DELIVERABLE No.	**D1.5**
DELIVERABLE TITLE	**Formulate recommendations for data management based on assessment of data completeness for regulation**
RESPONSIBLE AUTHOR	**Dana Kühnel, UFZ**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº814530 This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Executive Summary

The aim of D1.5 was to provide recommendations for a data management strategy for the use of non-guideline studies in the frame of the NanoRigo risk governance framework. It shall ensure that data from non-guideline studies (peer-reviewed scientific literature) is appropriate, complete, and of high quality for further use in the risk governance framework. As reviewed in D1.3, ´data quality´ is an open concept and highly context dependant. Hence, for the purpose of this deliverable, we defined data completeness as the status when the minimal testing requirements according to regulatory frameworks are fulfilled.

The development of quality criteria for data generated by non-guideline studies will be driven by endpoints as requested under regulatory frameworks. In a first step, these minimal testing requirements according to different regulatory frameworks for ecotoxicity and human toxicity testing were compiled (tests/assays and guidelines). Based on the endpoints that the respective tests/assays assess, and the quality criteria that are foreseen for these endpoints in the respective testing guidelines, quality criteria for non-guideline are proposed.

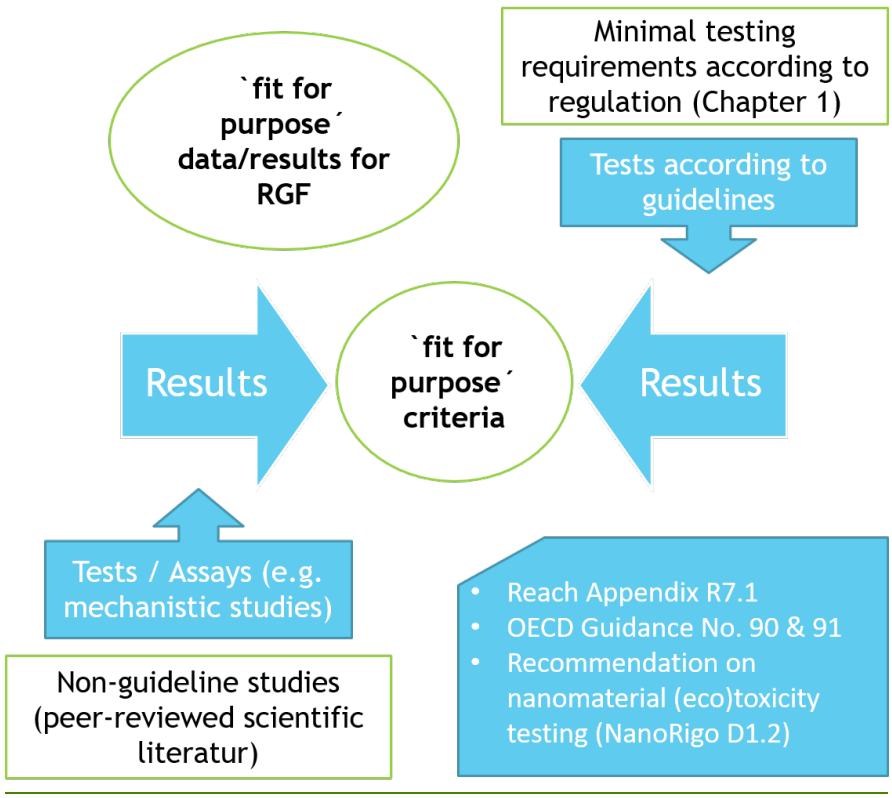

Fig. 1 Scheme for the deduction of ´fit for purpose´ criteria. The underlying question is: Is a result from a non-guideline study of sufficient quality to replace the result from a guideline study? In order to propose an evaluation of non-guideline studies in the frame of a future risk governance framework, minimal regulatory testing requirements (this deliverable), the recommendations made in D1.2 on nanomaterial (eco)toxicity testing and the recommendation from Reach App. R7.1 are taken into account.

Data gaps in regulatory decision making may be filled by data obtained in non-guideline studies. Here we focused on peer-reviewed scientific literature that according to Klimisch scoring should fullfil the conditions for score 2 (reliable with restrictions) but as it will depend on the quality of the peer review may also fall in score 3 (not reliable). In order to decide for (eco)toxicological studies not performed according to test guidelines whether the data is fit for purpose, criteria are proposed. These fit for purpose criteria allow a meaningful integration of data from non-guideline into the future risk governance framework.

Recommendations for the management of non-guideline studies in the frame of the future RGF are given. For this, we considered the regulatory contexts of nanomaterial hazard assessment. The major goals of data management are to compile data of known quality, to maintain the integrity of the data base, and to provide data to users.The ultimate aim of data management is to ensure that a given data repository or database best fulfills the needs for which it was created (Schöpfel, Ferrant et al. 2018, Higman, Bangert et al. 2019). For the purposes of NANORIGO, data management practices must enhance the reuse of non-guideline studies/data for regulatory purposes while reliably establishing their quality.

In order to enable the possibility of using non-guideline studies and data for regulatory decision-making, at the present time we therefore recommend:

- The use of FAIR-compliant (meta)data formatting to ensure interoperability (e.g., CSV, RDF, XML, or JSON formats)
- Study/data rating for each study or dataset using a suitable data quality assessment method identified from D1.3, such as the DaNa, GUIDEnano, caLIBRAte, NANOcred, or SciRAP criteria checklists
- Clear labelling that lists whether all ‘fit for purpose’ criteria are partially or fully fulfilled. This may take the form of a tiered approach, such as that used in the DaNa checklist, where certain data are considered as essential, and others are considered as optional

- Clear labelling that lists the extent of physicochemical characterization and whether 'mandatory' base-level characterization has been completed (this typically includes the substance name (or CAS-No) and form (powder or suspension), core chemical composition, shape, chemical composition (i.e. purity and contaminants), particle size and size distribution in suspensions, specific surface area, and surface chemistry/charge)

This is just a preliminary list since we expect that the recommended measures will change as a better conception of the needs for the NRGF are elaborated.

References

Higman, R., D. Bangert and S. Jones (2019). "Three camps, one destination: the intersections of research data management, FAIR and Open." Insights the UKSG journal **32**.

Schöpfel, J., C. Ferrant, F. André and R. Fabre (2018). "Research data management in the French National Research Center (CNRS)." Data Technologies and Applications **52**(2): 248-265.
Shoults-Wilson, W. A., O. I. Zhurbich, D. H. McNear, O. V. Tsyusko, P. M. Bertsch and J. M. Unrine (2011). "Evidence for avoidance of Ag nanoparticles by earthworms (Eisenia fetida)." Ecotoxicology **20**(2): 385-396.

NANORIGO

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº814530 This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it