CRUK Rubric - Basic and Clinical Research v2.0

Funder Template	CRUK - Basic and Clinical Research			
Purpose of rubric	Providing feedback to researchers			
Notes	CRUK have subject-specific guidance in the following areas: Discovery research (http://www.cancerresearchuk.org/sites/default/files/hands_on_data_sharing_advice_-_basic_science.pdf) Clinical research (http://www.cancerresearchuk.org/sites/default/files/hands_on_data_sharing_advice_-_clinical.pdf) Population research (http://www.cancerresearchuk.org/sites/default/files/hands_on_data_sharing_advice_-_population.pdf)			
Documents Used	Cancer Research UK Policy on Data Sharing and Preservation Cancer Research UK Practical guidance for researchers on writing data sharing plans (http://www.cancerresearchuk.org/funding-for-researchers/applying-for-funding/practical-guidance-for-researchers-on-writing-data-sharing-plans) Discovery research Subject-specific guidance (http://www.cancerresearchuk.org/sites/default/files/hands_on_data_sharing_advice_-_basic_science.pdf)			
Version control	v0 - basic document in development v1.0 - completed first draft for feedback v1.1 - revised and reordered in response to CRUK feedback v2.0 - formating edited for download purposes.			
	Performance Criteria	Performance Levels		
		Detailed	Addressed but incomplete / unsatisfactory	Not addressed
	What type of data will be in the final dataset?	Data types clearly defined. Eg imaging data, genotypic data, clinical measurements, survey data etc. It is important to clearly state which data can be shared, which data cannot be shared and why.	Data types mentioned for some of project / dataset but not all. No indication as to which data may or may not be shared. Reasons for data sharing suitability might be missing.	Data types are not mentioned.
	What scale / volume of data will be in the final dataset?	Clear estimate of dataset size given for each data type.	Dataset size given but not broken down by data type. Size not give for all data types. Dataset size is clearly unrealistic (not always possible to judge!).	Dataset size is not mentioned.
	What format of data will be in the final dataset?	Data formats clearly defined. Eg spreadsheets in .csv or .xlsx; micrographs in .tiff or .jgp; proprietary manufacturer formats where necessary.	Data formats are mentioned for some of dataset but not all.	Dataset formats are not mentioned.
	What are the restrictions on sharing the data publicly?	There is a clear assessment of any ethical, IPR or patient confidentiality concerns. Methods by which these can be mitigated are also discussed. Eg material transfer agreements (MTAs), restricted access, well-designed consent procedures. Alternatively, there is a clear statement that there are no restrictions on this dataset.	Data sharing restrictions or problems are mentioned, but without any details. Data sharing restrictions are mentioned, but no plans to mitigate these are discussed.	Data sharing restrictions are not mentioned.
	How will the dataset be stored during the project?	Clear description of data storage systems. Eg departmental server, on machine, on portable hard-drive. Ideally this section should include an assessment of the suitability of the storage and the security implications for the data.	Mention of data storage systems, but lacking detail or clearly inappropriate (could be difficult to judge).	Dataset storage is not mentioned.
	What is the long-term preservation plan for the dataset?	Clear strategy for long-term preservation of dataset, including retention period. Eg deposit in an appropriate responsible repository. If researchers are intending to use public repositories to archive their data, these should be specified. Alternatively, a clear statement that dataset won't be preserved / is not suitable for preservation.	Preservation is mentioned but strategy is not clear or lacks detail.	No mention of preservation of dataset.

	What standards will be used in the management of the data?	Data management methodology is clearly stated or referenced. eg file naming conventions, file architecture etc.	Methodology is only mentioned for a subset of the data to be collected or is lacking detail.	No methodology is mentioned.
	What metadata will accompany the dataset?	Clear outline of metadata strategy with references to existing good practice in the community eg MIAME for microarray or CDISC for clinical research. Detailed project-specific approach where community standards don't exist.	Some mention of metadata without detail about community standards or a project-specific approach.	Metadata is not mentioned.
	What supporting documentation will accompany the dataset?	Clear outline of documentation with references to existing good practice in the community or detailed project-specific approach where community standards don't exist.	Some mention of documentation without detail about community standards or a project-specific approach.	Accompanying documentation is not mentioned.
	How will the shared datasets be discoverable?	A discoverability strategy is clearly planned. This may include the use of data registries, repositories, data publications, word-of mouth, use of DOIs and data citations in research articles etc.	Data discoverability is mentioned, but strategy is not clear or is lacking detail.	No data discoverability strategy is mentioned.
	How will the data be shared?	Clear consideration of where, how and to whom the data will be made available. Strategy is in line with good practice in the area of research (if able to judge!). Assessment of specific access mechanisms if needed. Will data availability be advertised in publication via data citation? If researchers are intending to use public repositories to share their data, these should be specified.	Some mention made of how the data will be shared but details missing.	Data sharing is not mentioned.
	Under what conditions will data-reuse be permitted?	Clear indication how each subset of data will be available eg openly available, requirement for data sharing agreements, approval by data access committee. Indications of type of restrictions imposed by data transfer agreements. Logistics of data transfer may also be explained eg the use of secure file transfer if necessary for sensitive data.	Data re-use conditions are mentioned, but details are missing, either overall or for specific subsets of data.	Conditions for data re-use are not mentioned.
	Are you planning for a period to enjoy exclusive use of your study data?	A clearly defined period for exclusive use is given along with a justification for that period. This may depend upon the nature and value of the data and the way in which they are generated and used. Alternatively, there is a clear statement that a period for exclusive use is not required (this is unlikely, but possible).	A period for exclusive use is mentioned, but is not clearly defined, or is lacking justification.	A period for exclusive use is not mentioned.
	When will the datasets be publicly released?	A clear timescale is indicated eg no later than the publication of the main findings of the research (funder recommendation), at the end of the award, within 3 years of the generation of the dataset. Where a delay to release is indicated, reasons are given.	Timescale is mentioned but not clear or not clear for all datasets. Timescale is clearly not in accordance with funder expectations.	Timescales for data release are not mentioned.