

Building confidence to manage research data

Validated RNRRS Output.

Agricultural researchers in Uganda are now confident of their ability to manage data. Previously, although they had vast amounts of data it often wasn't organised properly. So, they couldn't use it effectively. Banana researchers checked and organised their data in a database. This means that they now have quality datasets that can be used fully, and have fewer problems with analysing data. Because they can extract findings quickly, they show decision makers that they can give prompt and reliable information and that investment in research is worthwhile. Good data management also helps managers apply best practices to data security and ownership. So, research is more useful and contributes more to improving the lives of the poor.

Project Ref: **CPP49:**

Topic: **7. Spreading the Word: Knowledge Management & Dissemination**

Lead Organisation: **University of Reading, UK**

Source: **Crop Protection Programme**

Document Contents:

[Description](#), [Validation](#), [Current Situation](#), [Current Promotion](#), [Impacts On Poverty](#), [Environmental Impact](#),

Description

CPP49

Research into Use

NR International
Park House
Bradbourne Lane
Aylesford
Kent
ME20 6SN
UK

Geographical regions included:

[Uganda](#),

Target Audiences for this content:

[Crop farmers](#),

A. Description of the research output(s)

1. Working title of output or cluster of outputs.

In addition, you are free to suggest a shorter more imaginative working title/acronym of 20 words or less.

Improving research throughput and effective use through capacity strengthening in data management and statistical applications

2. Name of relevant RNRRS Programme(s) commissioning supporting research and also indicate other funding sources, if applicable.

Forestry Research Programme funded preparation of a related booklet on protocol writing. ILRI and ICRAF (jointly, using own funds) commissioned follow-up consultancy that draws on these RNRRS-funded results. The joint ILRI/ICRAF Research Methods Group are now actively taking this forward.

3. Provide relevant R numbers (and/or programme development/dissemination reference numbers covering supporting research) along with the institutional partners (with individual contact persons (if appropriate)) involved in the project activities. As with the question above, this is primarily to allow for the legacy of the RNRRS to be acknowledged during the RIUP activities.

R8301 – *Archiving data from integrated pest and disease management projects within the Uganda Banana Research Programme*, was undertaken as a collaboration between the Statistical Services Centre at the University of Reading in the UK, and the Uganda National Banana Research Programme based at the Kawanda Agricultural Research Institute in Kampala. The named collaborator was Dr. W. Tushemereirwe, while key supporting staff were Yusuf Mulumba (Statistician) and Allan Rwakatungu (Research Assistant).

R8410 – *Increasing the effectiveness of Research within NARO, Uganda*, was undertaken at the specific request of the then Director General of NARO, Dr. W. Otim-Nape, with much of the work and collaborative efforts during project implementation being done in a collaborative effort between the Statistical Services Centre at the University of Reading in the UK, and Dickson Baguma, Director of the Monitoring, Evaluation and Planning Unit of the NARO Secretariat. The project was fully supported by the later Acting Director General of NARO, Dr. Denis Kyetere.

ZF0183 – *Writing Research Protocols: A Statistical Perspective* was the guidance material written with FRP programme level funds.

4. Describe the RNRRS output or cluster of outputs being proposed and when was it produced? (max. 400 words). This requires a clear and concise description of the output(s) and the problem the output(s) aimed to address. Please incorporate and highlight (in bold) key words that would/could be used to select your output when held in a database.

Project R8301 (May 2004) provided **capacity strengthening in data management** and best practice approaches in writing **research protocols** for **Uganda National Banana Research Programme (NBRP)**.

Output 1: research data management **training** for staff and production of a comprehensive archive of datasets from three IPM related research projects. This **data archive** was compiled, checked, documented with all related

meta-data, and organised so data could be fully and effectively utilized in future.

Output 2: Following lessons learned, the NBRP team produced a researchers' manual: *Guidelines and Procedures for Effective Data Management* [8902 words], plus appendices [7674 words] illustrating aspects of research protocols, for effective documentation and management of varied study types.

Output 3: The team developed an agreed **Policy for Research Management: with particular emphasis on research data management and statistical analysis** [5868 words], giving senior managers and research leaders organized structures to manage research more effectively, emphasising best practice approaches to **data security** and **ownership**.

These outputs aimed to address researchers' difficulties in effectively utilising expensively collected research data: they previously lacked expertise in managing, organising and utilizing **research data**. Good data management is neglected worldwide: the framework developed for a 'culture' of careful organisation and **good statistical practice** served to strengthen NARO to make sustained contributions to rural **poverty alleviation**.

R8301 was so successful that NARO's Director General requested a follow-up project (R8410) – to assess **capacity building** needs for scientific staff across the NARO system, and recruiting or developing **statistics** and data management specialists. R8410 (December 2005) aimed to facilitate removing bottlenecks in research processes, and thus increase capacity of researchers to extract findings systematically and quickly from data and demonstrate their effectiveness to those taking research into use, e.g. policymakers and other agencies who influence change.

Main outputs from R8410 were separate documents demonstrating a process to strengthen NARO research capacity in data handling, **study design**, use of statistical software and **statistical analysis** work. **Training Needs Analysis** reviewed training and support needs of NARO staff. An institutional assessment: *Developing Effective Statistical Support* concerned establishing a sustainable NAROWide statistical support service, to combat past problems of retention, availability, motivation and skills of statistical staff. *Research Capacity Strengthening Strategy* provided recommendations and justification for each strand of capacity development. A *Briefing Paper on Management Action Points* summarized what decision-makers need to DO, in time order, to take forward the suggested agenda in a constructive, organised way.

5. What is the type of output(s) being described here?

Please tick one or more of the following options.

Product	Technology	Service	Process or Methodology	Policy	Other Please specify
		X	X	X	

6. What is the main commodity (ies) upon which the output(s) focussed? Could this output be applied to other commodities, if so, please comment

Outputs are **cross-cutting** across the entire range of agricultural research that entails data collection. R8301

especially can be seen as an amplification of an often under-developed side of **Good Laboratory Practice** insofar as meticulous field or lab. notebook keeping is stressed to research students, but what happens to results thereafter is frequently inadequately conceptualised.

7. What production system(s) does/could the output(s) focus upon? Please tick one or more of the following options. Leave blank if not applicable

Semi-Arid	High potential	Hillsides	Forest-Agriculture	Peri-urban	Land water	Tropical moist forest	Cross-cutting
							X

8. What farming system(s) does the output(s) focus upon. Please tick one or more of the following options (see Annex B for definitions). Leave blank if not applicable

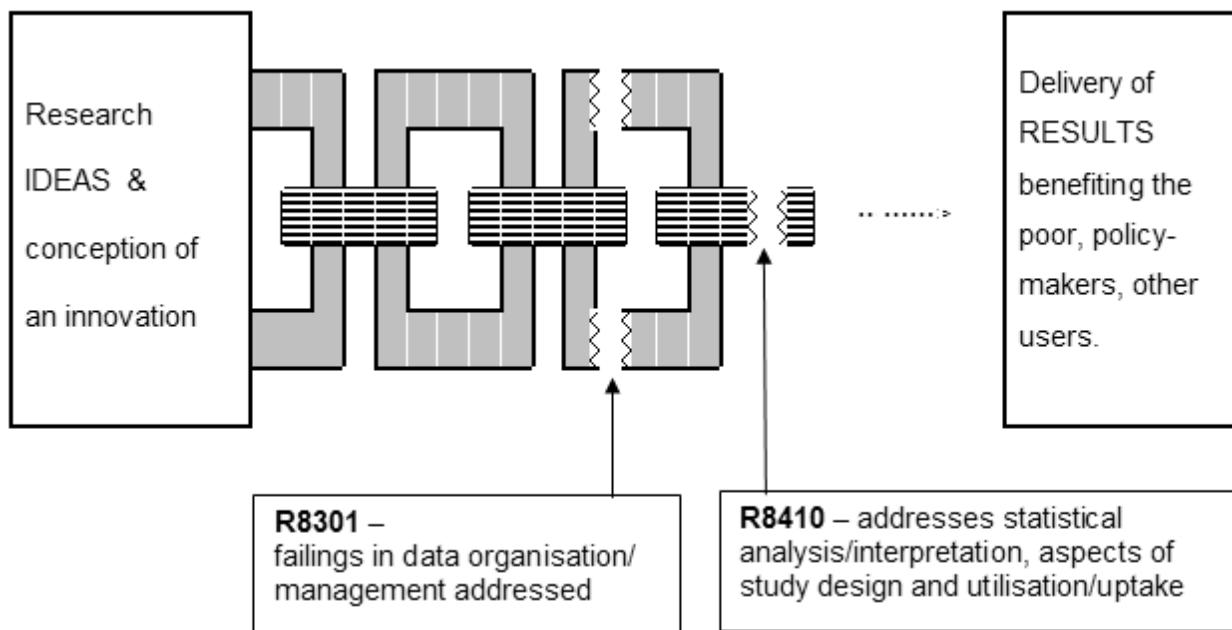
Smallholder rainfed humid	Irrigated	Wetland rice based	Smallholder rainfed highland	Smallholder rainfed dry/cold	Dualistic	Coastal artisanal fishing

As in Question 6 above outputs are cross-cutting across all farming systems.

9. How could value be added to the output or additional constraints faced by poor people addressed by clustering this output with research outputs from other sources (RNRRS and non RNRRS)? (**max. 300 words**).

Please specify what other outputs your output(s) could be clustered. At this point you should make reference to the circulated list of RNRRS outputs for which proformas are currently being prepared.

Any agricultural research based on data-collection or development of an organised evidence base involves an **information chain** of steps. If the information chain is broken because some steps malfunction, the process can't and won't work properly: research investment goes nowhere. Projects R8301 & R8410 deal with intermediate links in this chain.



These links in the chain are crucial to overall function, so would serve to add value to outputs from almost any other cluster which seeks to put their research into use.

Outputs are generic – training needs analysis and institutional assessment based on consultations across NARO Uganda, itself covering a wide variety of commodities, production and farming systems. Taking R8301 and 8410 into use in NARO could stand alone or be integrated with *most* other RIU offerings NARO might take forward. Furthermore, in questioning about 200 NARO staff, 95% highlighted *software, training and support in statistics and data management* as one of their top three constraints to conducting good research, ahead of *professional upgrading in their own discipline* (79%). Successfully implemented, project outputs empower researchers to achieve more effective results often, confidently, efficiently.

Unavoidable proviso: research teams must be competent in forging other links in the information chain e.g. technical writing/ communications, e.g. project management.

Any substantial agricultural research organisation regionally could benefit from similar inputs, given similar qualities to NARO – ability to articulate demand and foresee benefits, institutional buy-in, openness in acknowledging capacity-building needs, willingness to support statistics and data management staff posts plus training periods for suitable scientists and technicians.

Validation

B. Validation of the research output(s)

10. *How were the output(s) validated and who validated them?*

Please provide brief description of method(s) used and consider application, replication, adaptation and/or adoption in the context of any partner organisation and user groups involved. In addressing the “who” component detail which group(s) did the validation e.g. end users, intermediary organisation, government department, aid organisation, private company etc... This section should also be used to detail, if applicable, to which social group, gender, income category the validation was applied and any increases in productivity observed during validation (max. 500 words).

The outputs from R8301, namely the data archive of a cluster of banana projects, the *Guidelines on Data Management*, and the *Policy for Research Management*, were fully discussed, agreed and accepted by all research staff of the National Banana Research Programme (NBRP). The manual on data management was endorsed by the Director General of NARO, while the policy document for NBRP research was endorsed by the Director of the Kawanda Agricultural Research Institute (KARI). These two documents were also included on the NBRP website (www.banana.go.ug) with an invitation to external researchers to adapt and adopt these documents within their own research setting and farming systems. With strong support from the DG of NARO, hard copies of the manual on Data Management were also distributed to all research staff in NARO to encourage the use of these guidelines.

With respect to R8410, there were extensive consultations with managers, scientists and technicians across NARO, as well as a survey of all NARO research staff, so as to elicit their current levels of skills in statistics and data management and to learn about their future needs in these areas. 199 survey returns were received comprising approximately 50% of NARO's research cadre. The results contributed to the development of the *Research Capacity Strengthening Strategy*, which together with other outputs from R8410, was discussed and approved at the final NARO-wide workshop. The validation of the outputs was therefore done largely by senior management in NARO. This included the Director of the Monitoring, Evaluation and Planning Unit, and the Deputy Director General Outreach, both of NARO Secretariat (NAROSEC), other senior administrative and research officers in NAROSEC, Managers of Agricultural Research Development Centres (ARDCs), and Directors of Research Institutes.

11. *Where and when have the output(s) been validated?*

Please indicate the places(s) and country(ies), any particular social group targeted and also indicate in which production system and farming system, using the options provided in questions 7 and 8 respectively, above (max 300 words).

The outputs have been validated within Uganda, and in particular within Uganda's National Agricultural Research Organisation. Given the nature of the outputs from R8301 and R8410, and the limited scale and short duration of these projects (13 months and 9 months respectively), there was no scope for formal validation of outputs outside of Uganda. Outputs from R8301 were validated in the period during project activities from May 2003 to May 2004, while R8410 outputs were validated during the final workshop in November 2005 by NARO senior management at NARO Secretariat, and from ARDCs and Research Institutes.

The outputs from this cluster concern cross-cutting issues and therefore apply to all production systems and farming systems. The primary targets for the outputs were NARO researchers, whose adoption of these outputs as part of a complete research information chain (see 9 above) would benefit the users of NARO's research, e.g.

NGOs, Ministry of Agriculture through its extension services. The outputs were aimed at enhancing the research environment by improving the speed and efficiency of utilisation of field findings, and indeed often by making utilisation possible where it has been completely stymied by records unusable because they are inadequately checked, documented, synthesised and shared. This is expected eventually to improve delivery to the poor farmers for whose benefit the research was intended.

Current Situation

C. Current situation

12. *How and by whom are the outputs currently being used? Please give a brief description (max. 250 words).*

R8301 outputs are being used by NARO. The manual on Data Management received by scientists in NARO has appendices of examples of several activity protocols suitable for different research settings, e.g. for participatory and survey studies, lab and on-station studies, and for on-farm studies. Part II of the manual includes further guidelines specific to experimental projects, and additional materials relating to the disciplined use of spreadsheets for input of data and all associated meta-data (i.e. all background information relating to questions of what, when, where, who, why and how, associated with the numerical figures). The need for an integrative project level protocol is also emphasized and an example provided.

The outputs from R8410 were for senior management across the whole of NARO and those involved in the NARO re-structuring process, e.g. the old NARO Secretariat and staff of the Change Implementation Team based in Kampala, Uganda. During discussions with personnel of these groups and the then Director of KARI, it was suggested that in their plan to make Kawanda a service provider for lab-based and other services, a further service could be provided to support scientists in data management and statistics, via a specially set up unit, e.g. called the Statistics And Data Management Resource Centre (STADMARC), that could stabilize and coordinate technical aspects of statistical provision across PARIs (Public Agricultural Research Institutes, i.e. the National Agricultural Research Institutes (NARIs) and Zonal Agricultural Research Institutes (ZARIs)). This planning is currently underway.

13. *Where are the outputs currently being used? As with Question 11 please indicate place(s) and countries where the outputs are being used (max. 250 words).*

The outputs are currently being used in Uganda, largely within the National Agricultural Research Organisation (NARO). The short duration of the projects meant that time and resources were limited in promoting their use outside of NARO. Project R8410 was completed in December 2005 at a time when re-structuring processes were still on-going in NARO, so putting the outputs into use has not been possible until the re-structuring process has stabilised.

Related outputs from follow-on work with ICRAF and ILRI is being taken up internally in these CGIAR centres' HQ's in Nairobi.

Output 1 from R8301, namely the banana data and meta-data archive, is also being used to develop a database management system for banana relate projects, based on the “Logbook” prototype developed by the Research Support Group at ICRAF in Nairobi. Funding support for this work, in the form of a follow-on 2-year project, came from the Rockefeller Foundation as a direct result of work under R8301.

Early in R8410, the *Policy for Research Management* (Output 3 from R8301) was provided to Regina Musaazi of NAROSEC for modification, dissemination and adoption across NARO.

14. What is the scale of current use? Indicating how quickly use was established and whether usage is still spreading (max 250 words).

When R8301 project activities began, staff in the banana programme received an initial training programme in *Research Data Management*. This had an immediate impact on improved attention to data quality and data management issues. The training programme also included a few staff from the International Institute for Tropical Agriculture (IITA). Throughout project activities, on-the-job training to staff also happened in terms of writing good research protocols and better organization of research data. The management of the vast amounts of data collected within the Banana Programme became more streamlined with the result that there were fewer difficulties encountered by the researchers at the data analysis stage. At the end of the project in May 2004, the distribution of the manual on guidelines for data management to all researchers in NARO meant that the use of this output is still spreading, within the NARO system.

Not all those initially trained at Kawanda have remained in the same organization (NBRP or IITA), but the generic skills gained will have had an impact on improving research data management wherever they are, perhaps even outside of the NARO system. Certainly the key member of the NBRP team with respect to data management, i.e. the (then) NBRP statistician, is now working as the Statistician/Data Manager on a medical project within a research collaboration between Makerere University and Case Western Reserve University. The organizational skills acquired are now being applied there.

15. In your experience what programmes, platforms, policy, institutional structures exist that have assisted with the promotion and/or adoption of the output(s) proposed here and in terms of capacity strengthening what do you see as the key facts of success? (max 350 words).

In Uganda, the NARO Secretariat has been instrumental in supporting and promoting outputs from both projects in this cluster. For example, the Secretariat undertook the distribution of outputs from R8301. The institutional structure in NARO was responsible for undertaking the implementation of the survey of researchers current skills and needs with respect to statistics and data management. This was supported by Directors of the Research Institutes and Managers of the Agricultural Development and Research Centres.

In the key element of R8410, which concerned Statistics and Data Management Training and Recruitment Needs Analysis, there was also support in principle from staff of the Change Implementation team, although they were not in a position at that stage to commit to any suggested changes regarding the recruitment of statisticians and/or data managers since negotiations were still in process regarding positions of administrative and research staff within the new structure.

Capacity strengthening was integrated throughout project activities. As indicated under question 14 above,

training of about 30 researchers at the Kawanda Agricultural Research Institute, including staff from NBRP and IITA, led to a high level of awareness and acquisition of skills in ensuring good quality data and better management of data from their information collection exercises. This, together with (a) more attention to the documentation of research protocols to a high standard and (b) a better management system for large volumes of data, have been key successes in capacity building activities undertaken within these projects. Two staff members from NBRP also received further specialized short-term training in the UK.

Current Promotion

D. *Current promotion/uptake pathways*

16. Where is promotion currently taking place? Please indicate for each country specified detail what promotion is taking place, by whom and indicate the scale of current promotion (max 200 words).

Although the outputs from projects of this cluster are quite generic, there were no expectations in the project proposal or funding phases that these short-term projects would promote the outputs beyond the research environment in which they were envisaged. As such there are currently no promotion activities taking place. However, the lead author of the manual on *Guidelines for Data Management*, i.e. Dr. Charles Murekezi, is currently employed by IITA in Rwanda, and working closely with ISAR (the Agricultural Research Organisation in Rwanda). Dr Murekezi is attempting to promote statistics and data management developments in ISAR, with the expectation of securing funds from a suitable donor to take this further.

Project activities in Uganda have also been described by a colleague to the Director of the Directorate of Training, Documentation and Technology Transfer of the National Institute of Agricultural Research of Mozambique. The Director has shown high enthusiasm for similar work to be done in Mozambique, and has recently communicated with the said colleague to indicate her interest in pursuing this further.

17. What are the current barriers preventing or slowing the adoption of the output(s)? Cover here institutional issues, those relating to policy, marketing, infrastructure, social exclusion etc. (max 200 words).

The main barrier to adoption is the lack of local expertise in the research institutes themselves to push these generic skills forward and provide the necessary support with respect to statistics and data management activities. Although the guidelines manual helps, it does assume a certain degree of familiarity with facilities available in software packages such as Excel and the presence of some form of local technical support that NARO researchers currently do not have, since NARO has not been successful in the past in recruiting and retaining data managers or statisticians.

Our experience with NARO institutes has shown that large volumes of data continue to get collected with little emphasis on the need to make better use of the data before moving to the next project. For example, many new surveys continue to take place, because of the inaccessibility of past data to draw upon, to fulfil the needs of new research programmes. Resources are therefore wasted because of the lack of knowledge in data organization, management and statistical analysis.

18. What changes are needed to remove/reduce these barriers to adoption? This section could be used to identify perceived capacity related issues (max 200 words).

Changes needed to remove/reduce these barriers form the key outputs of Project R8410. They outline in considerable details steps that NARO (or the new National Agricultural Research Council) should take to set up and maintain a statistical and data management support service.

Recruitment to, and training in, a new area of focus will always need some special impetus as it is expensive. Where there is a fixed pot of money in a Medium-Term Expenditure Framework for a NARS, it is likely to threaten the amount earmarked for existing vested interests, and face resistance. Extra, external, funds may be required to kick-start such a development.

19. What lessons have you learnt about the best ways to get the outputs used by the largest number of poor people? (max 300 words).

The outputs here are not intended for use directly by poor people, but rather by those who conduct research to address the needs of the poor.

From our perspective, there is firstly a need to ensure the integrity of the whole information chain (see 9), including not neglecting weaknesses in the data management and statistics sections, but also including active pursuit of uptake pathways.

Secondly, there is a need to move away from purely activity-based incentives where NARS staff get generous allowances when in the field, but no similar incentive to stay in the office and analyse, write-up and promote their results. There is a need to make output-based incentives the norm.

Thirdly, the job descriptions, appraisal and reward mechanisms for research leaders and managers need to be tied in to their effectively dynamising the progress of research activities right along the information chain.

Impacts On Poverty

E. Impacts on poverty to date

20. Where have impact studies on poverty in relation to this output or cluster of outputs taken place? This should include any formal poverty impact studies (and it is appreciated that these will not be commonplace) and any less formal studies including any poverty mapping-type or monitoring work which allow for some analysis on impact on poverty to be made. Details of any cost-benefit analyses may also be detailed at this point. Please list studies here.

As indicated in 9 above, R8301 and R8410 are concerned with intermediate links in the information chain of agricultural research, not very directly linked to poverty impacts. The second project was completed in the last days of RNRRS funding. Where a causal chain might be established to increased effectiveness of agricultural

research, effects would only be visible over a substantial time period.

21. *Based on the evidence in the studies listed above, for each country detail how the poor have benefited from the application and/or adoption of the output(s) (max. 500 words):*

- *What positive impacts on livelihoods have been recorded and over what time period have these impacts been observed? These impacts should be recorded against the capital assets (human, social, natural, physical and, financial) of the livelihoods framework;*
- *For whom i.e. which type of person (gender, poverty group (see glossary for definitions) has there been a positive impact;*
- *Indicate the number of people who have realised a positive impact on their livelihood;*
- *Using whatever appropriate indicator was used detail what was the average percentage increase recorded*

See response to 20 above.

Environmental Impact

H. Environmental impact

24. *What are the direct and indirect environmental benefits related to the output(s) and their outcome(s)? (max 300 words)*

This could include direct benefits from the application of the technology or policy action with local governments or multinational agencies to create environmentally sound policies or programmes. Any supporting and appropriate evidence can be provided in the form of an annex.

There are no direct or immediate environmental impacts, but if R8301 and R8410 contribute to the development of a culture of evidence-based decision-making, and inter-institutional cooperation (see 23, sections 2 and 3), the effects some way down the line will be positive.

25. *Are there any adverse environmental impacts related to the output(s) and their outcome(s)? (max 100 words)*

No.

26. *Do the outputs increase the capacity of poor people to cope with the effects of climate change, reduce the risks of natural disasters and increase their resilience? (max 200 words)*

Having more organised and comprehensible information at hand to share is one *prerequisite* for more effective collaboration between e.g. agricultural researchers and agro-meteorological personnel. Given adequate investment of time, money and effort in the type of innovation platform mentioned above (see 23, sections 2 and 3), there would be more reason to hope that public institutions could sensibly contribute to policies, strategies and campaigns to help poor people to cope and adapt.