

Preliminary Guidelines for Empirical Research in Software Engineering

Kitchenham, B.A.; Pfleeger, S.L.;
Pickard, L.M.; Jones, P.W.;
Hoaglin, D.C.; El-Emam, K.;
Rosenberg, J.

Alunos: Mauricio Arimoto
Vânia de Oliveira Neves
Professor: Paulo Cesar Masiero



Introdução

- Necessidade de avaliar a qualidade dos estudos individuais incluídos em uma meta-análise;
- Enfoque nas diretrizes médicas;
- As diretrizes foram adequadas aos problemas de estudos empíricos de fenômenos da Engenharia de Software;
- A preocupação é com a criação de diretrizes para ajudar pesquisadores a evitar as principais armadilhas em suas atividades de pesquisa e apresentar suas pesquisas corretamente.



Introdução

- As diretrizes auxiliam no que “fazer” e no que “não fazer” em seis áreas temáticas básicas:
 - Contexto do Experimento
 - Projeto do Experimento
 - Condução do Experimento e Coleta de Dados
 - Análise
 - Apresentação dos Resultados
 - Interpretação dos Resultados

Contexto do Experimento

- Envolve três elementos:
 - Informações contextuais sobre as circunstâncias industriais em que um estudo empírico ocorre ou em que uma nova técnica é desenvolvida;
 - Análise das hipóteses da pesquisa e como elas foram obtidas;
 - Informações sobre pesquisas relacionadas.
- Objetivo:
 - Garantir que os objetivos da pesquisa estão definidos corretamente;
 - Garantir que a pesquisa fornece detalhes suficientes para outros pesquisadores e profissionais.

Contexto do Experimento

- **C1:** Certifique-se de especificar o máximo possível do contexto industrial. Em particular, defina claramente as entidades, os atributos e medidas que estão captando a informação contextual.
- **C2:** Se uma hipótese específica está sendo testada, indique-a claramente antes de realizar o estudo, e discuta a teoria a partir do qual ela é obtida, de modo que suas implicações sejam evidentes.
- **C3:** Se a pesquisa é exploratória, especifique de forma clara e, antes da análise dos dados, quais questões a investigação se destina a tratar, e como ela vai lidar com elas.
- **C4:** Descrever a pesquisa que é semelhante, ou tenha uma influência sobre a pesquisa atual e como o trabalho atual se relaciona com ela.



Projeto do Experimento

- O projeto do estudo descreve os produtos, recursos e processos envolvidos no estudo, incluindo:
 - A população a ser estudada;
 - A lógica e técnica de amostragem da população;
 - O processo de atribuição e gestão dos tratamentos;
 - Os métodos utilizados para reduzir o viés e determinar o tamanho da amostra.
- Objetivo:
 - Garantir que o projeto é apropriado para os objetivos do estudo.

Projeto do Experimento

- **D1:** Identificar a população a partir dos quais os sujeitos e objetos são obtidos.
- **D2:** Definir o processo pelos quais os sujeitos e objetos foram selecionados.
- **D3:** Definir o processo pelos quais os sujeitos e objetos são atribuídos a tratamentos.
- **D4:** Restringir-se a estudo de projetos simples ou, pelo menos, aos projetos que são completamente analisados na literatura. Se você não está replicando o projeto e análise de um estudo anterior, você deve consultar um estatístico para verificar se o seu é o projeto mais eficaz para o que você quer realizar.

Projeto do Experimento

- **D5:** Definir a unidade experimental.
- **D6:** Para os experimentos formais, realize um pré-experimento ou pré-cálculo para identificar ou estimar o tamanho mínimo da amostra necessária.
- **D7:** Utilizar níveis adequados de “blinding”.
- **D8:** Se você não puder evitar a avaliação de seu próprio trabalho, então torne explícito quaisquer interesses de investida (incluindo as suas fontes de apoio), e relate o que você tem feito para minimizar o viés.

Projeto do Experimento

- **D9:** Evitar o uso de controles sem menos ter a certeza da situação de controle poder ser claramente definida.
- **D10:** *Definir todos os tratamentos (intervenções).*
- **D11:** *Justificar a escolha das medidas de resultados em termos de sua relevância para os objetivos do estudo empírico.*



Condução do Experimento e Coleta de Dados

- A condução de um experimento envolve a coleta das medidas dos resultados experimentais.
 - Problema para experimentos relacionados ao software.
- Objetivo:
 - Garantir que o processo tenha sido bem definido para que o experimento possa ser replicado.



Condução do Experimento e Coleta de Dados

- **DC1:** Definir totalmente todas as medidas de software, incluindo a entidade, o atributo, a unidade e as regras de contagem.
- **DC2:** Para medidas subjetivas, apresente uma medida de concordância entre observadores, tais como a estatística kappa ou o coeficiente de correlação intra-classe para medidas contínuas.
- **DC3:** Descrever qualquer método de controle de qualidade utilizado para garantir a plenitude e precisão dos dados coletados.



Condução do Experimento e Coleta de Dados

- **DC4:** Para surveys, monitore e relate as taxas de resposta, e discuta a representatividade das respostas e o impacto das não-respostas.
- **DC5:** Para estudos de observação e experimentos, registre dados sobre indivíduos que abandonam (drop-out) os estudos.
- **DC6:** Para estudos de observação e experimentos, registre dados sobre outras medidas de desempenho que podem ser adversamente afetados pelo tratamento, mesmo que eles não sejam o foco principal do estudo.

Análise

- Há duas abordagens para analisar os resultados dos experimentos:
 - Análise clássica;
 - Análise bayesiana.
- Outra questão é a escolha entre a análise paramétrica e não paramétrica.
- Objetivo:
 - Garantir que os resultados experimentais obtidos são analisados corretamente.

Análise

- **A1:** Especificar os procedimentos utilizados para controlar múltiplos testes.
- **A2:** Considerar o uso de análise cega (blind).
- **A3:** Realizar análise de sensibilidade.
- **A4:** Garantir que os dados não violam as suposições de testes utilizados neles.
- **A5:** Aplicar os procedimentos de controle apropriados para verificar os resultados.



Apresentação dos resultados

- O leitor de um estudo deve ser capaz de compreender a razão, o projeto, a análise, os resultados e o significado dos resultados desse estudo.
- Também deve ser capaz de reproduzir ou replicar esse estudo.



Apresentação dos resultados

- **P1:** Descrever ou citar uma referência para todos os procedimentos estatísticos utilizados.
- **P2:** Relatar o pacote estatístico utilizado.
- **P3:** Apresentar resultados quantitativos assim como níveis de significância. Resultados quantitativos devem mostrar a magnitude dos efeitos e os limites de confiança.

Apresentação dos resultados

- **P3:** Apresentar resultados quantitativos assim como níveis de significância. Resultados quantitativos devem mostrar a magnitude dos efeitos e os limites de confiança.
 1. Relatar informações sobre as diferenças das pessoas quando se utiliza dados pareados.
 2. Relatar a magnitude de um tamanho de efeito.
 3. Relatar os limites de confiança para inferência estatística
 4. Relatar o nível alfa utilizado para os testes estatísticos
 5. Relatar se os testes foram bicaudais ou unicaudal.
 6. Relatar o valor da estatística t.
 7. Para a regressão, relatar a equação de regressão.
 8. Para a regressão, relatar o coeficiente de determinação.
 9. Para a regressão, se o modelo será usado para a previsão, relatar o processo de validação e os resultados.
 10. Para dar apoio aos requisitos de meta-análise, sempre relatar o erro padrão da variação média nos resultados medidos quando as medidas mudarem a partir da baseline para um momento posterior.

Apresentação dos resultados

- **P4:** Apresentar os dados brutos quando possível. Confirmar que eles estão disponíveis para revisões confidenciais pelos revisores e auditores independentes.
- **P5:** Fornecer estatística descritiva apropriada.
 1. Relatar o número de observações.
 2. Relatar todos os números com o grau adequado de precisão.
 3. Apresentar o numerador e o denominador de porcentagens.
 4. Com números pequenos, apresentar valores e não porcentagens
 5. Apresentar as medidas apropriadas de tendência central e dispersão quando resumir os dados contínuos.
 6. Não use o erro padrão no lugar do desvio padrão como uma medida de dispersão.
 7. Se os dados contínuos forem separados por "pontos de corte" em categorias ordinais, dar as razões para escolhê-los.
 8. Se os dados foram transformados, converter as unidades de medida de volta para as unidades originais para o relatório.

Apresentação dos resultados

- **P6:** Fazer o uso adequado dos gráficos.
 - *Erros comuns:*
 - Representação unidimensional de dados em duas ou mais dimensões;
 - Utilização de gráficos de pizza;
 - Escolha inapropriada da escala;
 - Omitir pontos distantes de gráficos de dispersão;
 - Omitir jittering em gráficos de dispersão, quando muitos pontos de dados se sobrepõem.



Interpretação dos resultados

- Todas as conclusões devem ser obtidas diretamente dos resultados.
 - Não se deve introduzir um novo material na seção de conclusões.



Interpretação dos resultados

- **I1:** Definir a população na qual a estatística inferencial e modelos preditos se aplicam.
- **I2:** Diferenciar entre significância estatística e importância prática.
- **I3:** Definir o tipo de estudo.
- **I4:** Especificar *as limitações do estudo*.



Conclusões e Discussões

- Foram apresentadas várias diretrizes
 - melhorar a qualidade da execução e da avaliação da pesquisa empírica.
- Importante porque:
 - Os pesquisadores de software muitas vezes cometem erros estatísticos;
 - Pesquisadores seniores estão pressionando por mais pesquisas empíricas para apoiar a engenharia de software.



Conclusões e Discussões

- Pesquisadores podem melhorar suas pesquisas;
- Revisores e leitores podem usar as diretrizes para avaliar a qualidade da pesquisa;
- Algumas diretrizes podem ter implicações éticas e/ou metodológicas;
- As diretrizes não são suficientes por si só.