

### Supplemental File 3: Grading rubric for evaluating 48 restricted health data sources

Attribute	Grade A	Grade B	Grade C
<p><b>Data Description</b> Description of the data itself, including summary describing the purpose, nature, and scope of the data collection, special characteristics of its contents, major subject areas covered, and major research questions.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Provides description about the data that supports understanding for a broad audience:               <ul style="list-style-type: none"> <li><input type="checkbox"/> Clear dataset description or abstract provided;</li> <li><input type="checkbox"/> How and when the data was generated, collected, and made available, including sources, methods, and type of data.</li> <li><input type="checkbox"/> Frequency and/or dates of collection, major topics/concepts covered, units of observation, geographic coverage, respondents, universe, sampling, etc.</li> </ul> </li> <li><input type="checkbox"/> Describes who is responsible for creating the data:               <ul style="list-style-type: none"> <li><input type="checkbox"/> Who or what organization is responsible for creating or providing the data</li> </ul> </li> <li><input type="checkbox"/> Describes the data's intended use or purpose:               <ul style="list-style-type: none"> <li><input type="checkbox"/> Data's intended audience, reuse examples, projects, purposes, and impacts.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Provides some description about the data, however, more is needed to facilitate understanding and selection for reuse;</li> <li><input type="checkbox"/> Describes some information about who is responsible for aspects of the data, however, the responsible party's role is unclear;</li> <li><input type="checkbox"/> Describes some information about the data's intended use or purpose, however, it is not specific enough to facilitate selection, understanding, or reuse.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Describes very little or nothing about the data;</li> <li><input type="checkbox"/> Describes very little or nothing about who is responsible for the data;</li> <li><input type="checkbox"/> Describes very little or nothing about the data's intended use or purpose.</li> </ul>
<p><b>Data Documentation</b> Detailed, structured information about the data itself that supports its interpretation and use.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Datasets are accompanied by multiple pieces of detailed, structured information about the data itself, as well as additional contextualizing information. Interpretation and reuse of the data is possible with the materials provided. (e.g., data dictionary, codebook, and user guide.)</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Datasets are accompanied by some detailed, structured information about the data itself. More may be required to facilitate interpretation and reuse. One of the following is provided: data dictionary, codebook, user guide, data collection instrument</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Datasets are not accompanied by information about their content and structure. Interpretation and reuse of the data is not possible with the available materials.</li> </ul>
<p><b>Searchability / Browsability</b> IF a data source has more than one dataset, there is the presence of mechanisms for searching and browsing them (e.g., full search interface, table, list, etc.).</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Datasets are searchable within the data source;</li> <li><input type="checkbox"/> Basic keyword search is available;</li> <li><input type="checkbox"/> There is a mechanism to browse datasets by one or more relevant facets or variables (i.e. topic, population);</li> <li><input type="checkbox"/> There is an advanced search interface allowing for boolean search and/or searching of specific metadata fields.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Datasets are searchable within the data source;</li> <li><input type="checkbox"/> Basic keyword search is available.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> There is no mechanism to search or browse the data source.</li> </ul>

**Supplemental File 3: Grading rubric for evaluating 48 restricted health data sources**

<p><b>Metadata Standards</b> The presence of metadata standards applied within the data source. For data sources that do not have individual datasets within them, this section can be applied to the data source as a whole.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Structured metadata is available and structured using one or more recognizable standards (e.g. Dublin Core, ISO 19115);</li> <li><input type="checkbox"/> Metadata elements are employed consistently. There is a set of core elements, including metadata needed to cite the dataset (e.g. title, date, description, keywords, publisher);</li> <li><input type="checkbox"/> Metadata is clearly presented within the web interface.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Structured metadata is available, but may not adhere to a recognizable standard;</li> <li><input type="checkbox"/> Metadata elements are employed consistently; there is a set of core elements that is usually present (e.g. title, description);</li> <li><input type="checkbox"/> Metadata may not be presented in a conventional interface for this tier.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Some metadata may exist but is very limited and unstructured (e.g., description only); OR</li> <li><input type="checkbox"/> There is no metadata.</li> </ul>
<p><b>Data Access Request Process</b> The presence, completeness, and clarity of information and/or content (e.g. application forms) required in order to successfully submit a data access request and to understand the data access request process.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Clearly explains all aspects of the data access request process, including: <ul style="list-style-type: none"> <li><input type="checkbox"/> All application forms needed for data access requests;</li> <li><input type="checkbox"/> Instructions or guidance for completing the application form(s);</li> <li><input type="checkbox"/> A time-line for application review and decision-making with a guaranteed decision-date;</li> <li><input type="checkbox"/> Each step involved in the request process, including the steps/processes undertaken by the data custodian;</li> <li><input type="checkbox"/> Information about how requests are reviewed, by whom, and on what basis.</li> <li><input type="checkbox"/> The technical process for submitting a request is explained in detail;</li> </ul> </li> <li><input type="checkbox"/> A statement indicating that support is available to assist with submitting data access requests with information on how to request support</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Clearly explains most aspects of the data access request process, including: <ul style="list-style-type: none"> <li><input type="checkbox"/> All application forms needed for data access requests;</li> <li><input type="checkbox"/> Indications of any restrictions on data usage;</li> <li><input type="checkbox"/> The time-line for application review and decision-making ;</li> <li><input type="checkbox"/> The technical process for submitting a request is apparent.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Indicates that data are available upon request but provides no clear instructions on: <ul style="list-style-type: none"> <li><input type="checkbox"/> How to submit a request; OR</li> <li><input type="checkbox"/> What a request should include; OR</li> </ul> </li> <li><input type="checkbox"/> Data access request form is not available.</li> </ul>
<p><b>Data Restrictions</b> The criteria for who are eligible to access the data. Data restrictions should inform a user whether or not they would be eligible to acquire the data, and indicate what</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> There is a section devoted to describing the restrictions placed on data;</li> <li><input type="checkbox"/> Specific populations or projects that are permitted to access the data are well described and include examples.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Restrictions on who/what can use the data are mentioned but not described.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> No information or the information provided is not clear enough to ascertain whether restrictions exist.</li> </ul>

**Supplemental File 3: Grading rubric for evaluating 48 restricted health data sources**

<p>parameters are necessary to meet eligibility.</p>			
<p><b>Pricing</b> Costs associated with acquiring and using dataset(s).</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Fees are explained clearly (A site that clearly identifies that no fees for access apply would be graded "A");</li> <li><input type="checkbox"/> A researcher could estimate their access cost based on the provided info;</li> <li><input type="checkbox"/> If a quote process is in place, the quote is expected to be provided early in the process based on a basic research plan submitted by the researchers:             <ul style="list-style-type: none"> <li><input type="checkbox"/> A contact person exists that could estimate the fee;</li> <li><input type="checkbox"/> Example projects exist with the fees charged to the project defined.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Access fees are clearly required;</li> <li><input type="checkbox"/> Enough information is provided to understand approximately the cost, but there is no way to estimate what the costs would be;</li> <li><input type="checkbox"/> If a quote process is in place, a great deal of work is required from the Requester before the pricing information will be made available to them.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> No information on fees is present; OR</li> <li><input type="checkbox"/> Access fees are clearly required, but a Requester cannot interpret the magnitude of these fees.</li> </ul>
<p><b>Contact Information</b> A contact person to support a data requester with understanding the data or completing the application process.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> There is a named individual and title/role as a designated contact <b>or</b> an email address designated for questions about data or access;</li> <li><input type="checkbox"/> This information is clearly displayed and accessible from within the portion of the website which displays information about access or the process of applying for access.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Information for a contact or mailbox for requests related to applications is provided;</li> <li><input type="checkbox"/> Requesters do not need to search for contact info elsewhere in the site.</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> No contact information; OR</li> <li><input type="checkbox"/> A generic "contact us" email as.</li> </ul>