Bayer Clinical Trial Data Transparency
Anonymization and Data Protection Procedures

The objective of this document is to describe the high level steps that are required to perform anonymization of all subject level data in response to requests for such information via Vivli.

General Approach
1. Data will be provided in the format utilized for analysis and stored within the sponsor company, e.g., SAS datasets.

2. Data will be provided for single studies in conformance with the criteria outlined in Bayer's scope statement as posted on the Vivli platform.

3. Anonymized data for all subjects will be included in the provided data, unless the subject has explicitly declined inclusion of their data. Such subjects must be flagged in the source data, so their data can be removed as part of the data anonymization process.

Data Anonymization
The following modifications will be applied to the data:

1. Subject numbers will be replaced with a different identification number in the anonymized data. The algorithm to translate subject numbers from the original number to the anonymized number will be deleted.

2. Age will be kept as reported unless the subject is greater than 89 years old. When the subject is greater than 89 years old, the age field will be made blank.

3. Age categories as they were used for analysis will be checked to ensure that the category breakdowns will not indicate specific ages greater than 89 years old. If necessary, age categories for subject greater than 89 years old may be collapsed together.

4. Individual month and day date parts for the date of birth will be made blank for all subjects 89 years of age and younger, while the birth year will be maintained as originally entered. All individual date parts (month, day, year) for date of birth of subjects greater than 89 years of age will be made blank.

5. Subject initials will be removed from the database, and the field left blank.

6. Any medication / medical device indicators reported with the exposure data, such as bulk patch identifier, lot numbers, or serial numbers, will be removed from the database, and the field(s) left blank.

7. Code and decode values of the coded investigator and site identification fields will be removed from the database, and the fields left blank.

8. SAS formats used to decode investigator names and site names and locations will be removed from the SAS format catalog that accompanies the data.

9. For all dates other than date of birth, the values of the individual date parts (month, day, year) will be removed from the database and the fields left blank. Associated relative days will remain, when populated. Such dates include but are not limited to: visit dates, adverse event dates, disposition dates, exposure dates, medical history dates, and concomitant medication dates.

10. Verbatim terms for adverse events, medical history, and concomitant medications will be removed from the database, and the fields left blank. Coded term values will remain populated.

11. Any comments dataset, e.g., CO or ADCO, will not be provided.

12. A knowledgeable medical person familiar with the study will provide additional study specific or subject specific criteria that should be considered in the anonymization process and removed from the study data. Such criteria may include spontaneous reporting of a rare disease as a medical history finding or adverse event, small number of a specific event reported in the study, strange events, unusual treatments, etc.

Quality control checks are performed and documented following creation of the anonymized data and accompanying metadata documentation.

The anonymized data are stored in a separate secure location from the original study datasets.

In addition, any requested documents will be provided following a similar anonymization process.

Any study, where the dataset cannot be anonymized, according to Bayer assessment, will not be transferred to the Vivli platform. Cases of doubt will be forwarded to the Bayer Data Protection officer for consultancy and final decision. If the data protection officer decided that a dataset cannot be anonymized, a written justification will be provided and will be publicly available on request.