2020 Public Release Research Dataset Frequently Asked Questions:

1. Are the data available as a download, rather than requesting a USB drive?
   Not at this time. The NEMSIS TAC is exploring options to distribute the data electronically as a download. Currently, the data are only available as an encrypted USB drive sent to you via FedEx.

2. Are the data available using a different encryption format than Bitlocker?
   Not at this time. Bitlocker requires Windows 10 software. We are unable to support MacOS or Linux. The TAC has tried to support operating system agnostic encryption software in the past, but it became clear the software was difficult to use to support a very limited number of requestors without access to Windows 10.

3. Are there special handling or storage requirements for the data?
   No. The data is deidentified, non-HIPAA data. However, the size of the dataset will require that you have a minimum of 16 GB of RAM and 140 GB of storage available.

4. I only need one geographic area / state / county; is this available?
   To comply with our state data use agreements the TAC cannot identify any single geographic area. If you require multiple areas, masked (or deidentified) data are available. Please contact nemsis@hsc.utah.edu to discuss your project.

5. How many years of data are available?
   NEMSIS version 3 data are available from 2017 to present. The 2017 data contains a decrease in submissions because of the transition from NEMSIS version 2 to version 3. Version 2 data are available from 2009 to 2016 but requires mapping and translation to version 3 data elements. Mapping and translation will require some additional effort. NEMSIS version translation tools are available at: https://nemsis.org/technical-resources/version-3/v3-element-mappingtranslation-documents/. Please contact nemsis@hsc.utah.edu to discuss your project if you require data prior to 2017.

5. Can I request multiple years at once?
   Our Institutional Review Board (IRB) requires that we send all requestors one year of data first. Once the data is received, decrypted, and running, please notify the NEMSIS TAC and additional years will be supplied.

6. Are the data available as a text file or Excel spreadsheet?
   No. Because of the dataset size and relational nature of the data tables, statistical programming, analysis, or database management software is required. Examples of suitable software include but are not limited to SAS, STATA, SPSS, R, Python, or SQL. Summary reports and dashboards are available on the NEMSIS website. Some dashboards allow the downloading of data as a spreadsheet. Please contact nemsis@hsc.utah.edu to discuss your project if you think summary data may meet your research needs.

7. How large is the annual dataset?
   For 2020, this event level dataset includes approximately 44 million EMS activations contained in 42 relational tables and requires between 115 and 135 gigabytes of storage once uncompressed. At least 16 GB of RAM will be required to manipulate the data. Some one-to-many tables such as PCRFACTVITALS are 128.5 million rows by 20 columns.

8. Does the annual dataset represent "patients" treated by EMS?
No, the unit of measure in the dataset is EMS activations. A single patient may be represented in multiple electronic patient care reports reported to NEMSIS because of the varying structure of EMS systems across the country. Please feel free to contact the NEMSIS TAC ([nemsis@hsc.utah.edu)](mailto:nemsis@hsc.utah.edu) to discuss your project if you have questions related to the unit of analysis for your research question.

9. I have a simple query/question; do I need to request the dataset?
No. The NEMSIS TAC provides a web-based "data cube" similar to a Microsoft Excel pivot table and a variety of dashboards that may meet your data needs. Basic information regarding the data cube is available at EMS Data Cube - NEMSIS. Please contact the NEMSIS TAC if you would like a 30-minute cube/dashboard tutorial.

10. How do I obtain the limited use / geographic identifiers?
The limited use / geographic variables are only available upon request and are not automatically included in the public release research dataset. You may respond to your data request approval email and request which limited use variables you require for your analysis and a brief explanation of how each is to be used, or why they are required. Each limited use variable or geographic identifier will be masked with a random alphanumeric hash so it is not identifiable before being provided as a supplemental table that can be joined into the public dataset. Please contact the NEMSIS TAC if you have questions about the limited use variables or masking process.