Supplementary Methods

**Issues of data completeness and consistency in developing the WTC Exposure Ranking Indices (WTC-ERIs)**

Quality assurance analyses of the exposure-relevant information available in the WTC-HPDC dataset identified two general issues:

1. suboptimal data resolution (for example, the use of only 4 areas/zones for location assignment in the questionnaire, while available information points to the existence of substantial environmental and microenvironmental - including occupational microenvironments - quality gradients within these areas), and
2. data “fuzziness” (for example, the value ranges of exposure-related attributes, e.g. time spent at a location, are not mutually exclusive, leading to multiple possibilities in potential exposure ranking estimates).

As a result, the following “Groups” were identified within the WTC-HPDC dataset, based on the specific data attributes that are listed in Supplementary Table 6:

Group A1: Sufficient information is available for deriving time, location, activity and PPE factors with minimal assumptions

Group A2: Information regarding activity in supplemental fields is different from reported WTC occupations.

Group B: Certain default assumptions are required to derive activity and location factors. (e.g., more than one activity is specified for one time period or location).

Group C: Imputation is required to complete missing entries in days and hours worked.

Group D1: Imputation is required for missing entries in PPE related fields.

Group D2: Insufficient data for deriving the time, location, activity and PPE factors; extensive use of assumptions is required and the uncertainties associated with estimates based on currently existing data are high.

Group D3: Imputation is required for missing time, location, and activity entries; assumption is required for data conflicts.

Group E1: Imputation is required for conflicting (and/or missing) time, location, activity and PPE entries.

Group E2: Insufficient data to perform imputation.

The EAQ data for 6,110 subjects in the Stony Brook database were used for classification according to the criteria of Table 10, resulting in the following subject numbers per group:

* Group A: 2625
* Group B: 1140
* Group C: 250
* Group D:1766
* Group E: 329

Methylation profiling data were collected for 650 of the above subjects in the Stony Brook database; the corresponding group classification , shown below, resulted in a total of 503 responders with both methylation data and robust (Group A, B and C) estimates:

* Group A: 307
* Group B: 172
* Group C: 24
* Group D: 146
* Group E: 1

Supplementary Table 1: Values of exposure factors for time-location (including contact with the 9/11 dust cloud)

|  |  |  |
| --- | --- | --- |
| **Category** | **Description** | **Factor** |
| Value of the exposure factor for September 11 Dust Cloud () | Directly in the cloud of dust from the collapse of the World Trade Center buildings | 500 |
| Exposed to significant amounts of dust but not directly in the cloud of dust from the collapse of the World Trade Center | 200 |
| Exposed to some dust but not in the cloud of dust from the collapse of the World Trade Center buildings | 100 |
| Present in lower Manhattan on September 11, 2001 but not reporting exposure to dust or being in the cloud of dust from the collapse of the World Trade Center buildings | 20 |
| Do not know | 0 |
| Not obvious, requires further analysis | 0 |
| Value of the exposure factor for date () | 09/11/2001 | 50 |
| 09/12/2001 | 10 |
| 09/13/2001 | 5 |
| 09/14/2001 - 09/17/2001 | 4 |
| 09/18/2001 - 10/10/2001 | 2 |
| 10/11/2001 - 11/30/2001 | 1 |
| 12/01/2001 - 02/28/2002 | 0.5 |
| 03/01/2002 - 05/31/2002 | 0.25 |
| 06/01/2002 - 09/12/2002 | 0.125 |
| Value of the exposure factor for location where worker spent majority of shift () | on the pile/in the pit; | 5.0 |
| adjacent to pile/pit; | 4.0 |
| landfill; | 3.0 |
| Barges/Loading piers; | 3.0 |
| Office of Chief Medical Examiner (OCME); | 0.5 |
| elsewhere, S. of Canal st. | 2.75 |
| Don’t know | -999.0 |

Supplementary Table 2: Values of the exposure factor for WTC Activities ()

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Code** | **Activity ID** | **Score** | **Categories of Activity** | **Shortened Activity Description** |
| 0 | 53-5021.00 | 9 | Barge Workers | Barge Workers |
| 1 | 47-2011.00 | 5 | Boiler makers | Boiler makers |
| 2 | 47-2051.00 | 5 | Carpenters | Carpenters |
| 3 | 47-2041.00 | 2 | Carpet, floor, and tile installers and finishers | Carpet, floor, & tile installers & finishers |
| 4 | 47-4011.00 | 10 | Construction and building inspectors | Construction & building inspectors |
| 5 | 47-2073.00 | 8 | Construction equipment operators | Construction equipment operators |
| 6 | 47-2031.00 | 5 | Dock builders | Dock builders |
| 7 | 47-2081.00 | 2 | Drywall Installers, ceiling tile installers and tapers | Drywall & ceiling tile installers |
| 8 | 47-2111.00 | 3 | Electricians | Electricians |
| 9 | 47-4021.00 | 4 | Elevator installers and repairers | Elevator installers & repairers |
| 10 | 17-2199.00 | 3 | Engineers, structural/other | Engineers, structural/other |
| 11 | 49-2022.01 | 8 | Equipment workers | Equipment workers |
| 12 | 47-2121.00 | 4 | Glaziers | Glaziers |
| 13 | 47-4041.00 | 10 | Hazardous materials removal workers | Hazmat removal workers |
| 14 | 49-9021.00 | 3 | Heating, air & refrigeration mechanics | Heating, air & refrigeration mechanics |
| 15 | 47-4051.00 | 2 | Highway maintenance workers | Highway maintenance workers |
| 16 | 47-2131.00 | 3 | Insulation workers | Insulation workers |
| 17 | 47-2221.00 | 8 | Ironworkers, structural and reinforcing iron and metal workers | Ironworkers: structural & reinforcing |
| 18 | 53-7062.00 | 9 | Laborers | Laborers |
| 19 | 37-3011.00 | 5 | Landscapers & groundskeepers | Landscapers & groundskeepers |
| 20 | 47-2021.00 | 5 | Mason: brick, block and stone | Mason: brick & stone |
| 21 | 47-2031.00 | 5 | Mason: cement, concrete finishers, segmental pavers and terrazzo workers | Mason: cement & concrete |
| 22 | 47-2161.00 | 5 | Masons: plasterers and stucco | Masons: plaster & stucco |
| 23 | 49-3011.00 | 6 | Mechanics | Mechanics |
| 24 | 47-2141.00 | 2 | Painters and paperhangers | Painters & paperhangers |
| 25 | 47-2071.00 | 2 | Paving, surfacing, & tamping workers | Paving, surfacing, & tamping |
| 26 | 47-2151.00 | 3 | Pipe Layers, plumbers, pipe fitters, and steam fitters | Plumbers & pipe/steam fitters |
| 27 | 47-2181.00 | 4 | Roofers | Roofers |
| 28 | 47-2211.00 | 3 | Sheet metal workers | Sheet metal workers |
| 29 | 53-3032.00 | 6 | Truck drivers | Truck drivers |
| 30 | 55-0000 | 4 | Military | Military |
| 31 | WBB1 | 4 | Body bag work | Body bag work |
| 32 | WBB2 | 4 | Bucket brigade | Bucket brigade |
| 33 | WCNM | 6 | Cable installation/repair/splicing (EXCLUDING work in manholes) | Cable work (excl. manholes) |
| 34 | WCM | 7 | Cable installation/repair/splicing (INCLUDING work in manholes) | Cable work (incl. manholes) |
| 35 | WCS | 2 | Canteen service | Canteen service |
| 36 | WCO | 2 | Counselor | Counselor |
| 37 | WCU | 7 | Custodian | Custodian |
| 38 | WDH | 5 | Dog handlers | Dog handlers |
| 39 | WDS | 10 | Dust suppression | Dust suppression |
| 40 | WEMT | 8 | EMT | EMT |
| 41 | WESC | 3 | Escorting | Escorting |
| 42 | WEXC | 9 | Excavation/ Confined space work | Excavation/ Confined space work |
| 43 | WFF | 9 | Fire Fighter | Fire Fighter |
| 44 | WIH | 6 | Industrial Hygienist | Industrial hygienist |
| 45 | WMW | 1 | Morgue work | Morgue work |
| 46 | WPO | 6 | Police Officer | Police Officer |
| 47 | WPS | 6 | Perimeter security | Perimeter security |
| 48 | WSA | 7 | Sanitation workers | Sanitation workers |
| 49 | WSR | 10 | Search and rescue | Search and rescue |
| 50 | WSNC | 10 | Sifting (EXCLUDING conveyer belt) | Sifting (excl. conveyer belt) |
| 51 | WSC | 10 | Sifting (INCLUDING conveyer belt) | Sifting (incl. conveyer belt) |
| 52 | WTW | 4 | Towing | Towing |
| 53 | WTL | 4 | Truck loading/unloading | Truck loading/unloading |
| 54 | WTR | 2 | Truck routing | Truck routing |
| 55 | 132-Other | 4 | Other (specify) | Other |

**Supplementary Table 3:** Values of PPE-related exposure factors

|  |  |  |
| --- | --- | --- |
| **Categories** | **Description** | **Factor** |
| PPE Type  () | Full face | 0.9 |
| Half face | 0.75 |
| Dust/surgical/disposable mask | -999 |
| Didn’t wear one | -999 |
| Other (Specify) | -999 |
| Don’t Know | -999 |
| PPE usage frequency  (,,) | Always | 1 |
| Most time | 0.7 |
| Some of the time | 0.3 |
| Rarely/Never | 0 |
| N/A | -999 |
| Don’t Know | -999 |
| PPE maintenance frequency  (,) | Daily | 1 |
| 1-2 times a week | 0.95 |
| 3-4 times a week | 0.98 |
| Weekly | 0.92 |
| Every 2 weeks | 0.90 |
| Monthly | 0.88 |
| Every other month | 0.86 |
| Don’t Know | -999 |
| Rarely/Never | 0.85 |
| When it was difficult to breathe through, dirty or damaged | 0.80 |
| PPE cleaning frequency  () | Daily | 1 |
| 3-4 times a week | 0.90 |
| 1-2 times a week | 0.85 |
| Rarely/Never | 0.80 |
| Don’t Know | -999 |
| PPE seal check frequency  () | Always | 1 |
| Most of the time | 0.90 |
| Some of the time | 0.85 |
| Rarely/Never | 0.80 |
| Don’t Know | -999 |

|  | **Group** | **A1** | **A2** | **B** | | | **C** | | | | **D** | | | | | **E** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Time | Are the first & last days subject worked on WTC-related activities specified? | ✓ | ✓ | ✓ | ✓ | ✓ | At least one ✗ | At least one ✗ | At least one ✗ | At least one ✗ | At least one ✗ | Irrelevant | Irrelevant | Irrelevant | Irrelevant |  |
| Is every work day accounted for in September? | ✓ | ✓ | ✓ | ✓ | ✓ |  |
| Are both the numbers of days worked and of average hours worked per work day valid for the remaining time periods? | ✓ | ✓ | ✓ | ✓ | ✓ |  |
| Is imputation for days and/or average hours worked possible? |  |  |  |  |  | ✓ | ✓ | ✓ | ✓ | ✗ |  |
| Activity | Did subject specify one activity code for all the time periods subject reported working? | ✓ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | Irrelevant | ✗ |  |
| Did subject specify more than one activity code for any of the time periods subject reported working? | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ |  |
| Location | Did subject specify the major location worked for all the time periods subject reported working? | ✓ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | Irrelevant | ✗ |  |
| If subject did not specify the major location of work, can the location be deduced from other fields for any time periods the subject reported working? | ✗ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ |  |
| PPE | Are entries for the mask fields valid? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Irrelevant | ✗ |  |
| Are entries for the respirator fields valid for all time periods subject reported working? | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Irrelevant | ✗ |  |
| Search & Rescue | Description of activity contains the words "search" and "rescue" but the subject did not select "search and rescue" as his/her main activity | ✗ | ✓ |  |  |  |  |  |  |  | Irrelevant |  |
| Data conflict | Conflicting data pertaining to exposure ranking calculation | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ |
| Estimated number of subjects in each group | | 2,625 | | 1,140 | | | 250 | | | | 1,766 | | | | | 329 |
| Estimated number of subjects with methylation data in each group | | 307 | | 172 | | | 24 | | | | 146 | | | | | 1 |
| ✓= Yes; ✗= No;  = N/A | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Supplementary Table 4: Completeness and classification of exposure data for Stony Brook subjects (in WTCGRC)

**Supplementary Table 5**: Top 10 CpG sites associated with ERI. Columns 6-8 report the difference between high and low exposure, estimated coefficient and estimated p-values, respectively.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ERI |  | |  | |  | |  | |  | |  | |  | |  | |
| CpG | Chr | | Coordinate | | Gene | | Location | | Diff | | Coef | | PV | |  | |
| cg22785561 | | chr4 | | 46923365 | | GABRA4 | | 3'UTR | | 0.0353 | | 0.235 | | 1.88E-05 | |
| cg20894465 | | chr6 | | 30688821 | | TUBB | | Body | | 0.00708 | | 0.247 | | 6.02E-05 | |
| cg08962081 | | chr5 | | 143978420 | | Intergenic | | Intergenic | | -0.0363 | | -0.224 | | 0.000109 | |
| ch.20.1062061F | | chr20 | | 47778018 | | STAU1 | | 5'UTR | | 0.0239 | | 3.17 | | 0.000111 | |
| cg07063463 | | chr20 | | 43932925 | | MATN4 | | Body | | -0.0495 | | -0.225 | | 0.000123 | |
| cg06233135 | | chr6 | | 42450495 | | Intergenic | | Intergenic | | -0.00446 | | -0.468 | | 0.000132 | |
| cg17600661 | | chr8 | | 94752611 | | C8orf39 | | Body | | 0.0129 | | 0.45 | | 0.000138 | |
| cg03691338 | | chr10 | | 50344050 | | Intergenic | | Intergenic | | 0.0379 | | 0.214 | | 0.000143 | |
| cg03490157 | | chr3 | | 6906061 | | GRM7 | | Body | | 0.0592 | | 0.279 | | 0.000174 | |
| cg02694715 | | chr11 | | 2019730 | | H19 | | TSS1500 | | 0.00756 | | 0.0342 | | 0.000186 | |

**Supplementary Figure 1**: Global methylation pattern from epigenome-wide ERI analysis. The volcano plot indicated approximately equal hyper- and hypo-methylation pattern in the association analysis with ERI. The p-values were computed from logit transformed beta values as response and ERI as predictor, adjusting for age, smoking status, race, and cell types. The black solid circles correspond to the top 500 CpGs associated with ERI. Red and blue horizontal lines correspond to p-value cutoffs 0.0001 and 0.05, respectively.

****