Towards New Metrics For Bioresource Use

Laurence Mabile, Paola de Castro, Elena Bravo, Barbara Parodi, Mogens Thomsen, Samuel Moore, Anne Cambon-Thomsen

Towards New Metrics For Bioresource Use

Laurence MABILE[a], Paola DE CASTRO[b], Elena BRAVO[b], Barbara PARODI[c], Mogens THOMSEN[a], Samuel MOORE[d] and Anne CAMBON-THOMSEN[a,e].

[a]*UMR1027 INSERM-Université Paul Sabatier Toulouse III, Toulouse, France*
[b]*Istituto Superiore di Sanita, Rome, Italy*
[c]*IRCCS AOU San Martino - IST,* Genova, Italy
[d] Ubiquity Press, London, United Kingdom
[e]BBMRI-ERIC, Graz, Austria.

Abstract. Abstract goes here.

The BRIF is an ongoing initiative that encompasses reflections and actions from various stakeholders (researchers, funders, industrials, editors) towards i/ standardised identification schemes and reporting for better visibility and tracing of bioresources on the web; ii/ incentive policies from hosting institutions; iii/ creation of tools allowing follow up of their use.
Tracing the use of bioresource is the first step in this process and for this purpose we have published the CoBRA (Citation of BioResources in journal Articles) guideline, launched the *Open Journal of Bioresources* and started developing new metrics.
The CoBRA guideline aims to standardise the citation of bioresources in scientific articles in order to trace their use on the web.
The *Open Journal of Bioresources (OJB)* was created in close collaboration with the open access publisher Ubiquity Press allowing both the resources and the OJB papers to be cited, and also providing authors with tools to get metrics on reuse and impact.
New better adapted metrics are being worked out in a dedicated BRIF working subgroup. A first list of relevant parameters to take into account in the impact measure of bioresources has been provided.
The tools proposed here foster easier access to samples and associated data as well as their optimised use, sharing and recognition for data producers. Input from the scientific editorial community would be highly appreciated at this stage.

Keywords. Bioresources citation; guideline; open access journal; metrics; impact

1. Introduction

For several years, the BRIF (Bioresource Research Impact factor)[1] [1] initiative has focused on specifying the framework to facilitate sharing of bioresources[2] through incentives and tools. The basis of the BRIF concept is that making feasible to trace the use of a bioresource and to calculate a corresponding impact factor should encourage institutions, researchers, bioresource managers and other actors involved in bioresource work, to share them. Sharing would then be seen as a gain rather than a loss of control or than an additional non recognised work, as often felt, so far. These issues are a concern in many biology and biomedical communities. Although the concept could be used in many areas (for example for primary resources in humanities and for ecological collections) we focus on human biological and biomedical resources because their very existence is depending directly on the willingness of patients and participants to give their samples and to allow the use of their data and there is an ethical imperative of making their contribution useful and recognised.

BRIF is an ongoing initiative that encompasses reflections and actions from various stakeholders (researchers, funders, industrials, editors) within dedicated working groups towards i/ standardised identification schemes and reporting for better visibility and tracing of bioresources on the web; ii/ incentive policies from hosting institutions; iii/ creation of tools allowing follow up of their use. Tracing the use of bioresource is the first step in this process and new tools have been or are being developed to make it feasible: the CoBRA guideline (Citation of BioResources in journal Articles), the Open Journal of Bioresources (OJB) and the BRIF metrics.

2. Citing bioresources: the CoBRA guideline

At present, bioresources are either cited in a confusing, heterogeneous way or they are not cited at all. The use of a bioresource in a research article is not retrievable systematically via PubMed or other bibliographic databases [2]. Traceability and visibility of bioresources in scientific literature or in other (online) sources would highlight their use. By being properly cited, bioresource use would be valued and their sharing thus encouraged. The CoBRA guideline [2] was hence developed to standardise citation of bioresources in scientific articles in order to trace their use on the web. This was achieved through close collaboration between the BRIF journal editors' subgroup with scientific journal editors, the EQUATOR [3] (Enhancing the QUAlity and Transparency Of health Research) network and the research community managing and/or using bioresources. It recommends mainly that each individual bioresource used to perform a research work should be mentioned in the Method section and should be cited as an individual "reference [BIORESOURCE]" according to a delineated format, using a unique identifier when possible. The detailed recommendation is given by the CoBRA checklist reported on the EQUATOR's website[4].

[1] http://gen2phen.org/groups/brif-bio-resource-impact-factor
[2] Bioresources are defined as any collection of biological samples with associated data, biological related databases independent of physical samples or other collections of biomolecular and bioinformatics research tools.
[3] http://www.equator-network.org/
[4] http://www.equator-network.org/wp-content/uploads/2015/03/Cobra-check-list.pdf

CoBRA needs now to be implemented and points to the necessity of integrating scientific editorial policies in the loop using several strategies. One way to enforce CoBRA use in articles is to include it in instructions to reviewers as part of the checklist used to process manuscripts. A second way is to add CoBRA in the list of reporting guidelines that is usually part of the instructions to authors. We also aim to obtain recommendation by the International Committee of Medical Journal Editors (ICMJE). In any such case though, compliance to the guideline is not guaranteed unless it is strictly verified by either reviewers or editorial staff (or made mandatory).

Associations of editors such as the European Association of Science Editors (EASE[5]) are of great help in reaching and empowering journal editors and authors of scientific publications. EASE Guidelines for authors and translators of articles to be published in English already include the necessity to mention in the methods section the origin and identity of experimental materials used and refer to the CoBRA guideline. The more key associations or committees of scientific journals editors will be aware of CoBRA, the more it will be applied. There is a need to go beyond the European dimension. Worldwide associations such as WAME (World Association of Medical Editors), AMERBAC, Canadian Editors Association and CSE (Council of Science Editors) must be informed of the existence of CoBRA and should promote it.

Other stakeholders are also key players in developing good practices and could contribute to the implementation of CoBRA. Institutions hosting bioresources as well as funding agencies can guide researchers in good reporting of bioresource use. In France, the National Institute of Science and Techniques Information (INIST[6] - CNRS) has been a great support in disseminating and promoting the guideline. The European Research Infrastructure of biobanking and biomolecular resources (BBMRI-ERIC[7]) has actively supported the BRIF initiative and included it in its 2015-2016 workplan to facilitate notably the implementation of CoBRA among its members. It will be added to the MTA/DTA and specified in publication policies. Other infrastructures could be interested in helping implementing CoBRA as one of the tools of their own strategy. As a matter of fact, "Research infrastructures in the biological and medical thematic area of the European Strategy Forum on Research Infrastructures (ESFRI[8]) roadmap are committed to provide access to the most advanced, unique, and large-scale biological resources, instruments and expertise in Europe to support research and development in all life sciences." On a global scale, consortia or scientific societies such as the Public Population Project in genomics and society (P3G[9]), the International Society for Biological and Environmental Repositories (ISBER[10]) and the European, Middle Eastern & African Society for Biopreservation & Biobanking (ESBB[11]) would help in extending these actions. Patient's associations could have a role in this too. Contributors to bioresources give importance to the fact that they are used and not sleeping resources. Thus accessing data on the use of such resources would be valuable for them too.

Over the last years, other initiatives throughout the world have flourished within the open access and sharing move to better identify and trace different types of resources (OpenAire, DataCite, CODATA, Force 11, ORCID and others). Among them, Research

[5] http://www.ease.org.uk
[6] http://www.inist.fr/
[7] http://bbmri-eric.eu/
[8] http://eu-openscreen.eu/index.php?id=130
[9] http://www.p3g.org/
[10] http://www.isber.org/
[11] http://www.esbb.org/

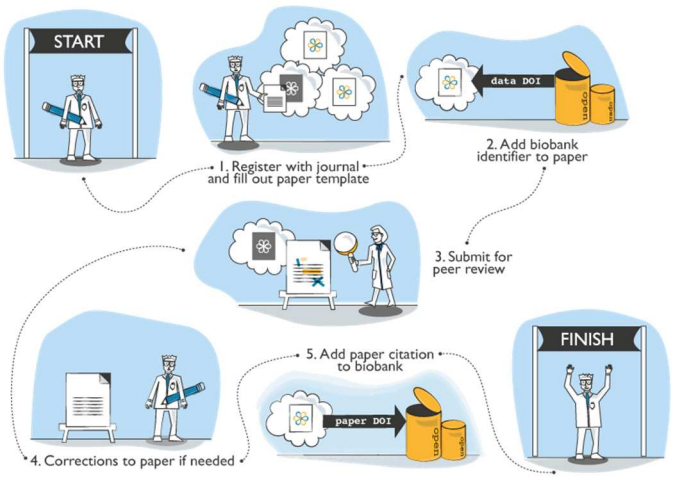
Data Alliance Working Group on Dynamic Data Citation has provided recommendations about making subsets of data citable. Connecting to these groups would certainly facilitate CoBRA implementation and foster a better granularity by using suitable identifiers. Such identifiers of subsets or combination of subsets of bioresources must first be worked out with the idea of keeping traceable their "genealogy" (origin of parental resources). In general, coordination of all these actions has become an urge if one wishes to develop standard citation tools and improve good reporting practices.

3. Publishing a bioresource: a new type of journal

The *Open Journal of Bioresources* is one journal in a suite of so-called 'metajournals' published by Ubiquity Press. These journals are dedicated to opening up and aiding the discoverability of all research elements involved in the research lifecycle, such as data, software, bioresources and hardware (forthcoming). The idea behind the metajournals is that researchers need to be able to discover and cite these research elements, but they also want credit for sharing them and the ability to track their impact. Given this, the metajournals offer credit – in the form of citation and altmetric data – for researchers making their resources permanently available and discoverable in accordance with community norms.

In the case of bioresources, the idea behind this journal is to provide a permanent marker paper so that users can definitively cite a bioresource they have accessed or referred to. The best way to do this is by integrating the bioresource into the traditional process for obtaining scholarly credit: the peer-reviewed journal article. In this way, users simply cite the bioresource as they would do any other journal article – and this is facilitated by the application of a digital object identifier (DOI) to all articles. This means that each article acts as a permanent marker for a bioresource and conforms to the standard processes for citing research.

OJB publishes bioresource papers, which are structured summaries of bioresources that are peer-reviewed to ensure they are accurately described. Papers are published in accordance with a structured template that describes the bioresource, outlines how it is preserved, the methods used in its creation, and how it can be accessed in the biobank.



These papers are not lengthy descriptions of bioresources but more akin to a short online form. Contents are therefore structured not by paragraphs, but by individual sentences and one-word answers. The result is a highly structured, objective description of a bioresource.
Because the bioresource paper is an objective description, so too is the peer review process. Importantly, OJB papers are not peer reviewed for their significance but rather that the information is accurately filled out and presented in accordance with the standards set by the CoBRA guidelines (see above). Because of this, the peer review process is relatively quick and articles can be published within a matter of weeks from submission. Articles are published open-access under the CC BY licence, ensuring anyone can access the final contents. For this, the journal charges a small APC of £100 – which is completely waivable if an author does not have access to funding for publication fees.
The published article then becomes a permanent marker paper for the described bioresource. Users cite the paper directly when they have accessed, used or simply referenced a bioresource. Citations are tracked and displayed on the article page alongside numbers of article views, tweets and Facebook likes. In this way, the bioresource paper allows authors to understand the true impact of their bioresource, which would not have been possible previously.
Articles are also sent to various scholarly indexes to aid discoverability, ensuring they become part of the permanent scholarly record. We have also been in discussion with PubMed about indexing articles there – which we're confident will happen in the future.

4. Towards a new metrics: the BRIFs

Once the bioresource is fully traceable and indexed, the impact of its use could be measured using the metrics tools offered on the net. Those tools are mainly based on citation indexes and assume that citation reflects the 'success' of the enterprise. But in the case of bioresources this is not sufficient. They do not reflect the full range of utility of a bioresource. For example, a clinical and biological collection of rare diseases will be used by a restricted community, whereas the resource has a high value, requiring a worldwide coordination effort and the contribution of different stakeholders. Other metrics are needed that take relevant parameters into consideration.
As part of the BRIF initiative, a dedicated working subgroup worked out this issue and provided a first list of relevant parameters to take into account in the impact measure. An online survey was sent to selected biobanks in order to assess those parameters in the evaluation of the impact of a bioresource. The answers from 28 biobanks (mainly from Italy and France) were used to classify parameters of scientific impact for bioresources. Several groups of parameters were defined according to their availability and to the feasibility of their retrieving for calculating the impact using one or several specifically designed algorithm(s). The main parameters relate to indicators of research productivity and sustainability; indicators of sample/data value; indicators of workflow and efficiency and indicators of collaboration and visibility. An extended study on various types of bioresources in more countries will allow refining the list and characteristics of such parameters.
On the basis of the selected parameters an algorithm will be proposed in close collaboration with BBMRI-ERIC IT service for measuring the use and impact of bioresources. It will be tested in the wider context of European biobanks covered by the

National Nodes of this European research infrastructure. A major step in this process is the proper identification of bioresources, including the physical resources; this point is presently being discussed between BBMRI-ERIC and DataCite.

The tools proposed here foster easier access to samples and associated data, their optimized use and sharing as well as the recognition of data producers. Input from the scientific editorial community would be highly relevant at this stage. This work could benefit from initiatives in other domains, in particular the long standing work performed in astronomy to provide mechanisms for quoting astronomical databases [5] and could serve as a reference for other communities, beyond human biological and medical bioresources.

Ackowledgements

We wish to thank Federica Napolitani, Anna Maria Rossi, Alessia Calzolari from Istituto Superiore di Sanità, Rome Italy, as well as Petr Holub, Jan Eric Litton, Outi Törnwall and Michaela Mayrhofer from BBMRI-ERIC, Graz, Austria for very fruitful discussions and constructive interactions.
The BRIF initiative was supported by funds from BBMRI-LPC project, grant agreement number 313010, theme FP7–INFRA-2012-1.

References

[1] Cambon-Thomsen, A., Carpenter, J., Dagher, G., Dalgleish, R., Deschênes, M., di Donato, J.H., Filocamo, M., Goldberg, M., Hewitt, R., Hofman, P., Kauffmann, F., Leitsalu, L., Lomba, I., Mabile, L., Melegh, B., Metspalu, A., Miranda, L., Napolitani, F., Oestergaard, M.Z., Parodi, B., Pasterk, M., Reiche, A., Rial-Sebbag, E., Rivalle, G., Rochaix, P., Susbielle, G., Tarasova, L., Thomsen, M., Thorisson, G.A., Zawati, M.H., Zins, M.; BRIF workshop group (2011) 'The role of a Bioresource Research Impact Factor as an incentive to share human bioresources', *Nature Genetics*, 43 (6), p. 503-504.

[2] Mabile, L., Dalgleish, R., Thorisson, G.A., Deschênes, M., Hewitt, R., Carpenter, J., Bravo, E., Filocamo, M., Gourraud, P.A., Harris, J.R., Hofman, P., Kauffmann, F., Muñoz-Fernàndez, M.A., Pasterk, M., Cambon-Thomsen, A.; BRIF working group (2013) 'Quantifying the use of bioresources for promoting their sharing in scientific research', *GigaScience*, 2 (7), p. 1-8. Available: http://www.gigasciencejournal.com/content/2/1/7 [7 Mar 2016]

[3] Bravo, E., Calzolari, A., De Castro, P., Mabile, L., Napolitani, F., Rossi, A.M. and Cambon-Thomsen, A. (2015) 'Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA)', *BMC Medicine*, 13 (33), p. 1-12. doi: 10.1186/s12916-015-0266-y
Available: http://www.biomedcentral.com/1741-7015/13/33 [7 Mar 2016]

[4] http://openbioresources.metajnl.com/

[5] Uhlir, Paul F. (2012) *For Attribution - Developing Data Attribution and Citation Practices and Standards: Summary of an International Workshop*, Washington, D.C: The National Academies Press.