**1. DATA SUMMARY**

**1.1 Description of data**

1. What is the origin of the data? (for each data type, source, project, funding)

2. What data types, dataset types and formats do you have?

3. What is the size of the data? (for datasets, number of variables and subjects, volume in MB/GB/TB)

**1.2 Data security levels**

4. Do you have different security levels for different data types/ dataset type or specific variable types, and for metadata o data variable description? If so, please specify the levels of data security (access levels) you consider.

**2. FAIR DATA MANAGEMENT**

**2.1 Data collection/ generation**

5. How is each data type collected? (methodology for depositing new cohort raw data into your appropriate repositories and methods to provide rich metadata to foster a quality reuse)

**2.2 Data storage**

6. Where is your data stored? (certified repositories? different location/repositories for each data type?)

7. What are the provisions for long-term secure storage during and after the ending of the project?

**2.3 Making data findable, including provisions for metadata**

8. Do you provide metadata for discoverability?

9. Is your data identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

10. What naming conventions do you follow? (e.g. for datasets names: Data\_<WPno>\_<serial number\_of\_dataset>\_<dataset title>.)

11. Do you provide search keywords to optimize possibilities for reuse?

12. Do you provide clear version numbers?

13. Do you have data overviews/ aggregate statistics?

14. What metadata do you consider for each data type, dataset type? (metadata standard). In case there is not a metadata standard for your variables, please outline what type of metadata that each of your data types/dataset type contains.

**2.4 Making data openly accessible**

15. What data, metadata do you keep openly accessible, if any (for each> raw data, variable names, variable names and number of data available per variable, variable aggregated statistics?)

16. Which data (data types, variables) should be accessible under an authenticated level, if any?

17. Which data (data types, variables) should be accessible under a controlled-access level?

18. What methods, technology/software tools are needed for access/ transfer of data?

19. Do you include the relevant software (e.g. in open source code) and its documentation?

20. Are there well described conditions for access? how is access provided? (DAC tools?)

21. Will you use euCanSHare´s centralized data access manager interface and tools? If not, describe methodologies and conditions for access (i.e. a machine-readable license)?

22. Will the metadata (and variable description) be made accessible by deposition in the euCanSHare´s centralized repository - centralized Mica-server)?

23. Will links to metadata (and variable description) in external repositories be provided?

**2.5 Making data interoperable**

24. Are your data interoperable, that is allowing data exchange and reuse between researchers, institutions, organizations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

30. What metadata you consider as “required metadata” for each data type you handle (minimal required metadata - think of it as the metadata required for between-cohort comparison and ultimately multi-cohort analysis)?

31. What data and metadata vocabularies, standards or methodologies do you follow to make your data interoperable?

32. In case it is unavoidable that you use uncommon or project-specific ontologies or vocabularies, do you provide mappings to more commonly used ontologies?

**2.6 Increase data reuse (through clarifying licenses)**

33. How is the data be licensed to permit the widest reuse possible?

34. How long is it intended that the data remains reusable?

35. What data quality assurance processes are in place?

**4. DATA SECURITY**

37. What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

**5. ETHICAL ASPECTS**

38. Are there any ethical or legal issues that can have an impact on data sharing?

39. Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?

**6. OTHER ISSUES**

40. Do you make use of other national/funder/sectorial/departmental procedures for data management? If so, which ones?