



Feasibility Study of Establishing a New Cohort of People  
Exposed to 9/11 as Children for 9/11-related Research

Feasibility Study of Establishing a New Cohort of People Exposed to 9/11  
as Children for 9/11-related Research, October 2020

Report to National Institute for Occupational Safety and Health

Feasibility Study of Establishing a New Cohort of People Exposed to 9/11  
as Children for 9/11-related Research

Principle Investigator: Robert M Brackbill, Research Director,

Co-PI: Erin Takemoto

World Trade Center Health Registry, New York City Department of Health and Mental Hygiene

Report authors:

Robert Brackbill

Mark Farfel

Erin Takemoto

Felix Ortega

Lennon Turner

Lucie Millien

Howard Alper

October 2020

Table of Contents

EXECUTIVE SUMMARY .....v

1.0 INTRODUCTION .....1

    1.1 Feasibility Study Goals and Specific Aims .....1

2.0 PROJECT PLANNING.....2

3.0 METHODS.....3

    3.1 Data Source.....3

    3.2 Definition of Exposed and Unexposed Zones.....3

    3.3 Denominator Estimates for Children in Exposed and Unexposed Zones.....4

    3.4 Selecting the Feasibility Study Sample from the Population of DOE Directory  
    Data.....4

    3.5 Power Considerations and Calculations.....5

    3.6 Methods for Conducting the feasibility Study.....6

        3.6.1 Aim 1: Assess the Feasibility of Tracing and Locating.....6

            3.6.1.1 Aim 1a: Review of DOE Directory Data.....7

            3.6.1.2 Aim 1b: Tracing and Locating and Follow-up Mailing.....7

            3.6.1.3 Aim 1c: Level of Effort.....10

        3.6.2 Aim 2: Assess the Level of Interest in Enrolling in a New Children’s  
        Cohort for Research.....10

            3.6.2.1 Aim 2a: Contact, Cooperation, and Response Rate..... 11

            3.6.2.2 Aim 2b: Representativeness.....11

4.0 RESULTS.....11

    4.1 Aim 1a: Review of DOE Directory Data..... 11

    4.2 Aim 1b: Tracing and Locating and Follow-up Mailing.....11

        4.2.1 Batch Tracing.....11

        4.2.2 Intensive Tracing Part 1.....12

        4.2.3 Intensive Tracing Part 2.....12

    4.3 Aim 1c: Assess Level of Effort.....13

    4.4 Aim 2a: Contact, Cooperation, and Response Rates.....14

    4.5 Aim 2b: Representativeness.....14

5.0 DISCUSSION.....14

    5.1 Strengths and Limitations.....17

6.0 CONCLUSION.....18

7.0 REFERENCES.....18

### Appendices

- Appendix A: Registry scientific outputs on children, adolescents, and young adults
- Appendix B: DOE Public Notice with opt-out option
- Appendix C: DOE Publics Notice dissemination materials (website, ads, social media)
- Appendix D: Tracing Protocol
- Appendix E: Feasibility Study Informational Brochure
- Appendix F: Study timeline
- Appendix G: Power calculations
- Appendix H: Details of findings at address level

### Figures

- Figure 1: Summary of Tracing Activities

### Tables

- Table 1: Exposed Zone Neighborhood Tabulation Areas (NTAs) and Corresponding Sociodemographic Characteristics (2000 US Census)
- Table 2: Unexposed Zone Neighborhood Tabulation Areas (NTAs) and Corresponding Sociodemographic Characteristics (2000 US Census)
- Table 3: Exposed Zone Neighborhoods, Zip Codes, and School Districts Included in Neighborhoods
- Table 4: Unexposed Zone Neighborhoods, School Districts, and Zip Codes
- Table 5: Populations by NTA for Exposed Zone Neighborhoods (2000 US Census)
- Table 6: Sampling Proportions for Exposed Zone for the Feasibility Study, Calculated from Population Values in
- Table 7: Sample Size by Gender, Age, Race/ethnicity Used for Sampling Based on DOE Directory Data for Exposed and Unexposed
- Table 8: Summary of Updates Received as a Result of Batch Tracing by Student
- Table 9: Intensive Tracing Part 1 Results
- Table 10: Intensive Tracing Part 2 Results

Feasibility Study of Establishing a New Cohort of People Exposed to 9/11  
as Children for 9/11-related Research, October 2020

Table 11: Results from Three Stages of Tracing and Locating and Three Subsequent Mailings  
of Brochures

Table 12: Undeliverable by Gender, Race/ethnicity, and Age for Exposed and Un-exposed

Table 13: Completed Returned Brochure by Race/ethnicity-gender Groups

## EXECUTIVE SUMMARY

The World Trade Center Health Registry (WTCHR) hosted by the New York City Department of Health and Mental Hygiene (DOHMH) was asked by the National Institute for Occupational Safety and Health (NIOSH) to assess the feasibility, 20 years after 9/11, of establishing a new cohort for research among people exposed to the 9/11/01 WTC disaster as children together with an unexposed comparison group. In collaboration with the NYC Department of Education (DOE), the project began in Aug. 2017. Data collection ended July 2020. The exposed zone was lower Manhattan and western Brooklyn within ~1.5 miles of the WTC site and the non-exposed zone was located  $\geq 6$  miles from the site. DOE directory information included ~160,000 people who were children at the time of 9/11. The project entailed three phases of tracing, each followed by contacting, a stratified random sample of 501 students attending schools in the exposed zone and 501 unexposed students with similar sociodemographic characteristics to the exposed children. Contacting consisted of mailing informational brochures to students/guardians and calling half of the sample members.

Overall, we found that DOE directory information was a viable source for locating former 9/11 students. DOE directory information was sufficiently complete for tracing all but one of the 1,002 former students. State-of-the-art tracing and locating methods were able to update or confirm a large proportion of contact information provided by DOE. For example, step 1 batch tracing using computerized data base searches updated 85% of the student records, and over half of the updates (58%) provided specific information on former students, now adults. Each round of intensive tracing with humans doing the data base searches improved the pool of addresses as evidenced by additional address updates and confirmations and reduced numbers of undeliverable addresses.

Brochures were mailed to a total of 4,761 addresses for students/guardians (1-3 brochures per address). Thirty-three former students (3%) responded to the brochure mailings (5.2% from exposed zone and 1.4% unexposed), of whom 31 expressed interest in a new child cohort. However, few racial/ethnic minority students responded, and the latter had a higher proportion of undeliverable addresses versus non-Hispanic whites. Females had a higher response than males. Phone outreach to a sub-sample of former students/guardians (n=489), resulted in nearly half being reached (contact rate=49%) and 136 of them agreeing to confirm or update their address (28%), for a cooperation rate of 56%. As expected, both rates were lower than those attained by the WTCHR in 2003-04. Only one of the 136 (0.7%) with confirmed addresses responded to the informational brochure mailed to them.

The overall low level of expressed interest in joining a new cohort for research, including the very low (1.4%) response from the unexposed zone, combined with the findings on differential response by race/ethnicity and gender and the contact and cooperation rates suggest that the formation of an epidemiologically useful cohort for research would likely require more extensive tracing and outreach efforts than those implemented in this feasibility study.

Based on the findings, tracing and contacting the entire DOE file of 153,000 former students would be expected to yield a sample of ~4,600 former students (3% overall response) interested in joining a new cohort across exposed and unexposed zones. The response from students from schools in the exposed zone suggests a level of interest potentially sufficient for creating a

cohort to study more common types of outcomes such as coronary heart disease. However, the prospective recruitment did not yield a representative sample, indicating a high likelihood of selection bias in any future recruitment for a children's cohort. This bias needs to be considered in any future recruitment strategy as it would limit the ability to study even common health conditions and, therefore, limit the scientific value of a new cohort for research.

Strengths of the project included availability of a large database on 9/11 former students and WTCHR infrastructure for carrying out the project. Limitations are that the project was not designed to determine (a) why former students did not respond, (b) the interest of guardians in participating in the research, or (c) whether former students who expressed interest would actually provide consent and join a cohort for research. Also, the project did not assess the feasibility of finding students who went to school in the target zones but lived elsewhere or attended private school; a future protocol might use strategies such as media outreach to do this. Lastly, use of a mailed brochure for assessing interest may have had limited appeal to some former students; however, using other methods such as phone calls would have been difficult because of the challenges involved in obtaining phone numbers. The cost for similarly tracing the entire DOE directory of 160,000 eligible students was estimated to be \$48 million.

These findings will help inform NIOSH's next steps related to the potential formation of a new cohort for 9/11-related research among people exposed as children. As a result of this project, we now know much more about the utility and completeness of the DOE student directory data, the ability of state-of-the art tracing and locating activities to obtain updated and confirmed addresses, contact and cooperation rates associated with phone outreach, the level of expressed interest in a new cohort among a sample of former students by zone and by demographics, and the costs of tracing and contacting activities.

## 1.0 INTRODUCTION

Given its experience developing and operating a large 9/11 exposure registry since 2002, the WTC Health Registry (WTCHR) hosted by the New York City Department of Health and Mental Hygiene (DOHMH) was asked by the National Institute for Occupational Safety and Health (NIOSH) to assess the feasibility of establishing a new cohort for research among people exposed to 9/11 as children. The Registry has been a valuable resource for 9/11 research on children based on its sub-cohort of 3,251 children age <18 years who lived or were enrolled in public or private schools (K-12) in lower Manhattan on 9/11/01. This sub-cohort was enrolled by their parents/guardians (hereinafter referred to as guardians) who completed the children's baseline Wave 1 health surveys as proxies. Child enrollees and those who aged into adulthood were included in Registry survey Waves 2, 3 and 4, which included separate pediatric surveys for guardians at Wave 2 and for guardians of young children and adolescents themselves at Wave 3. All child enrollees are now adults and included in the Wave 5 survey launched in April 2020.

The Registry has numerous publications on short- and longer-term 9/11-related physical health, mental health and behavioral outcomes among children and young adults. Outcomes include adverse birth outcomes, asthma, asthma control, substance use, school behaviors, PTSD symptoms, adolescent behavior, and risk of mental health problems in early adulthood, cancer, and mortality (see Appendix A for list of Registry scientific outputs on children).

Some researchers, community members, and other stakeholders have voiced concerns about the limitations of understanding the full health effects of 9/11 on children based on research using the currently available Registry data. Their concerns include (a) the degree to which children enrolled in the Registry represent the population of all WTC exposed children; (b) the absence of an "unexposed" comparison group of children; and (c) the size of the Registry's sub-cohort of children. Stakeholders raised the idea of developing a new cohort separate from the WTCHR with the goal of having a more complete and more representative cohort of children exposed to 9/11 and an "unexposed" comparison group for research on 9/11 health outcomes. All parties recognized that efforts to develop a new cohort of people exposed to 9/11 as children would face logistical and scientific challenges, limitations, and potential high costs. Among them is the challenge of locating decades after 9/11 members of the target population, many of whom may have relocated and/or changed their legal last names after marriage. Additionally, obtaining self-reported information on 9/11 exposure and health status is problematic and likely unreliable because of memory gaps, subsequent experiences and events, and the influence of historical revisionism regarding 9/11 on the recall of events. A question is whether interest in joining a new cohort for research many years post-disaster may be diminished, especially among those who were very young at the time of 9/11 or those in the unexposed comparison population.

### 1.1 Feasibility Study Goals and Specific Aims

The feasibility study was designed to assess these issues nearly two decades after 9/11 and provide guidance on the feasibility of developing a new 9/11 related children's cohort for research and thereby inform next steps by NIOSH regarding establishing a new cohort.

Aim 1 was to assess the feasibility of tracing and locating a sample of people exposed and unexposed to the 9/11 disaster in New York City as children (age <18 years) with specific sub-aims to: a) evaluate the completeness and accuracy of DOE directory data, a core source of information on children enrolled in NYC public schools at the time of 9/11; b) assess the degree to which tracing and locating procedures provides updated address information that is suitable for mailing informational brochures; c) assess the success rate and level of effort required for obtaining contact information on individuals who were children on 09/11/2001.



Aim 2 assessed the level of interest in enrolling in a new children’s cohort for research with specific sub-aims to: a) assess whether the contact, cooperation, and response rates would be adequate for developing a sufficiently large and scientifically valid cohort; b) determine if it was possible to recruit a reasonably representative sample of people who were children and attended school and/or resided in “exposed” and “unexposed” comparison zones in NYC at the time of 9/11.

## 2.0 PROJECT PLANNING

In August, 2017 representatives from the NYC Department of Education (DOE), NYC Department of Health and Mental Hygiene (DOHMH), Mayor’s Office, National Institute for Occupational Safety and Health, 9/11 Health Watch and the WTC Responder Steering Committee met to discuss the possibility of DOE sharing student directory information with DOHMH for the purpose of creating a 9/11 related children’s cohort. It was agreed in the meeting that DOE could provide the student directory information under an exception of the Family Educational Rights and Privacy Act (FERPA) regulations for releasing identifiable data for public good. The Registry then held sustained planning meetings with DOE and completed legal/administrative and Institutional Review Board (IRB) requirements prior to conducting this study [i.e., submitting a formal data request to DOE, obtaining DOHMH IRB approval of the study protocol, helping to develop and disseminate the mandatory DOE Public Notice, and executing a data use agreement (DUA) with NYC DOE].

The Registry submitted a formal request for student directory information to DOE in October 2017 describing the rationale and purpose of a potential future children’s cohort, data elements requested, definition of exposed and unexposed areas, and feasibility study aims, timeline and initial sampling plan. DOE approved the request and provided aggregated student directory data for the areas of New York City requested for planning purposes.

The study protocol was submitted to the DOHMH IRB in July 2018 and approved by the IRB in November 2018 following a comprehensive review. The protocol described the purpose and rationale for a 9/11 related children’s cohort, aims, design and methods for tracing and locating feasibility study sample members and assessing interest in joining a potential new 9/11 cohort, and criteria for consideration when initiating a potential children’s cohort. The IRB approved DOHMH to receive DOE directory information for the feasibility study sample, but not the directory information for the remaining ~150,000 children and their guardians in the full sample.

DOE provided directory information under a directory information exception that permits release of data for evaluation and research after disseminating a mandatory Public Notice with an option for individuals to opt-out of the data sharing. Concurrent with the IRB process and other discussions with DOE, DOHMH in collaboration with DOE developed the Public Notice (Appendix B, Public Notice with opt-out option) and a dissemination plan. The notice was translated into eleven languages and disseminated via the DOE website, two rounds of ads in 16 local newspapers in English, Spanish and Chinese, and a DOE social media campaign (Appendix C, Public Notice dissemination materials) in January 2019. In addition, DOHMH developed a page on the Registry’s website about the feasibility study for feasibility study sample members to access and verify that the study was a legitimate effort. The DOE public notice included all required legal elements, including providing at least 30 days for people to opt-out by calling an 1-800 number or mailing a signed reply form. The opt out processes was completed in March 2019 by DOE with fewer than 20 valid opt outs. A Data Use Agreement between DOE and DOHMH was executed in May 2019 to allow for transfer of data.

In planning for tracing work, the Registry created an “R-Tools” software module specifically designed for the tracking the locating effort. The Registry also developed a tracing vendor scope of work included in the DOHMH request for proposals, reviewed proposals, selected a vendor and worked with the vendor to develop a tracing protocol with detailed scripts for contacting students or guardians as part of intensive tracing. The tracing protocol (Appendix D) was approved by DOHMH and vendor Institutional Review Boards (IRBs). The required process of securing a vendor and IRB approvals and executing a DUA with the tracing vendor took nearly a year and the vendor began work in May 2019.

DOHMH developed a feasibility study outreach informational brochure (Appendix E) that was mailed to the sample population. The brochure described the potential new cohort for research and the various ways (mailed form, email, telephone) that the students could indicate his or her interest or not in participating in a future research project. The brochure was developed with input from the WTC Survivor Steering Committee, a Registry focus group of young adult enrollees, and DOHMH Communications during Jan.-June 2019, and received approvals from DOHMH IRB and Communications (see Appendix F for study timeline).

### 3.0 METHODS

#### 3.1 Data Source

The DOE student directory information was selected as the data source for the feasibility study as it provides the best-known denominator with the greatest known potential coverage of children who were residents in New York City at the time of 9/11 and in subsequent years. Directory information is provided and updated at the beginning of each school year by guardians and retained by DOE for children who were enrolled in NYC public schools. It is unlikely that other lists or databases exist that would enable the identification and enumeration of children in exposed and unexposed comparison zones on 9/11/01.

#### 3.2 Definition of Exposed and Unexposed Zones

The exposed zone was defined as lower Manhattan at or south of 14th Street and western Brooklyn within ~1.5 miles of the WTC disaster site. It includes the WTC Health Program eligibility catchment area and the area between Houston St and 14<sup>th</sup> Avenue in Manhattan. The latter provides the potential of considering a gradient of exposure to 9/11 disaster based on distance from the WTC site.

Criteria for the unexposed comparison zone were: (1) being geographically located  $\geq 6$  miles from the WTC site in order to minimize the likelihood of resultant environmental toxicant exposure and (2) having similar sociodemographic characteristics to the exposed children in lower Manhattan and western Brooklyn by race/ethnicity, home language, nativity, median household income, proportion above the poverty level, and education level. The 2000 U.S. census provided the most accurate and timely description of the sociodemographic characteristics of neighborhoods within the five boroughs of New York City. Based on the 2000 U.S. census, eight neighborhood tabulation areas (NTA) comprised the geographic area of the exposed zone (Table 1). With input from community and labor advisors and Registry staff, four non-exposed zone NTAs were identified which were socio-demographically like the exposed zone (Table 2) and met the geographic requirements.

Based on school address and child criteria, a comprehensive roster of children attending schools in the exposed zone described in Table 3 and the unexposed zone described in Table 4 was compiled by DOE using historical student directory information. The roster of children included those less than 18 years old at the time of 9/11/01 (born on or after 9/12/1983) who

were either: enrolled in a school any time in the 2001-2002 school year, or aged 0-5 years old on 9/11/01 and first enrolled in school between 2002 and 2007.

DOE directory information provided individual level data on 153,000 students attending schools within the selected exposed and unexposed zones during the 2001-2002 academic year and during the subsequent 5 years, to capture children who were too young to be in school at the time of the WTC disaster. Exposed children were enrolled in schools that served neighborhoods within ~1.5 miles of WTC site and unexposed children were enrolled in schools located more than 6 miles from the WTC site.

### 3.3 Denominator estimates for children in exposed and unexposed zones

The computation of a best denominator estimate for the exposed zone was based on use of the following three elements: (1) the 2000 census total of those aged 0 – 17 years for the exposed neighborhoods (n= 46,757); (2) the DOE total of former students in schools located in the exposure zone (n= 81,270); and (3) WTCHR data on the proportion of children who were both residents and enrolled in schools in the exposed zone (~24%) (see Thomas, 2007, Table 1). Combining this information yielded an estimated denominator for the exposed zone of 97,300 children (calculated as  $(46757+81270) * (1-0.24)$ ). The denominator for the comparison zone was based on 2000 census information for children aged 0-17 residing in the selected comparison zone neighborhoods (n=62,240). The estimated total number of children in the selected exposed and unexposed zones is 160,000.

### 3.4 Selecting the Feasibility Study Sample from the Population of DOE Directory Data

A stratified random sample of 1,002 student records was selected from DOE student directory information by DOE. The sample is comprised of 501 students from exposed areas and 501 from comparison areas. These sample sizes were deemed sufficient for assessing the effort and yield of locating former students while reducing costs of conducting intensive tracing to obtain updated address information, which ranged up to \$300 dollars per case.

The sample was selected from approximately 153,000 records in the DOE directory database of students who were enrolled in the NYC public schools on 09/11/2001 or were first enrolled between 2002 and 2007. The sample drew an equal number of students (n = 501) from exposed and non-exposed areas stratified by age group (0-4, 5-9, 10-13, 14-17 years), sex (male/female), and race/ethnicity (Asian, Black, Hispanic, White) and the comparison sample was selected to be comparable to those in exposed areas in terms of race/ethnicity.

The eight Neighborhood Tabulation Areas (NTAs) that comprised the exposed zone (Table 5) were collapsed to generate a demographic profile of the exposed zone. The proportions for each gender, age, and race/ethnic group strata (Table 6) were used to conduct sampling from both the exposed and unexposed zones. Table 7 provides the number in each cell by gender, age and race/ethnic group strata based on proportions in Table 6 which would be the ideal goal for attaining representation of interest in a new children's cohort for both exposed and un-exposed zones.

The student directory information DOE provided to DOHMH consisted of the following elements: Child's full name, sex of child, date of birth of child, place of birth of child (limited to NYC, NYS, or outside NYS), guardians names, guardian relationship to child (e.g. mother, father, grandparent), address history going forward in time from 9/11/2001 of child and guardians with respective years associated with addresses, available phone numbers and/or email addresses of guardians going forward in time from 9/11/2001, language spoken/written of guardians, schools enrolled in and period of enrollment in each school through the last school enrolled, the most recent educational institution attended and period of attendance.

In March 2019, a Registry epidemiologist reviewed the data on a DOE computer to assess the completeness of the database. On subsequent visits to DOE offices, the epidemiologist wrote and tested computer code for selecting a sample. A DOE analyst ran the code and produced the stratified (exposed and unexposed zones) sample. DOE transferred the 1,002 records to DOHMH using a secure data sharing application in June 2019. Separate files were transferred to obtain complete records for students and their guardians. A Registry data scientist reviewed the records and identified missing information for final resolution with DOE.

### 3.5 Power Considerations and Calculations

A combination of factors may result in a final cohort size that is substantially less than the estimated eligible population for the children's cohort. The scientific validity for making inferences about 9/11 exposure and long-term effects on health will also be related to the resources available for building a new children's cohort, the location rate, level of interest of participation in a potential new cohort, the prevalence of the health outcome of interest and the sample's representativeness of the exposure zone and comparison zone's demographic characteristics for the time period of interest.

Power calculations provide an estimate of the minimal sample required for statistical inference related to specific health outcomes and were included in the feasibility study protocol for selected outcomes (see Appendix G. Power computations). Power computations were based on expected associations reported in the literature. For instance, Thomas (2008) reported adjusted Odds Ratios (AOR) for dust cloud exposure and asthma among children ranging from 1.7 to 2.2. A conservative AOR of 1.8 was used for persons who would be younger than 21 at time of potential creation of a new children's cohort. Brackbill (2009) reported AORs ranging between 1.4 and 1.5 for intense dust cloud for adults older than 18. A conservative AOR was applied that declined with age for persons older than 21 years at time of possible creation of a new cohort. Regarding heart disease, Jordan (2011) reported statistically significant adjusted hazard ratios for heart disease for different 9/11 exposures for adult men and women that ranged between 1.3 and 2.1. The most recent report (Li et al, 2016) on cancer incidence and 9/11 exposure (for cancers diagnosed up to 2011) reported a significant Standardized Incidence Ratio (SIR) of 1.1 for non-rescue/recovery workers. Based on this information, sample estimates needed for a children's cohort range from almost 2 million people to detect an exposure/disease association for cancer, to 3,800 and 6,900 people for more common conditions such as asthma and coronary heart disease, respectively. In addition, if the final children's cohort sample under-represents specific categories such as zone, age groups, or race/ethnicity groups, then the sample required as indicated by power computations would need to be increased to offset the representation bias.

Extrapolating from these target sample size estimates, the feasibility study would need at least 2.0% of the 501 exposed children or 10 returns of interest from the exposure zone reply and 3.1% or 15 of 501 from unexposed children for asthma and about 18 and 28 responses from exposed and unexposed people, respectively, for coronary heart disease.

Because asthma has already been established as being associated with 9/11 exposure, a more relevant sample size is for outcomes such as heart disease that would require a total of about 50 responses for exposure and comparison areas. These criteria assume that most of the addresses are valid and that the target groups receive the brochures. For representativeness, these numbers would need to be proportionately distributed especially between age groups, sex, and race/ethnic groups to reflect a similar proportion in the underlying population.

### 3.6 Methods for Conducting the Feasibility Study

#### 3.6.1 Aim 1: Assess the Feasibility of Tracing and Locating

The completion of Aim 1 for the feasibility study involved three sub aims. The first was a review of the de-identified DOE directory data file for the children in the full roster and then then one submitted to DOHMH on the 1,002 students to assess the completeness/missingness of the data and to determine whether the level of missing information for certain elements would hinder the ability to obtain up-to-date address information. The second sub-aim involved the application of various strategies for correcting, updating, and verifying location information and assessing the degree to which tracing and locating would result in updated address information that was sufficiently complete for brochure mailing. The third sub-aim was to determine the level of effort required to acquire updated contact information for former students.

##### 3.6.1.1 Aim1a: Review of DOE Directory Data

The process of assessing DOE's ability to build the needed data files and reviewing DOE directory data was initiated by submitting requests to a DOE analyst to provide a file that included counts of students for the years and areas of New York City requested by DOHMH. DOE also provided counts on the various data elements requested by DOHMH to ascertain whether the entire file had potential for serving as a denominator for a new children's cohort. A Registry epidemiologist also reviewed the DOE data and performed searches on the data inside the DOE facility. After the sample for the feasibility study was transmitted to DOHMH, Registry staff reviewed the files for completeness and in some cases had to request corrected files from DOE.

For Aim1a, analysis consisted of first reviewing whether the file included the correct schools in the exposed and non-exposed zones. The review also assessed the presence of out-of-range or non-existent addresses and zip codes and the level of missing data for address variables such as street and apartment number. A 10% level or lower level was considered acceptable for missing information on specific variables.

##### 3.6.1.2 Aim1b. Tracing and Locating and Follow-up Mailing

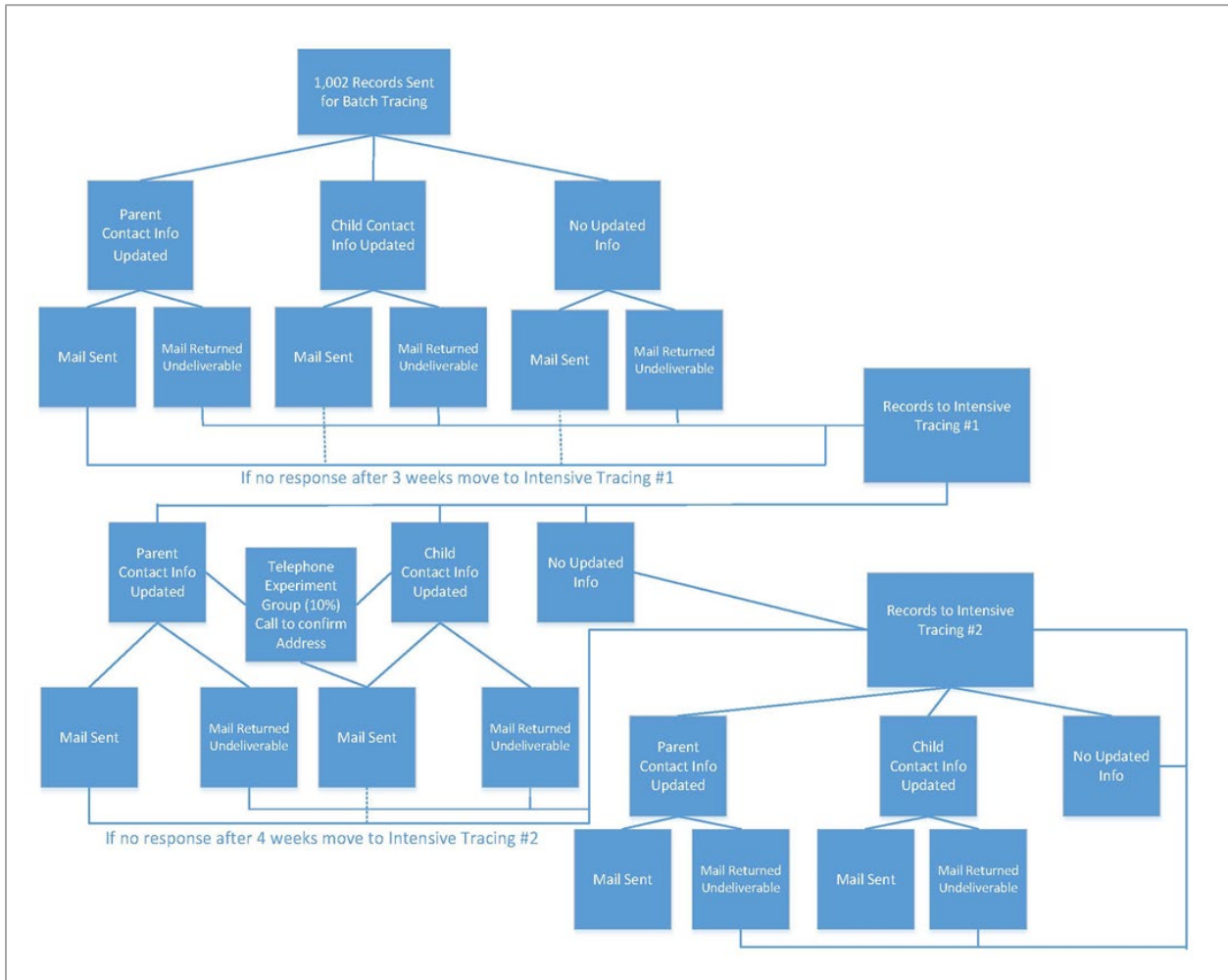
Tracing and locating were conducted by an experienced vendor in conjunction with DOHMH staff. This sub-aim was a core activity of the feasibility study and involved several iterative activities: a) conducting batch tracing (described below) using information contained in the sample of student directory data; b) mailing an outreach informational brochure using updated addresses for former students/ guardians for assessing interest in joining a new children's cohort for research; c) conducting intensive tracing (defined below) of individuals who could not be located by batch tracing or of individuals whose mailing was returned as undeliverable; e) re-mailing of brochures to persons whose contact information was updated through tracing. The tracing process used first name, last name, gender, date of birth of child, previous address, email address and telephone numbers for verifying the identity and obtaining updated contact information of the student or their guardians. A second round of intensive tracing involved contacting guardians or students by telephone (given that telephone number had been obtained in prior searches) and describing the study and confirming or updating the address that was available. Figure 1 depicts these activities. The following is a description of the batch tracing and two levels of intensive tracing and follow-up mailing of informational brochures.

Batch tracing was conducted from July 2019 to August 2019 using the National Change of Address (NCOA) database and a proprietary database which culls contact information from multiple databases. NCOA is a database consisting of change-of-address data submitted to the U.S. Postal Service. When batch tracing returned the same contact information or did not find the submitted contact information, the original DOE contact information was retained. The first

mailing of informational brochures was done in September 2019. Three weeks after the first mailing, Intensive tracing Part I (IT-1) was conducted from October 2019 through November 2019. This phase focused on two types of student cases: 1) those for which brochures were returned as undeliverable with no forwarding address, and 2) those who did not respond to the mailing sent after batch tracing. The aim of IT-1 was to obtain current mailing addresses for these students and/or their guardians to be used for a subsequent mailing.

Intensive tracing was designed to locate people through using all available information including searches of online databases (e.g., Lexus-Nexus). If someone could not be located within two separate one-hour sessions by a tracing specialist, they were considered un-locatable. When an identical address was returned from 2 out of 3 database searches, the case was coded as 'located' or 'confirmed'. A case deemed as having a located address could include either: 1) a new address listed in 2 of the 3 databases, or 2) the current address in the data file showed up in 2 of the 3 databases. If all 3 searches returned different contact information, then the contact information was coded as 'unconfirmed' (i.e., no consistent contact information was found). The second mailing was conducted in December 2019.

Figure 1. Summary of tracing activities



Intensive Tracing Part 2 (IT-2) began February 3, 2020, approximately four weeks after the previous batch of brochures was mailed and continued through August 2020. For IT-2, 490 student/guardian cases were randomly selected from 980 eligible cases. The subsample had equal distribution across students exposed and non-exposed to 9/11, and by gender, age (0-4, 5-9, 10-13, 14-17 years), and race/ethnicity (Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic Asian).

IT-2 focused on contacting students or one guardian (guardian 1) by telephone to update or confirm student contact information. Sampled cases for this tracing effort came from one of two sources: 1) brochures DOHMH mailed during IT-1 that were returned as undeliverable with no forwarding address, and 2) cases that did not respond to the IT-1 mailing after 4 weeks.

Enhanced intensive tracing was done if a current correct address could not be obtained by step 1 batch tracing or step 2 IT- 1. IT-2, however, required an available telephone number for guardians or students. The caller identified the individual whose name was in the directory information and inquired about a current address. The scripts for these calls were based on the experience and protocols developed by the Registry for recruitment, treatment referral outreach and studies involving encounters by telephone.

An average of an hour was devoted to each case. During this time, a tracer reviewed an individual case and called the student telephone number(s) starting with those identified as most likely current. When unable to reach a student, calls were next made to the telephone number(s) associated with guardian 1. During IT-2, tracers also conducted interactive tracing utilizing multiple databases such as Lexis-Nexis and Experian to search for more contact information when no current telephone number was working for either student or guardian 1. Up to three call attempts were made to each available working telephone number.

For Aim1b analysis, each stage of tracing and locating returned results on the outcome of the search for addresses. The outcomes included whether the address was found in the data bases searched and if found whether it was “updated” or “confirmed”. In some cases, the returned update included address specific for the former student and not for the guardian. Given that the addresses provided by DOE were only guardian addresses, this was new information. If the address was not found this was also recorded. Other key outcomes were based on the results of the mailings. These included a “returned” informational brochure with a reply, an informational brochure returned as “undeliverable” or returned with a “forwarding address”. “Undeliverable” indicates that the post office could not deliver the mailing because the address did not exist or the recipient’s name on the mailing was not at the address. If a forwarding address was available in the post office files, then a forwarding address was included on the returned piece. “No return” was also an outcome tracked after each mailing. The outcomes described above were for individual addresses for which most former students had multiple addresses. Both the results at address level and student level are provided in this report. The student level analysis has specific combinations such as student only, student and guardian, or guardian only that are indicated where appropriate.

An additional distinction is between “presumed valid” and “confirmed valid” addresses. Presumed valid addresses are addresses that have been found in data base searches and either updated or found and had not been returned as undeliverable. These addresses are presumed valid because it is unknown whether the person named on the address could receive the mailing. Confirmed valid are those that resulted in a returned brochure or were confirmed during IT-2 when a former student or their guardian was reached by phone and agreed to update or confirm the address on the call.

The results are presented first at the address level in terms of proportion of addresses, which represents multiple addresses for individual students. In addition, the results are also presented at the student level so that counts or proportions are among the number of student records. The denominator for the proportions is indicated where appropriate. The results are presented for batch, IT-1, and IT-2



separately and associated mailings and for all the activities combined. No inferential statistics were performed.

#### 3.6.1.3 Aim1c. Level of Effort

We assessed the yield achieved for a given level of effort expended in tracing and locating to inform the feasibility of a children's cohort if the tracing and locating were to be expanded in scope to form a new cohort. For this aim, we summarized the overall results of tracing and locating and mailing and addressed three questions: a) how much additional information on addresses was obtained by intensive tracing compared to batch tracing? b) did the number of un-deliverables decline after each stage of tracing and locating? c) what was the cost for obtaining confirmed and updated addresses and reducing un-deliverables? First, the number of updated and confirmed addresses after each stage of tracing and locating and number of returns and un-deliverables after each mailing were compared. Second, it would be expected that the number of un-deliverables would decrease with each stage of tracing and locating, and this was tested. Third, because tracers typically spent 45m to an hour on each case, we could estimate the total number of hours required to obtain updates and confirmed addresses over the course of the study.

3.6.2 Aim 2. Assess the Level of Interest in Enrolling in a New Children's Cohort for Research  
Aim 1 activities of updating contact information and mailing of brochures provided information for the assessing Aim 2.

#### 3.6.2.1 Aim2a. Contact, Cooperation, and Response Rate

This sub-aim was to assess contact, cooperation, and response rates based on the sample of 1,002 students for consideration in assessing the feasibility of a scientifically valid new cohort. IT-2 was the source of data for assessing contact and cooperation rates through the calls to a sample of former students with available phone numbers. The level of un-deliverable addresses was also informative about potential contact rates for a children's cohort. Response rates are based on the number of returned informational brochures after all the mailings were completed. The response rate was compared to the calculated response needed for an epidemiologically valid sample size for a new cohort based on power calculations. Demographic characteristics and exposure group (exposed/unexposed zone) status was used to assess the feasibility of obtaining a representative sample of children for participating in a children's cohort.

The level of interest or response rates was primarily based on the number of responses from former students who attended schools in the exposed and un-exposed zones. Information on contact and potential cooperation rates were derived from IT-2 based on the number of persons called who were reached, and how many provided updates or confirmations and how many refused to participate. In addition, response rate for addresses determined to be identified in IT-2 was also computed.

#### 3.6.2.2 Aim2b: Representativeness

First, this was indicated by the number of un-deliverables (unreachable) or presumed valid as students with no un-deliverable addresses for assessing which groups are less likely to be contactable in study. Second, the number of returns by gender, current age, race/ethnic group within the exposed and un-exposed zones was also indicative of which groups would be under-represented in a children's cohort.

## 4.0 RESULTS

### 4.1 Aim1a: Review of DOE Directory Data

DOHMH found errors in allocation of students to schools in exposed and un-exposed zones in the base file and DOE provided a corrected file. Among the 153,000 former students, there were no address data elements which had more than 10% missing data and many had less than 5% missing.

#### 4.2 Aim1b: Tracing and Locating and Follow-up Mailing

There were 2,522 addresses for 1,002 students in the DOE file available for tracing and locating and one to three addresses per student. Overall, 85% of the addresses were categorized as presumed valid and only 30% of students had at least one brochure returned undeliverable for any one of the addresses available. On a student level, 169 (17%) students had a confirmed valid address (33 student who responded to the brochure and 136 students with addresses confirmed by phone contact in intensive tracing part 2 (28% of the IT-2 sub-sample [136/489]). (See Appendix H. detailed results at address level).

##### 4.2.1 Batch Tracing

The 2,522 addresses provided by DOE were included in batch tracing. Most (72%) were updated (n=1561) or matched an address provided by DOE (n=266). The remainder (n=725) were primarily duplicate addresses. The updated and matched addresses were retained for use in the first mailing (see Appendix 9 for more detail). All but one student had sufficient address information for tracing (n=1,001). Contact information was considered updated when batch tracing returned new or additional contact information, including apartment number. Contact information was not considered updated when batch tracing returned the same address with a variation in the address format (e.g., Street vs. ST, Avenue vs. Ave).

Of the 1,001 student level addresses sent to batch tracing, information was updated for 855 cases (85.3%, see Table 8). Over half of the cases had an update that included student address information (n=581, 58.0%). This is important because the DOE directory included only addresses provided by guardians, and thus batch tracing yielded some specific information on former students who are now adults. For nearly a quarter of the cases, contact information was updated for both the student and guardian 1 (n=227, 22.7%).

Only 147 (14.7%) cases had no update. Of these, batch tracing returned no contact information for 112 (11.2% of total) cases and the same contact information as the directory for 35 cases (3.5% of total). In instances when the street address returned from batch tracing matched the directory address but lacked the directory apartment number the apartment number was retained, and an outcome code was assigned to reflect no update. Because apartment numbers are very important for New York City, extra effort was made to confirm that all contact information included apartment numbers when available.

##### 4.2.2 Intensive Tracing Part 1 (IT-1)

For IT-1, 1,586 (1797- 211) addresses were searched after excluding addresses which yielded responses from former students (n=13) or were associated with a brochure that was returned as undeliverable (n=198, known at the time that they were searched). Of these addresses, 85.1% received updates for a student or guardian (n=627) or matched an address in the directory or an address received from batch tracing (n=724). The remainder could not be located during IT-1.

At the student level, IT-1 resulted in an updated mailing address for 868 cases [480 updated+388 no change, or 87.5% of students) and 839 guardians (n=991, 84.7% of guardian 1) (Table 9). If the same or new addresses were identified in 2 of 3 databases used for intensive tracing, then the cases were designated as confirmed or updated, respectively. Nearly half of the students (n=480, 48.0%) and a quarter of guardians (n=241, 24.3%, guardian 1) had updated addresses. For another 39.1% of students and 60.3% of guardians, IT-1 returned the same address in the directory and no contact information for 12.5% of students and 15.3% of guardians.

IT-1 included searches for guardian 2 when no reliable contact information was found for student or guardian 1. For most (72%, 13/18) of the guardian2 cases, IT-1 returned no contact information.

As with batch tracing, an address was considered updated when IT-1 returned a new or additional mailing address information. When the address format varied between batch tracing data delivery and IT-1 data delivery, the outcome code did not reflect an update.

#### 4.2.3 Intensive Tracing Part 2 (IT-2)

Table 10 displays detailed results from IT-2 phone calls. Of the 490 students designated for IT-2, contact information (student or guardian 1) was confirmed via phone for 136 cases (27.8%). Refusals included 62 students and 44 guardians or another household member (12.7% and 9.0%, respectively). Almost half of cases were not reached by phone (n=238, 48.6%) and 10 (n=2%) were unable to respond for other reasons (e.g., deceased, institutionalized).

Among the 136 cases with successful phone confirmation of contact information either student or guardian, students were more likely to report the address was correct (94) or provide an update (29). All guardians reached indicated that their address matched the one in the directory (n=13).

Among the 238 cases not reached by phone, a valid phone number was not located for 25 cases (10.5%; neither the student nor guardian). The remaining 213 cases had a working phone number, but the tracer did not reach them after multiple attempts.

Although telephone outreach did not yield a large number of cases with updated contacted information, the process did serve to confirm some contact information, and provided an opportunity to introduce the possibility of participating in a new cohort for 9/11-related research and answer questions using information provided in the project's FAQs.

#### 4.3 Aim 1c Assess Level of Effort

This section relates the level of effort and cost of tracing and locating activities to the outcomes available for assessing cost and yield. Table 11 summarizes the results of the tracing and the follow-up mailings. Following batch tracing, DOHMH mailed brochures targeted to the 1,001 students using 1,797 addresses available for the students, and their guardian 1 and guardian 2. The mailing yielded 13 completed responses from former students (0.7%) (9 via postal mail and 4 via email), 203 brochures returned to DOHMH as un-deliverable (11.3%), and 22 with forwarding address (1.2%).

Following IT-1, 1,460 brochures were mailed to the 989 non-respondents of the previous mailing that included updated addresses from IT-1. This resulted in an additional 13 responses, 127 (8.7%) brochures returned as un-deliverable, and 22 with a forwarding address.

Following IT-2, 1504 brochures were to separate addresses representing 979 students. These consisted of mailings to addresses that had been updated by IT-2 (N=250), 419 for the ones not reached in IT-2, and 835 to those not included in IT-2 sample. This third mailing yielded 7 more responses from former students, 78 brochures returned as un-deliverable, and 30 with forwarding addresses.

After mailing a total of 4,761 brochures over the course of nearly one year, including multiple mailings to presumed valid and updated address across three rounds of mailings, the study received a total of 33 responses from former students or a 0.7% return for the mailed brochures. All but 2 responses expressed an interest in a new cohort. It was the case that across each stage of locating and tracing there was a notable decline in the percentage of brochures returned as un-deliverable (11.3% after batch tracing, 8.7% after IT-1, and 5.2% after IT-2), indicating that each stage of intensive tracing improved the pool of addresses received from batch tracing. Each subsequent mailing after the first mailing also yielded additional responses.

Overall, there was a 9.4% undeliverable rate (449 un-deliverables received/4761 brochures mailed). However, the proportion of addresses that resulted in undeliverable was 15.5% given that some updated addresses remained undeliverable (279/1797).

It appears that batch tracing alone may not have been sufficient, given that IT-1 still resulted in the updating of 39% of the addresses submitted, most of which had also been updated in batch tracing. IT-2 performed on a limited sample did result in 10% of the address being updated among those who cooperated and were receptive to confirming or updating their address (25 update/136 cooperate).

We related the cost of tracing to update contact information to the yield of the effort in several ways. First, the cost per response was about \$9,000 (not including the cost of mailing ~4,700 informational brochures) (\$300,000 tracing contract cost divided by 33 responses). Second, the cost of tracing per student was \$430 for students with no undeliverable address (\$300,000/770). Third, the cost was \$1,775 for each confirmed valid addresses, (\$300,000/(136 confirmed/updated by phone+ 33 replies).

#### 4.4 Aim 2a. Contact, Cooperation, and Response Rates

Based on 489 students in IT-2, the contact rate was 49% (240 connected calls/489) with a cooperation rate of 56% (136/240). The overall responses to the mailing of the informational brochures was 3.3% (33/1001). When the denominator is limited to students with presumed valid addresses, the response rate was slightly higher (4.4%). However, there were only one response from former students whose addresses were validated in IT-2. Another finding of note is that there were 33 (0.8%) responses from the 4,352 mailed brochures, excluding those returned undeliverable.

#### 4.5 Aim2b. Representativeness

Aim 2 was to assess the representativeness of presumed valid addresses and completed responses to the informational brochure by demographics and zone. This is important as it has a direct relationship to the likelihood of contacting and inviting former students to join a new children's cohort. These findings are presented in Table 12. The proportion of former students with presumed valid addresses was similar for former students from exposed and unexposed zones (68.9% and 71.7%, respectively) and by gender for both exposed and unexposed (e.g. exposed 69.2% vs. 68.5% and unexposed 72.3% vs. 70.9%). In contrast, former students whose current age was 32 to 35 had the highest proportion of presumed valid addresses compared to all other age groups for both exposed and unexposed (74.8% and 76.7%, respectively). Asian former students had the highest level of presumed active address (75.4%) followed by Whites (71.5%), Hispanics (68.6%) and Blacks (60.9%). Black students in the unexposed zone had the lowest level of presumed valid addresses (59.4%). The group with the highest proportion of presumed valid addresses were Asians from the unexposed zone at 78% (see Table 12).

A much higher percentage of former students from the exposed zone responded to the brochure compared to those in the un-exposed zone (5.2% (26/500) vs 1.4% (7/501), respectively (Table 13). Response also varied widely by race/ethnicity/gender, ranging from no black- male former students responding from either exposed or unexposed zones to 15% for white females and 10% black females from exposed zones. Several race/sex groups from the un-exposed zones had no response to the brochure including white-males, black-males and females, Hispanic males, and Asian females. The number of former students in these groups was 284 or 57% of the total un-exposed zone sample.

## 5.0 DISCUSSION

This study was initially conceptualized to assess whether it was feasible to build a new cohort of children twenty years after the 9/11 disaster. It was envisioned that a new cohort would potentially expand research among those who were children on 9/11/2001 beyond what is currently available through the World Trade Center Health Registry by including geographically broader eligibility areas and an unexposed comparison group. A decision to create a children's cohort two decades later required the resolution of two unknowns. First, was there a way to find and contact individuals who are potentially eligible for the cohort. Second, once contacted would the children, now adults, express interest in joining a new cohort for 9/11 research. Thus, the feasibility study was designed to provide an

estimate of what proportion of the total population by zone and by demographic characteristics may be locatable, contactable, potentially eligible and express an interest in being in a new children's cohort.

Aim 1 addressed whether: a) DOE directory data could be used as a basis for finding and contacting students by evaluating its completeness and accuracy; and b) state-of-the-art tracing and locating activities could confirm and update address information suitable for mailing informational brochures; as well as c) the cost and yield of conducting tracing and locating. Aim 2 assessed the level of interest in joining a children's cohort and the extent to which those who responded were representative of the target population, an important consideration in the formation of an epidemiologically useful children's cohort for research. The level of response to an informational brochure that was mailed to former students in the feasibility study sample is also key to informing a decision on forming a children's cohort, taking into account power calculation of estimated cohort sizes required to epidemiologically measure the association of 9/11 exposure with selected health outcomes.

The availability of contact information on former students from the New York City Department of Education provided a basis for considering the feasibility of creating a cohort of children exposed to 9/11. DOE through an exception to the regulation governing the identifiable information it possessed and a DUA with DOHMH provided individual records of addresses and other information gathered from parents and guardians at the beginning of each school year. Based on criteria provided by DOHMH, DOE formed a data file of 153,000 records from which 1,002 records were randomly selected for the feasibility study.

The feasibility study had two interlocking activities that included evaluating the degree to which tracing and locating could obtain updated addresses that when used for mailing to former students would reach the intended recipient. The process involved conducting three stages of tracing and locating with each stage followed by a mailing to the valid addresses available, including those that had been confirmed or updated. It should be noted that the address information provided by DOE was that of the parent/guardians at the time their child was enrolled in school and that the tracing and locating process used both the student name and the parental address in their search.

One key finding is that the DOE database had sufficient completeness for tracing and locating activities. We found that less than 10 percent of the information for key variables was missing and only one student record out of 1,002 had insufficient information for tracing. A second key finding is that state-of-the-art tracing and locating was able to update or confirm a large proportion of the contact information provided by DOE. The initial "batch" tracing, resulted in the updating of 85% of the student records; importantly, 58% of the updates obtained were student specific information, including telephone numbers. In the second stage of locating and tracing, 48% of student addresses were updated and 39% were confirmed, and the third stage provided additional address updates and confirmations.

Although brochures were mailed to presumed valid addresses following each round of tracing, and a large proportion of students (70%) had no brochures returned undeliverable at any of the available addresses for them, it is difficult to determine what proportion of mailings were received by the intended recipient. While many of the presumed active address may have been the correct addresses for the sample members, overall only 15 percent (n=156) of students had confirmed valid addresses, *i.e.* confirmed via the 33 student completed responses or the IT-2 phone calls that were made to half the feasibility sample which yielded 123 address confirmations or updates. We also found that each round of tracing improved the pool of addresses as evidenced by additional address updates and confirmations and the reduced number of undeliverable addresses.

Our findings of a low level (3%) of expressed interest in joining a new cohort for research after multiple rounds of brochure mailings combined with differential response by zone (exposed/unexposed) and demographics (race/ethnicity/sex) suggest that the formation of a scientifically valid cohort would likely require more extensive tracing and outreach efforts than those implemented in this feasibility study. An

important benchmark of the potential level of interest in joining a children's cohort is that there was only one response from students whose addresses were directly confirmed by phone (1/136, 0.7%). In addition, with as many as three information brochures mailed to them over a period of nearly a year, and assuming the addresses were correct and delivered and that the 85% of updated student addresses were current address, then a large proportion of former students may have declined repeated requests for a response.

While the response from students from schools in the exposed zone suggested a level of interest potentially sufficient for creating a cohort to study more common types of outcomes such as coronary heart disease, the potential ability to study common health conditions would limit the scientific value of a new cohort. Additionally, interest in joining a new cohort was highly biased by gender and race/ethnicity. For example, white females composed 42% of the completed responses and no responses were received from Black or Hispanic males. The very low level of interest (1.4%) among students in schools in the non-exposed zone suggests that it would be difficult to gather sufficient representation from the unexposed comparison zone. Other 9/11 studies of children have had difficulty forming non-exposed comparison groups, which required much additional time and effort (e.g., Dr. Trasande's study funded by NIOSH).

These findings suggest that the representativeness of a new cohort established two decades after 9/11 might be less than that of the sub-cohort of children in the WTCHR established in 2003-04. Among exposed children, the WTCHR had an overall coverage of 30% for children younger than 15 years and ~12% for children older than 15 years. The distribution of the Registry's children by race/ethnic group was ~44% non-Hispanic white, 9% non-Hispanic black, 19% Hispanic and 19% Asian and reflective of estimated 2000 census data in the Registry exposed area, especially for minority children area (black 4.2%, Hispanic 8.5%, Asian 18.3% white 69%).

The contact and cooperation rates for the feasibility study nearly two decades after 9/11 were as expected lower than those for the WTCHR when which was formed 2-3 years after 9/11. For instance, the Registry achieved an overall contact rate of 60% from over 200 lists of likely eligible persons (n=197,000) in comparison to the feasibility study contact rate of 49%. The cooperation rate for the Registry was 88% versus 56% for the feasibility study. The lower contact and cooperation rates may reflect a waning interest in requests related to 9/11 and other secular trends of less willingness of people in general to respond to phone calls. These findings reinforce the notion that formation of a scientifically valid cohort would likely require more extensive tracing and outreach efforts than those implemented in this feasibility study.

Based on the feasibility study findings, tracing and contacting the entire DOE file of former students would be expected to yield a sample of ~4,800 former students (3% overall response) interested in joining a new cohort across exposed and unexposed zones. The new sub-cohort of exposed children would possibly be slightly larger than the Registry's sub-cohort of exposed children. The monetary and management cost for locating and tracing 1,001 former DOE students was substantial with a tracing contract of \$300,000 and approximately 2 to 3 full-time equivalents of Registry staff devoted to management, tracking the incoming data and conducting the mailings and data analysis. Using the methods employed in the feasibility study to trace and contact the entire DOE file of 153,000 former students may cost ~\$46,000,000 (153 x \$300,000). Also, based on our findings, employing batch tracing alone to reduce costs may not be sufficient for updating addresses and contacting the sample. Another issue is that the highest undeliverable rates were among minority former students which suggests that additional tracing and locating efforts beyond what was done for the feasibility study would be required to reduce enrollment bias. If the enrollment for a children's cohort were to involve telephone contact, the results from IT-2 suggest that there would need to be substantial effort in acquiring telephone numbers and that such a study may encounter high refusal rates.

## 5.1 Strengths and Limitations

**Strengths:** First, DOHMH was able to successfully address a number of regulatory and administrative hurdles to acquire a contact information from DOE on a sample of former students who had attended schools in the designated target areas which likely comprised a large proportion of children who lived there as well. Second, the WTC Health Registry had the capability of building the infrastructure for managing this project including modifying an existing enrollee interface tool (R-tools) that was used for managing the DOE data and tracking the address information and events that occurred in the process. Third, the informational brochure was designed with input from stakeholders, including the Survivor Steering Committee and young adults in the Registry.

**Limitations:** First, the feasibility study was not designed to assess the interest of the guardians in participating together with their children in a new cohort for 9/11-related research. Second, for those former students who did express an interest in a new cohort, it is not known if they would enroll in a cohort after receiving informed consent and participate in surveys and research studies. Third, we don't know why nearly all former students in the sample did not respond (e.g., did not receive the brochure or were not interested). Fourth, due to the impracticality of identifying such children, the feasibility study did not include children who lived in the exposure or comparison zones who were enrolled in schools outside of these zones. A potential new children's cohort would likely consider including such children. Although DOE directory data does not include those in private schools or home-schooled, any future protocol for a new children's cohort might potentially include strategies such as media campaigns to encourage enrollment among children who were not included in the DOE public school directory list. We estimate that not including private school students and those who attended school outside the exposed and unexposed zones would under-cover about 25% of the residents under 18 years of age on 9/11/2001 living in those zones.

An additional limitation is gauging the level of interest solely by responses to a mailed informational brochure, even one designed with input from young adults and other stakeholders. Some former students may respond better to other forms of outreach such as social media and additional phone calls which were outside the scope of this feasibility study. However, the low rate of obtaining up-to-date telephone numbers, especially cell phone numbers, for former students limited our ability to contact them by phone. In addition, when people were contacted by phone, a large percentage did not answer or refused to talk when reached.

## 6.0 CONCLUSIONS

The findings will help inform NIOSH's next steps related to the formation of a new cohort for 9/11-related research among people exposed as children. This includes findings related to the completeness of the DOE directory data, the ability of state-of-the art tracing and locating activities to obtain updated and confirmed addresses, contact and cooperation rates, indicators of the level of interest in a new cohort by zone and by demographics, and the costs of tracing and contacting activities.

## 7.0 REFERENCES

Brackbill RM, Hadler JL, DiGrande L, Ekenga CC, Farfel MR, Friedman S, Perlman SE, Stellman SD, Walker DJ, Wu D, Yu S, Thorpe L. , 2009, Asthma and posttraumatic stress symptoms 5 to 6 years following exposure to the World Trade Center Terrorist Attack. *JAMA*;302(5):502-516.

Jordan HT, Miller-Archie SA, Cone JE, Morabia A, Stellman SD. 2011, Heart disease among adults exposed to the September 11, 2001 World Trade Center disaster: Results from the World Trade Center Health Registry. *Preventive Medicine*;53:370-376.

Li J, Brackbill RM, Liao TS, Qiao B, Cone JE, Farfel MR, Hadler J, Kahn AR, Konty KJ, Stayner LT, Stellman SD, 2016, Ten-year cancer incidence in rescue/recovery workers and civilians exposed to the September 11, 2001 terrorist attacks on the World Trade Center. *American Journal of Industrial Medicine*:59:709-721.

Thomas PA, Brackbill R, Thalji L, DiGrande L, Campolucci S, Thorpe L, Henning K., 2008, Respiratory and other health effects reported in children exposed to the World Trade Center Disaster of 11 September 2001, *Environmental Health Perspectives* 116: 1383-1390.



Table 1. Exposed Zone Neighborhood Tabulation Areas (NTAs) and Corresponding Sociodemographic Characteristics (2000 US Census).

NTA	Race			Ethnicity	Language			Native born (%)	Median household income	% above poverty	% college educated (highest degree)	N ≤ 18 years
	White (%)	Black (%)	Asian (%)	Hispanic (%)	English (%)	Spanish (%)	Chinese (%)					
Lower Manhattan	61.2%	6.5%	18.5%	8.9%	56.4%	7.8%	12.2%	68.6%	\$72,645	85.5%	37.6%	1843
Brooklyn Heights	82.9%	6.9%	4.7%	7.7%	79.5%	7.3%	3.7%	86.8%	\$74,558	84.6%	27.8%	2244
Chinatown	20.2%	7.4%	60.7%	17.7%	24.6%	13.3%	57.0%	42.1%	\$24,314	70.1%	10.9%	8850
Downtown Brooklyn	76.9%	9.6%	4.0%	8.6%	68.3%	11.9%	4.3%	80.8%	\$93,216	87.2%	41.4%	6179
East Village	68.9%	7.2%	13.1%	13.2%	61.9%	11.8%	8.8%	71.2%	\$45,841	81.1%	34.6%	2725
Lower East Side	39.6%	10.9%	24.1%	39.4%	35.8%	34.4%	21.2%	66.3%	\$24,267	69.6%	14.2%	15765
SoHo-Tribeca	62.8%	4.5%	24.6%	8.6%	61.3%	5.7%	20.8%	70.7%	\$70,097	84.7%	31.1%	4734
West Village	86.7%	2.6%	6.0%	5.5%	77.2%	5.3%	3.5%	81.8%	\$72,721	92.0%	39.2%	4417

Table 2. Unexposed Zone Neighborhood Tabulation Areas (NTAs) and Corresponding Sociodemographic Characteristics (2000 US Census).

NTA	Race			Ethnicity	Language			Native born (%)	Median household income	% above poverty	% college educated (highest degree)	N ≤ 18 years
	White (%)	Black (%)	Asian (%)	Hispanic (%)	English (%)	Spanish (%)	Chinese (%)					
Upper West Side	75.7%	9.6%	4.7%	16.0%	69.7%	14.4%	3.2%	79.0%	\$68,498	89.2%	31.0%	20,152
Flushing	30.4%	7.5%	49.5%	20.4%	25.3%	18.5%	41.0%	37.2%	\$34,962	80.3%	16.9%	14,204
Sunset Park East	33.6%	2.7%	34.2%	41.4%	18.4%	39.3%	26.3%	45.3%	\$27,322	68.7%	7.0%	19,324
Sunset Park West	40.2%	5.5%	9.6%	69.9%	20.9%	63.2%	8.6%	58.7%	\$31,598	73.5%	7.5%	14,797

Table 3. Exposed Zone Neighborhoods, Zipcodes and School Districts included in Neighborhoods

Neighborhood	School districts	Zip codes
Lower Manhattan	2	10004, 10005, 10006, 10007, 10038, 10280, 10282
Brooklyn Heights	13	11201
Chinatown	1, 2	10002, 10038, 10013
Downtown Brooklyn	13, 15	11201
East Village	1	10003, 10009
Lower East Side	1	10002, 10009
SoHo-Tribeca	2	10012, 10013, 10007
West Village	2	10014, 10011, 10003

Table 4. Unexposed Zone Neighborhoods, School Districts, and Zipcodes.

Neighborhood	School district	Zip codes
Upper West Side	3	10024, 10025
Flushing	25	11355, 11354
Sunset Park East	20	11232
Sunset Park West	20	11220

Table 5. Population by NTA for Exposed Zone Neighborhoods, 2000 U.S. Census.

	Non-Hispanic White				Non-Hispanic Black				Non-Hispanic Asian				Hispanic				Total N	NTA sum N
	Male		Female		Male		Female		Male		Female		Male		Female			
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%		
<i>Lower Manhattan</i>																		
0-4 years old	299	16.2%	295	16.0%	15	0.8%	13	0.7%	68	3.7%	76	4.1%	30	1.3%	34	1.8%	830	1,843
5-9 years old	181	9.8%	177	9.6%	14	0.8%	7	0.4%	44	2.4%	55	3.0%	25	1.1%	24	1.3%	527	
10-13 years old	100	5.4%	89	4.8%	5	0.3%	8	0.4%	26	1.4%	25	1.4%	12	0.5%	11	0.6%	276	
14-17 years old	59	3.2%	72	3.9%	8	0.4%	8	0.4%	14	0.8%	29	1.6%	9	0.4%	11	0.6%	210	
<i>Brooklyn Heights</i>																		
0-4 years old	353	15.7%	339	15.1%	17	0.8%	22	1.0%	15	0.7%	21	0.9%	37	1.6%	36	1.6%	840	2,244
5-9 years old	256	11.4%	239	10.7%	10	0.4%	22	1.0%	18	0.8%	10	0.4%	23	1.0%	20	0.9%	598	
10-13 years old	151	6.7%	170	7.6%	13	0.6%	13	0.6%	2	0.1%	15	0.7%	16	0.7%	16	0.7%	396	
14-17 years old	182	8.1%	143	6.4%	18	0.8%	16	0.7%	4	0.2%	10	0.4%	24	1.1%	13	0.6%	410	
<i>Chinatown</i>																		
0-4 years old	43	0.5%	30	0.3%	46	0.5%	43	0.5%	812	9.2%	737	8.3%	181	2.0%	187	2.1%	2,079	8,850
5-9 years old	31	0.4%	33	0.4%	48	0.5%	54	0.6%	860	9.7%	724	8.2%	260	2.9%	247	2.8%	2,257	
10-13 years old	29	0.3%	31	0.4%	40	0.5%	53	0.6%	863	9.8%	758	8.6%	215	2.4%	241	2.7%	2,230	
14-17 years old	35	0.4%	21	0.2%	53	0.6%	45	0.5%	959	10.8%	743	8.4%	218	2.5%	210	2.4%	2,284	
<i>Downtown Brooklyn</i>																		
0-4 years old	190	3.1%	197	3.2%	316	5.1%	326	5.3%	39	0.6%	36	0.6%	291	4.7%	255	4.1%	1,650	6,179
5-9 years old	153	2.5%	113	1.8%	381	6.2%	363	5.9%	19	0.3%	29	0.5%	376	6.1%	298	4.8%	1,732	
10-13 years old	73	1.2%	79	1.3%	338	5.5%	270	4.4%	26	0.4%	20	0.3%	293	4.7%	268	4.3%	1,367	
14-17 years old	95	1.5%	70	1.1%	329	5.3%	261	4.2%	22	0.4%	22	0.4%	340	5.5%	291	4.7%	1,430	
<i>East Village</i>																		
0-4 years old	185	6.8%	197	7.2%	18	0.7%	28	1.0%	84	3.1%	80	2.9%	121	4.4%	104	3.8%	817	2,725
5-9 years old	138	5.1%	141	5.2%	20	0.7%	33	1.2%	72	2.6%	59	2.2%	131	4.8%	156	5.7%	750	
10-13 years old	115	4.2%	122	4.5%	20	0.7%	22	0.8%	57	2.1%	50	1.8%	99	3.6%	99	3.6%	584	
14-17 years old	94	3.4%	91	3.3%	26	1.0%	20	0.7%	62	2.3%	54	2.0%	129	4.7%	98	3.6%	574	
<i>Lower East Side</i>																		

0-4 years old	253	1.6%	210	1.3%	267	1.7%	246	1.6%	414	2.6%	342	2.2%	1,058	6.7%	1,027	6.5%	3,817	15,765
5-9 years old	187	1.2%	168	1.1%	363	2.3%	288	1.8%	451	2.9%	366	2.3%	1,230	7.8%	1,232	7.8%	4,285	
10-13 years old	130	0.8%	124	0.8%	288	1.8%	298	1.9%	417	2.6%	363	2.3%	1,170	7.4%	1,033	6.6%	3,823	
14-17 years old	114	0.7%	109	0.7%	246	1.6%	281	1.8%	441	2.8%	373	2.4%	1,177	7.5%	1,099	7.0%	3,840	
<i>SoHo-Tribeca</i>																		
0-4 years old	443	2.8%	438	2.8%	21	0.1%	14	0.1%	236	1.5%	195	1.2%	54	0.3%	61	0.4%	1,462	4,734
5-9 years old	352	2.2%	375	2.4%	10	0.1%	13	0.1%	261	1.7%	225	1.4%	49	0.3%	62	0.4%	1,347	
10-13 years old	282	1.8%	259	1.6%	8	0.1%	12	0.1%	198	1.3%	155	1.0%	57	0.4%	36	0.2%	1,007	
14-17 years old	191	1.2%	192	1.2%	14	0.1%	23	0.1%	210	1.3%	184	1.2%	39	0.2%	65	0.4%	918	
<i>West Village</i>																		
0-4 years old	707	16	639	14.5%	13	0.3%	11	0.2%	26	0.6%	58	1.3%	55	1.2%	59	1.3%	1,568	4,417
5-9 years old	486	11	472	10.7%	16	0.4%	16	0.4%	30	0.7%	33	0.7%	50	1.1%	58	1.3%	1,161	
10-13 years old	366	8.3	347	7.9%	12	0.3%	9	0.2%	25	0.6%	14	0.3%	37	0.8%	41	0.9%	851	
14-17 years old	320	7.2	339	7.7%	15	0.3%	23	0.5%	23	0.5%	26	0.6%	52	1.2%	39	0.9%	837	

Table 6. Sampling Proportions for Exposed Zone for the Feasibility Study, Calculated from Population Values in Table 5.

Age group	Non-Hispanic White				Non-Hispanic Black				Non-Hispanic Asian				Hispanic				Overall
	Male		Female		Male		Female		Male		Female		Male		Female		Total
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
0-4	2,473	5.3%	2,345	5.0%	713	1.5%	703	1.5%	1,694	3.6%	1,545	3.3%	1,827	3.9%	1,763	3.8%	13,063
5-9	1,784	3.8%	1,718	3.7%	862	1.8%	796	1.7%	1,755	3.8%	1,501	3.2%	2,144	4.6%	2,097	4.5%	12,657
10-13	1,246	2.7%	1,221	2.6%	724	1.5%	685	1.5%	1,614	3.5%	1,400	3.0%	1,899	4.1%	1,745	3.7%	10,534
14-17	1,090	2.3%	1,037	2.2%	709	1.5%	677	1.4%	1,735	3.7%	1,441	3.1%	1,988	4.3%	1,826	3.9%	10,503
																	46,757

Table 7. Sample Size by Gender, Age, Race/ethnicity Used for Sampling Based on DOE Directory Data for Exposed and Unexposed Zones.

Age group	Non-Hispanic White				Non-Hispanic Black				Non-Hispanic Asian				Hispanic			
	Male		Female		Male		Female		Male		Female		Male		Female	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
0-4	26	5.3%	25	5.0%	8	1.5%	8	1.5%	18	3.6%	17	3.3%	20	3.9%	19	3.8%
5-9	19	3.8%	18	3.7%	9	1.8%	9	1.7%	19	3.8%	16	3.2%	23	4.6%	22	4.5%
10-13	13	2.7%	13	2.6%	8	1.6%	7	1.5%	17	3.5%	15	3.0%	20	4.1%	19	3.7%
14-17	12	2.3%	11	2.2%	8	1.5%	7	1.5%	19	3.7%	15	3.1%	21	4.3%	20	3.9%



Table 8. Summary of Update Received as a Result of Batch Tracing by Student (based on information provided by tracing vendor)

Total updated	855	85.3%
Student (S) only	139	13.9%
Guardian 1 (G1) only	143	14.3%
Guardian 2 (G2) only	39	3.9%
S, G1, G2	156	15.6%
S, G1	227	22.7%
S, G2	59	5.9%
G1 and G2	92	9.2%
No update	147	14.7%
Total	1,002	100.0%

Table 9. Intensive Tracing Part 1 Results  
(based on information provided by tracing vendor)

Result	N	%
<i>Total students</i>	992	100.0%
Students updated	480	48.4%
No change in contact information	388	39.1%
Nothing found	124	12.5%
<i>Total Guardian 1 (G1)</i>	991	100.0%
G1 updated	241	24.3%
No change in contact information	598	60.3%
Nothing found	152	15.3%
<i>Total Guardian 2 (G2)</i>	18	100.0%
G2 address updated	1	5.6%
No change in address information	4	22.2%
Nothing found	13	73.2%

Table 10. Intensive Tracing Part 2 Results  
 (based on information provided by tracing  
 vendor).

Event	N	%
Student or guardian updated or confirmed address	136	27.8%
Student or guardian no contact	238	48.6%
Student refused to confirm address	62	12.7%
Guardian refused to confirm address	32	6.5%
Other household member refused	12	2.4%
Student unable to participate/unavailable during study (institutionalized, incarcerated, out of country, mentally incapable, deceased)	10	2.0%
<b>Total</b>	<b>490</b>	<b>100.0%</b>

Table 11. Results from Three Stages of Tracing and Locating and Three Subsequent Mailings of Brochures.

Stage	Number searched	Updated with corrected addresses or phone numbers		Number mailed after each stage	Returned with response (either mail or email)		Returned undeliverable		Returned with forwarding address	
	N	N	%	N	N	%	N	%	N	%
1. Batch tracing	2,522	1,531	61%	1,797 <sup>a</sup>	13	0.7%	204	11.3%	22	1.2%
2. IT-1	1,586	627	39%	1,460 <sup>b</sup>	13	0.8%	127	8.7%	22	1.5%
3. IT-2	240 (available phone numbers)	25	10%	1,504 <sup>c</sup>	7	0.4%	78	5.2%	30	2.0%

a. 1,531 updated + 266 address found in batch tracing

b. 627 updated + 724 confirmed + 109 not found but not returned as undeliverable in first mailing

c. 250 not undeliverable or returned with forwarding address of 489 in IT-2 sample + 419 determined to be valid from previous mailings + 835 to 490 not included in IT-2

Table 12. Undeliverable and Completed Returned Brochure by Gender, Race, and Age for Exposed and Unexposed.

	Former students (exposed and unexposed)		Exposed				Unexposed				Completed Returned Brochure	
			Undeliverable (Batch, IT-1, IT-2)		Undeliverable (Batch, IT-1, IT-2)							
	N	%	Yes N	No %	No N	Yes %	Yes N	No %	No N	Yes %	N	%
<i>All students</i>	1,002	100.0%	156	85.7%	345	14.3%	142	29.7%	359	70.3%	33	3.3%
<i>Gender</i>												
Male	520	51.9%	80	30.8	180	69.2	72	27.7	188	72.3	9	1.7%
Female	482	48.1%	76	31.5	165	68.5	70	29.1	171	70.9	24	5.0%
<i>Current age</i>												
18-22	250	25.3%	39	31.7	84	68.3	35	27.6	92	72.4	7	2.8%
23-27	279	28.2%	48	33.8	94	66.2	42	30.7	95	69.3	12	4.8%
28-31	217	21.9%	36	33.6	71	66.4	35	31.8	75	68.2	6	2.8%
32-35	243	24.6%	31	25.2	92	74.8	28	23.3	92	76.7	8	3.2%
<i>Race/ethnicity</i>												
White	274	27.3%	40	29.2	97	70.8	38	27.7	99	72.3	20	7.3%
Black	128	12.8%	24	37.5	40	62.5	26	40.6	38	59.4	3	2.3%
Asian	272	27.2%	37	27.2	99	62.8	30	22.1	106	77.9	6	2.2%
Hispanic	328	32.7%	55	33.5	109	66.5	48	29.3	116	70.7	4	1.2%

Table 13. Completed Returned Brochure by Race/Gender Groups.

Race/gender group	Exposed				Unexposed			
	Returned		Not Returned		Returned		Not Returned	
	N	%	N	%	N	%	N	%
White/male	6	8.6%	66	91.4%	0	0.0%	70	100.0%
White/female	10	14.9%	57	85.1%	4	6.0%	63	94.0%
Black/male	0	0.0%	33	100.0%	0	0.0%	33	100.0%
Black/female	3	9.7%	28	90.3%	0	0.0%	31	100.0%
Hispanic/male	0	0.0%	71	97.3%	0	0.0%	72	98.6%
Hispanic/female	2	2.5%	60	95.2%	2	2.5%	63	100.0%
Asian/male	2	2.7%	84	100.0%	1	1.6%	84	100.0%
Asian/female	3	4.8%	78	97.5%	0	0.0%	78	97.5%
Total	26	5.2%	477	94.8%	7	1.4%	494	98.6%

