Epilepsy Research 129 (2017) 87–90

Contents lists available at www.sciencedirect.com

Epilepsy Research

journal homepage: www.elsevier.com/locate/epilepsyres

Special Communication

Common data elements and data management: Remedy to cure underpowered preclinical studies

Niina Lapinlampi[a], Esbjörn Melin[b], Eleonora Aronica[c], Jens P. Bankstahl[d], Albert Becker[e], Cristophe Bernard[f], Jan A. Gorter[c], Olli Gröhn[a], Anu Lipsanen[a], Katarzyna Lukasiuk[g], Wolfgang Löscher[h], Jussi Paananen[i], Teresa Ravizza[j], Paolo Roncon[k], Michele Simonato[k], Annamaria Vezzani[j], Merab Kokaia[b], Asla Pitkänen[a,*]

[a] Department of Neurobiology, A. I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, PO Box 1627, FI-70211 Kuopio, Finland
[b] Epilepsy Center, Wallenberg Neuroscience Center, Lund University, Lund, Sweden
[c] Academic Medical Center, Dept (Neuro)Pathology and Center for Neuroscience, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Stichting Epilepsie Instellingen Nederland, Heemstede, The Netherlands
[d] Department of Nuclear Medicine, Hannover Medical School, Hannover, Germany
[e] Translational Epilepsy Research Section, University of Bonn Medical Center, Bonn, Germany
[f] Université d'Aix Marseille, Marseille, France
[g] Laboratory of Epileptogenesis, Nencki Institute of Experimental Biology of Polish Academy of Sciences, Warsaw, Poland
[h] Department of Pharmacology, Toxicology, and Pharmacy, University of Veterinary Medicine, Hannover, Germany
[i] Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland
[j] Department of Neuroscience, Experimental Neurology, Mario Negri Institute for Pharmacological Research, Milan, Italy
[k] University of Ferrara and University Vita-Salute San Raffaele, Milan, Italy

ARTICLE INFO

Article history:
Received 16 October 2016
Accepted 19 November 2016
Available online 22 November 2016

Keywords:
Common data element
Database
Data management
Epileptogenesis
Epilepsy

ABSTRACT

Lack of translation of data obtained in preclinical trials to clinic has kindled researchers to develop new methodologies to increase the power and reproducibility of preclinical studies. One approach relates to harmonization of data collection and analysis, and has been used for a long time in clinical studies testing anti-seizure drugs. EPITARGET is a European Union FP7-funded research consortium composed of 18 partners from 9 countries. Its main research objective is to identify biomarkers and develop treatments for epileptogenesis. As the first step of harmonization of procedures between laboratories, EPITARGET established working groups for designing project-tailored common data elements (CDEs) and case report forms (CRFs) to be used in data collection and analysis. Eight major modules of CRFs were developed, presenting >1000 data points for each animal. EPITARGET presents the first single-project effort for harmonization of preclinical data collection and analysis in epilepsy research. EPITARGET is also anticipating the future challenges and requirements in a larger-scale preclinical harmonization of epilepsy studies, including training, data management expertise, cost, location, data safety and continuity of data repositories during and after funding period, and incentives motivating for the use of CDEs.

1. Introduction

Due to many failures in translating promising preclinical treatments into clinic, there is increasing concern that the pharmaceutical industries' interest in brain-related diseases will vanish. Consequently, there will be no novel, more efficient, and better tolerated treatments for neurological and psychiatric diseases, including epilepsy. The problems in translation have been related to models used, differences in pathophysiology of the disease between experimental models and humans, and importantly, the lack of statistical power and reproducibility of preclinical studies (Simonato et al., 2014; Steward et al., 2012). Small sample sizes have led to low statistical power, and consequently, overestimation of effect size and poor reproducibility (Button et al., 2013).

Part of the problem is that the data between laboratories are incomparable because of miscommunication, bias in reporting, lack of standardized data collection guidelines (Landis et al., 2012) and diverse experimental procedures. Furthermore, continuously increasing amount of data requires commonly shared scientific practices and good data management to ensure data integrity,

* Corresponding author.
E-mail address: asla.pitkanen@uef.fi (A. Pitkänen).

http://dx.doi.org/10.1016/j.eplepsyres.2016.11.010
0920-1211/

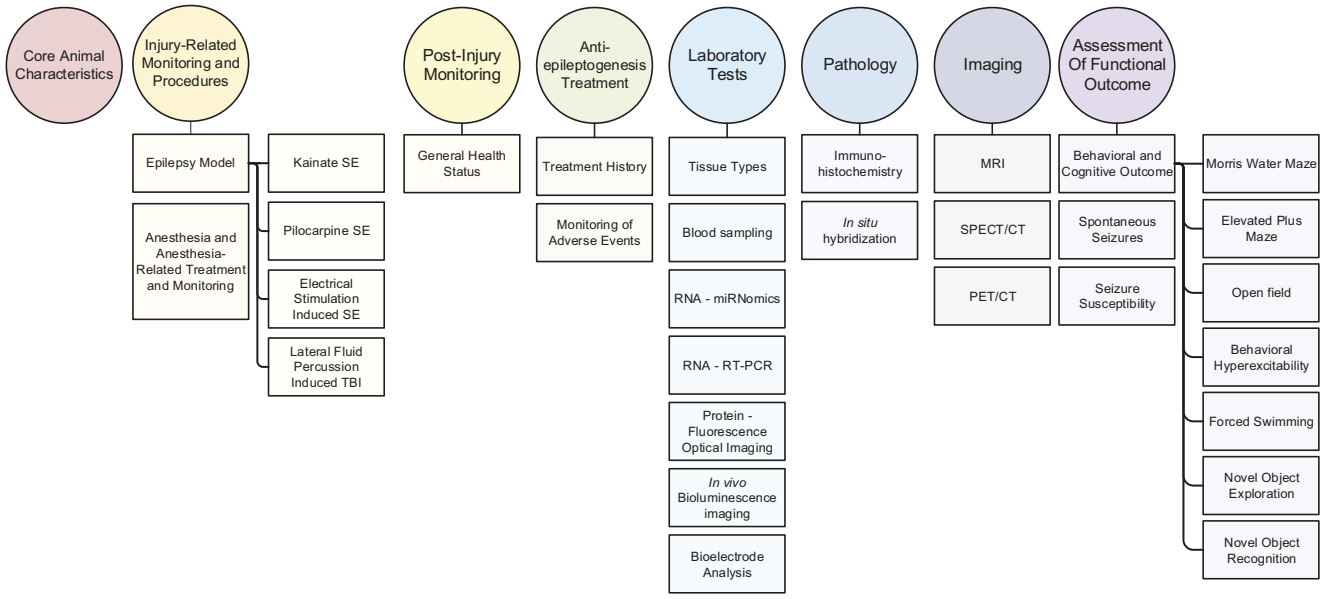

Fig. 1. EPITARGET Common Data Elements (CDEs) are constructed into eight main modules: Core animal characteristics, injury-related monitoring and procedures, post-injury monitoring, antiepileptogenesis treatment, laboratory tests, pathology, imaging, and assessment of functional outcome. The 8 main modules include multiple sub-modules. For example, behavioral and cognitive outcome/spontaneous seizures/seizure susceptibility sub-modules are under "assessment of functional outcome" module. Each sub-module has a specific case report form (CRF) which contains the CDEs related in a logical order.

findability, accessibility, interoperability, and reusability between researchers and research groups.

Until recently, there have been very few attempts for harmonization of practices in preclinical studies (Lemmon et al., 2014; Smith et al., 2015). In 2010, however, the National Institutes of Health (NIH) launched an initiative called National Institute of Neurological Disorders and Stroke (NINDS) Common Data Element project (Stone, 2010) that has led to the generation of common data elements (CDEs) and case report forms (CRFs) for more than 10 neurological diseases; these CDEs and CRFs can be used to harmonize clinical trials (https://www.commondataelements.ninds.nih.gov).

The need for standardization of preclinical studies between laboratories was recognized in the European Union (EU) 7th Framework (FP7) funded project "Targets and biomarkers for antiepileptogenesis" (EPITARGET), a consortium of 18 partners in 9 European Countries, 12 of which conduct preclinical studies. This led to the design of the first available CDEs for preclinical studies on epilepsy to help investigators to systematically collect, analyze, standardize and share preclinical epilepsy data. These CDEs and CRFs can now be downloaded from the EPITARGET web page (www.epitarget.eu). Here, we briefly summarize the procedures employed for generation of CDEs, the lessons learned, and the anticipated challenges ahead. We will also briefly discuss the good data management practices.

2. Methods

2.1. Terminology and generation of EPITARGET CDEs

A CDE can be defined as a basic unit which is common across all the study subjects. Examples include animal species, background strain, vendor information, and sex of an animal. CDEs can be divided into general core CDEs, disease-specific core CDEs, supplemental – highly recommended CDEs, supplemental CDEs, and exploratory CDEs ("Glossary," 2016).

Core CDE is a data element that collects essential information applicable to any study, including either those which span across all disease and therapeutic areas or those that are specific to one disease area. ***Supplemental-highly recommended CDEs*** are data elements which are essential based on certain conditions or study types. ***Supplemental CDEs*** are data elements which are commonly collected but whose relevance depends upon the study design or type of research involved. ***Exploratory CDEs*** are data elements that require further validation, but may fill current gaps in the CDEs and/or substitute for an existing CDE once validation is complete.

The CDEs describing the elements belonging to the same procedure (*e.g.*, CDEs for a given behavioral test such as Morris water-maze) are logically organized into a CRF. Next, CRFs are organized in modules, collating the CRFs related to the same entity (*e.g.*, "Imaging") (Fig. 1).

To tailor the EPITARGET CDEs according to the project needs, EPITARGET partners formed working groups to generate the CDEs, CRFs, and Guidelines in their areas of expertise, for example, in modelling of epileptogenesis, behavioral testing, blood analysis, or imaging. Documents underwent several iterations over a 1-y period, during which the working groups communicated via teleconferences and workshops.

EPITARGET CRFs were organized into eight main modules: core animal characteristics, injury-related monitoring and procedures, post-injury monitoring, antiepileptogenesis treatment, laboratory tests, pathology, imaging, and assessment of functional outcome. The main modules were further divided into multiple sub-modules, describing the variables tailored to represent the experimental designs of EPITARGET (Fig. 1).

2.2. Implementation

The preclinical EPITARGET CDEs are currently available at the EPITARGET webpage (http://www.epitarget.eu/cdes/). Based on the CDE collection, a data dictionary was built. A data dictionary is a metadata repository, defining every variable and their relationship to other variables. The metadata contained in the data dictionary was used to structure the data collection instruments for the EPITARGET database created in Research Electronic Data Capture (REDCap) (Harris et al., 2009).