



# Collection, verification, sharing and dissemination of data: the CONTRAST experience



Christopher F.L. Saarnak<sup>a,\*</sup>, Jürg Utzinger<sup>b,c</sup>, Thomas K. Kristensen<sup>a,d</sup>

<sup>a</sup> DBL, Department of Veterinary Disease Biology, University of Copenhagen, Thorvaldsensvej 57, DK-1871 Frederiksberg C, Denmark

<sup>b</sup> Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, P.O. Box, CH-4002 Basel, Switzerland

<sup>c</sup> University of Basel, P.O. Box, CH-4003 Basel, Switzerland

<sup>d</sup> School of Biological & Conservation Sciences, Faculty of Science and Agriculture, University of KwaZulu-Natal, South Africa

## ARTICLE INFO

### Article history:

Received 21 February 2013

Received in revised form 14 May 2013

Accepted 15 May 2013

Available online 22 May 2013

### Keywords:

CONTRAST

Data curation

Data sampling

Data sharing

Database management

Schistosomiasis

## ABSTRACT

The scientific community is charged with growing demands regarding the management of project data and outputs and the dissemination of key results to various stakeholders. We discuss experiences and lessons from CONTRAST, a multidisciplinary alliance that had been funded by the European Commission over a 4-year period, in order to optimize schistosomiasis control and transmission surveillance in sub-Saharan Africa. From the start, project partners from Europe and Africa set out an ambitious goal: to sample data following standard protocols at all field sites and then sharing the data in a way that would enable all project partners to have access through a password-protected Internet-based data portal. This required anonymous agreement on several common standardized sample forms, ranging from the mundane but important issue of using the same units of measurement to more complex challenges, for instance agreeing on the same protocols for double-treatment of praziquantel in different settings. With the experiences gained by the CONTRAST project, this paper discusses issues of data management and sharing in research projects in the light of the current donor demand, and offers advice and specific suggestions for similar interdisciplinary research projects.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

There are growing demands for the scientific community with regard to the management, sharing and dissemination of data and translation of key findings into policy. Although there are unprecedented opportunities of accessibility through global communication means in terms of Internet-based data sharing portals, ever more complex issues are being investigated, involving multidisciplinary research consortia. Standardization in research protocols is also coming more into focus to ascertain generalizability and to inform policy and practice (Hosking and Campbell-Lendrum, 2012). Fitting examples are projects pertaining to climate change and health (Costello et al., 2009) and the recently published Global Burden of Disease 2010 Study (Murray et al., 2012). Further to standardization comes the issue of data confidentiality and data sharing, which in large research projects with multiple partners and institutions, including the private sector, represents a major hurdle to overcome, even after the project has ceased to exist (Savage and Vickers, 2009).

The purpose of this paper is to discuss experiences and lessons from CONTRAST, a 4-year European Commission (EC)-funded multidisciplinary project to optimize schistosomiasis control and transmission surveillance in sub-Saharan Africa. Emphasis is placed on databases and data management issues, including protocol harmonization, data collection, verification, sharing and dissemination among project partners, the larger scientific community and other stakeholders.

## 2. Databases and data management

There is growing emphasis in collaborative research projects on database management, sharing of data and on pursuing a wide dissemination of results through open-access journals, other publicly available fora and social networks. Some funding agencies now require that detailed data management plans are submitted alongside research proposals. For instance, proposals submitted to the United States National Science Foundation (NSF) must, since January 2011, include a supplementary document – the project's "Data Management Plan" – which describes how the proposal will conform to the funding bodies' policy on dissemination and sharing of research results (see: <http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg.2.jsp#dmp>). The Massachusetts Institute of Technology (MIT) libraries have a team assisting their researchers with data management plans and data

\* Corresponding author. Tel.: +45 2962 4972; fax: +45 3533 1433.

E-mail address: [cls@sund.ku.dk](mailto:cls@sund.ku.dk) (C.F.L. Saarnak).

curation. In the United Kingdom, a so-called digital curation centre has been established to assist with the challenges of digital preservation, and digital curation and to lead and support services for higher education institutions in the United Kingdom (see: <http://www.dcc.ac.uk>).

The continuous build up of research data, their accessibility and their storage has been highlighted. Of particular note is a special issue recently published in the multidisciplinary journal *Science* (*Science Staff, 2011*). For a thorough analysis of the challenges of data storage, including digital vulnerability, the reader is referred elsewhere (*Carraway, 2011*). Presentation and dissemination of geographical distribution of schistosomiasis and other neglected tropical diseases has found new pathways following the advances of virtual Globe technology, such as Bing Maps (<http://www.bing.com/maps/>) and Google Earth (<http://earth.google.com>) (*Stensgaard et al., 2009*). Presenting disease distribution maps is a valuable tool for decision makers and scientists alike (*Yang et al., 2012; Porcasi et al., 2012*).

### 3. Experiences and lessons from CONTRAST

#### 3.1. CONTRAST – what is it?

The CONTRAST project is an interdisciplinary project integrating key skills and expertise to generate new knowledge on biological, environmental, molecular, spatial and socio economic factors relating to schistosomiasis transmission and control in sub-Saharan Africa. Funded by the EC as part of the Sixth Framework Programme, CONTRAST brought together 14 partner institutions, with a majority being universities and research institutions based in Central, East and West Africa (*Uttinger et al., 2013*). The project commenced in October 2006, ran for 4 years, and was coordinated by a team at the University of Copenhagen. An important feature of the CONTRAST project is the establishment and running of five specific research nodes across Africa with the following focus areas:

- (i) developing and validating innovative molecular tools to characterize schistosome parasites and their intermediate host snails (in Uganda);
- (ii) defining the importance of host-parasite dynamics across different ecological and epidemiological settings (in Cameroon);
- (iii) establishing a reference collection centre for schistosomes and intermediate host snails (in Kenya);
- (iv) advancing geographical information system and remote sensing platforms for integrated schistosomiasis risk mapping and prediction (in Zambia); and
- (v) encouraging and assessing novel local control interventions using social science approaches (in Tanzania).

With a strong focus on data, specifically on protocol harmonization, access, sharing and wide dissemination of data and key findings – including this special issue of *Acta Tropica* – we believe the experiences and lessons learned through the CONTRAST project are valuable to other large-scale research project endeavours, be they funded through EC or other funding bodies.

#### 3.2. The common CONTRAST database

When deciding to create a common database, one of the main challenges consists in streamlining data collections, and hence agreeing on a common format among the different partners. However, data sharing in science is not straightforward, and yet it is a key feature for effective research partnerships, thus fostering a fruitful learning culture that is based on mutual trust and jointly decided objectives, roles and responsibilities (*KFPE, 2012*). The following issues need careful consideration and were addressed

by CONTRAST. First, scientists and technicians need to use the same data tables, the same physical measurement units, and preferably the same approaches, tools and software for sampling, cross-checking, collating, compiling and analysing data over the course of the project. Second, the data forms have to be designed in a way that they serve the various disciplines involved in the project. Third, the data forms must be translated to fit with digital database standards, for instance a SQL database format (*Achimugu et al., 2010; Foley et al., 2010*).

Well into the project life cycle of CONTRAST, and after the sample forms had been created, the project coordinator and his team were approached by a commercial company (Fireflower Systems Ltd.; <http://www.fireflower.ca>), who proposed to create and host the multi-purpose database. The company offered to provide a web-based data management service, suitable for managing all kinds of data relating to schistosome parasites, intermediate host snails and the diseases caused by a chronic infection with blood flukes. A web-interface designed to match the CONTRAST field data collection was created. During the first annual meeting of the CONTRAST group, held in Yaoundé, Cameroon, in September 2007, all partner institutions in the project were taught on how to access and store data in the customized database. At the same time a manual explaining data entry and presentation was created (*Fireflower Systems Ltd., 2007*). With the web-interface, it was ascertained that every piece of data was placed in a unique cell in the database. The only possible sources of error were typing mistakes, which the interface, to some degree, could handle automatically by an auto-correction function. There were no ‘versions’ of the database. Instead, it evolved continuously, and could be accessed by several users simultaneously from different venues.

There were drawbacks as well. It soon became obvious that building a database online was time consuming, as each field had to be entered separately. The more straightforward solution would be to batch upload data. However, this was considered too vulnerable of an option, as it would need the user to fully understand the design, underlying structure and hierarchy of the database. It should also be noted that in the CONTRAST project, many of the partners had never used digital databases, and due to funding constraints, it was not possible to hire a professional database manager to oversee these processes over the course of the project.

The growth of the database was monitored by the project coordinator and his management team. When it became apparent that a considerable part of the data gathered in different field sites failed to make it into the database, partially explained by the aforementioned issues, it was decided that a technician should be appointed at the National Museums of Kenya, one of the partner institutions in East Africa and host of one of the research nodes. Subsequently, all partners were urged to forward their data, either as the raw paper forms or as data already entered electronically into digital spreadsheets. A small team at the National Museums of Kenya would then enter the data on behalf of all partners. Despite this initiative and active probing of the partners, many records were not made available in the database. Discussion with project partners revealed that sending the data files was considered a hassle, and that sharing of raw data was a concern, even within the remit of the CONTRAST project. At the time of writing this article in early 2013, the compiled data are gathered in a single large spreadsheet file, which can be accessed through the CONTRAST coordinator. However, the web interface is now defunct, which highlights the vulnerability of scientific databases: even with the best intentions it is difficult to locate the necessary funding for curating and maintaining them.

#### 3.3. Achievements made by CONTRAST

Among the various achievements made by CONTRAST (*Uttinger et al., 2013*), it is worth highlighting two outcomes that cover

Download English Version:

<https://daneshyari.com/en/article/3393886>

Download Persian Version:

<https://daneshyari.com/article/3393886>

[Daneshyari.com](https://daneshyari.com)