PEER REVIEWED

Data Management in Veterinary Clinical Studies

Data management tasks and procedures with their direct influence on data quality and compliance are an aspect of great importance for any clinical study. Standards and guidelines define basic requirements for data management but provide little detail on requirements for hardware equipment and specific software specification. There are multiple options and solutions available, all promising to provide the ultimate productivity and effectiveness for data management. However, the devil is in the detail depending on the objectives of a study, the environment to work in and the day-to-day challenges related to the conduct of veterinary clinical studies.

Data managers have to consider a lot of opportunities and solutions for setting up data management routines. Computerised systems, regardless of whether they are in the context of a veterinary clinical study or not, will consist of individual components, unless a complete (software) package offering all parts of such a system is used. For an electronic data capture (EDC) system, requirements regarding traceability and integrity are the same as for processing data 'conventionally' on paper. In the EU, major exigencies detailing requirements for data management procedures are provided in VICH GL9 'Guideline on good Clinical Practices (GCP)', that is linked to legislation via the European Commission Directive 2001/82/EC, as amended by Directive 2009/9/EC. In the US, VICH GL9 is also applicable. There the Food and Drug Administration's (FDA) 21 CFR part 11, amongst others, defines criteria for (electronic) records and requires mandatory signatures within clinical studies. Additionally, several associated 'Guidance for Industry' documents released by the FDA provide recommendations for various issues of data processing, including software validation. Other available guidance documents were initially issued for use in data management of human clinical studies, such as 'Good Practices for Computerized Systems in regulated 'GxP' environments' from the Pharmaceutical Inspection Co-Operation Scheme (PIC/S) and 'Good Clinical Data Management Practices' published by the 'Society for Clinical Data Management'. Their recommendations and guidance may be very useful to facilitate good data management practices for veterinary clinical studies. The European Commission's Guide to 'Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use, Annex 11: Computerised Systems' also provides appropriate guidance and is transferrable to use in veterinary clinical studies.

This article will briefly illuminate some aspects of data management procedures within a clinical study, from initial planning to providing the final data pack. Planning a clinical study, the team establishing the protocol and related procedures should always involve a well trained and experienced data manager from the very beginning.

Designing a clinical study, possibly the first topic with relevance for data management should be the decision of which medium for raw data collection will be used for data capture: shall it be the traditional way of recording the data on paper (paper-based) or is it justified to use an electronic data capture system for immediate recording (EDC-based). Decision-making criteria whether to use paper-based or EDC-based data capture will need to consider various aspects of the study design, but also feasibility and practical aspects. When deciding to use an EDC system, certain aspects need to be taken into account, as today there is a variety of EDC systems available on the market. A certain number of EDC solutions were initially designed for data capture in human clinical studies and can also be used in veterinary clinical studies: All allow the registration of individual animals, as this is comparable to the requirements of a Phase III study in human medicine. While very suitable for use in small and companion animal studies, adaption of such systems to the study environment on farms may sometimes be impossible due to the data structure these systems provide. So first of all, the location of generating raw data, in a veterinary practice, under laboratory conditions or on a farm, will make a relevant difference. Far away from urban infrastructure, internet availability and transfer rates may also become an issue when intending to use an EDC system in remote areas. Secondly, the study design plays an important role, e.g. when individual animals are enrolled one by one, the processes that basically all EDC systems support by their structure for individual subject registrations may be a good fit. But if cases are rather enrolled by larger groups or flocks at once, or where varying time points require high flexibility, this may reduce the options for use of an EDC system significantly. In certain livestock or poultry studies, animals are to be enrolled in batches, per pen or even as a complete barn/house together, followed by group/pen-wise treatment. Parallel to group registration, the individual animal's data will need to be captured precisely, as finally for data analysis the experimental unit will be the individual animal. As flexibility of some EDC solutions is limited, workarounds for pen- or batch-wise registration instead of individual animal registration may need to be taken into account. However, single EDC systems offer the flexibility to handle great animal figures and allow subgroup registration together with individual data capture. Any EDC solution that can cover all such options will definitely be of great help, but will indeed require a huge effort in setting it up accordingly prior to the start of such a study. If such a study is conducted in a multicentric setting, and workflows at each farm are different, this will complicate any use of an EDC system even further. Wherever an EDC system is intended to be used, before deciding on a solution, the available computer system as well as internet infrastructure at site, especially regarding data transfer rates, will require thorough checking to be able to ensure adequate performance during the course of the clinical study.

Nevertheless, there is no doubt of the attractiveness of using an EDC system providing the possibility to capture data immediately at site, without subsequent transcription or data entry, as a paper-based data capture model would require. While the setup time and costs may be higher for

an EDC system, it will help to save resources of all involved parties overall in many cases, the advantage of direct data availability for monitoring and plausibility checks included. When deciding to use paper-based DCFs, it has to be kept in mind that appropriate setup for (double) data entry and related quality control will have to be provided for.

After considering all these facts, it is strongly recommended to summarise all required data management tasks and their timelines in a data management plan. This allows for transparency and traceability, but even more for reliability of the agreed process. While the data management plan should define in sufficient detail all relevant processes, it must be shared with all parties involved in the clinical study to ensure the professional conduct of the clinical study and the collection and generation of true, accurate and integer data.

The first use of EDC solutions in veterinary clinical studies was made in the second half of the first decade of this century. These studies were all pet studies, allowing individual animal registration and data collection. In today's versions of certain EDC systems specifically designed for use in veterinary clinical studies, additional features necessary for use in studies on farm or under laboratory conditions are included. A lot of the initial challenges for using them under such conditions have been addressed in the meantime. In these more flexible EDC solutions, the data manager has the possibility to choose from different designs when setting up an individual study. For data capture, scenarios based on pens or batches are available and for data processing, with respect to analysis, the final focus is on the individual animal's data.

Having reviewed more than 10 available systems for their use in veterinary medicinal studies, the use in companion animal studies is possible in any EDC solution in order to handle multicentric field studies; they offer the possibility to define and keep separately different study sites, different roles of people to be involved, archiving and data transfer. It is obligatory that an EDC system provides options for randomisation and blinding to treatment towards any treatment.

A very useful functionality is an online randomisation tool. Different approaches are implemented; many EDC systems offer an upload of a predefined randomisation list. Based on that, following the completion of a key value, allocation to treatment group is done. In some EDC solutions, a real-time randomisation is implemented. Data managers have to consider carefully which solution is more appropriate for the chosen study, as both have their advantages. Real-time randomisation leaves aside the risk of upload errors or breaking the blinding code by inadvertently showing predefined randomisation information to anyone involved who must not know. Uploading complete randomisation lists prepared outside the EDC system, either individually for each centre (site randomisation) or together for all sites (global randomisation) leaves the option to provide the study sites with sealed printout copies that can be used as a fall-back position in case of e.g. internet failure.

Choosing the 'conventional' paper-based system, data managers will have to set up and provide appropriate data entry tools for digitalisation of data captured on hardcopy into electronic data by any kind of data entry. EDC systems may be used as data entry platforms for manual input from distant locations; some EDC solutions are already prepared for double data entry with immediate data comparison, with the results displayed in the (web) user interface. Apart from using existing EDC solutions for manual data entry from paper forms, various scenarios for transferring the data in digitalised form are possible. Simple or complex data entry spreadsheets are quite commonly used, as well as data entry masks representing a user-friendly front-end to a relational database. Automatic recognition systems with transfer of data in a database are also still used. No matter which data entry tool is finally used, it will have to be set up in a way that one can trust the process to correctly reflect the raw data: All setups need to be validated. This is applicable equally for a study-specific EDC setup, a spreadsheet or a database-associated solution. Usually, unless the underlying software is validated by the vendor, validation for the intended purpose is required. In all cases, study-specific setups with their project-related masks and functionalities will need to undergo testing and approval. Post-approval adaptions of study-specific setups will require change management procedures including appropriate documentation. All EDC systems known by the authors offer a full audit trail including change control; which allows the control of changes of data initially entered and modified, but also allows tracing of any later changes. Some also include in the audit trail the setup history of visits, forms and fields which provides the currently highest standard of transparency for documenting changes within an EDC system.

Any kind of electronic data capture system, whether an EDC solution or internally built components, should allow data managers and clinical monitors to check data under different aspects. Ideally, a system allows the implementation of immediate checks for presence/absence of values for mandatory parameters, and to set up checks for plausibility of entered values. Any implementation of checks should not prevent or restrict data being collected by the responsible person from entering such (realistic) data. Many EDC systems also provide communication tools to allow and facilitate direct communication between the person entering the data and the clinical monitor, the sponsor or any other person authorised; this also includes flagging or closing queries. Additional features that assist data managers and monitors in routine work (pre-defined scenarios) are offered. These features are, e.g., functionalities that allow the creation of individual data summaries, general and subject-specific study schedules, and complex checks of dependent or related parameters within a single data entry mask or amongst different forms and setup of email notifications. The latter may include notification of when a new subject has been enrolled, when an adverse event form has been completed, or outstanding next visits by patients. Performing 'traditional' monitoring and parallel or subsequent retrospective data entry from hardcopy may require covering the above (applicable) issues by individual components as time points in workflow may vary. Data managers and project managers or monitors will have to coordinate their actions accordingly. Once data in the database is clean, no further corrections are required due to the intensive quality control during the process, including checking for plausibility and solving all queries. Any data captured or digitalised in another language than English may need to be translated. Where required, coding of selected parameters may have to be performed using standardised coding dictionaries (ex. VeDDRA, ATCvet). Some EDC systems offer coding functionalities that already can be included in the data capture process and include common coding libraries. Data management solutions not

taking advantage of such libraries may require individual approaches.

There may also be other data produced which it may be valuable to have available in the central database. Most of the time, this data is generated by a laboratory, performing certain tests and also generating such test results electronically. Laboratory analysis data is usually part of statistical analysis, so it becomes necessary to import such data into the database before finalising the data pack. Other data that may be part of the final data pack may also be X-ray or ultrasound related files, or automatically recorded data like milk yield, cell counts or feed consumption data. Any data import needs to be planned well in advance, in fact prior to the collection of such data in systems that are not under the control of the data manager or a study monitor. Similar criteria applicable for validation of an EDC system are also applicable for any software collecting or generating raw data. It is highly recommended to address these topics during the planning phase of a study. Finally, after setup and approval of import routines, imported data should be handled like other study data and undergo all routines from plausibility checks to coding – whatever is required and useful.

After finalising the clean data pack, this data pack will have to be transferred into a format suitable for further processing by the statistician. Several EDC systems already offer the export of data to SAS, XML or other formats. Where no such prepared functionalities are available, data managers may need to develop appropriate routines. The ultimate objective will remain that data integrity and accuracy is maintained during data transfer, as well as security and confidentiality of the data needing to be assured. All of these topics are influenced by local IT infrastructure of the parties involved.

In summary, EDC solutions on the market offer a variety of options. It is worthwhile comparing systems for their scope of available features. They do offer easy and transparent options for data management in veterinary clinical studies. Data management needs to ensure accurately described processes that need thorough planning, transparent setup and validation before release. This includes a full audit trail for tracking of any changes made. Starting with careful consideration of which data needs to be captured (and what), a data management plan should always address the following questions: Who, Where and How it shall be done. It should also include all processes that ensure the objective of a clean data pack, and detail routines that allow safe transfer of correct, accurate and integer data to the partners for analysis. In general, implemented routine data management procedures should be transferable to different kinds of studies, assist involved parties in their tasks, and help to increase efficiency in the conduct of veterinary clinical trials.

Dr Brigitte Steiner

She is a veterinarian with over 14 years' experience in the animal health industry. After a couple of years in veterinary practice, she joined KLIFOVET as a project manager for veterinary clinical studies. In 2010, she moved into data management and focused on data management procedures, development and validation of EDC systems specifically developed for animal health. Since 2017, Brigitte has been heading the Data Management and IT group at KLIFOVET AG.

Email: brigitte.steiner@klifovet.com

Dr Klaus Hellmann

He is a veterinarian with over 23 years' experience in the animal health industry. Prior to starting Klifovet in 1997 as a CRO, he worked in veterinary practice, science and the animal health industry. He is a Board-certified veterinary pharmacologist, Diplomate of the European College of Veterinary Pharmacology and Toxicology and holds the title of an Auditor (European Organisation of Quality).

Email: klaus.hellmann@klifovet.com