated, use t-tests will be used to assess their statistical significance and to examine if the program’s impacts differed between primary list, secondary list, and high rate only sites. These analyses will help evaluate the impact of the SST program for all sites exceeding the high rate letter threshold and determine which sites are most likely to be affected by the program.

To improve the efficiency of the impact analyses, the evaluation contractor can also estimate a model that uses all sites in the evaluation sample. This model, which will have a similar format

as model 4, will include two treatment indicators (received the high rate letter only; received the high rate letter and an inspection) and interactions between the treatment indicators and indicators of site category (secondary inspection list; high rate letter only).

The advantage of using the whole evaluation sample to produce impact estimates is that one can detect program impacts with higher efficiency. Moreover, this approach would enable us to assess if the program's impact differed across primary list, secondary list, and high rate letter only sites. Our recommendation is that OSHA uses both approaches (i.e., separate analyses and combined analyses) in estimating program impacts. These would enable to identify program impacts with higher precision and test the sensitivity of the impact estimates to different estimation approaches.

Finally, the above models can be modified to estimate if the impact of the SST program differed across sites based on their observed characteristics. These analyses will enable us to assess if the SST program is more effective for sites with certain observable characteristics. Moreover, the above analyses will be produced to determine the SST program's impact on all available compliance outcomes.

2.3.3. Summary of Evaluation Design 3

Evaluation design 3 allows the assessment of the effectiveness of the SST program for all sites exceeding the high rate letter threshold: primary inspection list, secondary inspection list, and high rate only sites. Based on this design, OSHA will be able to:

- Estimate the likelihood of violation and other violation outcomes for sites exceeding the high rate letter threshold.
- Compare violation outcomes between primary list, secondary list, and high rate letter only sites.
- Identify site characteristics strongly related to compliance for sites exceeding the high rate letter threshold.
- Produce a preliminary estimate of the impact of the SST program (high rate letter, inspection, and high rate letter plus inspection) on the injury/illness rate of primary list, secondary list, and high rate letter only sites.

- Produce a reliable estimate of the impact of the SST program (high rate letter, inspection, and high letter plus inspection) on compliance for primary list, secondary list, and high rate letter only sites. Evaluation design 3 would enable us to detect if the SST program led to at least a 6.3 percentage-point reduction in the violation rate of primary, secondary, and high rate letter only sites.
- Assess whether there is a difference in the SST program impact between primary list, secondary list, and high rate letter only sites.

To successfully implement Evaluation Design 3, OSHA needs to:

- Avoid sending high rate letters in FY 2011 to 200 primary list sites in control group A, to 200 secondary list sites in control group B, and to 200 high rate letter only sites in control group C, despite the fact that they exceed the high rate letter threshold.
- Avoid FY 2011 inspections of primary list sites in control group A and treatment group A and of secondary list sites in control group B and treatment group C.
- Ensure the 200 primary list sites in treatment group B, the 200 secondary list sites in treatment group D, and the 200 high rate letter only sites in treatment group F are inspected in FY 2011.
- Inspect all 1,800 sites in the evaluation sample following FY 2011, regardless of their reported injury/illness rates.

2.4. Comparison of Evaluation Designs and Implications for the SST Program

Each evaluation design option described above has its own advantages and disadvantages in terms of the types of impact analyses that can be conducted and the implications for SST program operations. This section summarizes the three evaluation designs, providing OSHA with important insight on the value and the implications of implementing each design.

Evaluation Design 1 enables OSHA to assess the impact of the SST program for sites on the primary inspection list, which are the main target of the program. Using this design, the evaluation contractor can estimate the impact of the high rate letter, the incremental impact of the inspection, and the total impact of the SST program for employers that exceed the primary list

injury/illness rate threshold. Our recommendation is that OSHA includes 1,800 primary list sites in the evaluation. As shown in Exhibit 7, this would require the inspection of 600 primary list sites in FY 2011 and 1,800 primary list sites in FY 2012. These figures are well below the program’s regular inspection volume for primary inspection lists – in the past decade, about 2,150 primary list sites are inspected on an annual basis. So, Evaluation Design 1 would not create additional inspection burden on the SST program. The most important implication of Evaluation Design 1 is that it requires that 600 randomly selected primary list sites be inspected in two consecutive years (FY 2011 and FY 2012). Currently, sites are typically not inspected within two or three years of a previous inspection.

Exhibit 7: Summary of Proposed Evaluation Designs

	Primary List	Secondary List	High Rate Letter Only
Usual Number of SST Inspections per Year	2,150	750	0
Evaluation Design 1			
Inspections in FY 2011	600	--	--
Inspections in FY 2012	1,800	--	--
Evaluation Design 2			
Inspections in FY 2011	300	300	--
Inspections in FY 2012	900	900	--
Evaluation Design 3			
Inspections in FY 2011	200	200	200
Inspections in FY 2012	600	600	600

The main disadvantage of Evaluation Design 1 is that it does not enable OSHA to answer the question of whether a differently implemented SST program works better than the current program (i.e., to estimate the impact of expanding the SST program to inspect a higher number of secondary inspection list sites). Evaluation Design 2 is designed to estimate the impact of the SST program for primary list and for secondary list sites, enabling OSHA to assess whether the program had a differential impact for these employers with differing injury/illness rates. Based on an evaluation sample of 1,800 sites (900 primary list and 900 secondary list sites), Evaluation

Design 2 requires the inspection of 600 sites in FY 2011 (300 primary and 300 secondary list sites) and 1,800 sites in FY 2012 (900 primary and 900 secondary list sites).

The implementation of Design 2 requires that 600 (300 primary list and 300 secondary list) sites receive an inspection in two consecutive years (FY 2011 and FY 2012). Another implication is that, although the total number of inspections is within the SST program's current inspection volume, Evaluation Design 2 requires 900 secondary list sites be inspected in FY 2012. This number is slightly higher than the current annual average of secondary list sites inspected by the regular SST program (on average, approximately 750 secondary list sites are inspected annually). Based on the current level of SST inspections, Evaluation Design 2 will require that, at least in some OSHA areas, secondary list sites may need to be inspected even if the primary list is not exhausted.

Finally, Evaluation Design 3 expands the inspections to all sites that exceed the high rate letter threshold. Evaluation design 3 would enable OSHA to assess the impact of the SST program on primary list, secondary list, and high rate letter only sites. This would allow OSHA to identify which employers are likely to be affected the most by the SST program and assess if there is value in expanding the program to include the high rate letter only sites.

The total number of inspections under Evaluation Design 3 is well below the usual number of inspections conducted by the SST program on an annual basis. However, the design involves the inspection of 200 high rate letter only sites in FY 2011 and of 600 high rate letter only sites in FY 2012. High rate letter only sites are not typically inspected by the SST program – thus this design would reduce the usual number of inspections of primary list sites to accommodate the inspection of high rate letter only sites.

There are advantages and disadvantages in considering which evaluation design option to adopt. If OSHA's focus is to assess the program's impact on sites with the highest injury/illness rates with minimum disruption to the regular SST program, Evaluation Design 1 is the clear option. But, if OSHA is interested in assessing the value of expanding the number of secondary list sites inspected by the SST program, Evaluation Design 2 is a better option. Finally, Evaluation

Design 3 is obviously the most exhaustive design allowing OSHA to assess the implications of expanding the program to high rate letter only sites.

3. EVALUATION DESIGN IMPLEMENTATION

The evaluation design options presented above require the random selection of SST-eligible sites and their assignment into control and treatment groups. Depending on their assignment, sites will either: 1) receive the high rate letter and an inspection in FY 2011, 2) receive only the high rate letter in FY 2011 (no inspection), or 3) receive neither the letter nor the inspection. In FY 2012, all evaluation sites will be inspected.

One of the critical issues for successfully implementing the selected evaluation design is that a sufficiently high number of the evaluation sites are available for the follow-up inspections. If an adequate number of sites receive a follow-up inspection, the evaluation contractor will have sufficient statistical power to detect the impact of the SST program. In considering how to ensure an adequate number of evaluation sites receive follow-up inspections, one needs to take into account the fact that some sites may drop out of the sample following random selection. Sample attrition may take two forms: 1) a proportion of the randomly selected sites may be out of business at the time of random selection and 2) a proportion of the remaining sites may go out of business before the follow-up inspections. OSHA and the evaluation contractor must take into account both types of attrition when conducting the random assignment.

Regardless of which design OSHA chooses to implement, the total number of inspections conducted will be below the usual number of inspections conducted in a given year under the SST program. However, if Evaluation Design 3 is selected, OSHA would need to conduct 600 inspections of high rate letter only sites in FY 2012 (i.e., one year following random assignment). This indicates that nearly 30 percent of all inspections conducted in that year under the SST program would be for high rate letter only sites, which are not typically inspected by the program.

In this section, we provide an overview of the required process for successfully implementing random assignment of sites into control or treatment groups. Then, we discuss how sample attrition may affect the successful implementation of the selected evaluation design. To overcome this issue, we propose a random assignment strategy which leads to sufficiently high

sample sizes for producing statistically significant estimates of the impact of the SST program. We also present three options for implementing Evaluation Design 3, which roll out the evaluation over multiple years, providing OSHA with flexibility regarding the number of high rate letter only site inspections conducted in a given year.

3.1. Random Assignment

Regardless of the design option selected, it is critical to the validity of the evaluation to ensure establishments are randomly assigned into the various evaluation design groups (i.e., treatment and control groups) that determine which sites receive which interventions.¹⁴ Random assignment of sites into the control and treatment groups ensures there are no differences between the three groups in terms of site characteristics and their compliance behavior. As a result, observed differences in violation outcomes between the three groups during the follow-up inspections will be attributed to the SST program.

Random assignment will be conducted once OSHA determines, based on the reported injury/illness rates, which sites are to be included in the primary inspection list, the secondary inspection list, or the high rate letter only list. Once these lists are produced, a rigorous method must be used to randomly select sites for the evaluation and assign them into the control or treatment groups. We recommend that the evaluation contractor should perform the random assignment, both to minimize any burden on OSHA and to ensure selection is performed randomly and uniformly throughout the evaluation.

3.2. Sample Attrition

The successful implementation of the selected evaluation design requires that the random assignment process takes into account sample attrition. Once the evaluation sites are selected, there is a chance that many of them would not be in operation or that they would be ineligible for the SST program. Furthermore, among the remaining sites, some are likely to go out of business by the time the follow-up inspections are conducted. Random assignment must take into account

¹⁴ The term “random assignment” refers to the random selection of eligible sites in the control and treatment groups for the evaluation. This should not be confused with the random selection process built into the SST website programming to select sites to be included in each inspection cycle.

sample attrition to ensure that a sufficiently high number of selected sites are available for the follow-up inspections. This will enable the evaluation contractor to estimate the impacts of the SST program and have the required statistical power to determine if the impacts are statistically significant.

Below, we discuss the two types of sample attrition in more detail and present a random assignment strategy which will help overcome attrition and ensure sufficiently high sample sizes are achieved.

3.1.1. Initial Attrition

Initial attrition occurs once sites are randomly selected and assigned into the control and treatment groups. This type of attrition can occur for a number of reasons, including but not limited to:

- *Sites received recent inspection* – Sites that had previously received a comprehensive inspection within the time frame specified by the SST directive would be ineligible for the SST program in FY 2011.
- *Sites out of business or that changed ownership* – Some sites may have gone out of business or may have changed ownership, making them ineligible for the SST program.
- *Office-only or government sites* – Sites that are office-only locations or government facilities are not included in the SST program, so they cannot be included in the evaluation.
- *Sites participating in other OSHA programs* – Some sites participate in other OSHA programs, including the Voluntary Protection Program and the Consultation Program. Sites in the programs specified by the directive are exempt from the SST program for some period of time, so they cannot be included in the evaluation.

Interviews with Area Office personnel that implement the SST program suggest that about 30 percent of sites on the primary or secondary inspection lists may be ineligible for inspection for one or more of these reasons. The most common reason for SST ineligibility is that the site had a recent prior inspection.

3.1.2. Attrition Over Time

Besides initial attrition, a number of sites may drop out of the sample between random selection and the follow-up inspections in subsequent years. Attrition over time generally occurs because establishments go out of business during the course of the evaluation period. Area Office personnel that implement the SST program indicated that this type of attrition could affect approximately 10 percent of the sample in a given year.

The combined effect of initial attrition and of attrition over time may significantly reduce the evaluation sample, compromising the effectiveness of the design to produce statistically significant impact estimates. To illustrate this point, consider Evaluation Design 1, which requires 600 primary list sites be randomly assigned into each of the control group, treatment group A, and treatment group B. If the evaluator draws exactly 600 sites for each group, about 180 sites (or 30 percent) will be dropped from the sample due to initial attrition (e.g., ineligible for SST program, out of business, etc.) – this will leave about 420 sites in each group. Furthermore, of the remaining 420 sites, an additional 42 sites (or 10 percent) will be dropped from the sample by the time the follow-up inspections are conducted due to closings and other circumstances. As a result, by the time follow-up inspections are conducted, there will be 460 sites in each group for a total of 1,380 sites instead of the original 1,800. This level of attrition may severely compromise the statistical power of the design and make it impossible to produce statistically significant impact estimates.

3.3. Process for Minimizing the Effects of Attrition

To ensure there is an adequate sample by the end of the study period, the evaluation contractor should implement a random assignment strategy that accounts for sample attrition. We propose the evaluation contractor implement a strategy that includes random selection with oversampling and sample screening to maximize the number of eligible sites in the initial sample. Below, we present the steps of our proposed approach.

Step 1 – Randomly select a sufficient number of sites to offset expected attrition

Random assignment will be conducted once OSHA determines which sites are included in the primary inspection list, the secondary inspection list, and the high rate letter only list. Using an

appropriate random assignment process, the evaluation contractor will use these lists to randomly select sites into the control or treatment groups.

In randomly selecting sites for the evaluation, the contractor should calculate the number of sites that must be selected to achieve the required sample size after attrition. For example, in Evaluation Design 1, the desired final sample size in each group is 600 sites, for a total of 1,800 sites. To account for attrition, the contractor should randomly select 954 sites into each group (i.e., control group, treatment group A, and treatment group B). Assuming 30 percent are lost to initial attrition, there would still be 668 sites in each group, for a total of 2,004 sites. These sites, depending on the group in which they are assigned, they will receive the specified interventions from the SST program in FY 2011. Next, assuming 10 percent of the 668 sites in each group are lost to attrition over time, there will still be 600 sites remaining in each group (for a total of 1,800 sites) for the follow-up inspections in FY 2012. A similar random assignment approach should be adopted if OSHA decides to implement Evaluation Design 2 or Evaluation Design 3.

Step 2 – Identify ineligible sites prior to the FY 2011 inspections

Having selected a sufficient number of sites to account for attrition, the evaluation contractor should work with OSHA to help identify the sites that are ineligible for inspection, went out of business, or changed ownership to be able to screen these sites out of the sample. Currently, this screening takes place in each Area Office. However, having the evaluation contractor identify these sites will remove this burden from the Area Offices and help ensure this initial attrition can be tracked and monitored. Moreover, this will ensure that eligibility criteria are applied consistently across all Area Offices to avoid sample bias.

To identify the sites in the evaluation sample that were recently inspected and thus ineligible for SST in 2011, OSHA would need to provide the evaluation contractor recent inspection records. The evaluation contractor will then need to match these to the evaluation sample to identify sites in the evaluation sample that received an inspection recently enough to disqualify them from SST. These sites would be removed from all treatment and control groups. The evaluation contractor can further improve the sample and minimize subsequent attrition by performing screening on the businesses in the evaluation sample to identify any that have gone out of

business or changed ownership since the ODI was conducted. This screening can be performed via web-searches and telephone calls to employers to validate that the establishments in the evaluation sample are still in operation. Those that are not in business or had a recent ownership change can be removed from the sample.

Step 3 – Additional random selection and screening to obtain necessary samples

To ensure an adequate sample, the evaluation contractor will examine the sample remaining after removing ineligible sites and those that have gone out of business or changed ownership. If the rate of initial attrition found through the screening actually exceeded the 30 percent estimate, then there will not be enough sites in the sample to off-set attrition over time. In this case, the evaluation contractor can select additional sites and repeat the random assignment and screening process until a sufficient sample size has been reached. Returning to the example in step 1 above, if the initial screening process leaves fewer than 668 sites in each condition, the contractor can randomly assign and screen additional sites until each group has 668 sites, enough to offset the expected attrition over time and still leave 600 sites in each group for the follow-up inspections.

Using this three-step process for random selection, the evaluation contractor will account for the effects of initial attrition on the evaluation sample for each group, and ensure there is a sufficient sample available to adjust for expected attrition over time. By accounting for this attrition, the contractor will help OSHA to implement the evaluation in a manner that will provide a sufficient number of sites in the final sample after the follow-up inspections to allow for sufficient statistical power to detect program impacts.

3.4. Timing of Inspections

The processes for random assignment and accounting for attrition described above are critical issues for the validity and statistical power of the evaluation. Another important consideration is to ensure that the implementation of the selected Evaluation Design does not impose a burden on the regular SST program. All three designs require 600 inspections be conducted in FY 2011 and 1,800 inspections be conducted in FY 2012; these figures are well below the 2,900 inspections conducted by the SST program in a typical year. However, Evaluation Design 3

requires OSHA to inspect 200 high rate letter only sites in FY 2011 and 600 high rate only sites in FY 2012. As high rate letter only sites are not typically inspected by the SST program, inspecting 600 high rate only sites in FY 2012 would require OSHA to inspect 600 (or 28%) fewer primary inspection list sites in that year.

Below, we discuss three options for conducting the initial and follow-up inspections for Evaluation Design 3 and examine how these affect the total number of inspections of high rate letter only sites:

- ***Option 1 – Initial Inspections in FY 2011, Follow-Up Inspections in FY 2012***

This option requires all the initial inspections be conducted in FY 2011 and all follow-up inspections be conducted in FY 2012. As shown in Exhibit 8, if this option is used to implement Evaluation Design 3, 223 high rate letter only sites will be inspected in FY 2011 and 600 high rate letter only sites will be inspected in FY 2012. This option is perhaps the simplest way to implement Evaluation Design 3, while minimizing the effects of attrition and other external conditions on the evaluation.

- ***Option 2 – Initial Inspections in FY 2011, Follow-Up Inspections in FY 2012-2013***

One way of reducing the number of inspections of high rate letter only sites in a given year is to spread the follow-up inspections over two years. As shown in Exhibit 8, using this option to implement Evaluation Design 3 requires 223 initial inspections of high rate letter only sites in FY 2011 and 300 follow-up inspections of high rate letter only sites in each FY 2012 and FY 2013.

- ***Option 3 – Initial Inspections in 2011- 2012, Follow-Up Inspections in 2012-2014***

This option reduces the number of inspections of high rate letter only sites in a given year by spreading both the initial and the follow-up inspections over a two-year period. To do so, option 3 splits the evaluation sites into two samples, with the evaluation of one sample starting in FY 2011 and the evaluation of the second sample starting in FY 2012. Specifically, option 3 requires that half the initial inspections are conducted in FY 2011 and the remaining half in FY 2012. Additionally, the follow-up inspections are conducted over two subsequent periods – the follow-up inspections for the FY 2011 sample are conducted in FY 2012 and FY 2013, while the follow-up inspection for the FY 2012 sample are conducted in FY 2013 and FY 2014. As shown in Exhibit 8, if this

option is used to implement Evaluation Design 3, the inspections of high rate letter only sites will be spread over a four-year period (112 sites in FY 2011; 262 sites in FY 2012; 300 sites in FY 2013; and 150 sites in FY 2014).

Exhibit 8: Inspections of High-Rate Letter Only Sites under Evaluation Design 3

	Number of Inspections of High-Rate Letter Only Sites			
Option 1	FY 2011	FY 2012	FY 2013	FY 2014
Initial Inspections	223	0	0	0
Follow-up Inspections	0	600	0	0
Total Inspections	223	600	0	0
Option 2	FY 2011	FY 2012	FY 2013	FY 2014
Initial Inspections	223	0	0	0
Follow-up Inspections	0	300	300	0
Total Inspections	223	300	300	0
Option 3	FY 2011	FY 2012	FY 2013	FY 2014
Initial Inspections	112	112	0	0
Follow-up Inspections	0	150	300	150
Total Inspections	112	262	300	150

Note: These estimates are based on a 10 percent attrition rate between initial and follow-up inspections.

If we compare options 2 and 3 with option 1, we see that options 2 and 3 both reduce the overall number of follow-up inspections of high rate letter only sites that would be required in a single year. Furthermore, options 2 and 3 would enable OSHA to estimate how the timing of the follow-up inspection affects employer compliance. Specifically, a comparison of the program's impact between sites inspected in FY 2013 and sites inspected in FY 2012 will help assess if the program has an immediate impact on employer compliance or if the program's impact takes more time to be observed. We should note, however, that spreading the follow-up inspections evaluation over two years may result in much higher sample attrition than conducting all follow-up inspections in the year following initial inspections. If OSHA decides to adopt Evaluation Design 3, our recommendation is OSHA uses option 1 to implement the evaluation; options 2 and 3 should be considered only if inspecting a high number of high rate letter only sites in a given year is considered a primary concern.

4. EMPLOYER BEHAVIOR

One of the objectives of the evaluation is to identify the best practices and measures to reduce future occupational injuries and illnesses among employers. In addition, OSHA has expressed interest in learning whether SST interventions may encourage employers to change their practices to improve workplace safety and health. This section describes a means for obtaining this information through the implementation of one of the evaluation designs presented above.

Each of the three evaluation designs described will provide impact estimates of the SST program interventions on a number of measures of employer compliance, such as the likelihood of violation, number of violations, and financial penalties. The impact analyses will provide evidence on whether the SST interventions led to lower injury/illness rates and to improved employer compliance. These impacts are presumably due to some actions the employer took to abate hazards and improve safety and health at the workplace. However, the specific types of actions taken will essentially be unknown. There are several questions that could be answered with more direct measurement of employer behavior change such as:

- Do employers take corrective actions, or different types of corrective actions, as a result of a high rate letter or inspection?
- Are employers who received an SST intervention more likely to seek assistance from OSHA or other sources than those who did not receive this intervention?
- What types of employer behaviors are related to the largest reductions in injury and illness rates, violations, and serious violations? Do these differ by industry type?

With the evaluation design options presented, it is possible to obtain more direct measures of employer behavior change as a result of inspection or high rate letters with some additional data collection. OSHA can work with the evaluation contractor to develop a set of questions to capture employer behavior. These questions would specifically assess employer behaviors at the sites related to safety and health over the prior year. For example, these questions can ask the employers about:

- Changes to the safety and health program
- Safety or health related training programs conducted
- Changes to equipment or equipment added to improve safety and health

- Changes to policy or procedures made to improve safety and health
- Assistance sought from OSHA (e.g., through the Consultation program)
- Assistance sought from other sources (private consultants, insurers, etc.), and
- Hiring of additional safety and health personnel.

One option for data collection could be to include these questions as part of the follow-up inspections conducted for all sites. In this option, Compliance Safety and Health Officers (CSHOs) could be trained to include these questions as part of the follow-up inspection (e.g., as part of the opening conference). OSHA would then provide the answers to these questions for each site in the sample to the evaluation contractor along with the site's DUNS number. This would enable the contractor to link the reported employer behavior to the site's injury and illness rates, as well as to inspection outcomes for analysis. The strengths of this approach are that it would be efficient, taking advantage of inspections that will be conducted as part of the evaluation designs presented. In addition, the response rate would be expected to be very high. However, this approach also has a number of potential disadvantages. For example, it adds an additional information collection burden to the inspection which OSHA may find infeasible. It would also involve training CSHOs to ask the questions and record the answers consistently and then provide the data back to the contractors. The main disadvantage of this approach is that the questions are being asked as part of the inspection. This could bias the results since employers may feel the need to exaggerate their improvements to satisfy the inspectors.

A second, more appropriate, option is to collect employer behavior information through a survey of all sites in the evaluation sample. Working with OSHA, the evaluation contractor would develop a survey instrument including the questions about employer safety and health-related behavior over the prior year. The contractor would send the survey to all sites in the evaluation sample near the end of the follow-up period. This approach avoids burdening the CSHOs, changing the inspection, and potential bias involved with collecting data as part of the inspection. The data collection would be independent of inspections and come directly from the evaluation contractor. However, this approach has some potential disadvantages. There is the added expense of developing and conducting a survey and the likelihood of a lower response rate than the former approach. In addition, a survey will involve obtaining Office of Management and

Budget (OMB) clearance which can be a lengthy process. Nevertheless, if OSHA is interested in this approach, the evaluation contractor can begin developing the survey and progressing through the OMB clearance process early in the first year of the evaluation (e.g., FY 2011), and would likely be able to conduct the study corresponding with the subsequent follow-up inspections.

Regardless of the data collection method, these data would be used to examine whether there is a relationship between the number and type of employer behaviors reported and any change in the injury/illness rates for the site or violation outcomes found in the follow-up inspections. It would also be possible to examine whether the number or type of employer behaviors related to reductions in injury/illness rates or violations differed by industry or other site characteristics.

Moreover, using similar analyses to those described in Section 2 for assessing the relationship between OSHA interventions and inspection outcomes, it would be possible to compare the number and type of employer behaviors reported by sites in the treatment and control groups to determine whether there is a relationship between the SST interventions and employer behaviors. For example, it would be possible to determine whether sites receiving a high rate letter, inspection, or both in the prior year reported engaging in more or specific types of behaviors to improve safety and health compared to those that received no interventions.

Collecting the employer behavior information provides a means of assessing whether employers engage in more or different types of behaviors as a result of OSHA intervention, and what behaviors are related to reductions in injury/illness rates, whether resulting from an OSHA intervention or not. Given the advantages and disadvantages of the two data collection methods, we recommend gathering the data via a survey.

5. CONCLUSION AND RECOMMENDATIONS

The main objective of this project is to identify ways to evaluate the effectiveness of the SST program on employer compliance. This report presented a number of evaluation design options that can be used to assess if the SST program's interventions are effective in improving employer compliance with federal workplace safety and health standards. Our three proposed evaluation designs are based on a random assignment experimental design, which is considered the most effective way to evaluate program impacts.

Specifically, this report presented the following evaluation design options:

- ***Evaluation Design 1*** – Uses an experimental design to estimate the impact of the SST program on compliance for employers on the primary inspection list. Under this design, OSHA will be able to assess if the SST program is effective in improving compliance for employers with the highest reported injury/illness rates, which are the main focus of the SST program.
- ***Evaluation Design 2*** – Uses an experimental design to estimate the impact of the SST program on compliance for employers on the primary *and* on the secondary inspection list. This design expands the scope of Evaluation Design 1 in that it enables OSHA to not only estimate the impact of the SST program for primary list sites, but also for secondary list sites. This would enable the agency to determine if there is value in expanding the SST program to conduct a higher number of secondary list sites.
- ***Evaluation Design 3*** – Uses an experimental design to estimate the impact of the SST program on compliance for all employers exceeding the high rate letter threshold (primary list, secondary list, and high rate letter only sites). This design would enable OSHA to evaluate the impact of the SST program for all sites receiving the high rate letter and assess whether there is value to expanding the program to conduct more inspections of secondary list and high rate letter only sites.

Each of these designs has implementation requirements that may affect the regular SST program operations. Evaluation Design 1 is the least burdensome design as it requires only the inspection

of primary list sites. Evaluation Design 2 requires the inspection of a slightly higher than normal number of secondary inspection list sites. The most burdensome design is Evaluation Design 3, as it requires OSHA to inspect employers who exceed the high rate letter threshold but are not on the two inspection lists. As these sites are not currently targeted by the SST program, implementing Evaluation Design 3 would reduce the SST program's regular number of primary list site inspections to accommodate inspections of high rate letter only sites.

This report also provides a detailed discussion of two important issues that OSHA and the evaluation contractor must consider in implementing the selected design. The first issue is evaluation sample attrition – a proportion of sites selected for the evaluation may be ineligible for the SST program or may go out of business during the implementation period. We suggested a process that requires the implementation contractor to conduct random selection and screening of sites in a way that would offset concerns over sample attrition.

The second issue is that, if Evaluation Design 3 is selected, OSHA would need to inspect a high number of follow-up inspections of high rate letter only sites in the year following random assignment. We discussed three options for spreading the evaluation period to reduce the number of inspections of high rate letter only sites in a given year: Option 1 – conduct all initial inspections in FY 2011 and all follow-up inspections in FY 2012; Option 2 – conduct all initial inspections in FY 2011 and split follow-up inspections between FY 2012 and FY 2013; and Option 3 – conduct initial inspections in FY 2011 and in FY 2012 and conduct follow-up inspections from FY 2012 through FY 2014. The advantage of Options 2 and 3 is that they allow OSHA to spread the evaluation period and reduce the number of high rate letter only inspections conducted in a given year. However, extending the evaluation period may produce higher sample attrition. Our recommendation is that OSHA consider using Option 2 or Option 3 only if inspecting a high number of high rate letter only sites in a given year is a primary concern.

In conclusion, this document provides OSHA with a number of options in implementing a rigorous evaluation of the SST program. We recommend that OSHA consider adopting Evaluation Design 2 to rigorously evaluate the impact of the SST program on employer

compliance for both primary list and secondary list sites. While requiring only modest program changes and a feasible sample size, such an evaluation would provide important insight on the effectiveness of the program and would provide critical information in determining the focus of the SST program in the future. If the SST program is more effective for primary list employers, then the program should continue to focus more on targeting employers with the highest injury/illness rates. However, if the program is at least as effective for secondary list employers, it may be valuable for OSHA to consider providing more emphasis on the secondary list employers.

We also recommend that OSHA consider implementing a survey of all employers in the evaluation to gather information on employer behavior related to improving their workplace safety and health. This information will enable OSHA to examine how the SST program's interventions affected employer compliance and to examine which behaviors led to lower injury/illness rates. Our recommendation is that OSHA works closely with the implementation contractor to develop a survey instrument that can be submitted for OMB clearance at the start of the evaluation period. This will enable OSHA and the evaluation contractor to implement the survey near the end of the follow-up inspections.