**Appendix F: Proposed data queries to assist in outbreak investigations**

The following table includes examples of data queries that can be used to assist in outbreak investigations. These may not be exactly replicable or relevant in your system. Rather, these examples are meant to illustrate the many ways in which routinely collected data can be used in outbreak investigation. All data queries are intended to show both cases and contacts but may be filtered for cases only if needed.

Some of the data queries provided in this list can be used to look for emerging clusters, such as the ‘Cluster Investigation’, ‘Exposure Investigation’, or ‘Regional Investigation’ data queries. A common method to find emerging clusters from these would be to scan for locations or specific events appearing in the exposure details fields and then search the rest of the list for similar or matching locations of events. Additionally, these data queries in particular can be sorted alphabetically by employer name to quickly find emerging workplace clusters.

Other data queries listed here can be used to continue or broaden ongoing investigations such as ‘Cluster + Demographics’, ‘Address Search’, or ‘Task Search’ (where possible). These data queries are good to run regularly using keywords or demographic information from ongoing cluster investigations to identify more potential members of the cluster.

Note: additional data queries can also be used to help monitor data quality of key variables, which will further improve the usefulness of data queries used for outbreak investigations.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of Data Query** | **Variables Included** | **Filterable Variables** | **Purpose** |
| Line List Generator | Cluster ID, Case ID, First Name, Last Name, Date of Birth, Address, City, State, Mobile Phone, Symptom Onset Date, Test Date, Employer Name, Employer Phone, Hospitalized, Vaccination Status | Case ID, Cluster ID | Creates a list of individuals linked to a known cluster for summary reporting. |
| Cluster Investigation | Case ID, Opened Date, Interviewed Date, Closed Reason, City, Suspected Cluster, Case Description, Exposure Source Category/Date/Details, Employer Name | City, Case Description, Exposure Source Category, Exposure Source Free Text, Employer Name | This is the primary data query used to begin cluster investigations. This query allows the team to search the data by any of the variables included. |
| Cluster + Demographics | Case ID, First Name, Last Name, Opened Date, City, Age, Date of Birth, Language, Test Date, Employer Name, Employer City, Exposure Source Category/Date/Details | All | This is a secondary data query to the cluster investigation one which includes an expanded demographics section which can be useful for investigating clusters occurring among demographic communities (frequently place of worship, funerals, and social event clusters) |
| Exposure Investigation | All exposure variables (consider narrowing down if too much detail or specificity) | All | This is a secondary data query to the cluster investigation one that includes more detail of the exposure information on the case record which can be useful for finding tentative links to the cluster or searching for emerging clusters |
| Address Search | Case ID, Opened Date, Address, City, State | Address, City, State | This is the primary data query used to identify cases or contacts that share an address or parts of an address. This is particularly useful for multigenerational families, apartment-based clusters, or multi-family houses. Searches can be as specific as street number or as broad as city level. |
| Individual Search | Case ID, First Name, Last Name, Mobile Phone, Test Date, Symptom Onset Date, Employer Name, Employer Address, Employer Phone | All | This is the primary data query used for identifying cases or contacts based on partial information collected through contact tracing. Searches are possible on any part of a name (including phone number). |
| Task Search  (if records have attached tasks or similar functions) | Associated Case ID, Date Created, Name of Staff that Created, Phone Number, Comments, City | All | This data query allows the team to search through call logs and escalations on case records. It can be searched based on keywords in the text or based on the person the call log or escalation is associated with. This provides insight into the complete interaction CIs have with cases and contacts, beyond what is easily captured in the data fields. |
| Regional Investigation | County, Case ID, Opened Date, Interviewed Date, Closed Reason, City, Suspected Cluster, Case Description, Exposure Source Category/Date/Details, Employer Name | County, City, Case Description, Exposure Source Category, Exposure Source Free Text, Employer Name | This data qury is very similar to cluster investigation but sorts cases based on what geographical region (county-level) they are from. This can be helpful for detecting emerging clusters, especially among workplaces or social groups. |

Special/Less Common Data Queries

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of Data Query** | **Variables Included** | **Filterable Variables** | **Purpose** |
| Testing Location Search | Case ID, Opened Date, Interviewed Date, Closed Reason, City, Suspected Cluster, Case Description, Exposure Source Category/Date/Details, Employer Name, Testing Facility | City, Case Description, Exposure Source Category, Exposure Source Free Text, Employer Name, Testing Facility | This gives the ability to search for cases by the date and location of their positive test. This can be useful if working a business cluster that has had onsite testing or looking for social groups that may have been tested together. |
| Hospital Search | Case ID, Opened Date, Hospital Name, Date Hospitalized | All | This data query is used for searching individuals admitted to specific hospitals on certain dates, this can be useful for investigating facility-based clusters or trends in hospitalizations |
| K-12 Summary | Case ID, Age, Exposure Source Free Text, School Program Type, School Type, School Name | All | This data query allows to search for minor cases and contacts based on the school, daycare, or childcare programs they are associated with |
| Unknown Exposures | Case ID, Opened Date, Interviewed Date, Closed Reason, City, Suspected Cluster, Case Description, Exposure Source Category/Date/Details Employer Name | City, Case Description, Exposure Source Free Text, Employer Name | This data query is a variation of the cluster investigation one that is filtered to only include cases that report unknown or uncertain exposures. This can be useful in times of low case counts to flag cases for further investigation by specialized case investigators or other outbreak team members. |
| Specialized Keyword Data Queries | Any, as needed | None | When investigating large clusters or settings with known ongoing exposures, establishing a data query that will automatically capture any instance of a keyword or phrase can save a lot of investigative time. For example, any mention of a particular business name, event, or setting |
| Priority Task Search | Associated Case ID, Date Created, Name of Staff that Created, Phone Number, Comments, City | All | Similar to the Specialized Keyword Data Query, depending on the number of cases and time of year, it can be helpful to establish queries to search call logs for mention of certain locations or activities, such as holidays |