Tradução do artigo : Lessons from applying the systematic literature review process within the software engineering domain

Autores: Pearl Brereton a,\*, Barbara A. Kitchenham a, David Budgen b, Mark Turner a, Mohamed Khalil

**Resumo:**

Uma consequência do crescente número de estudos empíricos em engenharia de software é a necessidade de adotar abordagens sistemáticas para avaliar e agregar resultados de pesquisa, a fim de fornecer um resumo equilibrado e objetivo das evidências de pesquisa para um determinado tema. O artigo relata experiências com a aplicação de uma dessas abordagens, a prática de revisão sistemática da literatura, aos estudos publicados relevante para tópicos do domínio de engenharia de software. O processo sistemático de revisão da literatura é resumido, uma série de são descritas revisões realizadas pelos autores e outros e algumas lições sobre a aplicabilidade dessa prática ao software engenharia são extraídos.

O processo sistemático básico de revisão da literatura parece apropriado à engenharia de software e à preparação e validação de um o protocolo de revisão antes de uma atividade de revisão é especialmente valioso. O artigo destaca áreas onde alguma adaptação do processo para acomodar as características específicas do domínio da engenharia de software, além de áreas nas quais as melhorias nas a infraestrutura e as práticas de engenharia de software aumentariam sua aplicabilidade. Em particular, o suporte à infraestrutura fornecido pelo software os bancos de dados de indexação de engenharia são inadequados. Além disso, a qualidade dos resumos é ruim; geralmente não é possível julgar a relevância de um estudo a partir de uma revisão do resumo sozinho.

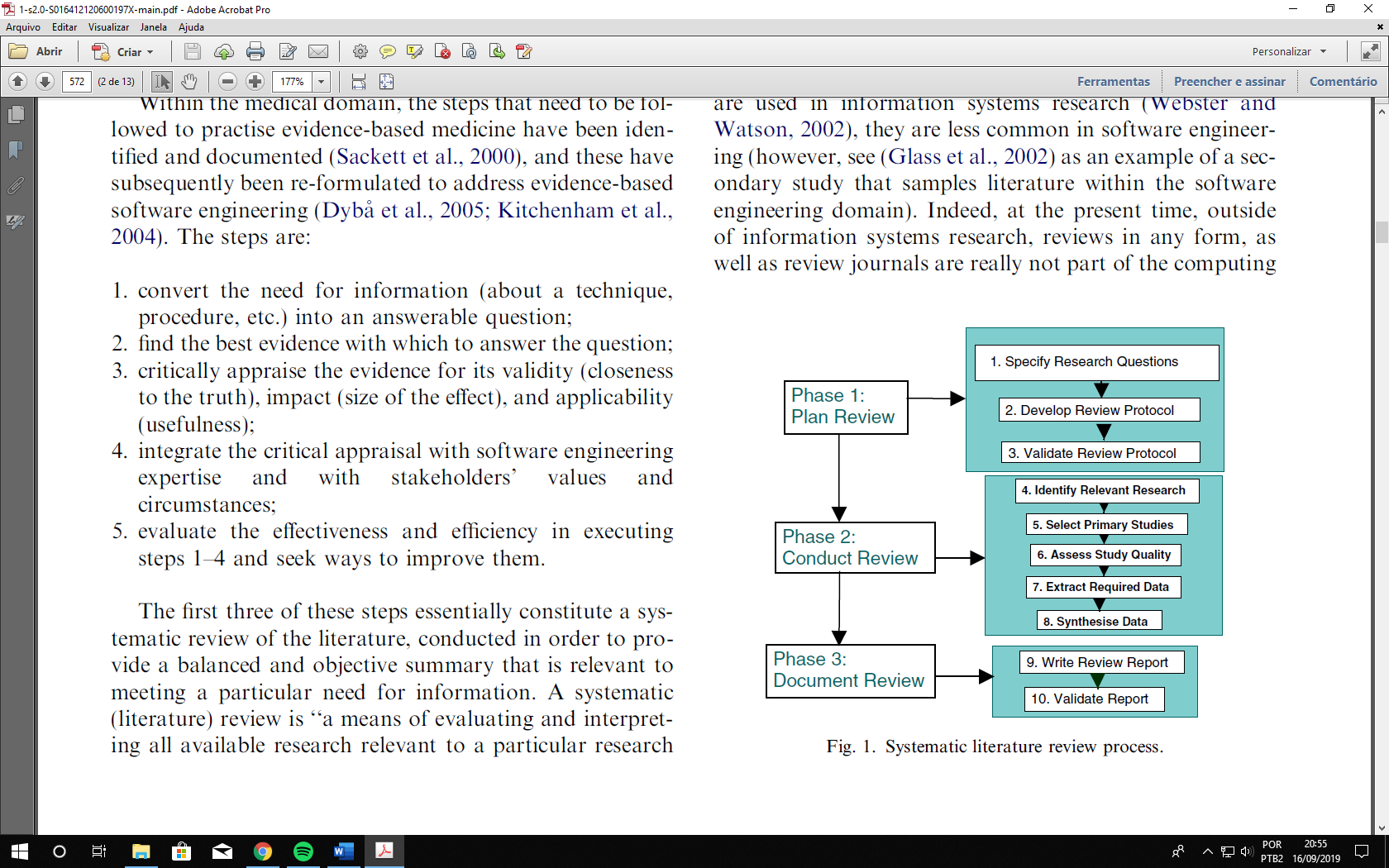
**Introdução:**

Estudos empíricos estão agora sendo realizados com mais frequência, como meio de investigar uma ampla gama de fenômenos na engenharia de software. No entanto, como cada inevitavelmente, seu escopo é limitado, pesquisadores e os fabricantes precisam ser capazes de rigorosa e sistematicamente localizar, avaliar e agregar os resultados de todos estudos empíricos relevantes relacionados a um tópico específico de interesse, a fim de fornecer um resumo objetivo da evidência relevante. Esta necessidade foi tratada dentro de um número de outras disciplinas, incluindo medicina clínica, política social, educação e sistemas de informação, através de a aplicação do paradigma baseado em evidências. A evidência-paradigma baseado em direitos preconiza a avaliação objetiva e síntese de resultados empíricos relevantes para um determinado questão de pesquisa através de um processo de literatura sistemática revisão e integração dessas evidências em prática profissional (Sackett et al., 2000).

Uma das críticas que podem ser feitas aos pesquisadores em engenharia de software e ciência da computação, particularmente em contraste com os sistemas de informação, é que eles fazem pouco ou nenhum uso dos métodos e experiências disponíveis de outras 'disciplinas de referência' (Glass et al., 2004). Nós portanto, neste artigo, apresente e discuta nossas experiências de aplicar o processo sistemático de revisão de literatura à domínio de engenharia de software com o objetivo de identificar os aspectos do processo que são transferidos "como estão" para o software engenharia, aqueles que precisam ser adaptados às particularidades características do domínio e áreas em que melhorias na infraestrutura atual de engenharia de software e práticas são necessárias para obter maior vantagem do processo. A seção 2 fornece uma introdução à literatura sistemática revisão destacando as principais fases e estágios da o processo. Isto é seguido, nas Seções 3-5, respectivamente, por visões gerais de três revisões sistemáticas relacionadas ao serviço baseados em sistemas, o modelo de aceitação de tecnologia e diretrizes para conduzir revisões sistemáticas da literatura. Lições aprendemos de nossas experiências ao realizar essas revisões são então descritos e discutidos e, finalmente, algumas conclusões são desenhados.

**2. Revisão sistemática da literatura**

O contexto mais amplo deste estudo é o de investigar o uso do paradigma baseado em evidências na engenharia de software. A possibilidade de aplicar o paradigma baseado em evidências para o campo de engenharia de software foi criado, discutidos e entusiasticamente apoiados no ICSE 2004 (Kitchenham et al., 2004). O objetivo do software baseado em evidências engenharia (EBSE) é resumida por Kitchenham et al., como sendo: provide provide fornecer os meios pelos quais as atuais as melhores evidências da pesquisa podem ser integradas a práticas experiência e valores humanos na tomada de decisão processo de desenvolvimento e manutenção de Programas''.



Dentro do domínio médico, as etapas que precisam ser seguidas para praticar medicina baseada em evidências foram identificados documentados (Sackett et al., 2000), e estes têm posteriormente reformulado para tratar de evidências engenharia de software (Dyba˚ et al., 2005; Kitchenham et al., 2004). Os passos são:

1. converter a necessidade de informações (sobre uma técnica, etc.) em uma pergunta respondida;

2. encontre a melhor evidência para responder à pergunta;

3. avaliar criticamente as evidências quanto à sua validade (proximidade verdade), impacto (tamanho do efeito) e aplicabilidade (utilidade);

4. integrar a avaliação crítica à engenharia de software conhecimento e com os valores e circunstâncias;

5. avaliar a efetividade e eficiência na execução etapas 1 a 4 e procure maneiras de melhorá-las.

As três primeiras etapas constituem essencialmente uma sistemática revisão da literatura, realizada com o objetivo de fornecer um resumo equilibrado e objetivo que seja relevante para atendendo a uma necessidade específica de informações. Uma sistemática revisão (literatura) é um meio de avaliar e interpretar toda a pesquisa disponível relevante para uma pesquisa específica questão ou área de tópicos ou fenômeno de interesse "(Kitchenham, 2004). Os trabalhos de pesquisa resumidos no revisão são referidos como estudos primários, enquanto a revisão em si é um estudo secundário. A acumulação de evidências através de estudos secundários pode ser muito valioso em oferecer novas idéias ou na identificação de onde um problema pode estar esclarecido por estudos primários adicionais. Por exemplo, um estudo de excedentes de custo de software mostrou que os resultados relatados em um estudo altamente influente realizado no início dos anos 90 (relatório CHAOS) estavam significativamente passo com os relatados em outros estudos (Jørgensen e Moløkken-Østvold, 2006). Uma avaliação crítica de o relatório CHAOS de Jørgensen e Moløkken-Østvold identificou vários problemas metodológicos. Outro Por exemplo, onde surgiram novos insights, é uma sistemática revisão do poder estatístico em experimentos de engenharia de software (Dyba˚ et al., 2006). Aqui, os resultados mostram que o poder estatístico de experimentos de engenharia de software cai substancialmente abaixo das normas aceitas, bem como dos níveis encontrado na disciplina relacionada de sistemas de informação pesquisa''. Os autores continuam fazendo recomendações sobre como pesquisadores empíricos de engenharia de software pode resolver as deficiências relatadas.

A realização de uma revisão sistemática envolve várias atividades, que podem ser agrupadas em três fases principais: planejamento; conduzir a revisão; e relatar a revisão. A Fig. 1 ilustra o processo geral de revisão em 10 etapas.

As revisões sistemáticas da literatura estão principalmente preocupadas com o problema de agregar evidências empíricas que pode ter sido obtido usando uma variedade de técnicas, e em contextos (potencialmente) amplamente diferentes - o que é geralmente o caso da engenharia de software. Enquanto eles são usados ​​na pesquisa de sistemas de informação (Webster e Watson, 2002), são menos comuns na engenharia de software (no entanto, veja (Glass et al., 2002) como um exemplo de estudar amostras de literatura dentro do software domínio de engenharia). De fato, atualmente, fora pesquisa de sistemas de informação, revisões de qualquer forma, bem como revistas de revisão realmente não fazem parte da computaçãocultura de pesquisa, que se concentra quase inteiramente na publicação estudos primários.

Para entender o papel da evidência, precisamos reconhecer que, em um amplo espectro de disciplinas de estudo, existe um requisito comum para encontrar práticas objetivas que possam ser empregado para agregar os resultados de diferentes estudos empíricos de maneira consistente. O alcance de formas e questões é muito amplo: por um lado, agregando (digamos) estudos experimentais que medem a massa do elétron é basicamente uma questão de usar matematicamente transformações para ajustar as variações experimentais condições; considerando que reunir os resultados de uma conjunto de pesquisas, que podem ter empregado diferentes conjuntos de perguntas e foi administrado a populações bastante diferentes, apresenta um problema muito menos matematicamente tratável. Uma das principais questões subjacentes a essa diferença é a papel do ser humano no processo de coleta de dados: no ex, o único envolvimento é como observador externo, enquanto neste último, o humano é participante do tratamento em si.

Esse processo de agregação é fundamental para qualquer abordagem baseada em evidências que busca fornecer objetivos resumos de dados empíricos. A área da medicina ocupa uma posição intermediária entre os dois exemplos acima. A medicina clínica (pelo menos) é capaz de fazer uso extensivo de ensaios clínicos randomizados (ECR) paradigma experimental e nos ECRs utilizados em clínicas estudos, o papel do humano é como sujeito, sendo receptor do tratamento experimental. Isso torna possível realizar meta-análise para agregação dos resultados, e isso, juntamente com a natureza de alguns dos resultados, ajudou a consolidar o sucesso da medicina baseada em evidências. Na engenharia de software, no entanto, ECRs e metaanálises 1 raramente são possíveis e precisamos estabelecer se os resultados de uma revisão sistemática são úteis, apesar de a ausência desses procedimentos.

Um elemento essencial na condução de um estudo secundário como uma revisão sistemática da literatura, é estabelecer um protocolo para o estudo (durante a fase de planejamento). O protocolo visa minimizar o viés no estudo, definindo em avançar como a revisão sistemática será realizada, e é ele próprio um documento que deve ser revisado. O protocolo incorpora o plano detalhado da revisão, especificando processo a ser seguido, quaisquer condições a serem aplicadas quando seleção de estudos primários, quaisquer condições de contorno, qualidade etc. Também pode especificar detalhes de quem empreenderá as várias tarefas e subtarefas. O protocolo pode ser um documento bastante substancial (que para o segundo caso estudo descrito neste artigo tem cerca de vinte páginas) e deve estar sujeito ao controle de versão. Em particular, os motivos de quaisquer alterações que ocorram após o protocolo foi acordado deve ser gravado. Como consequência, o protocolo também deve ser utilizável como base para documentar uma revisão.

Para investigar a aplicabilidade da literatura sistemática análises ao domínio de engenharia de software, o autores (com outros) se comprometeram ou estão em processo da empresa, uma série de revisões que visam abordar uma série de questões de engenharia de software. Estes revisões são resumidas nas seções a seguir, usando os cabeçalhos abstratos estruturados (contexto, objetivos, métodos, resultados, conclusões) que fazem parte das recomendações por relatar revisões sistemáticas (Khan et al., 2001). Além disso, para cada uma das revisões, há uma breve introdução e uma descrição do seu estado atual. O as avaliações usaram as diretrizes da Kitchenham, conforme descrito em (Kitchenham, 2004) ou, em alguns casos, em versões anteriores do relatório. Na Seção 6, discutimos as lições aprendidas dessas revisões em relação ao modelo de processo mostrado em Figura 1.

Ressaltamos que os resultados desses as avaliações são ou serão relatadas em outros lugares. Para os propósitos Neste artigo, estamos preocupados em relatar o processo de realização de revisões sistemáticas em software engenharia e avaliação da sua praticidade. Enquanto não todas as revisões descritas aqui foram concluídas em completo, todos eles progrediram o suficiente para fornecer experiência das duas primeiras fases da revisão processo.

**3. Sistemas baseados em serviços: uma revisão sistemática de problemas (R1)**

Esta revisão foi realizada com o objetivo de identificar questões de pesquisa e, portanto, difere do "tradicional" visão centrada na hipótese de uma questão de pesquisa – embora provavelmente é bastante típico de toda uma classe de pesquisa questões que possam surgir em um assunto de base tecnológica, como engenharia de software. O tópico de modelos de serviço de software é ele próprio um que ainda não levou a nenhum empírico significativo estudos e ainda está sendo mapeado. Assim aqui, o objetivo da revisão era identificar os dois questões sendo levantadas e também até que ponto as questões levantados em diferentes trabalhos foram consistentes. A revisão é essencialmente completo, embora o relatório final ainda esteja para ser validado externamente.

*3.1 Contexto*

Em 2000, Brereton e Budgen publicaram uma classificação dos problemas que precisavam ser resolvidos para componentes sistemas (CBS) para atingir todo o seu potencial (Brereton e Budgen, 2000). Desde aquela época, houve um considerável crescimento em pesquisa e prática nos campo de sistemas baseados em serviço (SBS), levando à visão (de autores e outros) que uma classificação semelhante de problemas para o SBS seriam úteis. Uma equipe de pesquisadores de engenharia de software da King's College London e as Universidades de Durham, Keele e Manchester realizou a revisão.

*3.2 Objetivos*

Os objetivos da revisão foram:

• identificar os principais problemas que precisam ser abordados se O SBS deve ser implementado com sucesso e amplamente adotado;

• identificar as soluções propostas para abordar as questões levantadas;

• identificar os métodos de pesquisa utilizados para investigar propostas soluções;

• fornecer uma estrutura para o posicionamento de novas pesquisas Atividades;

• identificar as lacunas na pesquisa atual.

*3.3 Métodos*

Seis periódicos de arquivo datados de 2000 foram utilizados como dados fontes e papéis foram selecionados para revisão se tivessem um termo-chave (ou sinônimo) no título, resumo ou palavras-chave Lista. Todos os trabalhos selecionados foram incluídos na revisão (ou seja, nenhum outro critério de qualidade foi aplicado, com base que estes eram artigos de jornal que já teriam cuidadosamente revisado). Durante o processo de extração de dados (estágio 7) os revisores foram auxiliados pelo uso de um formulário de registro de dados e notas de orientação.

*3.4 Resultados*

Os principais problemas identificados como precisando ser resolvidos abordado se o SBS deve ser implementado com sucesso e amplamente adotados foram mudança, seleção e coordenação e as soluções apresentadas foram focadas principalmente em tecnologias. Os métodos de pesquisa utilizados foram principalmente os de implementação de conceito e análise conceitual. UMA estrutura baseada na estrutura CBS desenvolvida anteriormente foi proposto e as lacunas identificadas incluíram tópicos relacionados a negócios e questões orientadas para as pessoas.

*3.5 Conclusões*

Os objetivos da revisão sistemática foram bem-sucedidos alcançadas e experiências de aplicação da metodologia a um tópico de engenharia de software também são relatados (Brereton et al., 2005).

**4. Revisão sistemática da literatura sobre aceitação de tecnologia modelo e seus recursos preditivos (R2)**

Esta revisão é de natureza mais "convencional" do que R1, em que existe um conjunto claro de perguntas de pesquisa que só podem ser respondido apelando a resultados empíricos. Além do que, além do mais, Embora, no momento da redação, a revisão esteja incompleta, um A revisão preliminar do piloto já foi realizada pela um de nossos alunos, levantando questões suficientes para motivar essa estudo mais completo.

A revisão está no estágio 7 (ver Fig. 1), ou seja, o protocolo foi desenvolvido e revisado, estudos primários têm selecionados e avaliados e os dados necessários estão sendo extraído. Conforme recomendado em todas as orientações médicas, o O protocolo de revisão foi intensivamente testado. Isso significa que, embora o estudo ainda esteja formalmente na fase intermediária, a maioria dos elementos necessários para conduzir uma revisão já foi explorado.

*4.1 Contexto*

O estudo está sendo conduzido pelos autores deste trabalho como parte da Engenharia de Software Baseada em Evidências projeto e outros (<http://evidence.cs.keele.ac.uk>) financiado por EPSRC.2 O modelo de aceitação de tecnologia (TAM) foi proposto por Davis (1989) e Davis et al. (1989) como instrumento para predizer a probabilidade de um novo tecnologia adotada dentro de um grupo ou organização, e houve várias variações subsequentes do modelo original da TAM, incluindo TAM2 (Venkatesh e Davis, 2000). Sempre que o TAM é validado para consistência interna, obtém uma pontuação muito alta contra qualquer medida (Davis, 1989; Szajna, 1994; van der Heijden, 2003). Como conseqüência, os resultados da aplicação do TAM são frequentemente aceitos como preditores de uso e adoção muito além do ponto de validação. No entanto, o uso real da tecnologia, em vez da intenção de uso, raramente é monitorado. Este estudo tem como objetivo investigar se o TAM é um preditor confiável de uso real, em vez da intenção de usar, com base no emprego de subjetivo (autorreferido) e objetivo (registros do computador) medidas.

*4.2 Objetivos*

Os objetivos da revisão são:

• identificar até que ponto o TAM e suas revisões são capazes de fornecer previsões confiáveis ​​do real uso de uma tecnologia em oposição à intenção de uso;

• determinar se o tipo de medida de uso (subjetivo ou objetivo) afeta a precisão das previsões da TAM;

• identificar se a versão do TAM ou a tecnologia sendo avaliado, afeta a precisão das previsões.

*4.3 Métodos*

A estratégia de pesquisa para a revisão é direcionada principalmente para encontrar artigos publicados (revistas de arquivo, conferências procedimentos ou relatórios técnicos) do conteúdo de cinco bancos de dados eletrônicos, embora cada um tenha identificado A fonte foi verificada quanto a outras referências relevantes. Um número de cadeias de pesquisa foi construído usando informações relevantes termos baseados nas questões de pesquisa e a busca foi restrito a artigos publicados entre 1989 e o presente dia. O ano de 1989 foi escolhido como linha de base, pois quando o primeiro artigo a descrever o TAM foi publicado (Davis, 1989).

A seleção das fontes primárias foi inicialmente baseada em revisão do título, das palavras-chave e dos resumos, embora isso tenha sido estendido para incluir a seção de conclusões nos casos onde o título, as palavras-chave e o resumo forneceram insuficiente em formação. Todos os estudos selecionados foram revisados contra um conjunto detalhado de critérios de inclusão criados para identificar se um estudo pode ou não ajudar a responder às especificações questões de pesquisa. Em particular, nosso objetivo foi identificar se um estudo descreve uma investigação empírica em qual o TAM foi usado, se a variável de uso real é medido, se o estudo inclui particular Variáveis ​​TAM e se ele relata o relacionamento com uso real. O processo de extração de dados está sendo conduzido usando um formulário padronizado de gravação eletrônica de dados.

*4.4 Resultados e conclusões*

Inicialmente, 208 artigos foram identificados como potencialmente relevantes às questões de pesquisa; no entanto, após a aplicação critérios de inclusão / exclusão, 59 artigos permaneceram no conjunto de papéis relevantes. O número de falsos positivos em o conjunto inicial (trabalhos que podem ter sido relevantes, mas em investigação detalhada acabou não sendo assim) foi decepcionante Alto.

**5. Revisão sistemática da literatura de diretrizes para a realização de revisões sistemáticas da literatura (R3)**

Este estudo é mais uma "meta-revisão", embora ainda usando os conceitos de revisão sistemática. Está no planejamento fase e estamos no processo de desenvolvimento do protocolo (estágio 2) que está sendo pilotado atualmente.

*5.1 Contexto*

As diretrizes usadas para conduzir literatura sistemática as revisões variam entre os diferentes domínios que comumente use essa abordagem. O objetivo deste estudo, que também faz parte do projeto EBSE, é revisar as diretrizes existentes para revisões sistemáticas da literatura em vários domínios, identificando as semelhanças e diferenças entre eles, na tentativa de propor diretrizes apropriadas para realizar revisões sistemáticas da literatura no domínio da engenharia de software.

5.2 Objetivos

Os objetivos da revisão são:

• pesquisar e explorar outros domínios semelhantes onde diretrizes podem ser encontradas e determinar se quaisquer diretrizes são mais adequadas para engenharia de software do que orientações médicas;

• encontrar critérios de qualidade adequados para avaliar os tipos de estudos realizados em engenharia de software.

*5.3 Métodos*

Para esta revisão, uma investigação preliminar está sendo realizada para identificar domínios de pesquisa semelhantes ao domínio de engenharia de software. Um domínio de pesquisa será considerado semelhante ao domínio da engenharia de software se formas semelhantes de estudos empíricos forem realizadas domínio. Esta investigação preliminar é baseada em reuniões /entrevistas com vários especialistas dentro de um intervalo disciplinas começando com as localizadas em Keele e Universidades de Durham e, em seguida, acompanhando os estes (Budgen et al., 2006).

De acordo com os resultados, publicações específicas para cada domínio considerado suficientemente "próximo" ser alvo. Esta pesquisa incluirá livros didáticos junto com bases de dados especializadas em publicações de pesquisa nos domínios escolhidos. A literatura reunida será então qualificada contra critérios de qualidade predefinidos. Esses critérios são projetado para avaliar as diretrizes em termos de quem desenvolveu eles e como eles foram validados, a fim de decidir se o as diretrizes propostas devem ser incluídas na revisão. UMA está sendo realizado um estudo piloto antes da revisão final conduzido.

*5.4 Resultados e conclusões*

Até o momento, os domínios da educação, criminologia e enfermagem e obstetrícia foram identificadas como tendo práticas empíricas mais alinhadas com o software Engenharia.

**6. Lições aprendidas**

Nesta seção, revisamos nossas experiências com o processo de conduzir revisões sistemáticas da literatura. Nós temos estruturamos nossa discussão em torno do modelo de 10 etapas que foi introduzido no §2. Embora dois de nossos estudos ainda tenham para ser concluído, realizamos um amplo teste de ambos os protocolos de pesquisa. Isso envolveu pilotar a maioria dos os elementos da revisão, exceto a síntese final de, e documentação dos resultados sistemáticos da revisão da literatura, e, portanto, nossas lições se baseiam nos três estudos. Para facilidade da leitura, relatamos nossos resultados contra os mais relevantes estágio, embora as lições aprendidas dos estudos R2 e O R3 é parcialmente baseado em nossas experiências de pilotar alguns dos estágios durante o desenvolvimento do protocolo.

*6.1 Etapa 1: especificar perguntas de pesquisa*

A especificação da (s) pergunta (s) de pesquisa é a mais crítico elemento de uma revisão sistemática. As questões de pesquisa são usados ​​para construir sequências de pesquisa para pesquisas automatizadas, eles determinam os dados que precisam ser extraídos cada estudo primário e restringir o processo de agregação. As questões de pesquisa fazem parte do protocolo que não deve ser alterado depois que o protocolo for aceito. Nas três análises, as perguntas foram refinadas como resultado de pilotar o protocolo. Isso foi particularmente perceptível para o R3, que começou com duas perguntas:

Pergunta 1: Quais são os elementos de uma sistemática revisão de literatura em uma amostra de domínios?

Pergunta 2: Quais são as diretrizes que podem ser adotadas em engenharia de software?

Após revisar o protocolo e investigar algumas das fontes de informação conhecidas, as expandimos para um conjunto de perguntas mais detalhadas:

Pergunta 1: Quais são os elementos essenciais da uma revisão sistemática da literatura nos domínios estudou?

Pergunta 2: Quais são os elementos essenciais da revisão sistemática da literatura em domínios que são semelhantes à engenharia de software?

Pergunta 3: Quais são os critérios para avaliar a qualidade de estudos primários?

Pergunta 4: como as pesquisas são organizadas? Essa questão leva a duas sub-perguntas separadas:

Q4.1. Quais estratégias são usadas para identificar as principais estudos relacionados ao tema da pesquisa?

Q4.2. Quais critérios são usados ​​para avaliar a integridade da estratégia de pesquisa?

Pergunta 5: Quais procedimentos são usados ​​para combinar evidências de diferentes estudos primários outros que a metanálise formal (por exemplo, classificação sistemas para tipos de estudo)?

Uma maneira de abordar a dificuldade do escopo da pesquisa perguntas é através da introdução de uma pré-visualização sistemática estudo de mapeamento do tópico de pesquisa. Esta abordagem é obtido no EPPI-Center (The Evidence for Policy and Informações Práticas e Centro de Coordenação3), que realiza e publica análises relacionadas à intervenção social. Um mapa sistemático é usado para descrever os tipos das atividades de pesquisa realizadas relacionadas a uma pergunta de pesquisa. O mapa descreve fatores como distribuição de estudos, faixa etária coberta por estudos e o número de estudos que avaliam especificidades políticas e práticas. O mapeamento é semelhante à extração de dados na medida em que é feito para cada estudo inserindo detalhes um formulário. A diferença é que o mapeamento é feito o mais rápido possível, para um grande número de estudos iniciais, e descreve os estudos, em vez de extrair detalhes específicos. É diferente síntese, pois não há interpretação das descrições. feito em um mapa. No entanto, fornece um contexto para a síntese posterior. O objetivo desse mapa é fornecer um contexto para a revisão, para ajudar na interpretação da síntese, restringir a questão da síntese e inclusão / exclusão critérios e reduzir o número de estudos em potencial.

*6.1.1 Lição aprendida*

L1: Espere revisar suas perguntas durante o protocolo desenvolvimento, como sua compreensão do problema aumenta.

L2: Um estudo de mapeamento pré-revisão pode ajudar no escopo questões de pesquisa.

6.2 Etapa 2: desenvolver protocolo de revisão

Conforme descrito anteriormente, o protocolo fornece detalhes do plano para a revisão, incluindo, por exemplo, a especificação do processo a serem seguidas e as condições a serem aplicadas ao selecionar estudos primários, as métricas de qualidade a serem aplicadas estudos primários e a alocação de revisores a determinados Atividades. No estudo R1, a natureza distribuída do equipe levou a alguns problemas durante o desenvolvimento de o protocolo de revisão. No final, o protocolo foi desenvolvido por apenas dois dos revisores, o que significa que esses revisores tornaram-se mais familiarizados do que os outros processo sistemático de revisão da literatura e, em particular, mais familiarizado com o papel do próprio protocolo. Este causou problemas mais tarde, quando um membro da equipe não seguiu o processo de pesquisa especificado no protocolo e outros membros não entenderam completamente os dados requisitos de extração.

No estudo R2, a pilotagem do protocolo se mostrou essencial na revelando problemas de extração e agregação de dados. No especial:

• Os membros da equipe que assumiram o papel de extratores de dados foram não está familiarizado com termos estatísticos como o membro da equipe quem definiu o formulário de extração de dados. Pilotando a extração processo revelou vários mal-entendidos sobre a natureza das constantes de correlação e regressão.

• O piloto revelou que estudos primários frequentemente relatavam vários testes para um único estudo e a extração de dados processo precisava ser refinado para garantir que os resultados de cada estudo não foram contadas duas vezes.

• Os estudos primários usaram diferentes meios de relatar seus resultados e a extração e agregação de dados necessário alterar o processo para atender a essas diferenças.

• Os estudos primários relataram subconjuntos das informações que significavam que precisávamos definir um processo para gerenciar valores ausentes.

No R3, o exercício piloto identificou a necessidade de informações a serem coletados no nível do domínio, bem como no primário nível de estudo. As informações no nível do domínio foram 3 necessário fornecer uma avaliação de quão semelhante domínio é a engenharia de software. A pilotagem também revelou que a revisão R3 é diferente de outra engenharia de software analisa onde os pesquisadores podem procurar estudos em bancos de dados eletrônicos, periódicos ou anais da conferência. Em vez disso, é baseado na literatura publicado em domínios amplamente fora dos autores área de especialização, o que significa que executar uma ampla busca de diretrizes tem sido problemática, devido à falta de conhecimento do domínio. Isso nos levou a revisar o escopo de nosso estudo e estender a metodologia do além de uma revisão sistemática da literatura para incluir entrevistas com especialistas em domínio.

*6.2.1 Lições aprendidas*

L3: Todos os membros da equipe de revisão sistemática precisam fazer uma parte ativa no desenvolvimento do protocolo de revisão.

L4: A pilotagem do protocolo de pesquisa é essencial. Vai encontrar erros nos procedimentos de coleta e agregação de dados. Também pode indicar que você precisa alterar a metodologia que você pretende usar para abordar o questões de pesquisa.

*6.3 Etapa 3: validar o protocolo de revisão*

O protocolo é um elemento crítico de uma revisão e os pesquisadores precisam especificar e executar procedimentos para sua validação. Para R1, validação do protocolo de revisão incluiu a execução de uma execução piloto da extração de dados processo (e formulários) em quatro artigos usando dois dos revisores. Como resultado disso, os formulários foram ligeiramente modificados e algumas notas para a extração de dados foram preparadas. UMA aspecto particularmente útil das notas foi a inclusão de definições para os métodos de pesquisa que provavelmente ser usado nos papéis selecionados, o que ajudou bastante na uniformidade de codificação. O protocolo também foi revisado informalmente Barbara Kitchenham, mas os resultados disso foram não adequadamente (ou formalmente) considerado pela equipe de revisão.

Para o R2, a validação foi realizada através de um procedimento formal. processo de revisão usando dois revisores externos que estavam solicitado a preencher um questionário abordando a integridade e qualidade dos itens de revisão. Suas observações levaram para algumas revisões do protocolo.

*6.3.1 Lições aprendidas*

L5: A extração de dados é auxiliada por ter definições de dados e diretrizes de extração de dados do protocolo gravado em um pequeno documento separado.

L6: É necessário haver um processo de validação acordado separadamente da atividade de pilotagem de protocolo. Idealmente, externo os revisores devem realizar esse processo de validação.

*6.4 Etapa 4: identificar pesquisas relevantes*

Uma vez finalizado o protocolo e a revisão sistemática entra na fase de execução, os pesquisadores devem executar o estratégia de busca definida no protocolo. Estudos R1 e R2 adotou estratégias de pesquisa muito diferentes.

Para R1, durante a fase de planejamento, ficou evidente que centenas de artigos foram publicados em conferências recentes e workshops sobre tópicos relacionados ao SBS. O opções de amostragem desses trabalhos ou de aplicar algumas critérios de qualidade foram debatidos, mas finalmente foi decidido extrair fontes primárias apenas de periódicos de arquivo, em grande parte na base pragmática que não fomos capazes de determinar critérios claros de qualidade para decidir quais conferências e oficinas para incluir. Isso significava que o estratégia de pesquisa foi restrita a especificar o arquivo diários a serem utilizados e os anos que seriam pesquisados. O ano inicial foi identificado a partir da primeira publicação que referenciou o termo software baseado em serviço. Em retrospecto, a relativa imaturidade da tecnologia significava que havia havia apenas alguns artigos de periódicos disponíveis e, claramente, identificando pesquisa relevante durante o período em que uma tecnologia O desenvolvimento rápido pode ser um pouco problemático. (Esse é provavelmente um dos aspectos da engenharia de software que não tem equivalente nos outros domínios que temos examinado.)

Por outro lado, o estudo R2 precisava obter um conjunto completo possível, porque procura os resultados de um processo de avaliação raramente usado (ou seja, validação do TAM contra uso real em vez de validação contra intenção de usar). Por esse motivo, uma pesquisa restrita foi não é possivel. A construção da estratégia de busca foi baseada sobre a abordagem sugerida pelas normas médicas:

• A questão da pesquisa foi decomposta em elementos relacionados à tecnologia (aceitação da tecnologia modelo), o tipo de estudo (avaliação) e os medida de resposta (correlação com o esforço real) para obtenha os principais termos de pesquisa.

• Palavras-chave obtidas de estudos primários conhecidos foram avaliado para outros termos principais.

• Sinônimos para os principais termos foram identificados.

• As strings de pesquisa foram construídas usando o valor booleano 'AND' para aderir aos termos principais e "OU" para incluir sinônimos.

Além disso, o protocolo de pesquisa R2 especificou o início apontam para buscas em 1989, quando o TAM foi introduzido pela primeira vez, identificaram uma revisão sistemática anterior da TAM avaliação e observou que a lista de referência de todos os estudos seriam procurados para procurar outros candidatos estudos.

As diretrizes médicas recomendam a busca em vários fontes. Identificamos a seguinte indexação Serviços:

• IEEExplore;

* Biblioteca digital ACM;

• Google scholar (<scholar.google.com>);

* Biblioteca Citeseer (<citeseer.ist.psu.edu>);

• biblioteca eletrônica da Keele University (<opac.keele. ac.uk>);

• Inspec (<www.iee.org/publish/inspec/>);

• ScienceDirect (<www.sciencedirect.com>);

• EI Compendex (<www.engineeringvillage2.org/controller/ servlet / athensservice>).

No entanto, a pilotagem das pesquisas encontrou problemas imediatamente. Os termos de pesquisa propostos no protocolo foram aplicado em dois principais bancos de dados eletrônicos: IEEExplore4 e o Portal ACM.5 Descobrimos que os mecanismos de pesquisa de os dois bancos de dados estão organizados em torno de completamente diferentes modelos. Portanto, é impossível fazer uso direto do mesmo conjunto de termos de pesquisa para os dois mecanismos. Para Por exemplo, o ACMPortal não suporta lógica complexa combinação, embora seja suportado pelo IEEExplore.

Além disso, descobrimos que, em alguns mecanismos de pesquisa, a avaliação de uma cadeia de pesquisa booleana depende da ordem dos termos, independente de colchetes. Por exemplo, os dois Expressões booleanas ilustradas nas Figs. 2a e 2b fornecem o mesmo número total de resultados no IEEExplore, mas a ordem em que os resultados aparecem é significativamente diferente.

Isso é particularmente importante, pois os resultados da O mecanismo de pesquisa IEEExplore é classificado em ordem de relevância. Outros mecanismos de pesquisa também exibem inconsistências semelhantes. Por exemplo, o Citeseer (http://citeseer.ist.psu.edu/) trata várias palavras como um termo booleano e procura instâncias de todas as palavras juntas. O uso da expressão booleana "E" neste mecanismo procura todas as palavras, mas não necessariamente juntos. Como exemplo, o valor booleano pesquise 'modelo de aceitação de tecnologia' (sem aspas) procura por instâncias do "modelo de aceitação tecnológica" (todas três palavras juntas) e não encontra documentos. O termo de pesquisa alterado usando booleano "e" (ou seja, "tecnologia e aceitação e modelo ") retorna 210 documentos. Contudo, como nenhum dos resultados retornados inclui todos os termos juntos, a relevância dos resultados é muito baixa. O mecanismo também oferece a capacidade de pesquisar no banco de dados do site usando o Google. Uma pesquisa no Google do Citeseer usando o termo de pesquisa "modelo de aceitação tecnológica" (com citações e, assim, procurando a frase exata) encontra quarenta documentos. Por esse motivo, decidimos usar diferentes conjuntos de termos de pesquisa para cada banco de dados, com cada um os termos de pesquisa derivados dos termos originalmente proposto no protocolo.

*6.4.1 Lições aprendidas*

L7: Existem estratégias de pesquisa alternativas que permitem que você para atingir diferentes tipos de critérios de conclusão da pesquisa. Você deve selecionar e justificar uma estratégia de pesquisa que seja apropriado para sua pergunta de pesquisa.

L8: Precisamos pesquisar muitos tipos diferentes de eletrônicos fontes; nenhuma fonte única encontra todos os principais estudos.

L9: Os atuais mecanismos de pesquisa de engenharia de software não são projetado para apoiar revisões sistemáticas da literatura. Ao contrário dos pesquisadores médicos, a engenharia de software os pesquisadores precisam executar tarefas dependentes de recursos pesquisas.

*6.5 Etapa 5: selecionar estudos primários*

A seleção de estudos primários é geralmente uma etapa processo:

• O título e resumo dos estudos identificados pela primeira as pesquisas são revisadas (de preferência por pelo menos duas pesquisadores) e artigos irrelevantes são rejeitados. Este revisão deve errar por precaução, se os pesquisadores Se não concordar, o artigo deve ser incluído.

• Cópias completas dos trabalhos não rejeitados anteriormente são obtido. Estes artigos são revisados ​​por dois ou mais pesquisadores contra os critérios de inclusão / exclusão definido no protocolo para obter uma lista final dos principais estudos. Os dois pesquisadores devem resolver qualquer desacordo (se necessário, com a ajuda de um árbitro).

Esse processo pode repetir se o protocolo de pesquisa especificar que a lista de referência dos estudos primários deve ser revisada procurar outros possíveis estudos primários candidatos.

Para R1, como a pesquisa foi realizada manualmente, e oconjunto de papéis era relativamente pequeno, essa etapa não apresentava dificuldades reais.

Para R2, houve um problema ao decidir se os papéis retornados pelo mecanismo de busca foram relevantes para o estudo. O protocolo especificou que a seleção inicial do primário fontes seriam puramente baseadas em uma revisão do título, palavras-chave e resumo. No entanto, na prática, foi muito difícil determinar se um trabalho era um candidato em potencial para o estudo TAM usando apenas títulos, resumos, e palavras-chave. Na medicina, ao contrário da engenharia de software ciência da computação, os resumos são estruturados e contêm informações que são, juntamente com títulos e palavras-chave, geralmente o suficiente para determinar o conteúdo de um artigo. No contraste, em engenharia de software e ciência da computação, as palavras-chave não são consistentes entre diferentes periódicos importantes e entre organizações como ACM e IEEE. Na maioria dos casos, também foi necessário ler o conclusão antes de tomar a decisão de incluir ou excluir um estudo primário específico. De fato, quando as decisões foram tomadas feitos exclusivamente no resumo dos trabalhos, vários estudos foram excluídos que eram conhecidos por serem importantes antes da o piloto. O resultado foi que o protocolo foi alterado afirmar que as conclusões serão incluídas no processo inicial de identificar estudos primários.

*6.5.1 Lição aprendida*

L10: O padrão de resumos de engenharia de TI e software é muito ruim para se confiar ao selecionar estudos primários. Você também deve revisar as conclusões.

*6.6 Etapa 6: avaliar a qualidade do estudo*

É importante avaliar a qualidade dos estudos primários em para apoiar o processo de inclusão / exclusão e as alocação da ponderação para estudos específicos durante os dados estágio de síntese. Não há definições universais de estudo qualidade, mas foi sugerido que a qualidade está relacionada à até que ponto o viés é minimizado e externo e interno a validação é maximizada (Khan et al., 2001).

No R1, a equipe de pesquisa não avaliou a qualidade dos estudos individuais baseados na lógica de que, restringindo sua revisão em artigos publicados em revistas de arquivo; eles já havia garantido que os papéis fossem de aceitável qualidade.

No R2, a equipe de revisão estabeleceu critérios de qualidade baseados sobre a integridade dos dados, com documentos de "problemas" sendo classificado como questionável ou incompleto. Questionável estudos são estudos que fazem alegações sobre o relacionamento eles encontraram entre o TAM e o uso real, mas não forneça dados ou testes estatísticos para apoiar as reivindicações. Estudos incompletos relatam alguns, mas não todos, dos requisitos em formação. Por exemplo, um artigo que anota um relacionamento entre os principais elementos do TAM (ou seja, utilidade percebida e facilidade de uso percebida) e uso real, mas não diga se o relacionamento vale para cada elemento separadamente seriam classificados como incompletos. Se a informação ambíguo ou ausente, é importante entrar em contato com o autores e peça informações adicionais. E se se não houver informações adicionais, a qualidade a classificação será usada na análise de sensibilidade, p. avaliando se os resultados dos estudos incompletos ou ambíguos afetar os resultados gerais.

Para R3, a equipe de revisão identificou critérios de qualidade para avaliar a adequação e autenticidade de quaisquer diretrizes identificado. Definiu uma série de perguntas para quantificar a fonte. Exemplos dessas perguntas são: ‘‘ um único indivíduo ou um grupo desenvolve as diretrizes? '' e "O grupo incluiu indivíduos de diferentes países?" A classificação de qualidade desenvolvida para as diretrizes cobertas um resultado varia de 0, quando as diretrizes foram preparado por uma única pessoa sem validação formal, para 5, para diretrizes produzidas por um grupo multinacional com uma validação formal por revisores independentes. Esta classificação deve ajudar a determinar a confiabilidade da fonte de As diretrizes.

*6.6.1 Lições aprendidas*

L11: Todas as normas médicas enfatizam que é necessário avaliar a qualidade dos estudos primários. Contudo, depende do tipo de literatura sistemática revise você está realizando.

L12: É importante ter certeza de como a avaliação da qualidade será usado na agregação de dados subsequente e análise.

6.7 Etapa 7: extrair os dados necessários

O objetivo desta etapa é usar formulários de extração de dados registrar com precisão as informações obtidas pelos pesquisadores dos estudos primários. Para reduzir a oportunidade de viés, os formulários de extração de dados devem ser definidos e pilotado quando o protocolo do estudo é definido.

Para R1, a equipe de pesquisa descobriu que indivíduos que também realizou a revisão piloto produzida na maior parte completa, dados de boa qualidade. Outros às vezes não tinham certeza sobre o que fazer e às vezes não seguiu o protocolo, embora não tenham solicitado esclarecimentos a qualquer momento. Possível razões para isso são:

• as instruções podem não ter sido claras;

• algumas semanas se passaram entre as instruções enviado aos revisores e as revisões sendo realizadas;

• a justificativa para ser sistemático e o processo geral foram documentados na forma de um rascunho, em vez de do que como um "protocolo" formal.

Para o R2, quando o processo de extração de dados foi realizado, os dois revisores usaram duas abordagens diferentes. O primeiro Essa abordagem exigia que cada revisor extraísse dados de um papel individualmente e, em seguida, compare as duas extrações de dados formulários e discutir quaisquer divergências. Isto é o método recomendado pelas orientações médicas. O segundo A abordagem era ter um revisor extraindo os dados e o segundo revisor atuando como verificador. Nós achamos isso não houve grande diferença entre as duas abordagens em termos de esforço e tempo. No entanto, ter um revisor agindo como um verificador foi um pouco mais rápido e pode ser vale a pena considerar se há um grande número de papéis rever.

*6.7.1 Lições aprendidas*

L13: Ter um leitor atuando como extrator de dados e um ato como verificador de dados pode ser útil quando houver uma grande número de artigos a serem revisados.

L14: Os membros da equipe de revisão devem garantir que eles entendam o protocolo e o processo de extração de dados.

*6.8 Etapa 8: sintetizar dados*

Uma vez extraídos, os dados devem ser sintetizados de maneira adequada para responder às perguntas da revisão. No medicina, a maioria das revisões sistemáticas da literatura visa na meta-análise formal de dados quantitativos. Dois dos revisões, R1 e R3, baseiam-se na coleta de informações qualitativas dados. A agregação de dados qualitativos é baseada em simples formatos tabulares.

Portanto, apenas uma das três revisões sistemáticas discutidas neste trabalho (R2) considerou o uso de meta-análise. No entanto, tentar isso para o R2 revelou imediatamente o problema de que os coeficientes de regressão não podem ser agregados da mesma maneira que os coeficientes de correlação podem (Lipsey e Wilson, 2001). O exercício piloto revelou que alguns pesquisadores relataram correlações, enquanto outros relataram o coeficiente de correlação múltipla ou a regressão coeficientes. A equipe de revisão decidiu que não poderia basear agregação apenas na meta-análise, mas também precisaria resumir dados em forma de tabela simples.

Quando os dados são tabulados (em oposição a sujeitos a meta-análise), pode não estar claro se o as perguntas de pesquisa da revisão foram respondidas. O os pesquisadores podem precisar explicar como os dados resumidos aborda as questões de pesquisa.

6.8.1 Lições aprendidas

L15: É provável que as revisões sistemáticas de engenharia de software ser de natureza qualitativa.

L16: Mesmo ao coletar informações quantitativas, pode não é possível realizar meta-análise de software estudos de engenharia porque os protocolos de relatório variam muito de estudo para estudo.

L17: Tabular os dados é um meio útil de agregação mas é necessário explicar como os dados agregados realmente responde às perguntas da pesquisa.

*6.9 Etapa 9: Escrever relatório de revisão*

Uma vez concluída a revisão sistemática e as perguntas respondida, a revisão sistemática deve ser documentada. As orientações médicas sugerem que o protocolo pode ser usado como base para o relatório final.

Somente R1 atingiu a fase 3 e a documentação de revisão está assumindo a forma de um artigo para envio a um revista acadêmica de engenharia de software. A equipe enfrentou dois desafios nesta fase. Um relacionado ao limitado quantidade e qualidade da manutenção de registros. As decisões foram distribuídas através de muitas trocas de e-mail e as anotações que foram tomadas em uma reunião de quatro dos cinco revisores. Embora tenhamos conseguido construir o processo necessário informação teria sido melhor se tivéssemos mantido um log de projeto mais detalhado e formal. O outro desafio que surgiu foi a restrição de comprimento imposta (ou pelo menos recomendado) por alguns periódicos de engenharia de software (e ainda mais restrito por muitas conferências de engenharia de software). Documentar completamente uma revisão dentro do máximo contagem de palavras de muitos periódicos é um desafio e seria talvez seja aconselhável fazer alguns detalhes de um revisão disponível em um site confiável. Esta abordagem foi tomada pelo EPPI-Center, que publica análises em três níveis de detalhe: breves introduções, resumos e revisões completas.

*6.9.1 Lições aprendidas*

L18: As equipes de revisão precisam manter um registro detalhado das decisões feitas durante todo o processo de revisão.

L19: A comunidade de engenharia de software precisa estabelecer mecanismos para publicação de literatura sistemática revisões que podem resultar em documentos mais longos do que aqueles tradicionalmente aceitos por muitos softwares lojas de engenharia ou que tenham apêndices armazenados em repositórios eletrônicos.

*6.10 Etapa 10: validar relatório*

Uma vez documentada a revisão sistemática, o médico diretrizes sugerem que o documento deve ser independentemente revisado. Publicamos um relatório técnico interno relatório de R1 (Brereton et al., 2005) e estão estendendo esse para validação externa (por meio de envio a uma avaliação Diário).

**7. Discussão**

Nesta seção, analisamos os objetivos do estudo e discutir como as lições que aprendemos ajudam a abordar eles. Os objetivos são:

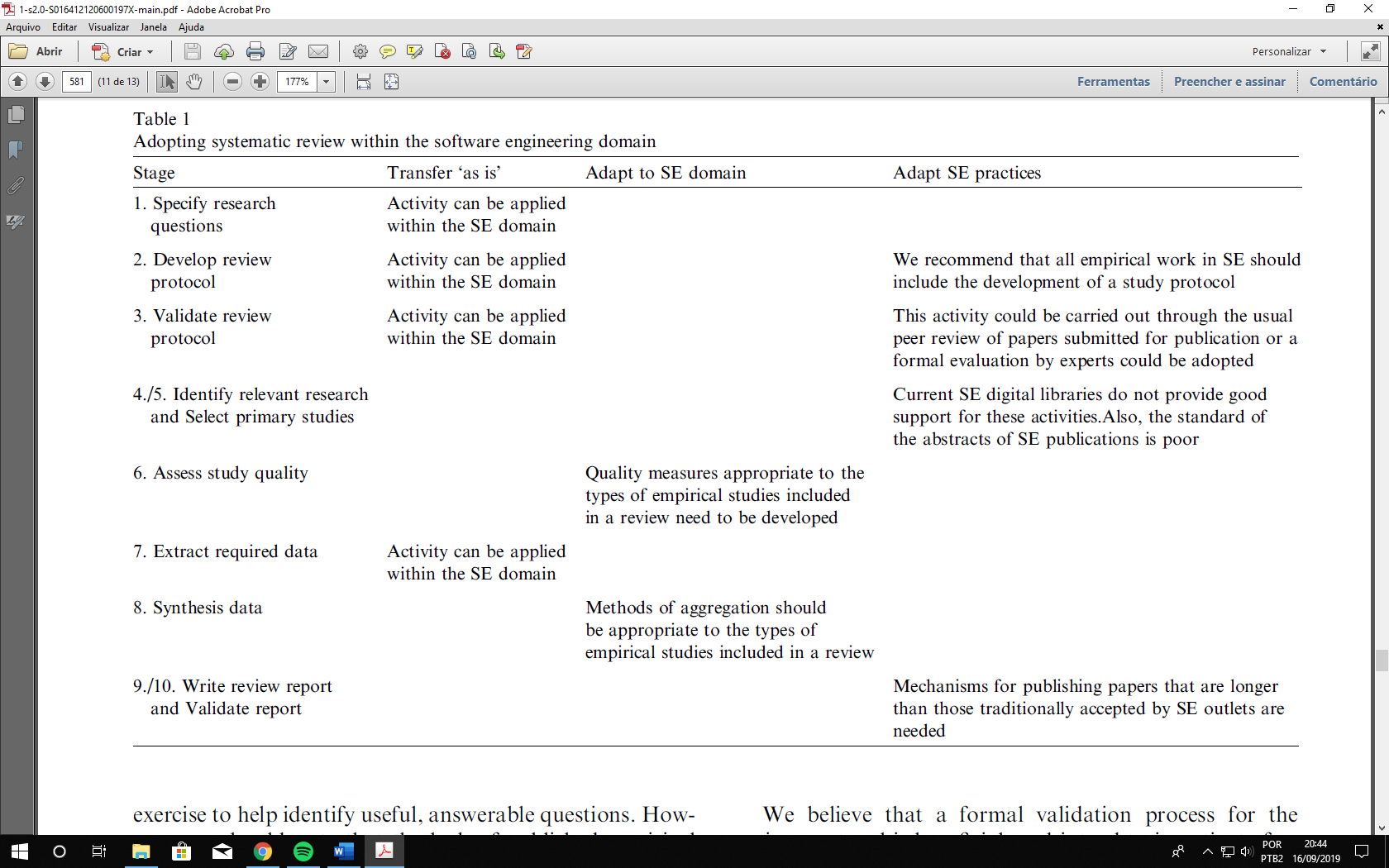
• identificar aspectos do processo que são transferidos "como estão" para Engenharia de software;

• identificar aspectos que precisam ser adaptados às particularidades características do domínio de engenharia de software;

• identificar áreas nas quais melhorias no software atual infraestrutura e práticas de engenharia são necessárias para obter a maior vantagem do processo.

A discussão está organizada em torno das etapas do processo de revisão e está resumido na Tabela 1.

Etapa 1 (especificar perguntas de pesquisa) do processo essencialmente transporta "como está". As orientações médicas não impedem a revisão das questões de pesquisa durante o protocolo desenvolvimento ou introdução de um escopo de pré-revisãoexercício para ajudar a identificar perguntas úteis e respondíveis. Contudo, devemos notar que a falta de publicações empíricas publicadas estudos em engenharia de software, comparados à medicina clínica, e a natureza desses estudos, que utilizam uma ampla procedimentos empíricos (comparados aos extensos ECR em estudos clínicos) provavelmente contribuirão para maior grau de dificuldade em estabelecer questões de pesquisa.

Todos os estudos se beneficiaram da produção de um protocolo. Os estudos R2 e R3 pilotaram o protocolo extensivamente e identificou uma série de problemas práticos como resultado. Encontramos a criação e pilotagem de um protocolo de estudo grande benefício e agora recomendaria que todos trabalho de pesquisa empírica deve começar com o desenvolvimento de um protocolo de estudo. No entanto, é fundamental que todos os membros de uma equipe de pesquisa participe ativamente da construção e a pilotagem do protocolo ou alguns dos benefícios podem estar perdido. Staples e Niazi também destacam o valor da pilotagem protocolo, mas suscitam uma preocupação adicional sobre saber quando terminar o processo de pilotagem e mover para a fase de execução (Staples e Niazi, 2006). Reeves et al. salientar a importância de tarefas iniciais, como esclarecimentos das perguntas que a revisão está tentando responder, definição termos e critérios de inclusão / exclusão (dica 2 do (Reeves et al., 2002)) e adicione uma palavra de cautela para que pode ser difícil convidar novos membros para sua revisão grupo após esta fase, pois é provável que novos membros encontrem difícil 'acompanhar' e entender as normas e significados realizada em conjunto pelo grupo ''. Este último ponto apoia nossa ver que todos os membros da equipe de revisão sistemática precisam uma parte ativa no desenvolvimento do protocolo de revisão (L3).

Acreditamos que um processo formal de validação para o protocolo de revisão é benéfico e pretende investigar mais o valor da avaliação independente de ambas as revisões protocolos e relatórios finais de pesquisa nos estudos R2 e R3.

Nossas tentativas de identificar pesquisas relevantes e selecionar os principais estudos destacaram uma série de domínios específicos problemas Em particular, os principais bancos de dados de indexação on-line não fornecem apoio adequado para procurando porque eles usam diferentes modelos subjacentes e nem todos suportam pesquisas booleanas complexas. Este precisa criar diferentes cadeias de pesquisa para diferentes bancos de dados foi criado por outras pessoas (veja, por exemplo, a Dica 6 in (Reeves et al., 2002) e também (Kitchenham et al., 2006) e (Staples e Niazi, 2006)) e é claramente um problema a ser resolvido. Além disso, as palavras-chave são padronizados e os resumos são de baixa qualidade para pesquisadores para determinar se os trabalhos são relevantes para questões de pesquisa específicas. Portanto, precisamos melhorar a qualidade dos papéis que nós mesmos escrevemos. Resumos pretendem ser um resumo independente completo de um papel e pode ser que tenhamos de considerar advogar o uso de resumos estruturados, conforme usado em alguns revistas.

Identificamos e aplicamos vários critérios de qualidade estudos primários selecionados para R2 e R3, no entanto há necessidade de mais trabalho nesta área, em particular para estabelecer quais critérios são adequados para tipos específicos de revisão, tipos específicos de estudo empírico e tipos específicos de análise de dados.

A extração de dados necessários usando dados padronizados os formulários de gravação parecem funcionar bem quando os revisores adquiriu alguma experiência em fazer isso e apreciamos que é uma atividade qualificada e demorada.

Conforme discutido anteriormente, abordagens empíricas usadas em software engenharia incluem uma prevalência de métodos qualitativos, ea variação nos protocolos de relatórios, mesmo quando dados quantitativos disponíveis sugerem que a metaanálise formal pode não ser possível. Os comentários também destacam a necessidade de explicar como os dados resumidos tratam questões de pesquisa.

Para as 2 etapas finais (escreva o relatório de revisão e valide revisão), há necessidade de engenharia de software jornais e conferências para acomodar artigos mais longos ou permitir o acesso a informações adicionais através de um site confiável. Alguns periódicos aceitaram isso. Por exemplo, o Journal of Information and Technology6 fornece informações em seu site sobre procedimentos para realizando revisões e ‘‘ incentivar [s] autores em potencial identificar áreas para revisões sistemáticas, redigir trabalhos e enviá-los '' (Dyer et al., 2005). Os problemas associados com a publicação de estudos empíricos, especialmente em prestigiadas conferências como a Conferência Internacional de Software Engenharia (ICSE) também foram destacados por Basili e Elbaum em uma palestra convidada no ICSE 2006 (Basili e Elbaum, 2006).

**8. Conclusões**

Nossas experiências de tentativa de literatura sistemática análises confirmaram que as etapas básicas da sistemática processo de revisão parece relevante para a engenharia de software como eles fazem para a medicina. No entanto, algumas modificações nossas práticas normais poderiam melhorar significativamente sua valor como ferramenta de pesquisa e como fonte de evidência para praticantes. Observamos particularmente que o relato de estudos empíricos em engenharia de software têm vários deficiências e falta de conformidade, especialmente em termos de facilidades de busca, entre os aplicativos digitais bibliotecas também é um obstáculo para revisores sistemáticos de literatura na disciplina.

*Reconhecimentos*

Gostaríamos de agradecer o apoio do EPSRC por financiar o projeto EBSE (EP / C51839X / 1) e agradecer Stuart Charters e Stephen Linkman por suas contribuições para discussões gerais. Agradecemos também ao Servicebased Co-revisores de sistemas: Nicolas Gold, King's College Londres; Keith Bennett, Universidade de Durham e Nikolay Mehandjiev, Universidade de Manchester. Michael Rusca realizou uma revisão inicial da literatura da TAM como um projeto de graduação da Keele University e usamos seu trabalho como ponto de partida para o estudo deste tópico. Tore Dyba˚ e Tracy Hall atuaram como revisores do protocolo utilizado em R2 (estudo da TAM).