Instructions to Apply for
Vital Statistics Data

The Bureau of Vital Statistics (BVS) is in the New York City Department of Health and Mental Hygiene (DOHMH) and is responsible for legal registration of Vital Events (births, deaths and terminations of pregnancy) that occur in New York City (NYC). The Bureau's Office of Vital Statistics oversees the data cleaning, analyses, reporting and disclosure of data for research, government and public health use. To access NYC vital statistic data, please see below.

Vital Statistics can be provided in several data forms:
1. Publically accessible data is available on the DOHMH Website:
    a. [Annual Summary of Vital Statistics](#)
    b. [DOHMH Epi Query - Interactive Web based data tools](#)
2. If the data you are seeking is currently unavailable on our website, please email **vsdata@health.nyc.gov** about the following options:
    a. *Limited* datasets – line item datasets that include the variables specified in the hyperlink provided for each event during a given year. The dataset does not include direct identifiers (name, address, SSN) although combinations of variables provided may permit re-identification. Access to this limited use dataset requires completion of this application and the administration of a data use agreement.
    b. *Customized dataset* – line-item datasets requiring DOHMH resources to define and prepare datasets for the requester. Access to a customized dataset requires completion of this application and the administration of a data use agreement. Fees for data preparation requiring two or more hours of DOHMH personnel labor may apply.
    c. *Certified certificate* copies for research and other *proper purposes* in accordance with the New York City Health Code may be obtain. Please contact **vsdata@health.nyc.gov** to determine whether this form is applicable. Fees for certified copies and redaction activities will apply.

What is individually identifiable information?

New York City Health Code §3.25 defines *individually identifiable information* as "any information or data which by itself or in combination with any other publically available information, could, in the opinion of the Department, be used to identify a particular individual."

Some examples of *individually identifiable vital information*:

Examples of identifiers include:
- First or last name
- Certificate number
- Office of the Chief Medical Examiner (OCME) number
- Medical record number
- Social Security number
- Address
- Exact date (e.g. Date of Birth, Date of Death)

Examples of identifiable variables that put individual records at risk of re-identification
- Aggregate data by zip code or smaller geographical levels (e.g. census tracts/blocks)
- Exact measures (e.g., using birth weight rather than birth weight category)
- Write-in items including cause of death "literals"

Record-level datasets that exclude direct *identifiers* (i.e. *Limited* Datasets) or aggregated data in small geographic areas and/or time periods are often deemed identifiable as [opportunities](#) for individual re-identification exist.

When to use this application:
Use this application and its respective affidavits to obtain *individually identifiable* birth, death, and spontaneous termination of pregnancy (fetal deaths) *information* from the New York City (NYC) Department of Health and Mental Hygiene (DOHMH) for research, public health practice, or administrative purposes meeting the requirements of the NYC Health Code.

Do not use this application for certified copies of birth and death certificates to be used for travel, passport, proof of citizenship, social security, driver's license, school registration, personal identification and other legal purposes. Please see Birth and Death Certificates.

> If you were previously provided NYC vital event data for an approved non-research purpose you may not use the data for research without a separate, approved application.

> If you are seeking to renew an expired approval or to obtain already approved data, please email **vsdata@health.nyc.gov** with your previously approved OVS log number. Reapplication may not be required.

> Note, Multi-year projects that intend to access Vital Statistics data not yet available at the time of the application, may require updates to IRB determination letters.

DOHMH Review of Submitted Application:
This application will be reviewed by DOHMH personnel to ensure the proposed data use is compliant with the NYC Health Code §3.25 requirement to protect individually identifiable information that may compromise the health, safety or privacy of any person if disclosed and the NYC Administrative Code and NYC Health Code (*See* Ad Code §17-169 and HC §3.25, §201.07, §203.07, §205.07, and §207.11) requirements for proper purpose use. The DOHMH is compelled to "impose reasonable conditions as to the use and re-disclosure of information (data)" and "may limit access to the minimum necessary (data elements) to fulfill the purpose for which information is requested" (New York City Health Code §207.11).

Data Use Agreements
A data use agreement governs the sharing of data between parties. Should a data use agreement be required for your access, the application for record level data submitted may serve as an attachment to the data use agreement. Please take care when completing the application as the use you define in the application will be legally binding.

Special Note for Graduate Students (i.e. Dr. PH and PhD Candidates)
Graduate students applying for record-level data must designate a faculty member and/or research mentor to serve as either Principal Investigator (PI) or co-PI for the project. The selected faculty/mentor will have supervisory oversight of the research agenda, the methods of analysis and data security given the student's access to confidential, HIPAA-sensitive and individually-identifiable vital statistics data. Please include corresponding contact information in Section B of your application. When your final application is approved, a Confidentiality Affidavit will be requested for the applicant and his/her faculty mentor.

Where to Find Commonly Requested data:
New York State Vital Statistics Data:
NYC DOHMH maintains files for all births, deaths, and spontaneous terminations of pregnancy occurring in New York City *only*. If you are seeking vital data for the rest of New York State, please contact:

> Bureau of Production Systems Management, Vital Records – IRB Research Unit
> New York State Department of Health, PO Box 2602
> Albany, New York 12220-2602
> **bio-info@health.state.ny.us**
> (518) 486-9012

Geocoding resources:
> Single addresses: http://a030-goat.nyc.gov/goat/Default.aspx
> Batch-download the following:
> http://www.nyc.gov/html/dcp/html/bytes/applbyte.shtml#geocoding_application

Confirmation/Fact of Death:
The National Death Index (NDI) can provide information —for health, medical, and research purposes only—on deaths occurring anywhere in the US, including date of death, US State of death, and cause of death since 1979. Underlying and multiple cause of death information is available from DOHMH once the study is approved by NDI plus. DOHMH refers all requests for fact of death since 1979 to the National Death Index. For further information, consult the National Death Index.

Application Contents

Section A: Study Information
Section B: Principal Investigators and contact information
Section C: Project Protocol and Objectives
Section D: List of Vital Statistics Data Elements Request. (Includes written justification for individual identifiers; use data dictionaries/file layouts provided by OVS)
Section E: Security of Confidential Data
Section F: Vital Data Linkage or Matching (To be completed only if your request requires data linkage and/or matching)
  Affidavit of Principal Investigator
  Affidavit of the Co-principal Investigator
  Affidavit of Organization Official Authorized to Execute Agreements
  Supplemental Affidavit of Sponsor/Contractor/Consultant *(required for all studies funded via*
contracts, cooperative agreements, interagency agreements of non-Federal grants)

E-mail **VSdata@health.nyc.gov** or call the Data Use and Disclosure Office at 646-632-6733 for questions about preparation of the application.

Checklist for complete submission of
Vital Statistics Record-level Data Application:

☐   Complete all questions in application form.

☐   Submit 1 completed PDF copy of the application to **vsdata@health.nyc.gov.** This application may be reviewed and returned to you with questions that require your amendments. Access to data requires all required amendments are finalized.

☐   Once application is finalized, Data Use Office will issues a Data Use Agreement requiring signatures of personnel accountable for data security and use (*DOHMH employees are exempt from this application and data use agreement process).*

☐   Submit copies of IRB determinations. See Section B4.

☐   Submit comprehensive list of data elements requested. See Section C. Include justification for access to individually identifiable data items requested.

☐   Matching/Linkage project?

　　☐ Submit Section F of Application

　　☐ Once approved, *securely* submit identifiable data to DOHMH (see instructions in Section D for secure submission) or contact **vsdata@health.nyc.gov**.

Matching/Linkage - General Information

The Office of Vital Statistics (OVS) and the DIITT Data Matching Center provides data matching services for approved research, public health practice and evaluation as well as administrative purposes requiring DOHMH data that fall under the guidelines of NYC Health Code. Data matching is a lengthy process that involves the use of identifying characteristics such as Name, Date of Birth (DOB), and Social Security Number (SSN) to link members of a disease exposure or administrative registry with their corresponding records in the vital statistics files. Such matches are often conducted to identify causes-of-death among registry members or to determine the birth outcomes among women who share certain risk factors, exposures or experiences

Core to the data matching process is defining which variables in the respective datasets will be used to match the data. Matching projects use different datasets that include different identifying information, and as a result may use different variables to match. Match projects may have different definitions of what constitutes a match between a submitted data record and a vital data record. The minimum criterion for a match between registries is often an exact match on Full Name and Date of Birth. However, people exist with the same name and date of birth. For this reason, additional data items to match such as address, phone number, or social security numbers are preferable. These additional variables allow matches to be found even if there are small discrepancies in name and date of birth.

The requester must provide identifying data of a type and format that improves the specificity of the match. For a birth match, Mother's Maiden Name, Infant's DOB and Gender, SSN (only last 4 digits), Address and Telephone Number (where available) greatly improve the chance that a true match will be found. For a death match, Decedent's Full Name, Name Alias, DOB, Gender, SSN (entire or last four digits), Address, and Date of Last Contact provide an ideal set of identifiers. When submitting addresses, the street name, house number, apartment number, city, and zip code should be formatted as separate data fields. All identifiable data must be submitted by a secure method approved by the DOHMH Division of Informatics, Information Technology and Telecommunications.

The data submitted must be provided in formats requested and de-duplicated to the extent possible by the requester. Matches produced are standardized and reviewed to remove false positives introduced through spelling and naming conventions. The remaining matches are then scored according to the likelihood that they represent true matches. Almost any programmed definition of a match will miss some real matches and for this reason, a manual review may be required to identify matches that cannot be identified exclusively through automated processes. DOHMH may request the applicant's assistance in reviewing identified linked records to improve the yield of true matches. At the completion of the process, the identifying variables used to conduct the match are usually removed from the dataset and replaced with a unique identifier, unless otherwise required and permitted.