

Cancer in World Trade Center Responders: Findings From Multiple Cohorts and Options for Future Study

Paolo Boffetta,^{1,*} Rachel Zeig-Owens,² Sylvan Wallenstein,³ Jiehui Li,⁴ Robert Brackbill,⁴ James Cone,⁴ Mark Farfel,⁴ William Holden,¹ Roberto Lucchini,³ Mayris P. Webber,² David Prezant,^{2,5} and Steven D. Stellman^{4,6}

Background Three longitudinal studies of cancer incidence in varied populations of World Trade Center responders have been conducted.

Methods We compared the design and results of the three studies.

Results Separate analyses of these cohorts revealed excess cancer incidence in responders for all cancers combined and for cancers of the thyroid and prostate. Methodological dissimilarities included recruitment strategies, source of cohort members, demographic characteristics, overlap between cohorts, assessment of WTC and other occupational exposures and confounders, methods and duration of follow-up, approaches for statistical analysis, and latency analyses.

Conclusions The presence of three cohorts strengthens the effort of identifying and quantifying the cancer risk; the heterogeneity in design might increase sensitivity to the identification of cancers potentially associated with exposure. The presence and magnitude of an increased cancer risk remains to be fully elucidated. Continued long-term follow up with minimal longitudinal dropout is crucial to achieve this goal. Am. J. Ind. Med.

© 2015 Wiley Periodicals, Inc.

KEY WORDS: *World Trade Center; cancer; epidemiology; cohort study; surveillance*

INTRODUCTION

The terrorist attacks of September 11, 2001 (9/11), that caused the destruction of the World Trade Center (WTC)

towers and numerous surrounding buildings resulted in intense exposures of tens of thousands of individuals to known or suspected carcinogens, leading to concern about development of excess cancers in this population [Lioy et al.,

¹Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York

²New York City Fire Department, New York, New York

³Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, New York

⁴New York City Department of Health and Mental Hygiene, New York, New York

⁵Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, New York

⁶Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York

Contract grant sponsor: CDC/NIOSH; Contract grant number: BAA 2011-Q-13340; Contract grant sponsor: Cooperative Agreement; Contract grant number: U50/ATU272750; Contract grant sponsor: National Center for Environmental Health (NCEH); Contract grant sponsor: Cooperative Agreement; Contract grant number: 1U50/OH009739; Contract grant sponsor: Centers for Disease Control and Prevention (CDC), National Institute for Occupational Safety and Health (NIOSH); Contract grant numbers: 200-2011-39377; 200-2002-0038; 5U10 OH008232; Contract grant sponsor: American Red Cross Liberty Fund, The September 11th Recovery Program, The Bear Stearns Charitable Foundation, The September 11th Fund, and The Robin Hood Foundation Relief Fund; Contract grant sponsor: Department and National Program of Cancer Registries of the CDC; Contract grant number: U58/DP000783.

*Correspondence to: Paolo Boffetta, MD, MPH, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1130, New York, NY 10029. E-mail: paolo.boffetta@mssm.edu

Accepted 7 December 2015

DOI 10.1002/ajim.22555. Published online in Wiley Online Library
(wileyonlinelibrary.com).

2002; McGee et al., 2003; Landrigan et al., 2004]. The World Trade Center Health Program, established in 2011 under the James L. Zadroga 9/11 Health & Compensation Act, has led to publicly funded monitoring and treatment services for a number of cancers. To date, three cohort studies have reported point estimates of total cancer rates 6–14% in excess of background rates, with 95% confidence intervals that overlapped with 1.00, but with statistically significant excess rates for selected cancer sites including thyroid and prostate cancer [Zeig-Owens et al., 2011; Li et al., 2012; Solan et al., 2013], which may be due at least in part to participation in the medical surveillance programs offered to the participants in the monitoring programs.

Despite noteworthy consistency among studies, there are several reasons to undertake a detailed examination of each study and compare results. First, all three cohorts continue to be monitored for cancer, and the findings may affect policy responses to future disasters, as well as continued 9/11-related public policy. Second, the studies differ with respect to recruitment, sources and demographic characteristics of cohort members, methods of assessment of WTC and other occupational exposures and confounders, methods of follow-up, and statistical analysis, including assessment of latency. Third, there is some overlap in the populations so that the same cancers have likely been reported in more than one study. Finally, although some cancers were consistently in excess, findings for some other cancers varied among the three studies.

Eligibility of specific cancers for treatment under the Zadroga Health and Compensation Act, as well as continuing public concerns about future cancer incidence as the exposed population ages, require periodic assessment of the knowledge base and of the ongoing studies on which policy decisions will continue to be made. Thus, it is essential to understand the similarities and differences among studies along with their individual strengths and weaknesses.

This report represents a systematic effort by investigators from the three published cohort studies to compare their methods and findings in order to lay the groundwork for forthcoming follow-up studies and understand any differences that may emerge. A second goal is to determine the feasibility of conducting parallel analyses with common methods and objectives or, in a best-case scenario, to pool data for a single comprehensive analysis that includes all known 9/11 responders.

METHODS

The three published studies from Fire Department of New York (FDNY) [Zeig-Owens et al., 2011], the World Trade Center Health Registry at the New York City Department of Health and Mental Hygiene (DOHMH) [Li et al., 2012], and the World Trade Center Health Consortium

(WTCHC) [Solan et al., 2013] were reviewed. The key features of their design are summarized in Table I. We focused on first responders to the WTC attacks, since two of the three cohort studies [Zeig-Owens et al., 2011; Solan et al., 2013] deal exclusively with this population. The major differences between the studies are enrollment strategies, heterogeneity of the populations, and the length of the follow-up period. The WTCHC is an open cohort (i.e., it continues to accrue new subjects), whereas the other two are closed (Table I). The WTCHC and DOHMH cohorts are racially and ethnically heterogeneous, whereas the FDNY study sample is mostly non-Hispanic white (94%) and entirely male firefighters (Table I).

At the time of the publication [Zeig-Owens et al., 2011], the FDNY study had not completed matches to some cancer registries, in particular New Jersey, so this study also relied on medical records for case ascertainment although fewer than 5% of the FDNY cases were obtained this way ($n=17$); the other two studies only included cancer cases obtained from matches to state cancer registries. For FDNY, similar results were obtained with or without inclusion of these 17 cases.

The results of the analysis on cancer incidence in the three cohorts are summarized in Table II. All three studies included cancer data until 2008 and showed modest elevations of standardized incidence ratio (SIR) for all cancers (Table II). The point estimates for the SIRs for all cancer sites combined were similar and ranged from 1.06 to 1.14 across studies, with 95% confidence intervals that overlapped with 1.00. The three studies showed similar magnitudes for the SIR for thyroid (SIR range 2.02–3.12) and prostate cancer (SIR range 1.23–1.49). In terms of dissimilar results, FDNY reported significant increases in melanoma (not restricted to melanoma of the skin), and non-Hodgkin lymphoma. In contrast, DOHMH reported elevated results for multiple myeloma. All three studies showed a reduced SIR for lung cancer (SIR range 0.42–0.65), statistically significant in two of the three studies.

Methodological Issues in the Comparison of the Studies

A number of methodological issues should be taken into consideration when comparing the findings of the three cohort studies.

Study populations

The WTCHC is an open cohort with ongoing recruitment. Recruitment for the FDNY and the DOHMH cohorts are based on a closed design, that is, it is limited to individuals who fulfilled the inclusion criteria at a given date (or within a given time period). Open enrollment in a cohort of volunteers has two contrasting sources of bias.

TABLE I. Overview of Three Published WTC Cancer Papers on Responders

	WTC Cohort		
	FDNY firefighters	DOHMH, World Trade Center Health Registry	World Trade Center Health Consortium
Reference	Zeig-Owens et al. [2011]	Li et al. [2012]	Solan et al. [2013]
Study population			
Study sample restricted to responders ^a (Total # in publication)	8,927 (N = 9,853)	21,850 (N = 55,778)	20,984 (N = 20,984)
Source of subjects	Employment roster	Lists, self-referral, and outreach ^b	Lists, self-referral, and outreach
Type of cohort	Closed	Closed	Open
Inclusion criteria for the cancer papers	Active (not retired) firefighters on 1/1/1996 with known WTC-exposure status and were less than age 60 on 9/11/2001	New York State residents on 9/11/2001 without pre-enrollment reportable cancer	No restrictions
Gender distribution (in the cancer papers)	100% male	81% male	85% male
Residence in New York State	100%	100%	89%
Age on 9/11/2001 (years)	Mean: 44.0 median: 43.7	Mean: 39.6 median: 39.0	Mean: 39.0 median: 38.0
Race/Ethnicity	94% non-Hispanic white, 3% black, and 3% Hispanic	66% non-Hispanic white, 11% black, 16% Hispanic	59% non-Hispanic white, 13% black, 25% Hispanic
Approximate sample overlap	32% with DOHMH <1% with WTCHC	13% with FDNY 19% with WTCHC	20% with DOHMH <1% with FDNY
Main occupational groups	Firefighters (100%)	Protective services (21%); firefighters (13%); construction (9%); sanitation (8%)	Protective services (42%); construction (24%)
Study follow-up period			
Start of person-time	9/11/2001 ^c	Date of enrollment (9/2003–11/2004) for early period; 1/1/2007 for later period.	6 months after enrollment
End of person-time	Censored at the earliest of first cancer diagnosis, death or 12/31/08. For participants living in a state that was not linked to a cancer registry, person-time was censored at last date of contact with FDNY (~4%).	Censored at the earliest of first cancer diagnosis, death, or 12/31/2008.	Censored at the earlier of death, or 12/31/2008.
Total person years	61,884	41,280	153,077
Average follow-up	6.9 years	2.0 years for later period	3.6 years
Cancer case identification			
Linkage with state cancer registries	Yes (NYS, PA, FL, NC, VA)	Yes (NYS, PA, NJ, CT, CA, FL, MA, NC, OH, TX, WA)	Yes (NYS, PA, NJ, CT); sensitivity analysis for FL, NC
Medical records used	Yes (<5% of cases)	No	No
WTC exposures			
Other	Time of first arrival to the WTC site	DOHMH exposure summary index (time of first arrival to the WTC site, duration on site, worked on the pile, being in the WTC site before noon on 9/11)	Dust cloud exposure, duration on site, location of work; and an exposure index based on dust, duration and location.
Main data analysis—SIR	SIR, adjusted for age, race/ethnicity and study period ^d	SIR, adjusted for age, sex, race/ethnicity and study period ^e	SIR, adjusted for age, sex, and race/ethnicity
Reference population	SEER13	NYS	NYS, NJ, and CT; and national data (SEER17) for PA
Cancer site for SIR	First primary site diagnosed between 9/12/2001 and 12/31/2008 (multiple cancers reported in Fig. 1 for all cancers as sensitivity analyses)	First primary site between enrollment and 12/31/2008	Multiple primary sites
Lag time	2 years delay in diagnosis date for cases that might have been detected by FDNY screening	5 years since 9/11	None, other than inherent in waiting 6 months from registration
Overall SIR for all sites combined	Yes	Yes	Yes
# cancer sites presented	15	23	19
Other data analysis—internal comparisons			
Primary analysis for exposures	Cox models and SIR ratios estimated using Poisson regression models	Cox model	Poisson model with offsets
Secondary analysis for exposures		Cox model adjusted for lag time	Cox model

^aOnly referred to exposed FDNY firefighters in FDNY cohort, rescue/recovery workers and volunteers in DOHMH cohort, and all responders in WTCHC cohort. Cohorts included in the analyses of the referenced papers might vary slightly.

^b[Murphy et al., 2007].

^cStart of post-9/11/2001 period. FDNY study included a pre-9/11/2001 period, not included in this table.

^dFDNY cohort restricted to males only.

^eDOHMH cohort restricted to known race/ethnicity.

First, later-arriving enrollees may differ in terms of health at baseline from earlier-arriving enrollees, for example, they may enroll because of symptoms of conditions possibly related to exposure. Second, open cohorts may suffer from

survivor bias, that is, potential cohort members who died from conditions of interest before they were able to enroll, and preferential enrollment of individuals with or without symptoms of the diseases of interest (related or unrelated to

TABLE II. Standardized Incidence Ratios by Cancer Site and Study Site

Cancer site	SEER code	# Observed	FDNY (N = 8,927)		DOHMH (N = 20,991) first primary cancers; 2007–2008 (later period) ^a		WTCHC (N = 20,984) multiple primary cancers; 2002–2008	
			first primary cancers; 2001–2008		# Observed	2007–2008 (later period) ^a	# Observed	multiple primary cancers; 2002–2008
			2001–2008	2007–2008				
All sites combined		263	1.10 (0.98–1.25)	223	1.14 (0.99–1.30)	302	1.06 (0.94–1.18)	
Oral cavity and pharynx	20010–20100	N/A	N/A	≤5	0.77 (0.25–1.80)	10	1.00 (0.48–1.84)	
Esophagus	21010	≤5	0.58 (0.15–2.32)	≤5	1.16 (0.24–3.39)	7	1.77 (0.71–3.65)	
Stomach	21020	8	2.24 (0.98–5.25)	≤5	0.91 (0.19–2.67)	7	1.33 (0.53–2.74)	
Colorectal ^b	21041–21052	21	1.52 (0.99–2.33)	21	1.24 (0.77–1.90)	25	0.97 (0.63–1.43)	
Liver	21071, 21072	N/A	N/A	6	1.22 (0.45–2.67)	6	0.86 (0.32–1.88)	
Pancreas	21100	≤5	0.78 (0.29–2.09)	≤5	0.70 (0.14–2.03)	≤5	0.49 (N/A)	
Lung	22030	9	0.42 (0.20–0.86)	13	0.65 (0.35–1.12)	18	0.62 (0.37–0.98)	
Soft tissue sarcoma	24000	N/A	N/A	N/A	N/A	≤5	1.58 (N/A)	
Skin melanoma ^c	25010	33	1.54 (1.08–2.18)	10	1.32 (0.63–2.43)	12	1.01 (0.52–1.77)	
Breast ^d	26000	N/A	N/A	18	1.39 (0.82–2.20)	11	0.74 (0.37–1.32)	
Prostate	28010	90	1.49 (1.20–1.85)	67	1.43 (1.11–1.82)	82	1.23 (0.98–1.53)	
Testis	28020	≤5	0.86 (0.36–2.06)	≤5	0.36 (0.01–1.98)	7	1.27 (0.51–2.62)	
Urinary bladder	29010	11	1.01 (0.56–1.83)	8	0.94 (0.41–1.85)	15	1.18 (0.66–1.94)	
Kidney and renal pelvis	29020	10	0.86 (0.46–1.60)	12	1.38 (0.71–2.41)	17	1.34 (0.78–2.14)	
Brain and other nervous system	31010, 31040	N/A	N/A	≤5	0.68 (0.08–2.45)	7	1.34 (0.54–2.77)	
Thyroid	32010	17	3.07 (1.86–5.08)	13	2.02 (1.07–3.45)	26	3.12 (2.04–4.57)	
Hodgkin lymphoma	33011, 33012	0	0 (N/A)	≤5	2.47 (0.67–6.31)	≤5	1.25 (N/A)	
Non-Hodgkin lymphoma	33041, 33042	21	1.58 (1.03–2.42)	11	1.15 (0.57–2.06)	13	0.85 (0.45–1.45)	
Multiple myeloma	34000	≤5	1.49 (0.56–3.97)	7	2.85 (1.15–5.88)	≤5	0.54 (N/A)	
Leukemia	35011–35043	9	1.40 (0.73–2.70)	6	1.25 (0.46–2.72)	≤5	0.53 (N/A)	

N/A, not available.

^aExcluded those who developed cancer or died in early period (2001–2006) and those with unknown or undefinable race/ethnicity.^bFDNY only reports colon cancer not colorectal.^cFDNY includes all melanoma but the other two sites only report melanoma of the skin.^dFor DOHMH, includes only females; for WTCHC, includes males also.

exposure) may occur. The effects of these biases may be assessed in sensitivity analyses.

The WTCHC and the DOHMH cohorts were based on voluntary participation of individuals identified through employer-generated lists and outreach activities, resulting in an unknown degree of self-selection, which may have been correlated with opportunity or severity of exposure. The FDNY cohort is based on employment rosters requiring routine physical check-up, and is not affected by this potential bias. This aspect of the design would not affect the validity of analyses within each cohort (although it might affect their precision) but would likely introduce bias of uncertain direction and magnitude in comparisons between cohorts.

A specific issue of the three WTC cohorts is the fact that the same individuals might be included in more than one program. Overlap is estimated at approximately 3,000 FDNY individuals who are also in the DOHMH cohort; 4,000 individuals in both the DOHMH and WTCHC; and

fewer than 100 individuals in both the FDNY and WTCHC cohorts. The overlap complicates comparisons and precludes any formal combination of the published results. However, when DOHMH removed FDNY from their analyses, similar results were obtained [Li et al., 2012].

Assessment of WTC exposures

When the WTC towers collapsed on the morning of 9/11 the surrounding area was immediately engulfed by tons of dust and debris [Lioy et al., 2002]. For nearly 3 months after 9/11, fires caused by the attack burned where the towers had stood [Lioy et al., 2002]. These events exposed rescue and recovery workers and others to many compounds such as glass, silica, cement, benzene, asbestos, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), cadmium, and polychlorinated furans and dioxins [Lioy et al., 2002; McKinney et al., 2002; Landrigan et al., 2004; Lioy and Georgopoulos, 2006; Lorber et al., 2007]. These

known exposures [OSHA, 2002 at: <https://www.osha.gov/nyc-disaster/summary.html>] were obtained from air and dust samples collected days or even weeks after 9/11 and were taken from locations throughout the disaster site. Individual and area sample results were obtained by the Occupational Safety and Health Administration [OSHA, 2014] and National Institute for Occupational Safety and Health [NIOSH, 2002]. Unfortunately, individual level monitoring was not available for use in any of the three studies.

Although workers from all three WTC cohorts worked at the site for days, and in many cases for months, there was no official documentation of the time spent working at the WTC sites. When it was documented, records were often incomplete and inconsistent [Banauch et al., 2002; Weakley et al., 2011]. As a result, WTC studies have characterized exposures using self-reported information from questionnaires. However, the information collected varied by WTC cohort. After reviewing the collected information, investigators from the three WTC cohorts established some common exposure variables, including exposure level on 9/11/2001 and tasks completed during the rescue/recovery work [Weakley et al., 2011]. The FDNY study included this common exposure variable in their first cancer study; however, it did not find an association; results based on the common exposure variable were not reported in the other two studies, which reported no or weak association with their exposure classifications.

The three cancer studies reviewed here analyzed all levels of WTC exposure combined (i.e., any WTC exposure), and compared rates in WTC exposed with rates in the general

(non-WTC exposed) population. However, some sub-analyses investigated the presence of an exposure gradient based on the study-specific WTC exposure categories [Zeig-Owens et al., 2011; Li et al., 2012; Solan et al., 2013]. Table III describes the way each cohort categorized WTC exposure in relation to the other studies. To highlight some differences, the exposure information obtained by FDNY was related to the time and day a firefighter first worked at the WTC site and did not specifically ask about exposure to the dust cloud on 9/11, although it is a reasonable assumption that FDNY members who worked at the WTC site during the morning of 9/11 were caught in the dust cloud. Two studies (DOHMH, WTCHC) used a specific question about being caught in the dust cloud to classify WTC exposure. Furthermore, both non-FDNY cohorts included personnel who worked in locations outside of the WTC pile. Although nearly all FDNY firefighters worked on the WTC pile, questions about work in other locations were not included in the FDNY questionnaires. As noted in Table III, DOHMH and WTCHC used different overall scales, combining exposures in different ways: (i) time of arrival and being caught in dust cloud; (ii) working on the pile or in the close vicinity; (iii) length of time worked; and (iv) different cut-offs differentiating high and low levels of these metrics.

Assessment of other occupational exposures

Exposures to other known, probable, or possible environmental and occupational carcinogens in addition to

TABLE III. WTC Exposure Descriptions by the Three WTC Cancer Studies

Exposure gradient	FDNY	% of cohort	DOHMH	% ^a of cohort	WTCHC	% ^a of cohort
High	Workers who arrived at the WTC site the morning of 9/11 before the collapse.	18	Workers who were in Manhattan, south of Chambers Street, between the time of the first plane impact and 1,200 hr on 9/11 (encompassing the collapse of WTC towers) AND either worked on the WTC dust and debris pile on 9/11 or worked at the WTC site for more than 90 days, starting before 9/18/2001	13.5	Workers who were caught in the dust cloud and also worked on the pile and worked on the site for ≥ 90 days.	3
	Workers first arrived at the WTC site between the afternoon of 9/11 and 7/25/2002, when the site closed for FDNY. Nearly all FDNY workers worked on the WTC pile.	82	All other workers not classified in the other levels.	76.5	Workers who were caught in the dust cloud on 9/11 and did not qualify for the highest group	17
Low	Did not capture exposure information about work in locations not on the WTC pile for this study.	<1	Workers who began work at the WTC site after 9/17/2001, did not work on the WTC debris pile, worked fewer than 30 days at the WTC site, and were not present south of Chambers Street between the first plane impact and 1,200 hr on 9/11.	6.8	All other workers not classified in the other levels.	65
			Worked exclusively at Staten Island site.	2.2	Workers who were not directly exposed to the dust cloud, did not work on the pile and worked for <40 days.	13

^a1.1% from DOHMH cohort and 2% from WTCHC cohort with missing or unknown exposure status.

the 9/11 environment are unfortunately common among the types of rescue and recovery workers that responded to 9/11, particularly firefighters, police, construction workers, and laborers. These and other worker groups have been the focus of various cancer incidence and mortality studies that, in some cases, have identified increases in overall or site-specific cancers [Howe and Burch, 1990; Stern et al., 1995; Violanti et al., 1998; LeMasters et al., 2006]. For example, a recent NIOSH study of a pooled cohort of 29,993 US firefighters found excess cancer mortality (SMR = 1.14, 95% CI 1.10–1.18) and incidence (SIR = 1.09, 95%CI 1.06–1.12) [Daniels et al., 2014]. Statistically significant increases in incidence were noted for digestive and respiratory cancers, and mesothelioma in particular (SIR = 2.29, 95%CI 1.60–3.19).

None of the studies considered occupational and environmental sources of exposures that were unrelated to the WTC attacks, especially occupational carcinogen exposures that may have occurred prior to 9/11/01 (potential confounders). Obtaining these data retrospectively is difficult and subject to recall bias, as well as the fact that various occupations have differing levels of effectiveness and availability of personal protective equipment over time.

Cancer Screening

Excess incidence of several cancers reported in the three studies may be due at least in part to surveillance bias. Screening for early detection of cancer has been linked to rising incidence rates for thyroid, prostate, and breast cancers, particularly within the past decade [Welch and Black, 2010; Davies and Welch, 2014; Uppal et al., 2015]. The DOHMH cohort did not have data available on cancer screening among its enrollees [Li et al., 2012], and, therefore, indirect assessments were made in two ways: by comparing the proportion of stage 1 cancers to corresponding cancers in the reference population and by comparing the frequency of reported routine physical check-ups within the preceding 12 months between those with and without subsequent cancers. The DOHMH study has included questions on cancer screening activities in its 2015 follow-up survey. Routine screening procedures in the FDNY cohort include complete blood counts and test for prostate-specific antigen (PSA) at each visit, scheduled every 12–18 months, as well as chest x-rays scheduled at two year intervals. If indicated, chest x-rays may be followed up with chest CT scans. To account for this, additional analyses were conducted in which the diagnosis date by 2 years for cancer detected these routine screening procedures [Zeig-Owens et al., 2011]. In the WTCHC cohort, full medical exams and screening tests for cervical (Pap smear), breast (mammography), colon (colonoscopy), and lung cancer (low-density CT scan) are offered, as recommended by the US Preventive Services Task Force.

Collection of Biological Samples

There were no biological samples collected in the DOHMH and the FDNY cohorts. Serum samples have been collected from a subset of members of the WTCHC cohort; the use of these samples for research studies is currently being explored.

Follow-up

Ascertainment of incident cancer cases in all three cohorts was carried out through the use of cancer linkage with state cancer registries or review of medical records. The FDNY cohort used both methods, whereas the other two cohorts used cancer linkage with state cancer registries. The Registries of New York and Pennsylvania were used in all three studies; the WTCHC and DOHMH also included Connecticut and New Jersey, in which a relatively large proportion of members lived at the time of enrolment; the FDNY and DOHMH included two additional states in which cohort members may have retired (Florida, North Carolina), and some studies included a few additional states comprising a small number of members (e.g., Texas, California, Washington). Lack of inclusion of relevant cancer registries would result in some degree of under-ascertainment of cancer cases. Although the use of a common set of cancer registries would increase the comparability of results across the three cohorts, the fact that the distribution of states of residence varies between the three studies complicates the choice of an efficient strategy aimed at maximizing both the completeness of follow-up of each cohort and their comparability. This problem will become less relevant in the next follow-up, since the list of cancer registries interrogated by each cohort has expanded. Although only accounting for 17 cancer cases, FDNY also used medical records obtained from hospitals and treating physicians to identify cancer cases not reported by the registries, and to supplement the data obtained from the registries. The additional cancer information identified through review of medical records, when it is available, can be used to estimate the potential under-ascertainment of cancer cases but may overestimate the point estimates if included in the external comparisons.

In addition, cases of multiple primary malignancies were analyzed in two of the studies presented [Zeig-Owens et al., 2011; Solan et al., 2013]. A reason to include multiple primary malignancies is that some of the cancers at high incidence in the WTC cohorts have a relatively good prognosis (e.g., thyroid, prostate), and these patients remain at increased risk of a multiple primary malignancies. A reason not to include multiple primary malignancies is the increased risk of second or multiple additional primary malignancies among those with a history of a first primary malignancy [Ng and Travis, 2008].

Statistical analysis

All three cohorts conducted an external analysis, comparing the incidence of cancer in the cohort with that of an external reference population, and an internal analysis, in which cohort members were classified according to some index of exposure (see above). Different reference populations were used to calculate expected cases of cancer for external comparisons based on the SIR: rates from the National Cancer Institute Surveillance, Epidemiology and End Results Program (SEER)-13 in the FDNY study, New York State population rates provided by the New York State cancer registry in the DOHMH study, and a combination of cancer registry and SEER data in WTCHC. Different statistical approaches have also been used in internal analyses conducted in the three cohorts: poisson regression with expected cases used as offset in FDNY and WTCHC, Cox regression in FDNY, and DOHMH.

Latency and lagged analyses

Lagged analyses ignored the period immediately following exposure, based on multi-stage models of carcinogenesis which require relatively long time intervals between exposure and effect [Checkoway et al., 1990; Lagiou et al., 2005]. Analyses based on different periods since WTC exposure were conducted in the FDNY cohort (up to 2004, 2005, and later) and the DOHMH (up to 2006, 2007, and later). In the WTCHC cohort (open enrollment), the first 6 months after registration were excluded from the primary analysis to address the possibility of selection bias from preferential enrollment of subjects with preclinical cancer (or related symptoms).

Coordination Between the Studies

The coordination between the three studies can lead to considerations for future parallel analyses of the cohorts. This review of the methodological issues inherent in the comparison of the three studies leads to several options which would increase the comparability of the findings and simplify their interpretation, as in the following examples. Provisions to avoid double-counting should be developed to enable any future analysis of the combined WTC population, whereas complying with the strict provisions for data protection in place for each cohort. One solution already in place is that DOHMH routinely removes these 3,000 potential FDNY overlap individuals from their analyses without any significant impact on their findings. Data protection issues prevent each cohort from identifying individuals included in other cohorts; a more general solution could be the identification of a “neutral” center (e.g., the New York State Cancer Registry) in which the

identified data of the cohort members can be centralized to avoid overlaps. Common sets of cancer registries and reference populations could be selected, resulting in more comparable SIR results. Similarly, all cohorts could use the same reference population for a specific set of years, to make comparisons between cohorts more uniform, and should describe the rationale for the choice of different reference groups. Common latency and exposure windows could be applied to investigate biologically relevant events. A protocol for enhancing comparability of data management and analysis, and reporting of results could be developed to address these and other sources of heterogeneity.

The three cohorts have already developed a framework for common exposure definitions, especially for the highest exposure levels, based on key exposure indicators described above: (i) time of arrival and being caught in dust cloud, (ii) working on the pile or in the close vicinity, and (iii) length of time worked. Despite loss of exposure specificity, it may be possible to use these definitions, perhaps with further refinement, in future studies. For instance, all three cohorts can classify responders as arriving on the scene on 9/11 and a direct comparison can be made using this criteria. Further differentiation of exposure could be achieved by comparing common variables between two studies (e.g., those who worked on the pile on 9/11 in FDNY and DOHMH, or persons who were caught in the dust cloud in WTCHC and DOHMH). Length of time working at the WTC site is another key exposure variable that has been linked with adverse health effects following 9/11 and should be included, if possible, as an exposure indicator.

Assessment of occupational exposures outside the WTC experience could be performed in the DOHMH and the WTCHC, possibly within the framework of nested case-control analyses (the FDNY study has no information on jobs outside the Fire Department). Ultimately, only with long-term follow-up studies based on periodical cancer surveillance will we be able to provide strong and clear evidence on cancer risk among WTC responders and for which specific cancer types it applies to.

It is fortuitous that we currently have three studies that can be used to cross-check the cancer outcomes resulting from the WTC disaster. This situation was the result of specific historical circumstances. The FDNY cohort began its program within weeks of 9/11 because external funding was not initially required; the exposed subjects were already known through work records, an internal health delivery and analytic team already existed; and leadership quickly worked out a collaborative approach with a single IRB that agreed to prioritize this work as a national imperative. The other two cohorts needed to identify funding, recruit the exposed, establish or expand relationships with labor and other groups, establish infrastructure, obtain approvals from numerous IRBs, and determine whether their aim was to provide monitoring, treatment, clinical research, or all three,

resulting in delays in commencement of recruitment. In addition, initial questionnaires from the FDNY and WTCHC cohorts were developed independently. Later instruments were developed through collaboration, although differences remained due to territorial, cultural, educational, language, and exposure differences between and within cohorts. Nevertheless, this unusual situation has achieved the advantage that the three largely independent results can be checked against each other, with the result, as summarized in this paper, that even after allowance for the differences in approaches the results are largely mutually consistent.

The goal of creating a single responder cohort for cancer analyses using a consistent set of exposure metrics and a unified analytical approach has been discussed by the authors. Although this approach would be highly desirable from a research viewpoint, there are substantial difficulties to achieve this seemingly simple goal, related in part to the different ways in which informed consent was obtained, to the eligibility and recruitment strategies used by the three groups, and the quantity and quality of data collected by each. Even if human subject issues could be easily resolved, analyses would still be faced with challenges of harmonizing pre-exposure health status, WTC and non-WTC exposures, eligibility, recruitment, and selection issues, as well as the problem of having widely different questionnaire content and access to participants' medical records, as discussed in this review. The increase in numbers, a laudable goal, would be at the expense of an increase in heterogeneity.

A phased approach, now being pursued, is based on the fact that the update of cohort outcomes is an on-going process in all three groups. This will provide an opportunity to carry out separate but coordinated analysis on an extended follow-up, whereas continuing to explore means for actual pooling of records. In the simplest approach, parallel analyses would be carried out in each cohort using the same reference population and statistical methodology. The ultimate, most ambitious approach would consist of pooling individual data from all three cohorts, but many of the impediments already discussed still remain. An interesting intermediate approach, requiring fewer resources, would be to pool cases of specific cancer sites with reported elevated SIRs, such as thyroid or prostate cancer, and to draw a small set of matching controls from their respective cohorts in a series of nested studies.

DISCUSSION AND CONCLUSIONS

Cancer SIRs in the three cohort studies reported so far, and summarized in Table II, are remarkably consistent, in spite of methodological differences. All three studies showed modest elevations of the point estimates of risk of all cancers combined, thyroid cancer, and prostate cancer, and a decreased risk of lung cancer. Elevated SIRs for melanoma

and non-Hodgkin lymphoma were observed in the FDNY study only, and multiple myeloma in the DOHMH. It is worth noting that an increased risk of prostate cancer, melanoma, and non-Hodgkin lymphoma has been observed in other studies of firefighters [LeMasters et al., 2006; Tsai et al., 2015].

The analysis of cancer risk in the three WTC cohorts represents an effort to derive valid epidemiologic results about an important health question in spite of the limitations of the underlying data. The greatest limitation is the lack of high quality, comparable, exposure information for all cohort members. Each cohort attempted to obtain exposure information via questionnaires administered at time of entry into to the cohort but the task is challenging given the passage of time, the numerous exposure variations (type of exposures, exposure times, and durations, type and use of respiratory protection), and different languages and educational levels. For future disasters, objective exposure information should be collected and workers and volunteers should at the very least be equipped with electronic identification cards to track times present and specific locations (via GPS monitoring). And perhaps one day the use of respiratory protection could be monitored electronically as is currently done for indoor ventilators.

Given the public interest in understanding early cancer effects, reported results for all three cohorts were based on data through the end of 2008, so that no cases diagnosed more than seven years after 9/11 were accounted for. Although the period of development of solid tumors can extend to decades [Lagiou et al., 2005], results of analyses with extended follow up, which are currently in progress, will provide more definite answers, whereas taking into account the increasing rates of cancer due to the aging of the cohorts.

The presence of three studies of WTC responders provides unique strength to the effort of identifying and quantifying the risk of cancer in this population. The heterogeneity in several aspects of the design might increase the overall sensitivity to identify cancers potentially associated with WTC exposure, and the availability of three largely independent studies allows for replication and reduces the likelihood of false positives. In this respect, coordination between the three studies, as shown in this review, is an appropriate approach to maximize their scientific and public health yield, address potential bias, and sources of heterogeneity, and minimize the opportunity for misinterpretation of the results.

The presence and magnitude of an increased risk of cancer among WTC-exposed individuals remain to be fully elucidated in future analyses of the three WTC cohorts. The systematic comparison of methods and findings of the three cohorts provided in this report is intended to be a contribution to maximization of the yield of the research effort on cancer for the benefit of WTC responders and of

other groups experiencing similar exposures. Although improving collection and analytic methods for this and future disasters is always of interest, we must not lose sight of the single most important requirement for meaningful future work—the need for long-term follow up of these cohorts with minimal longitudinal dropout. To date, the three cohorts have been able to maintain robust datasets on occupational and non-occupational risk factors, including lifestyle habits, and co-morbidities. The continuation of the WTC Health Program will further extend the current results beyond the implications for cohort members in terms of surveillance, treatment and compensation, and will inform the field of environmental and occupational cancer in general.

The collaborative analyses described in this report serve several important purposes. First, they provide an evidence base for treatment and compensation programs, most notably the World Trade Center Health Program that was established and funded by the James Zadroga 9/11 Health and Compensation Act of 2010. The Act provides medical monitoring and treatment for emergency responders, recovery, cleanup workers, and volunteers who worked at the World Trade Center site, the Pentagon, and the Shanksville, PA, crash site, and health evaluations and treatment for other persons who were exposed to the aftermath of the attacks in the New York City disaster area.

The list of cancers eligible for treatment or compensation under the Act is reviewed periodically and may be updated based on new findings. The three studies described here are the largest and most comprehensive cohorts of 9/11 exposed individuals, and so are the most likely source of potential new results. The findings covered in this review are based on observation periods through 2008, or only a little over seven years since the initial exposures. Any new conditions such as cancers with lengthy latency periods are more likely to be observed with the further passage of time and aging of the cohorts. All three study groups have now updated their case files through 2011, thereby extending the follow-up period by 3 years. Therefore, a second goal of this collaborative analysis has been to lay the groundwork for forthcoming analyses that will take advantage of the lessons learned in these comparisons.

Finally, although any future large-scale disaster would be expected to elicit a public health response specially crafted to its unique circumstances, exposures, and populations, it is most likely that epidemiological studies of long-term health consequences will be undertaken with cohort studies which may vary in detail but whose overall features do not differ greatly from those described here. Our experiences suggest that such investigations may proceed with a measure of assurance that comparable and valid conclusions can be drawn using different approaches.

AUTHORS' CONTRIBUTIONS

Each author has made substantial contributions to the conception or design of the work, and to the drafting the manuscript. Each author has approved the version to be published and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

This work was partially supported by CDC/NIOSH contract BAA 2011-Q-13340. The DOHMH study is supported by Cooperative Agreement U50/ATU272750 from the Agency for Toxic Substances and Disease Registry of the Centers for Disease Control and Prevention, which included support from the National Center for Environmental Health (NCEH); Cooperative Agreement 1U50/ OH009739 from the National Institute for Occupational Safety and Health; and the New York City Department of Health and Mental Hygiene. The WTCHC study is supported by the Centers for Disease Control and Prevention (CDC), National Institute for Occupational Safety and Health (NIOSH) (contracts 200-2011-39377 and 200-2002-0038; grant 5U10 OH008232), the American Red Cross Liberty Fund, The September 11th Recovery Program, The Bear Stearns Charitable Foundation, The September 11th Fund, and The Robin Hood Foundation Relief Fund. Cancer surveillance at the New York State Department of Health is supported by the Department and by the National Program of Cancer Registries of the CDC (cooperative agreement U58/DP000783).

ETHICS REVIEW AND APPROVAL

Not relevant.

DISCLOSURE BY AJIM EDITOR OF RECORD

Steven Markowitz declares that he has no conflict of interest in the review and publication decision regarding this article.

REFERENCES

Banauch G, McLaughlin M, Hirschhorn R, Corrigan M, Kelly K, Prezant D. 2002. Injuries and illnesses among New York City Fire Department rescue workers after responding to the World Trade Center attacks. In Centers for Disease C, Prevention, editors. MMWR Morb Mortal Wkly Rep. 51 Spec No:1-5.

Checkoway H, Pearce N, Hickey JLS, Dement JM. 1990. Latency analysis in occupational epidemiology. *Arch Environ Health* 45:95–100.

Daniels RD, Kubale TL, Yiin JH, Dahm MM, Hales TR, Baris D, Zahm SH, Beaumont JJ, Waters KM, Pinkerton LE. 2014. Mortality and cancer incidence in a pooled cohort of US firefighters from San Francisco, Chicago and Philadelphia (1950–2009). *Occupat Environ Med* 71:388–397.

Davies L, Welch HG. 2014. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg* 140:317–322.

Howe GR, Burch JD. 1990. Fire fighters and risk of cancer: An assessment and overview of the epidemiologic evidence. *Am J Epidemiol* 132:1039–1050.

Lagiou P, Adami HO, Trichopoulos D. 2005. Causality in cancer epidemiology. *Eur J Epidemiol* 20:565–574.

Landrigan PJ, Lioy PJ, Thurston G, Berkowitz G, Chen LC, Chillrud SN, Gavett SH, Georgopoulos PG, Geyh AS, Levin S, et al. 2004. Health and environmental consequences of the world trade center disaster. *Environ Health Perspect* 112:731–739.

LeMasters GK, Genaidy AM, Succop P, Deddens J, Sobeih T, Barriera-Viruet H, Dunning K, Lockey J. 2006. Cancer risk among firefighters: A review and meta-analysis of 32 studies. *J Occup Environ Med* 48:1189–1202.

Li J, Cone JE, Kahn AR, Brackbill RM, Farfel MR, Greene CM, Hadler JL, Stayner LT, Stellman SD. 2012. Association between World Trade Center exposure and excess cancer risk. *JAMA* 308:2479–2488.

Lioy PJ, Weisel CP, Millette JR, Eisenreich S, Vallero D, Offenberg J, Buckley B, Turpin B, Zhong M, Cohen MD, et al. 2002. Characterization of the dust/smoke aerosol that settled east of the World Trade Center (WTC) in lower Manhattan after the collapse of the WTC 11 September 2001. *Environ Health Perspect* 110:703–714.

Lioy PJ, Georgopoulos P. 2006. The anatomy of the exposures that occurred around the World Trade Center site: 9/11 and beyond. *Ann N Y Acad Sci* 1076:54–79.

Lorber M, Gibb H, Grant L, Pinto J, Pleil J, Cleverly D. 2007. Assessment of inhalation exposures and potential health risks to the general population that resulted from the collapse of the World Trade Center towers. *Risk Anal* 27:1203–1221.

McGee JK, Chen LC, Cohen MD, Chee GR, Prophete CM, Haykal-Coates N, Wasson SJ, Conner TL, Costa DL, Gavett SH. 2003. Chemical analysis of World Trade Center fine particulate matter for use in toxicologic assessment. *Environ Health Perspect* 111:972–980.

McKinney K, Benson S, Lempert A, Singal M, Wallingford K, Snyder E. 2002. Occupational exposures to air contaminants at the World Trade Center Disaster Site—New York, September—October, 2001. *Morbidity and Mortality Weekly Report*. 51:453–456. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5121a1.htm> (Accessed 11/9/2015).

Murphy J, Brackbill RM, Thalji L, Dolan M, Pulliam P, Walker DJ. 2007. Measuring and maximizing coverage in the World Trade Center Health registry. *Stat Med* 26:1688–1701.

National Institute for Occupational Safety and Health. World Trade Center Response. 2002. Summary Report to the New York City Department of Health: NIOSH Air Sample Results for the World Trade Center Disaster Response (February 2002). <http://www.cdc.gov/niosh/nas/rdrp/appendices/chapter10/a10-4.pdf> (Accessed 9/15/2014).

Ng AK, Travis LB. 2008. Subsequent malignant neoplasms in cancer survivors. *Cancer J* 14:429–434.

Solan S, Wallenstein S, Shapiro M, Teitelbaum SL, Stevenson L, Kochman A, Kaplan J, Dellenbaugh C, Kahn A, Biro FN, et al. 2013. Cancer incidence in world trade center rescue and recovery workers, 2001–2008. *Environ Health Perspect* 121:699–704.

Stern F, Schulte P, Sweeney MH, Fingerhut M, Vossenbach P, Burkhardt G, Kornak MF. 1995. Proportionate mortality among construction laborers. *Am J Ind Med* 27:485–509.

Tsai RJ, Luckhaupt SE, Schumacher P, Cress RD, Deapen DM, Calvert GM. 2015. Risk of cancer among firefighters in California, 1988–2007. *Am J Ind Med* 58:715–729.

Uppal A, White MG, Nagar S, Aschebrook-Kilfoy B, Chang PJ, Angelos P, Kaplan EL, Grogan RH. 2015. Benign and malignant thyroid incidentalomas are rare in routine clinical practice: A review of 97,908 imaging studies. *Cancer Epidemiol Biomarkers Prev* 24:1327–1331.

US Department of Labor, OSHA. 2002. OSHA Sampling Results Summary as of 10/08/2002. <https://www.osha.gov/nyc-disaster/summary.html> (Accessed 9/15/2014).

US Department of Labor, OSHA. 2014. “Additional World Trade Center Information”—Detailed results: Asbestos, silica, metals, organics. <https://www.osha.gov/nyc-disaster/wtc-additional.html> (Accessed 9/15/2014).

Violanti JM, Vena JE, Petralia S. 1998. Mortality of a police cohort: 1950–1990. *Am J Ind Med* 33:366–373.

Weakley J, Maslow C, WTC Working Group. 2011. Defining common categories of exposure among four cohorts of rescue/recovery workers who responded to the World Trade Center disaster. abstract presented at Council of State and Territorial Epidemiologists (CSTE) Annual Conference; 6/13/2011, Pittsburgh, PA.

Welch HG, Black WC. 2010. Overdiagnosis in cancer. *J Natl Cancer Inst* 102:605–613.

Zeig-Owens R, Webber MP, Hall CB, Schwartz T, Jaber N, Weakley J, Rohan TE, Cohen HW, Derman O, Aldrich TK, et al. 2011. Early assessment of cancer outcomes in New York City firefighters after the 9/11 attacks: An observational cohort study. *Lancet* 378:898–905.

Institutions at which the work was performed: Icahn School of Medicine at Mount Sinai, New York, NY; New York City Fire Department, New York, NY; New York City Department of Health and Mental Hygiene, New York, NY.