Protocol for a Tertiary study of Systematic Literature Reviews and Evidence-based Guidelines in IT and Software Engineering

Barbara Kitchenham, Pearl Brereton, David Budgen, Mark Turner, John Bailey and Stephen Linkman

Background

At ICSE04, Kitchenham et al. (2004) Suggested software engineering researchers should adopt "Evidence-based Software Engineering" (EBSE). EBSE aims to apply an evidence-based approach to software engineering research and practice. The ICSE paper was followed-up by a paper at Metrics05 (Jørgensen et al., 2005) and an article in IEEE Software (Dybå et al., 2005).

Following these papers, staff at the Keele University School of Computing and Mathematics proposed a research project to investigate the feasibility of EBSE. This proposal was funded by the UK Economics and Physical Science Research Council (EPSRC). The proposal was amended to include the Department of Computer Science, University of Durham when Professor David Budgen moved to Durham. The EPSRC have now funded a joint Keele and Durham follow-on project (EPIC).

The purpose of the study described in this protocol is to review the current status of EBSE since 2004 using a tertiary study to review articles related to EBSE in particular articles describing Systematic Literature reviews (SLRs)

Evidence-based research and practice was developed initially in medicine because research indicated that expert opinion based medical advice was not as reliable as advice based on scientific evidence. It is now being adopted in many domains e.g. Criminology, Social policy, Economics, Nursing etc. Based on Evidence-based medicine, the goal of Evidence-based Software Engineering is:

> "To provide the means by which current best evidence from research can be integrated with practical experience and human values in the decision making process regarding the development and maintenance of software." (Dybå et al., 2005)

In this context evidence is defined as a synthesis of best quality scientific studies on a specific topic or research question. The main method of synthesis is a Systematic Literature Review (SLR). In contrast to an ad hoc literature review, an SLR is a methodologically rigorous review of research results.

Research Questions

The research questions to be addressed by this study are:

- How much EBSE activity has there been since 2004?
- What research topics are being addressed?
- Who is leading EBSE research?
- What are the limitations of current research?

Search Process

The search process is a manual search of specific conference proceedings and journal papers since 2004. The nominated journals and conferences are:

Table 1 Sources to be Searched

Source	Responsible
Information and Software Technology (IST)	Kitchenham
Journal of Systems and Software	Kitchenham
IEEE Transactions on Software Engineering	Kitchenham
IEEE Software	Kitchenham
Communications of the ACM (CACM)	Brereton
ACM Surveys	Brereton
Transactions on Software Engineering Methods (TOSEM)	Brereton
Software Practice and Experience	Budgen & Kitchenham
Empirical Software Engineering Journal (ESEM)	Budgen
IEE Proceedings Software (now IET Software)	Kitchenham
Proceedings International Conference on Software Engineering (ICSE 04, 05, 06,07)	Linkman, Brereton, & Kichenham
Proceedings International Seminar of Software Metrics (Metrics04, Metrics05)	Kitchenham
Proceedings International Seminar on Empirical Software Engineering (ISESE 04, 05, 05)	Kitchenham & Brereton

Specific researchers will also be contacted directly:
Dr Magne Jørgensen
Professor Guilherme Travassos.

Inclusion criteria

Articles on the following topics, published between Jan 1st 2004 and June 30th 2007, will be included

- Systematic Literature Reviews (SLRs) i.e. Literature surveys with defined research questions, search process, data extraction and data presentation
- Meta-analyses (MA)

Exclusion Criteria

The following types of papers will be excluded

- Informal literature surveys (no defined research questions, no search process, no defined data extraction or data analysis process).

- Papers discussing process of EBSE.

When an SLR has been published in more than one journal/conference, the most complete version of the survey will be used.

Primary study selection process

The results will be tabulated as follows:
- Number of papers per year per source
- Number of candidate papers per year per source
- Number of selected papers per year per source.

The relevant candidate and selected studies will be selected by a single researcher. The rejected studies will be checked by another researcher. We will maintain a list candidate papers that were rejected with reasons for the rejection.

Quality Assessment

Each SLR will be evaluated using the York University, Centre for Reviews and Dissemination (CDR) Database of Abstracts of Reviews of Effects (DARE) criteria (http://www.york.ac.uk/inst/crd/crddatabase.htm#DARE). The criteria are based on four questions:

- Are the review's inclusion and exclusion criteria described and appropriate?
- Is the literature search likely to have covered all relevant studies?
- Did the reviewers assess the quality/validity of the included studies?
- Were the basic data/studies adequately described?

The questions are scored as follows:

- Question 1: Y (yes), the inclusion criteria are explicitly defined in the paper, P (Partly), the inclusion criteria are implicit; N (no), the inclusion criteria are not defined and cannot be readily inferred.
- Question 2: Y, the authors have either searched 4 or more digital libraries and included additional search strategies or identified and referenced all journals addressing the topic of interest; P, the authors have searched 3 or 4 digital libraries with no extra search strategies, or searched a defined but restricted set of journals and conference proceedings; N, the authors have search up to 2 digital libraries or an extremely restricted set of journals.
- Question 3: Y, the authors have explicitly defined quality criteria and extracted them from each primary study; P, the research question involves quality issues that are addressed by the study; N no explicit quality assessment of individual papers has been attempted.
- Question 4: Y Information is presented about each paper; P only summary information is presented about individual papers; N the results of the individual studies are not specified.

The scoring procedure is Y=1, P=0.5 and N or Unknown=0.

The data will be extracted by one researcher and checked by another.

Data Collection

The data extracted from each paper will be:

- The source (i.e. the conference or journal that published the paper).
- The year when the paper was published. Note if the paper was published in several difference sources both dates will be recorded and the first date will be used in any analysis. This is necessary in order to track the EBSE activity over time.
- Classification of paper
    - Type (Systematic Literature Review SLR, Meta-Analysis MA, Evidence-based Guideline EBG). Note a study may have multiple types, for example a study reporting an SLR may also report evidence-based guidelines based on the SLR results.
    - Scope (Research trends or specific research question).
- Main software engineering topic area.
- The author(s) and affiliation (organisation and country).
- Research question/issue.
- Whether the study referenced an EBSE paper or the SLR Guidelines (Kitchenham, 2004).
- How many primary studies were analysed.
- Whether the study proposed practitioner-oriented guidelines.
- Summary of paper.
- Quality score for the study.

The data will be extracted by one researcher and checked by another.

Data Analysis

The data will be tabulated (ordered alphabetically by the first author name) to show the basic information about each study. The number of studies in each major category will be counted.

The tables will be reviewed to answer the research questions and identify any interesting trends or limitations in current EBSE-related research as follows:

- Question 1 How much EBSE activity has there been since 2004? This will be addressed by simple counts of the number of EBSE related papers per year.
- Question 2 What research topics are being addressed? This will be addressed by counting the number of papers in each topic area. We will identify whether any specific topic areas that have a relatively large number of SLRs.
- Question 3 Who is leading EBSE research? We will investigate whether any specific organisation of researches have undertaken a relatively large number of SLRs.
- Question 4 What are the limitations of current research? We will review the range of SE topics, the scope of SLRs and the quality of SLRs to determine whether there are any observable limitations. We will also investigate whether the quality of studies is increasing over time by plotting the quality score against the first publication date, and whether the quality of studies has been influenced by the SLR guidelines (by comparing the average quality score of SLRs that referenced the guidelines with the average score of SLRs that did not reference the guidelines).

Dissemination

The results of the study should be of interest to the software engineering community as well as researchers interested in EBSE. For that reason we plan to report the results on a Web page. We will also document the full result of the study in a joint Keele University and University of Durham technical report. A short version of the study will be submitted to IEEE Software.

References

1. Barbara Kitchenham, Tore Dybå and Magne Jørgensen. (2004) Evidence-based Software Engineering. Proceedings of the 26th International Conference on Software Engineering, (ICSE '04), IEEE Computer Society, Washington DC, USA, pp 273 – 281 (ISBN 0-7695-2163-0).
2. Kitchenham, B. Procedures for Performing Systematic Reviews. Joint Technical Report, Keele University TR/SE-0401 and NICTA 0400011T.1, July 2004.
3. Tore Dybå, Barbara Kitchenham, and Magne Jørgensen. Evidence-based Software Engineering for Practitioners, IEEE Software, Volume 22 (1) January, 2005, pp58-65.
4. Magne Jørgensen, Tore Dybå, and Barbara Kitchenham. Teaching Evidence-Based Software Engineering to University Students, 11th IEEE International Software Metrics Symposium (METRICS'05), 2005, p. 24.