

Quality management for robust and reliable research^{*}

Eva Giesen^{**}

Institut National de la Santé et de la Recherche Médicale (Inserm), Directorate General, 101, Rue de Tolbiac, 75013 Paris, France

Received: 2 October 2015 / Accepted: 13 November 2015

Abstract. Faith in scientific publications is based on the reliability of scientific data, sound assessment of risk and its management and high quality of management of a research project. Different tools enable researchers to detect possible weak points in their study set-up and their data, validation of methods and equipment, quality control of products and sub-products contribute to ensuring reliability. Quality management and project management can be used (and are useful) in bio-medical research.

Keywords: Quality management, quality control, risk, ISO 9001: 2015 standard

1 Introduction

Research that fails to provide reliable results is worthless for scientists and citizens alike. The same can be said of a machine, instrument or method. And what about data that are lost, misclassified or impossible to find? The scientific community readily agrees that robust, reliable methods and reliable results are an absolute imperative for research and the very basis of the scientific endeavour. The need for reliability is obvious and “goes without saying”; this obviousness however can impede the implementation of specific measures to ensure reliability. In nearly thirty per cent of cases, the reason for retraction of an article is a scientific error or the inability to demonstrate the veracity of the results and that may only be the tip of the iceberg [1]. Scientists themselves are not always aware of the problem and often insufficient precautions are taken to teach the concepts and methods of validation, verification and conservation of methods, data and results. This is particularly dangerous in a research landscape ever more dependent on contracts and ever poorer in human resources and stable structures.

If citizens, funding bodies and even scientists themselves cannot have confidence in scientific results and articles, how can knowledge advance and scientific results be exploited to arrive at therapeutic, societal, political and legal decisions which may have consequences far beyond the sole domain of scientific research? Although scientific fraud and misconduct are probably rare, non-robust and irreproducible results may be more common than one might think. Systematic investigations of the reliability of methods and experiments are relatively scarce despite their being encouraged by scientific institutions and journals, because difficult to accomplish in complex fields such

as life sciences. When the issue has been specifically studied, the rates of good reproducibility have been shown to be low [2, 3].

Confidence in the quality of a research project requires that, when the various stages of a study are repeated by other scientists at different times, the same results are obtained. Apart from making the most of research results, the purpose of publication is to make it possible for other scientists to use reliable information for further experiments in order to “contribute to therapeutic progress, improve care and prevention systems, help decision-making processes in the political arena (the precautionary principle) and industrial production” [4]. To ensure this, ability and honesty when it comes to describing their work accurately are essential attributes for scientists.

What is really meant by “robustness” and “reliability”? How to ensure these parameters? How can an ISO 9001 Quality Management system [5] help? This is the focus of this article.

2 Some definitions

- (1) *Repeatability* means that measurements made by a given investigator on a given sample in the same experimental conditions (method, equipment, etc.) do not vary beyond certain pre-defined limits [6]. This is often expressed as a standard deviation.
- (2) An instrument, method or data point is *reliable* if, in a set of given conditions, no deviation is observed over time within a given time frame. Experimental data are unreliable if they drift over time even if the animal's species, weight, gender or age have been kept constant in the experiments. In identical conditions, an experiment carried out at a given moment in time ought to give the same result when it is repeated later [6].

^{*} List of total number of contributing authors at the end.

^{**} Correspondence: eva.giesen@inserm.fr

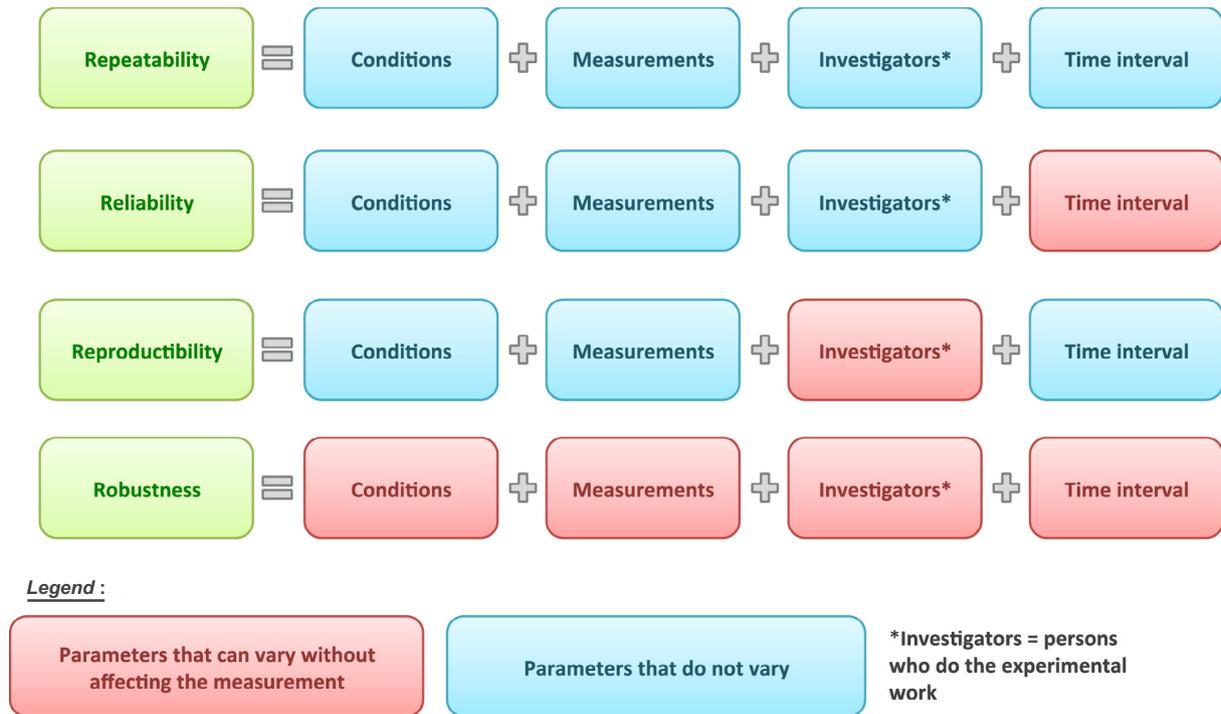


Fig. 1. Comparison of the characteristics of an instrument or method.

- (3) *Reproducibility* means that measurements made by different investigators on a given sample in the same experimental conditions (method, equipment, etc.) do not vary beyond certain pre-defined limits. This approach is exploited in compliance with accreditation standards like ISO [7, 8], and in round-robin tests in which the same sample is tested by different laboratories in strictly defined conditions. Inter-laboratory reproducibility is confirmed if the results do not vary beyond certain pre-defined limits. The confidence interval for the reproducibility can never be smaller than that for the repeatability.
- (4) An instrument or method is *robust* if – whatever the conditions – no deviation is observed. The more resistant an instrument or method is to variations in uncontrolled external conditions, the more robust it is [9]. It is therefore understood that robustness is an estimate that depends on the number of parameters investigated which had no effect on the instrument or its measurements. Of course, the conditions in which an instrument or method is to be used have to be established at the beginning and then specified and adhered to.

Figure 1 outlines relationships between these definitions.

3 Critical points and associated risks

For a method, for an instrument or for a process, every step does not require the same level of monitoring. For this reason, it is important to define sensitive steps (critical

points) as well as methods, instruments and accessories which might entail risk.

Is not always easy and requires both analysis and expert assessment to identify exhaustively all the parameters that might affect a method, using an instrument or running a process. If many parameters are identified, they should be classified by order of importance.

Once the criticality of a method, instrument or information is identified, it is important to quantify the risk (in terms of repercussions) that would result from deviation.

Here are some examples of risks:

- Having to multiply the number of experiments, samples or animals to detect statistically significant differences; with a robust, reliable method, fewer experiments will be required.
- Generating an erroneous result in a medical test that will be used to decide how to treat a patient.
- Taking a decision on the basis of incorrect information.
- Having to retract an article because the results prove wrong or the raw data were not well recorded.
- Losing a customer or collaboration because the results fail to meet the customer's requirements in terms of reliability and robustness.

Mapping of a process and modelling how it works in a flowchart help illustrate the various steps involved. Critical points can be identified on the basis of the investigator's experience together with relevant specifications obtained in the literature or information provided by a supplier, etc. Once critical points are identified, a risk analysis will be performed.

Some examples for critical points are the incubation temperature in a chemical reaction, the room temperature

Table 1. Possible risks and suggested preventive actions.

Possible risk	Preventive action (action which prevents possible problems or non-conformities to arise)
Retraction of an article because of lack of documented evidence	<ul style="list-style-type: none"> • Scrupulously document all results with their context (i.e. all raw and processed data, working conditions, any controls or standards used, etc.) in a well-organized logbook (signed), property of the research institution/university/company.
Loss of documented evidence	<ul style="list-style-type: none"> • Organize and record data and all pieces of evidence for a publication (patent, contract, project) in a special file with well-defined structure. Pay special attention to archiving and future accessibility of the documents.
Loss of a customer (failure to meet specifications)	<ul style="list-style-type: none"> • Intensive and well-documented analysis of customer's needs.
Too much scatter in the results necessitating duplication of the experiment or testing of more samples/animals	<ul style="list-style-type: none"> • Determine and follow conditions of robustness, repeatability, reliability and statistics (also during the design of study).
Making a decision on the basis of incorrect information	<ul style="list-style-type: none"> • Map out the decision-making process and all the contributory data. Check the reliability of data.
Drift in measurements in the course of an experiment	<ul style="list-style-type: none"> • Check any parameter that might cause drift. Make "control" readings.
Expiration date passed	<ul style="list-style-type: none"> • Monitor stocks of critical products.

in an enzyme-catalysed reaction such as reverse transcriptase, the ozone concentration and illumination conditions might affect a fluorophore, or the expiration date which rules out the use of a particular batch of product.

The tolerance of deviation depends on individual requirements and has to be specified. Different pieces of laboratory equipment may be subject to different requirements and stringencies of precision, e.g. some automatic pipettes are particularly closely monitored to guarantee the reliability of their precision, whereas others might be reserved for less stringent applications: the key here is clear identification of the different classes of pipettes, e.g. by using dots of different colours. Moreover, uncontrolled variability can be reduced by specified and controlled working conditions (e.g. animal feed (depending on availability of ingredients), tissue culture media (depending on the batch of serum), biological reactions (depending on the temperature of the room).

As a rule, certain steps of an experiment or process are riskier than others. A flowchart (Fig. 2) helps identify critical points. A critical point is a step at which a problem might result and lead to the generation of a defective product. Table 1 shows some possible risks and suggested preventive actions.

The same approach is possible for critical equipment which needs to be monitored closely.

Critical equipment is used in all areas of experimental research. If a piece of equipment is critical, it has to be qualified [10, 11]. Qualification – to demonstrate that a device is suitable for its intended use and is being adequately maintained and calibrated – is essential before a method can be validated. It involves assessing the device's performance on the basis of a series of criteria (specificity/selectivity, accuracy, precision, and linearity, limits of detection and quantification, application interval, robustness). Qualification can be performed with reference

products, standard weights etc. If reference products are not available, calibration of products, weights, etc. has to be done with the help of other available standards. The following have to be qualified (and fully documented) for a device:

- its performance with respect to the requirements of the research work (especially environmental conditions) (use of kits or reference products);
- calibration details (e.g. tests: pipette calibration; constancy of the balance's intrinsic weighing error; temperature uniform between the top and bottom of a -80°C freezer); recalibration frequency (periodic or continuous), etc. should be stipulated for every instrument;
- maintenance: a logbook should be kept up to date;
- usage conditions (quality control of the appropriate training of users).

Scrupulous recording is vital with respect to the conditions in which a method or device is reliable. If deviations occur, deviations and their context have to be recorded and analysed. Users have to be informed of the deviations and their possible consequences on research data.

Documentation of results covers also the conditions in which critical measurements were made (e.g. humidity measurements inside the laboratory; incubation times; centrifugation speed), equipment and materials used, reference to the computer file containing the results (e.g. storage directory and the name of the file containing a record, for example the photograph of the electrophoretic gel).

The validation process (and its documentation) of any adjustment or the development of a project (what? by whom? when? to replace what? and for its validation: by what means? when? by whom? what result?); as well as for its consequences (need for training? need for more staff?

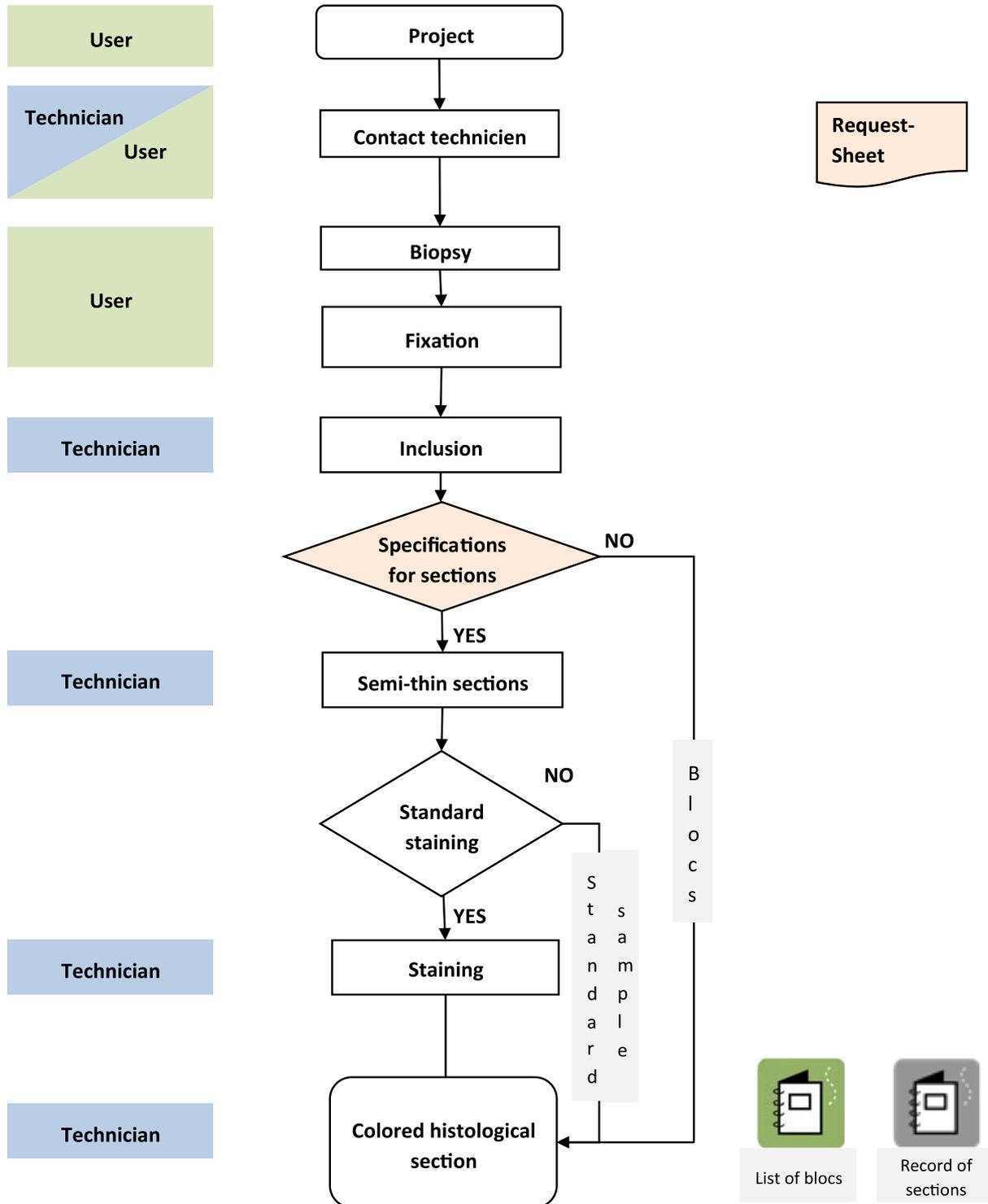


Fig. 2. Preparation of histological samples (critical steps in yellow).

Table 2. Precautions to be taken and quality tests to be carried out as a function of risk and the necessary degree of confidence.

Constraint (from low to high)	Example	Special attention	Precautions/quality tests
<i>Routine work</i>	Assaying a routine sample	Standard Operating Procedures and protocols Supplier's kit	Follow usage instructions and SOPs, etc. In-house tests: control sample
<i>Evolution and new technology</i>	Replacement of a device; new method	Parameters changed one at a time Documentation Define critical points	Establish usage conditions for a new method (usage window) and implement them. Reference sample from another method
<i>Laboratory skills and scientific experience</i>	Doctorate and continuation of the work after the student has left	Standard Operating Procedures and protocols verified, compiled and documented, skills transferred	Traceability of operations to ensure continuity of work. Management of knowledge and skills (hands-on teaching)
<i>Credibility of publications</i>	Reproducible results from a publication in other laboratories	Reliability, reproducibility Methodological validation Validation of equipment	Provide sufficient information in the article
<i>Scientific networks</i>	Joint flow cytometry services in cancer research	Consensus on study design and experimental conditions	Inter-laboratory tests: compare results with a standard Adherence to an accepted standard
<i>Experiment on a rare, precious or expensive sample, subject or consumable</i>	Tissue biopsy	Attention to critical points in the experiment	Flowchart and risk analysis Special precautions and tests
<i>Research on humans</i>	Refrigerated transport of clinical trial samples	Legal texts	Regulatory tests, e.g. check seals during refrigerated samples

review of critical points? rescheduling? modification of reliability and robustness parameters for certain data points or the entire project?) should be documented; this point can become necessary, for example, after the emergence of new knowledge about a method (new publications).

Suppliers have to be told about the quality required of their products. Before the project begins, the Project Team informs them about their imperatives in terms of characteristics (chemical, physical, mechanical, etc.), logistics (volume, shelf-life, same batch, delivery deadlines, etc.) and documentation (certificates, attestations, etc.).

4 Quality control

Reliability and robustness are assessed by checking measurements against plausible, validated or expected results (quality control). A quality test gives a snapshot at a time point “*t*” of a sample. Tests repeated over time (reliability) on diverse samples and in a range of different working conditions (robustness) make it possible to check that a device, method or data point is reliable and robust. The type of test and effort depends on the criticality of work. Table 2 gives a few examples of precautions to be taken and quality tests to be carried out as a function of risk and the necessary degree of confidence.

Quality controls require that (see Fig. 3):

- all conditions (parameters) that might have impact are identified;
- the risk has been evaluated;

- critical points in the method, use of an equipment or running of a process have been defined;
- levels of reliability, repeatability and, if relevant, reproducibility are specified.

Quality control reduces the risk to obtain products that fail to fulfil the specifications.

NB: the idea of product compliant (conform) or non-compliant (non conform) in the spirit of ISO 9001 does not exclude aspects of discovery or innovative scientific research. In practice, it is a “product” of scientific research, either a material product (a synthetic substance, a breed of animal, etc.), an immaterial product (a discovery) or a combination of the two (an experimental result, a dissertation, a report, etc.) which are compliant/conform with specifications compiled by the scientists themselves: triplicate assays, experimental conditions, time-to-thesis defence, etc. A non-compliant product is therefore defined on the basis of those characteristics, i.e. assays not carried out in triplicate, etc. Table 3 gives some examples of non-compliant products.

Documenting the occurrence of a non-compliant product permits to analyse later the underlying cause and to correct it. A non-compliant product has to be processed, either by accepting it as is (immediate release), rejecting it (release refused) or correcting it (release after correction). It is important to start by defining who is qualified to release a non-compliant product and who is not.

Quality control tests to check the reliability of a device, method or data point sometimes involve use of references (standards, calibrators or controls); these may be internal

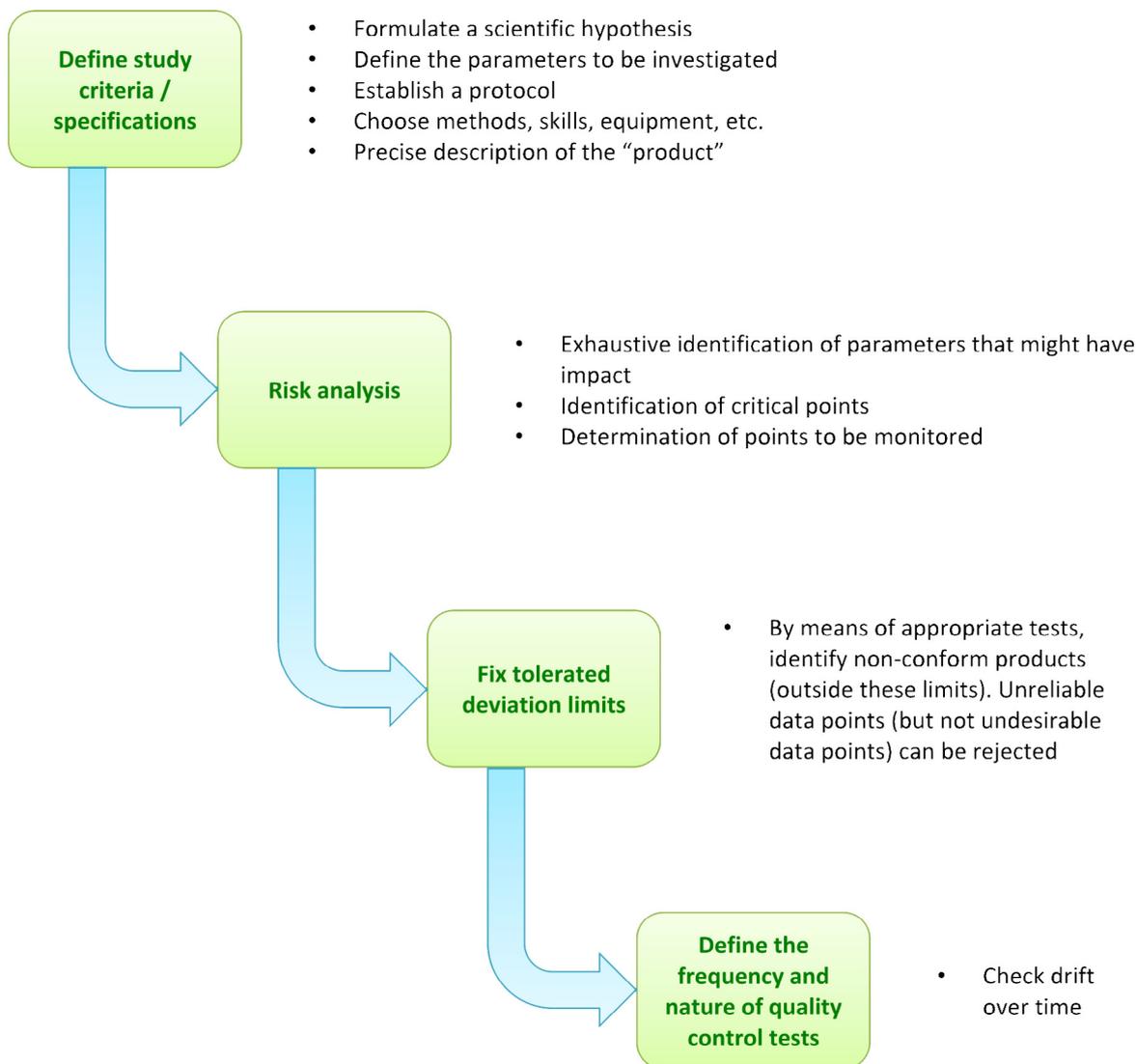


Fig. 3. Organisation of quality control in a research project. The degree of formalisation of the procedures will depend on the research context.

Table 3. Non-compliant products: concrete examples.

Product	Quality control	Non-compliant product
Dissertation	Thesis Committee	Failure to follow the defence schedule
Sensitivity of a method	Defined sensitivity limits	Reading outside of the sensitivity limit
Efficacy of an anti-cancer agent	Defined experimental conditions	Measurements taken outside of permissible conditions

references, nationally or internationally recognised references or references that have been compared to an international standard or calibrator. A standard is a control that has been identified by a group of experts; a calibrator is officially recognised.

Appropriate metrology is always important, e.g. a laser heats up as the cytometric flow equipment runs and data can be deviant over time, but this can be corrected by using internal controls.

5 Investigating causes of poor quality data

Data are reliable, when a series of measurements, checks or comparisons, formalised and recorded, show that they do not drift over time according to tightly defined criteria (*see the definition of “reliable data”*).

A reliable result is obtained by controlling risk in the measurement process. All aspects of measurement have to be controlled. There are various ways of identifying causes.

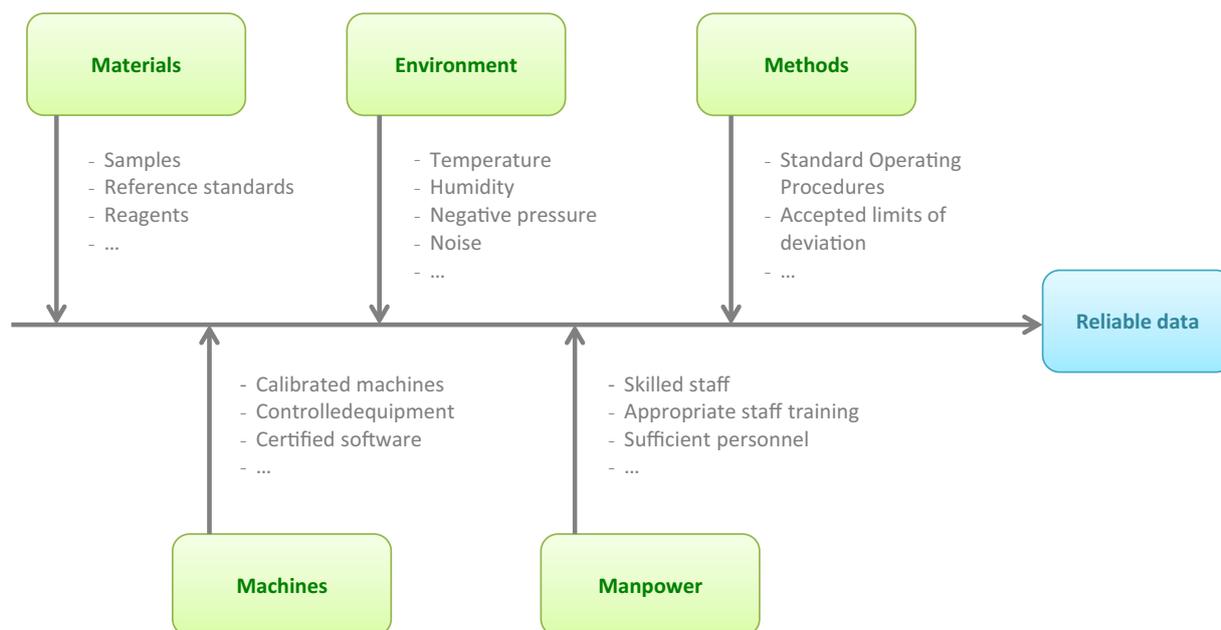


Fig. 4. The “5-points” questioning scheme to identify the sources of defaults and non-conformities.

Root cause analysis is a way of identifying the causes of a malfunction and establishing factors that have to be controlled in a given process. These factors are: materials, machines, manpower, Environment, methods, and, in some methods, management. The factors have to be questioned as shown in Figure 4 in order to identify possible risks. The same questioning scheme is used to determine the origin of a given default or non-conformity or when cause analysis has to be formalized.

6 Management of the different stages of a project

A research project is often carried out as in collaboration between different teams or even in a network [12]. In some cases, the research is conducted on behalf of another party. In these cases as in other circumstances, the work will benefit and risk can be reduced when the work is conducted in Project Management mode (Fig. 5: Conducting research in Project Management mode). The project may involve routine work or research.

The steps and management conditions for both types of work share a number of features but they will be precisely pre-defined and less flexibility will be allowed in the Project Management mode. A Project Leader is appointed and a project schedule established. As seen before, the confidence in this research work will depend on each of the different steps of the project.

Thus, from the design phase on, it is essential to:

- Confirm the scientific relevance of the question being asked.

- Define the project’s legal and regulatory framework, e.g. getting authorisations or rulings from the qualified agencies and/or an Ethics Committee who guarantee the project’s ethical and regulatory compliance. Some journals reject publications about projects that have not been approved by the appropriate authorities.
- Clearly establish needs (e.g. specifications) and methodology (notably the statistics used, sample sizes, etc.).
- Identify resources and methods (e.g. machines, products and reagents) that will fulfil the requirements. Every decision should be justified (is the method known (publication) or validated (standard)? Why select one product over another? Why use a certain animal species?). Also think about the premises, e.g. storage, product conservation (products purchased and produced in house, used and unused), access.
- Identify the necessary skills and qualifications. Set up complementary training programmes.
- Establish a schedule for the project in the light of the above points and also the availability of personnel, machines, products and premises (think about storage, restrictions, archiving, etc.) and the possibility of replacing someone in the event of unforeseen absence or seek replacement for a defective machine.
- Define critical points, associated risks (see above) and what to do in the event of the detection of a non-compliance (Action Plan). Think about risks and restrictions related to other projects: contamination of tissue cultures, confidentiality vis-à-vis competing customers, data safety, availability of premises and machines, etc.
- define parameters for establishing the reliability and robustness of data: What kind of quality control?

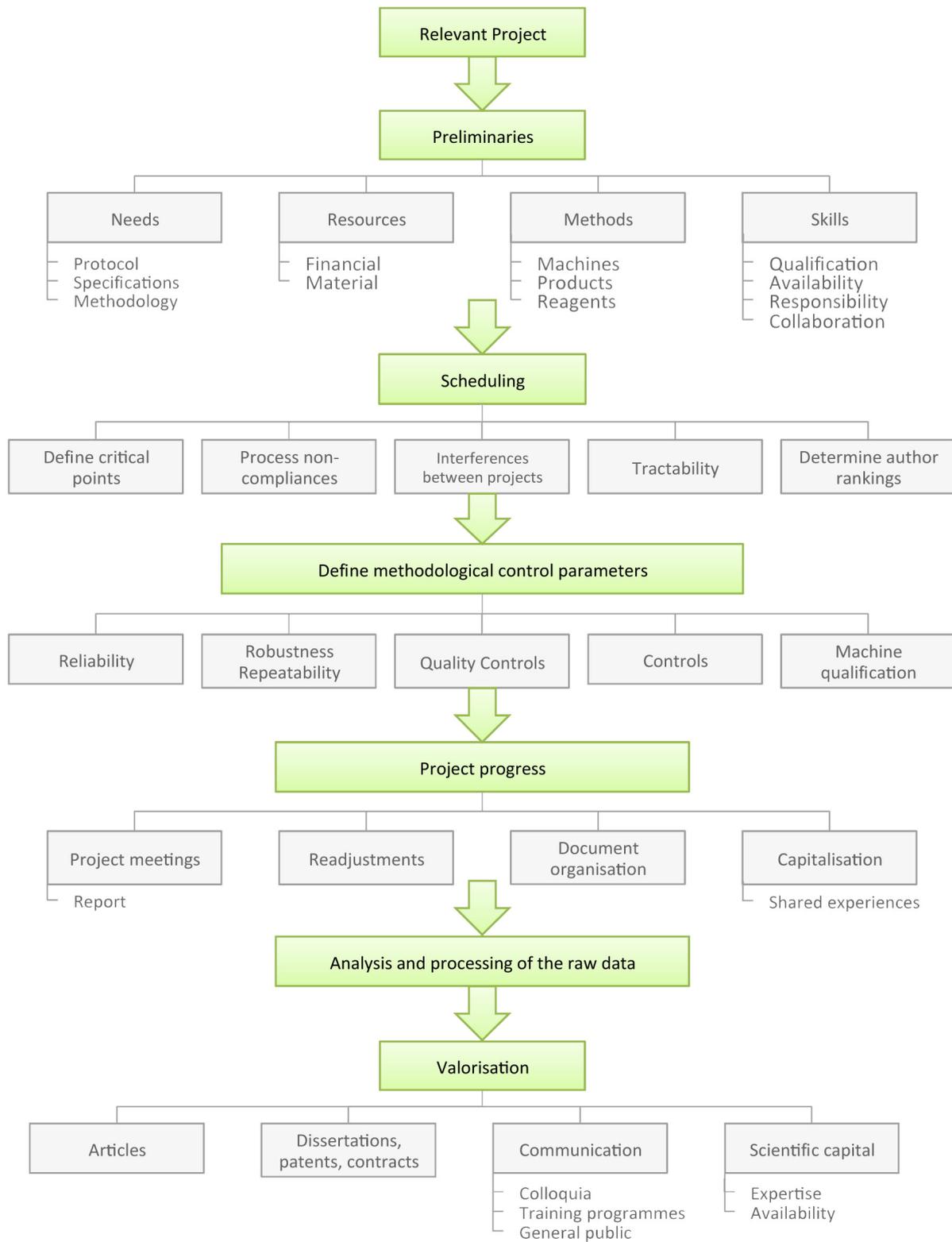


Fig. 5. Conducting research in Project Management mode.

By whom (individual tests or cross-testing – reproducibility)? On what (type of data, sampling, accepted variability)? How many times (repeatability)? What should be documented?

- Identify the responsibilities of all those involved, whatever level they are involved in . . . and broadcast them.
- Define monitoring modalities (progress of the project): scheduling the project is a way of readjusting how the project is run, motivating the team and reassuring potential customers by guaranteeing fulfilment of specifications.

Insofar as a project does not just involve a single person or team but a group or a network, it is essential for the reliability of the project that design and scheduling are established by all those involved (or at least the heads of the teams involved and the beneficiary) in order to cover all requirements/restrictions/possibilities/availabilities. Project management meetings ought to be organised throughout the progress of the work and recorded.

For each partner, tests, measurements and monitoring modalities for each machine/method to be used, ought to be defined in a coordinated fashion (techniques, frequencies, sampling, repeatability and criticality); this will help establish possible risks more precisely. In any case, all critical points and parameters should be discussed at Project Meetings to ensure globally effective organisation (no loss of information).

When the project is underway, reliable, robust data are guaranteed by:

- Reviewing critical points (matching them to the methods, materials and resources being used) and action plans if a risk appears.
- Up-dating of work instructions: a single way of proceeding whatever the number and qualifications of investigators. All instructions should be present (training) and known (perhaps with an examination) by users.
- Coordination of the qualification of equipment.
- Control of reference products and biochemical kits: characteristics fit for purpose; checking of expiration dates/validity, labelling, etc.
- Each partner is responsible for compiling documents on the reliability of the methods and the recording of results. If the same methods are used by different teams or if some type of experiment is shared, “homogenization trials” for approaches and data can be done as described in inter-laboratory tests.
- Any adjustment should be traced by the participating team; if the adjustment has impact on the work of other teams, all participants should be informed.

All these points apply to both, the Project Team and any sub-contractor, working on the project (points presented in the service provision specifications).

When a given method or type of measurement is used by different teams, the reproducibility of results must be ensured; inter-laboratory tests may be a solution.

Other interesting information can be found in two ISO standards, in particular for testing and calibration laboratories [7] and for conformity assessment [8].

7 Discussion and recommendations

If the question “Are specific measures to ensure reliability worth the trouble” was easy to answer, there would be no debate about it. Nevertheless, certain favourable aspects and certain difficult points emerge from an analysis.

Difficulties which have to be overcome

The effort to improve robustness and reliability may be regarded as a restriction that could compromise individual creativity. It could also be perceived as a criticism or a need for justification of the investigator, since he/she has to agree to be transparent about the protocols used, the results obtained, mistakes committed, etc., recognize mistakes and accept judgement and criticism.

Verification, validation, metrology and a quality approach require time and effort on the part of the manager, the Project Leader and the investigators. Defining methods, means, validity parameters, compile and follow specifications, etc. are an investment which only pays in the future.

Globally, the effort for more reliability may generate additional stress to provide results, which fulfil expectations as well as the customer’s and collaborator’s requirements and are obtained within a deadline. Therefore the reasons and sense of the actions designed to increase robustness and reliability have to be explained and accompanied by the manager.

With better traceability, the investigators lose the possibility to “cook”, or “massage” data (if ever they intended to do so) in order to produce the desired result; this may lead to a feeling of increased control and lack of freedom, but this “lack of freedom” is certainly on the desirable side.

Validation and verification clearly require time and effort and staff runs a risk of frustration if the work towards reliability is not valued. Usually, the customer/collaborator does not ask to see validations of methods, results, tolerated variability, etc. and hierarchy does not always value validations and checks when only the number of publications counts. Moreover, the investigators’ professionalism, rigour and efficacy are not readily discerned from a publication or a report.

In favour of more robustness and reliability is a very high degree of confidence

Confidence in oneself may translate into personal development and professional achievement. By virtue of rigour and validation of the used methodology and monitoring of equipment, the investigator can build on the findings for future ventures. With high confidence in his/her results, the investigator is in a position to assess the results and prove their reliability on an objective footing, thereby enhancing his/her credibility and standing with internal and external partners as well as bolstering customers’ and partners’ confidence. Confidence in the use of equipment and experimental methods increases recognition as an expert and demonstrates added value.

A moral commitment to provide reliable results to establish relationships with confidence, transparency, mutual respect and honesty is related to good scientific practice.

The confidence of partners will enhance the laboratory's reputation, renown and, therefore, stability. Customers and collaborators consider word of mouth the most reliable when it comes to their confidence.

Reliable results mean savings of time, energy, and cost and provide more efficiency and safety. Harmonisation of procedures (when appropriate), data analysis and reliability of equipment help cut down the frequency of mistakes and the need for replication thereby reducing stress.

Therefore, since the scientific community, institutions, funding bodies and society as a whole require high-quality, reliable, reproducible and honest research, state-of-the-art verifications and frank communication, work to ensure reliable investigations and more globally an ISO 9001-compatible quality management system are precious tools and a step forward. Consciousness that quality of research is precious is rising in the scientific community and a number of efforts, as also demonstrated by the increasing number of ISO 9001 certifications in research are under way. Hopefully, these "silent" initiatives will in the future benefit from better recognition.

8 Conclusion

Reliability of scientific data is of prime importance for scientific progress and trust into a research group of collaborators, funding bodies and the general public. Individual scientists, but also institutions, as recently pointed out [13], are responsible for the quality of research. Therefore, quality management and good record keeping are promoted now by institutions. Specific measures which we recommend to use to ensure reliability are presented here. These measures are coherent with an ISO 9001 quality management system. The newly revised version (September 2015) of the ISO 9001 standard contains among other new requirements, an enforced risk management of processes. Particularly in experimental sciences the above discussed measures should be useful, since experimental sciences rely strongly on accurate measurements and problems involving reliability and reproducibility of data have been debated lately.

List of contributing authors

Marion Alvarez, Cermep, Lyon, France
 Catherine Chevalier, Inserm US16, Nantes, France
 Valérie Gall, Cerimed, Marseille, France
 Elisabeth Moreau, Inserm US 019, Villejuif, France
 Martine Lelièvre-Pégorier, Inserm U 1138, Paris, France
 Isabelle Remy-Jouet, UMR 1096, Rouen, France
 Séverine Rémy, Inserm U 1064, Nantes, France
 Corinne Schiltz, Inserm DRH, Paris, France

References

1. R. Van Noorden, The trouble with retractions, *Nature* **478**, 26–28 (2011)
2. C.G. Begly and L.M. Ellis, Raise standards for preclinical cancer research, *Nature* **483**, 531–534 (2012)
3. F. Prinz, T. Schlange, K. Asadullah, Believe it or not: How much can we rely on published data on potential drug targets? *Nat. Rev. Drug Disc.* **10**, 712–713 (2011)
4. [https://www.rh.inserm.fr/INSERM/IntraRH/RHPublication.nsf/%28vDocumentsParIDB0%29/172CF2989D8F4FCAC1256BF900370D31/\\$file/Guide_metiers.pdf](https://www.rh.inserm.fr/INSERM/IntraRH/RHPublication.nsf/%28vDocumentsParIDB0%29/172CF2989D8F4FCAC1256BF900370D31/$file/Guide_metiers.pdf)
5. Quality management systems: Requirements 2015, ISO 9001 (ICS: 03.120.10)
6. Afnor FD X60-5000 Terminology relating to reliability, maintainability and availability
7. General requirements for the competence of testing and calibration laboratories, 2005, ISO 17025 (ICS: 03.120.20; 19.020)
8. Conformity assessment – General requirements for proficiency testing, 2010, ISO 17043 (ICS: 03.120.20)
9. A. Rivet, B. Brasme, M. Carret, P. Méras, I. Jeacomine, I. Paintrand, E. Facchin, S. Boulanger, *La Démarche Qualité d'une plate-forme technologique* (2013) CERMAV-CNRS (ISBN 978-2-9546225-0-7)
10. Medical laboratories – Particular requirements for quality and competence: ISO 15189
11. Guide on equipment qualification: <http://www.edqm.eu/en/quality-management-guidelines-86.html>
12. Management des activités de recherché, 2014 NF X 50-553 (ICS: 03.100.40; 03.120.10)
13. C.G. Begley, A.M. Buchan, U. Dirnagl, Robust research: Institutions must do their part for reproducibility. *Nature* **525**, 25–27 (2015)