

Evidence-Based Software Engineering

Barbara Kitchenham
Tore Dybå
Magne Jørgensen

1

Agenda

- The evidence-based paradigm
- Evidence-Based Software Engineering (EBSE)
 - Goals
 - Procedures
- Comparison with evidence-based medicine
- Conclusions

2

The Evidence-Based Paradigm

- Evidence-based medicine has changed research practices
 - Medical researchers found
 - Failure to organise existing medical research cost lives
 - Clinical judgement of experts worse than systematic reviews
- Evidence-based paradigm adopted by many other disciplines providing service to public
 - Social policy
 - Education
 - Psychiatry

3

Evidence-Based Software Engineering (EBSE)

- Research question
 - Is evidence-based paradigm feasible for Software Engineering?
- Methodology
 - Analogy-based Comparison
 - Evidence-based paradigm in medicine v software engineering

4

Importance of Question

- Software has a central place in everyday life
 - Public & private services
 - E-business
 - Medical data and equipment
 - Taxation
- Plans to increase dependence on software
 - E-government
 - Life-critical systems
 - Drive-by-wire Cars
 - Wearable monitoring devices

5

Dependability issues

- System failures
 - Waste public money
 - May harm individual citizens
- To gain benefit from new systems
 - Must be dependable
 - Perceived to be dependable
- High level process models (CMM) & individual certification of engineers
- Need reliable production processes/methods
 - Understanding limitations, risks and costs

6

Goal of EBSE

- Adapted from Evidence-Based Medicine
 - 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
- Provides
 - Common goals for research groups
 - Help for practitioners adopting new technologies
 - Means to improve dependability
 - Increase acceptability of software-intensive systems
 - Input to certification process

7

Practicing EBSE

1. Convert information need into answerable question
2. Track down best evidence
3. Critically appraise evidence
4. Integrate critical appraisal with SE expertise and stakeholder requirements
5. Evaluate and improve above steps

8

Comparison with medicine - 1/5

- Convert information need into answerable question
 - Medicine
 - Construct a question based on:
 - Treatment
 - Patient group
 - Outcome measures
 - SE
 - No concept of starting research from an existing body of knowledge

9

Comparison with medicine – 2/5

- Tracking best evidence
 - Medicine
 - Medline (4600 biomedical journals)
 - 6 evidence-based journals specialising in systematic reviews
 - Cochran collaboration
 - Database of systematic reviews (CRT-based)
 - SE
 - No primary source for abstracts
 - One journal dedicated to surveys

10

Finding the literature

- Automated search methods
 - Breakdown research question into components
 - Build search strings
 - Identify & use synonyms & wild cards
- Snowballing (following citations)
- Handsearching
 - Specific question across defined journals
 - Critical conference proceedings
- Finding unpublished studies
 - Write to experts

11

Comparison with medicine - 3/5

- Critically appraise evidence
 - Medicine
 - Cochran collaboration & Australian NHMRC
 - Publish standards for clinical trials & other forms of empirical study
 - Publish standards for systematic reviews
 - SE
 - No agreed standards for empirical studies
 - A proposal for formal experiments/surveys
 - Nothing for qualitative or observational studies
 - No agreed standards for systematic review

12

Systematic Review

- Minimise publication bias
 - Need to ensure that negative results are found
 - Medical sources keep clinical trial registers
 - Can follow all trials not just published ones
 - Contact research groups and experts
 - Review conference proceedings
- Look for duplicate publication
 - Only count each study once
 - No matter how many published reports

13

Comparison with medicine – 4/5

- Integrate critical appraisal with SE expertise and stakeholder needs
 - Medicine
 - ANHMRC Guidelines
 - Using the evidence
 - Centre for Reviews and Dissemination (University of York)
 - Getting research into practice
 - SE
 - Research results seldom used in industry

14

Integrating evidence

- Assessing the evidence
 - Determine strength of evidence (type of study)
 - Size of effects (practical not just statistical)
 - Relevance (appropriateness of outcome measures)
- Applying the evidence
 - Applicability to other settings
 - Summarise benefits & harms
- Present the evidence to stakeholders
 - Balance sheet of evidence & harms

15

Comparison with medicine – 5/5

- Evaluate & Improve
 - Medicine
 - Individual track their own performance against checklists
 - SE
 - Concept fits well with
 - Quality Improvement Process (QIP)
 - General process improvement concepts

16

Scientific Issues- 1/2

- The skill factor
 - SE methods usually require a trained individual
 - Cant blind subject to treatment
 - Cant control for experimenter and subject expectations
- Need to improve protocols
 - Use blinding whenever possible
- Need to qualify our experiments appropriately
 - Strength of evidence will be less

17

Scientific Issues –2/2

- The lifecycle issue
 - Techniques interact with other techniques over along period of time
 - Difficult to determine causal links between techniques and outcomes
 - Intermediate outputs of a specific task may not be meaningful to practitioners
 - Improved reliability cant be demonstrated in a design document

18

Addressing Lifecycle issues

- Experiments on techniques in isolation
 - Still have problem that outcomes are not practitioner relevant
- Large-scale empirical studies
 - Hard to generalise because context is critical
 - Quasi-experiments similar to experiments but without randomisation
 - Need arguments to justify causality
 - Benchmarks based on data from a variety of projects
 - Difficulty with representativeness

19

Conclusion

- ESBE lacks the infrastructure required to support evidence-based paradigm
 - Would need financial support to put in place appropriate infrastructure
- Scientific problems more intractable
 - Need to develop appropriate protocols for SE studies
- Some aspects of EBSE easy to adopt
 - Systematic review
 - Requirement of every PhD student
 - Procedures easy to adopt from medicine
- EBSE needs to be tested on real problems

20