SOP38c AD2: Guidance on completing the Database Validation Form

No.	Database information
1.	Describe the intended use of the database and justification for choice *Does the intended use meet the criteria of the study protocol?* *What data will be captured?*
2.	What company developed the database software? *Off the shelf database?* *Study specific configuration?* *Bespoke?*
3.	What are the requirements of a database for this study? Consider what conditional formatting will be used (e.g. Only permit values between 16-99 in “Age” field). *Drop down lists* *Cell formatting* *Formulae* *Ranges* *Fields accept the correct data type* *Field length* *Field only accept valid answers and flag errors* *How to differentiate between unknown and missing values*
4.	What Standard Operating Procedures (SOP’s), manuals, or guidance documents are available for this data processing system? *List all available and relevant documentation associated with the database*

No.	Database information
5.	Is any installation documentation from the manufacturer or developer available? *Approved specifications* *User Manual* *Testing documentation from developers (Verification)* *Testing documentation from users (Validation)*
6.	Will the system interact with any other software (e.g. data imported from external sources, or exported to other software)? If so, describe how these systems will interact and be tested. *Dummy data transfer/merger testing* *Test v Live database field*
7.	Where will the database be hosted? If it will not be hosted on Barts Health NHS Trust or Queen Mary University of London ICT servers, please justify the reasons. *Consider security, encryption, cloud usage*
8.	Will a data management plan be put in place? If not, please justify why one is not needed. *Simple studies may not require a dedicated data management plan. In such cases, the research team should be ensuring that activities like data entry, data verification, data lock and analysis are fully described in the protocol.*
9.	If blinding is required, how will this be maintained? *Consider a Data Management personnel access chart*
10.	How will users be trained in using the database? *Training records, competency records, continuous assessment, re-training* *Designated training officer*

No.	Database information
11.	How will data queries be managed? *Consider designated data management personnel*
12.	How will an audit trail be maintained? It should be possible to trace when any changes are made to the data, who made them, and when. *Will autosave be activated?* *What is considered a change?* *Will this be a real-time database?* *Will a copy or backup of the database be saved separately after data input?*
13.	Is a maintenance plan / schedule in place? *Will the database be re-validated following maintenance?*
14.	Is there technical customer support available? *Customer support* *24-hour support* *Is there a contingency plan in place?*
15.	How will access to the data processing system be managed? Who will be responsible for managing access? *Consider a Data Management personnel chart*
16.	How will version changes to the database be managed? *Define version control* *When will version changes take place?* *Will the database be re-validated following change?*

No.	Database information
17.	How will the data processing system be backed-up? *Consider accidental loss, environmental factors*
18.	What security measures will be in place to secure the database and data? *Password protection* *Data management chart*
19.	How will the database be locked when all data input and cleaning is completed? *Describe what process you will follow*
20.	How will data be exported from the database for analysis? How will it be sent to the statistician? *Consider what export function the database will have*
21.	How will the database and data be archived at the end of the study? *Local digital archiving facilities* *External digital archiving companies* *Transfer data to disk* *Be mindful that data should be archived but remain accessible for 25 years.*