

Paper presentation: “Preliminary Guidelines for Empirical Research in Software Engineering” (Kitchenham et al. 2002)

Who? Eduardo Moreira Fernandes

From? FEDERAL UNIVERSITY OF MINAS GERAIS
Department of Computer Science
Belo Horizonte, State of Minas Gerais, Brazil

When? April 06, 2015

Prologue

Paper's name	Preliminary Guidelines for Empirical Research in Software Engineering
Authors	Kitchenham, B., Pfleeger, S., Pickard, L., Jones, P., Hoaglin, D., El Emam, K., and Rosenberg, J.
Published in	IEEE Transactions on Software Engineering (TSE), Vol. 28, No. 8, Pages 723-734, August 2002

Overview

Paper goals

- To stimulate discussion about empirical research practices in software engineering,
- To support the design, conduction and evaluation of empirical studies,
- To assist in improving of empirical studies and stimulate the critical evaluation of works, and
- To support the combination of different and related empirical researches.

Work's basis

- Review of guidelines for medical researches, and
- Authors' experience in reviewing empirical researches in software engineering.

Paper's motivation

The problem

Accordingly to the authors' view:

- *The standard of empirical software engineering research [case studies, surveys and others] is poor.*
- This is **not** a particular problem in software research.

An example

High rate of error in medical and clinical empirical researches (as in 1987, 1988 and 1993):

- Methodologic errors,
- Statistical errors, and
- Lack of details about the applied statistical techniques.

Paper's motivation

In TSE papers

Observed deficiencies (just like in medical researches):

- Poor experimental design,
- Inappropriate use of statistical techniques, and
- Discordant conclusions from the obtained results.

Approach to
solve the
problem

Focus in medical research guidelines: medical statisticians are aware of the poverty of statistical analysis in journals.

Paper's content

Problems' origin

The origin of statistical problems in software research is justified by:

- Difficulty of applying standards of statistical procedures in software experiments, and
- Lack of statistical skills by empirical research people.

The solution

Proposal of **preliminary guidelines** covering: experimental context, experimental design, conduct of the experiment and data collection, analysis, presentation of results, and interpretation of results.

Experimental context

Composition

The authors discuss about:

- 1 Background information about the context of the industry **or** of the developed software engineering technique,
- 2 Discussion of the research hypotheses, and
- 3 Information about related research.

Goals

To ensure a sufficient definition of the research objectives **and** a sufficient description of the research.

Experimental context

- C1. *Be sure to specify as much of the industrial context as possible. [...] clearly define the entities, attributes, and measures that are capturing the [...] information.*
- C2. *If a specific hypothesis is being tested, state it clearly prior to performing the study and discuss the theory from which it is derived [...].*
 - Without the link theory-hypothesis, empirical results cannot contribute to a wider body of knowledge.

Experimental context

- C3. *If the research is exploratory, state clearly and, prior to data analysis, what questions the investigation is intended to address and how it will address them.*

Many subsidiary analyses **vs** single purpose study

- C4. *Describe research that is similar to, or has a bearing on, the current research and how current work relates to it.*
- It is the integration of the body of knowledge.

Experimental design

Composition

It describes the products, resources and processes involved in the study, such as:

- 1 The studied population, the rationale and technique for population sampling,
- 2 The intervention process, and
- 3 The methods applied to reduce bias and determine sample size.

Goals

To ensure that the design is appropriate for the study objectives.

Experimental design

- D1. *Identify the population from which the subjects and objects are drawn.*
- It is necessary to be drawn inferences from the experiment results (**e.g.**, generalization of the results to a defined population).
- D2. *Define the process by which the subjects and objects were selected.*
- Random sampling or another method that should be explained and justified.
 - Definition of inclusion and exclusion criteria.

Experimental design

- D4. *Restrict yourself to simple study designs or [...] to designs that are fully analyzed in the statistical literature. If you are not using a well-documented design and analysis method, you should consult a statistician [...].*
- D7. *Use appropriate levels of blinding.*

Conducting the experiment and data collection

Goals

With respect to:

- **Data collection:** to ensure a well-defined data collection process to support the experiment replication.
- **Experiment conduction:** to identify any deviations from the experimental plans.

Conducting the experiment and data collection

- DC1. *Define all software measures fully, including the entity, attribute, unit and counting rules.*
- This will provide: a sufficient understanding of different measurements **and** the determination of translation among measurement schemes.
- DC3. *Describe any quality control method used to ensure completeness and accuracy of data collection.*

Conducting the experiment and data collection

- DC4. *For surveys, monitor and report the response rate and discuss the representativeness of the responses and the impact of nonresponse.*
- DC5. *For observational studies and experiments, record data about subjects who drop out from the studies.*

Analysis

Main goal To ensure the correct analysis of the experimental results. Data analysis in accordance with the study design.

A1. *Specify any procedures used to control for multiple testing.*

- To avoid problems with data overuse (when the data set is small and new data is hard to be obtained).
- There are methods for dealing with multiple tests on the same data set.

Analysis

A2. *Consider using blind analysis.*

- In order to avoid the tendency to “fishing for results” .

A5. *Apply appropriate quality control procedures to verify your results.*

- “Don’t manipulate your data to favor your hypothesis.”

Presentation of results

Main goal

A sufficient presentation of results provides:

- The understanding of all aspects of the study,
- The replication of the study, and
- The combination of the results in a future meta-analysis.

P1.

Describe or cite a reference for all statistical procedures used.

Presentation of results

- P3. *Present quantitative results as well as significance levels. Quantitative results should show the magnitude of effects and the confidence limits.*
- P4. *Present the raw data whenever possible. Otherwise, confirm that they are available for confidential review [...].*
- P6. *Make appropriate use of graphics.*

Interpretation of results

Main goal

To make conclusions that follow directly from the obtained results.

1. *Define the population to which inferential statistics and predictive models apply.*
2. *Differentiate between statistical significance and practical importance.*
 - Researches may show a statistical significance in some result, but there may be no practical importance.

Interpretation of results

13. *Define the type of study.*

- In order to establish the reliance that readers should put on the conclusions of the study, and
- To suggest the suitability of the developed study for future meta-analysis.

14. *Specify any limitations of the study.*

- The researchers are responsible for discussing the limitations of their study.

Paper review & Discussion

The good Sufficient coverage on the topic of empirical research and balance between generalization and specificity.

The bad It is hard to find bad points in this paper because of its solidity.

The bottom line This paper is a good starting point for those looking to venture into empirical research in software engineering.