

# Qualidade de Software do Maior Sistema de Teletoxicologia Utilizado no Brasil

João M. Alves<sup>1,2</sup>, Christiane G. von Wangenheim<sup>1</sup>, Alexandre Savaris<sup>2</sup>,  
Aldo von Wangenheim<sup>1,2</sup>

<sup>1</sup> Programa de Pós-Graduação em Ciência da Computação  
Universidade Federal de Santa Catarina (UFSC)  
Florianópolis, Santa Catarina, Brasil

<sup>2</sup>Laboratório de Telemedicina  
Instituto Nacional para Convergência Digital (INCoD)  
Florianópolis, Santa Catarina, Brasil

{joao,savaris}@telemedicina.ufsc.br, {gresse,awangenh}@inf.ufsc.br

**Abstract.** *Teletoxicology Systems are information systems that help physicians and toxicologists to attend cases of patient exposure to poisonous agents. They provide information about poisoning and poisoning case management and support toxicovigilance activities. There is still no consensus on how to assess the quality of such systems. This work performs a case study applying a customized software quality evaluation model – the AdEQUATE model – in order to assess the software quality degree of a teletoxicology system widely used in Brazil, the DATATOX information system. Analysing the specificities of DATATOX we customized AdEQUATE, a model for evaluating the quality of telemedicine systems based on the ISO/IEC 25010 standard through the use of a questionnaire. The acquired data were analysed employing descriptive statistics. Overall, the DATATOX information system has been perceived as having a good quality, specifically for effectiveness, efficiency, and satisfaction. Nevertheless, few characteristics have been considered of low quality, including reliability and compatibility. This work also attests the feasibility on using AdEQUATE as a tool for software quality evaluation.*

**Resumo.** *Sistemas de teletoxicologia são sistemas de informação que auxiliam médicos e toxicologistas a registrar casos de exposição de pacientes a agentes tóxicos. Tais sistemas fornecem informações sobre intoxicações, gerenciam casos de intoxicações e providenciam atividades de vigilância toxicológica. Contudo, ainda não há um consenso em como avaliar a qualidade de tais sistemas. Este trabalho realiza um estudo de caso aplicando um modelo de avaliação de qualidade customizado – o AdEQUATE – com o objetivo de determinar o grau de qualidade de software de um sistema de teletoxicologia largamente usado no Brasil, o sistema DATATOX. Analisando as especificidades do DATATOX, customizou-se o AdEQUATE, um modelo para avaliação de qualidade de sistemas de telemedicina baseado na norma ISO/IEC 25010 através do uso de um questionário. Os dados coletados foram analisados usando estatística descritiva. No geral, o sistema DATATOX foi considerado como tendo uma boa qualidade, especificamente em termos de eficácia, eficiência e satisfação. Contudo, algumas características foram consideradas como de baixa qualidade,*

*incluindo confiança e compatibilidade. Este trabalho também comprova a viabilidade de usar o AdEQUATE como ferramenta para avaliação de qualidade de software.*

## 1. Introdução

O número de intoxicações reportadas aos Centros de Informação Toxicológica (CITs) do Rio Grande do Sul e de Santa Catarina aumentou em 46% entre 2005 e 2013, sendo que mais de 147.000 casos foram causados por drogas de abuso ou animais venenosos [CIT/RS 2015, CIT/SC 2015]. No Reino Unido, anualmente, aproximadamente 170.000 das entradas no *National Health Service* são relacionadas a intoxicações. Esse volume é bastante expressivo e sobrecarrega o pessoal de hospitais e de centros de emergência [Perry et al. 2014].

Para auxiliar pacientes e centros hospitalares, os CITs têm a responsabilidade de fornecer informações e aconselhamento, gerenciamento de casos de intoxicação ou envenenamento, serviços de análise clínica, vigilância, pesquisa, educação e treinamento na prevenção de intoxicações [World Health Organization et al. 1997]. Muitos países além do Brasil mantêm CITs que funcionam em regime 24/7, como os Estados Unidos [Mowry et al. 2015] e as Ilhas Britânicas [Tormey and Moore 2013]. Fornecendo serviços para uma vasta gama de *stakeholders* (de um médico em uma unidade de tratamento intensivo a um fazendeiro que acabou de ser picado por uma cobra em seu sítio), a maior parte das interações ocorre por telefone; contudo, alguns CITs providenciam interação *on-line*, como o *Web Poison Control* [Poison Control Centers 2015] que fornece um auxílio limitado para vítimas de intoxicações via Internet.

Na ausência de um auxílio toxicológico imediato, teletoxicologia – uma especialização da telemedicina que trata do fornecimento de serviços toxicológicos à distância – é uma opção válida [Skolnik 2013]. Contudo, não existe um consenso sobre o conjunto de funcionalidades que constitui um sistema de teletoxicologia. Alguns sistemas como o TOXBASE [Bateman et al. 2002] fornecem ontologias para a classificação de agentes tóxicos; outros, como o *American National Poison Data System* [Seifert et al. 2013], registram casos de exposição de pacientes a agentes tóxicos.

Nesse contexto, avaliar a qualidade de sistemas de teletoxicologia é de extrema importância. Contudo, uma avaliação usando modelos tradicionais que consideram qualidade como um todo (como a norma ISO/IEC 25010) é complexa; a mesma é genérica e pode não considerar necessidades específicas de sistemas de toxicologia. Ademais, métricas de avaliação de qualidade como as da norma supracitada demandam muito tempo e intervenção direta no sistema. Por exemplo, a coleta via *logs* para contabilizar quantos erros por tarefa um usuário comete geraria um *overhead* para a equipe de desenvolvimento de um *software* que pode já estar em uso.

De outro modo, o modelo AdEQUATE (*questionnaire for Evaluation of Quality in Telemedicine systems*) é uma abordagem customizada da norma ISO/IEC 25010 [Alves et al. 2015] que foi criado considerando especificidades de sistemas de telemedicina, decompondo o objetivo de avaliação em métricas e sendo operacionalizado por um questionário para permitir a análise da qualidade segundo a percepção dos usuários. Por usar um questionário, esse modelo de avaliação de qualidade de *software* é mais rápido e facilmente aplicável que modelos tradicionais, e como customização de uma norma

reconhecida mantém uma abrangente gama de características de qualidade.

Este trabalho tem como objetivo avaliar a qualidade de *software* do sistema DATATOX – um sistema de teletoxicologia desenvolvido sob a observação da Associação Brasileira de Centros de Informação e Atenção Toxicológica (ABRACIT), utilizado por diversos CITs em nove estados do Brasil – usando o modelo AdEQUATE em 11 dimensões. Além disso, pretende-se mostrar a viabilidade de se adotar modelos de avaliação de qualidade sistematicamente criados a partir de normas internacionais de qualidade aplicados a sistemas de teletoxicologia, focando em suas especificidades enquanto mantendo a generalidade dos modelos internacionais usados como base.

A seção seguinte descreve sistemas de informação toxicológica, o processo de customização do modelo AdEQUATE e suas aplicações via questionário. Os dados coletados são analisados e discutidos nas sessões 3 e 4, respectivamente. A última seção apresenta a conclusão e a validade do estudo.

## 2. Materiais e Métodos

As subseções seguintes descrevem o objeto de estudo – o sistema DATATOX – e o modelo de avaliação proposto – o AdEQUATE.

### 2.1. O sistema DATATOX

O sistema DATATOX é um exemplo de sistema de informações toxicológicas implantado em 2013, baseado nas necessidades da ABRACIT de um sistema interativo, *on-line* e robusto. Usado por múltiplos CITs em nove estados do Brasil em um regime 24/7, o sistema já gerenciou mais de 81.000 ocorrências de intoxicações por drogas de abuso, pesticidas (tanto domésticos quanto industriais), venenos de rato, produtos de limpeza, produtos cosméticos e de higiene, produtos industriais, metais, plantas e fungos e animais peçonhentos. Médicos são os requisitantes de informação mais comuns, seguidos pelos parentes dos intoxicados e enfermeiras; residentes de medicina e toxicologistas de plantão coordenados por um médico são os atendentes das ocorrências.

A Figura 1 ilustra uma interação usando o sistema DATATOX. Um requisitante em uma localidade A qualquer inicia o contato com uma ligação telefônica, recebida pelo atendente (1) em um CIT de uma localidade B. Durante essa etapa o atendente colhe informações sobre o paciente e o agente de intoxicação, e paralelamente as registra no sistema (2). Dependendo do caso o atendente pode recorrer a outras fontes de conhecimento para aconselhar o requisitante, seja utilizando bibliografia local ou bases de dados externas (3.a); outra opção é consultar o médico supervisor (3.b), que também pode consultar material externo (4), dando *feedback* ao atendente (5). O final da interação se dá com a instrução do atendente ao requisitante (6), podendo incluir requisições de exames ou mesmo o encaminhamento para tratamentos específicos.

### 2.2. O modelo AdEQUATE

AdEQUATE é um modelo de qualidade desenvolvido com base na norma ISO/IEC 25010, cujo objetivo é avaliar a qualidade de *software* de sistemas de telemedicina e telessaúde sob a perspectiva dos seus usuários finais [Alves et al. 2015]. Como [Behkamal et al. 2009], o modelo sistematicamente decompõe as características de qualidade de uma norma internacional e a estende criando medidas, e pré-define um questionário para operacionalizar a coleta de dados. O questionário<sup>1</sup> é composto de 70

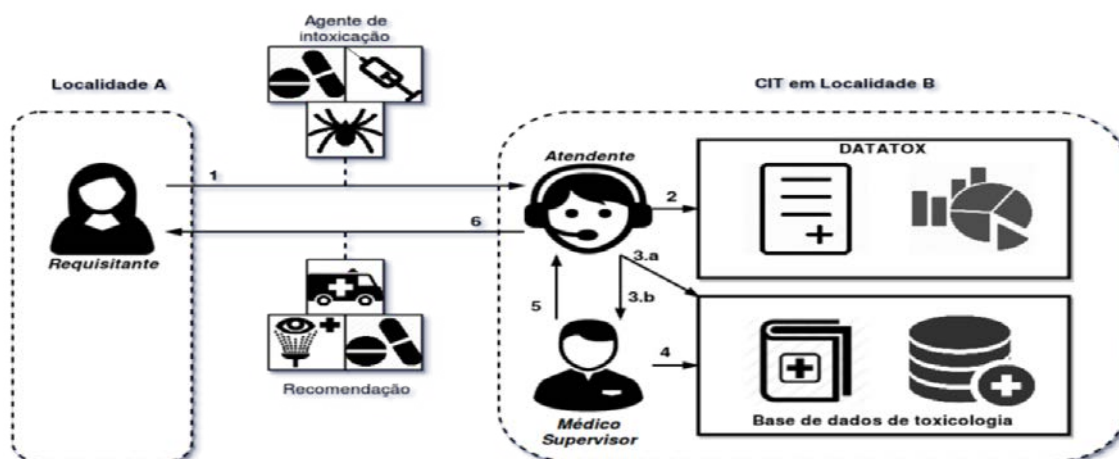


Figura 1. *Workflow* de um atendimento no DATATOX.

itens: dois itens demográficos sobre o papel dos usuários e a experiência dos mesmos, e 68 itens sistematicamente derivados das características e subcaracterísticas da referida norma. Todos esses 68 itens são estruturados (não permitindo respostas em texto livre), e são formulados de maneira positiva para evitar erros de entendimento e/ou interpretação [Sauro and Lewis 2011]. As opções de resposta seguem uma escala Likert com quatro pontos, mais três opções adicionais: "Não se aplica", "Não sei" e "Não entendi o item". Todos os itens são acompanhados por contra-exemplos.

Contudo, embora não seja convencional avaliar qualidade de software por meio de questionários – dados possíveis vieses estatísticos – o anexo C da norma ISO/IEC 25010 define que todas as subcaracterísticas e características de qualidade podem ser avaliadas através do uso de medidas externas. Ademais, a norma ISO/IEC 25022 propõe o uso de questionários como método de verificar qualidade [International Standard Organization 2003].

Para avaliar a qualidade de *software* do sistema DATATOX no contexto dos CITs e da ABRACIT, este trabalho apresenta um estudo de caso; o mesmo é caracterizado na seção 2.3, com os resultados obtidos sendo analisados na seção 3.

### 2.3. Definição do Estudo de Caso

Baseado no estado da arte referente a sistemas de teletoxicologia e na experiência de mais de 10 anos do Laboratório de Telemedicina no desenvolvimento de sistemas de telemedicina, foram consideradas relevantes todas as características e subcaracterísticas de qualidade propostas pelo modelo AdeQUATE<sup>1</sup>. Para coletar os dados com maior precisão, todos os contra-exemplos foram revisados e adaptados para o contexto da rotina de sistemas de toxicologia.

### 2.4. Execução do Estudo de Caso

De novembro a dezembro de 2015 foram coletados dados usando o questionário AdeQUATE disponível *on-line* pelo sistema *web* LimeSurvey<sup>2</sup>. Durante esse período

<sup>1</sup>O conjunto completo das características, subcaracterísticas e itens do questionário pode ser acessado em <https://arquivos.ufsc.br/f/61d4095f6a/>

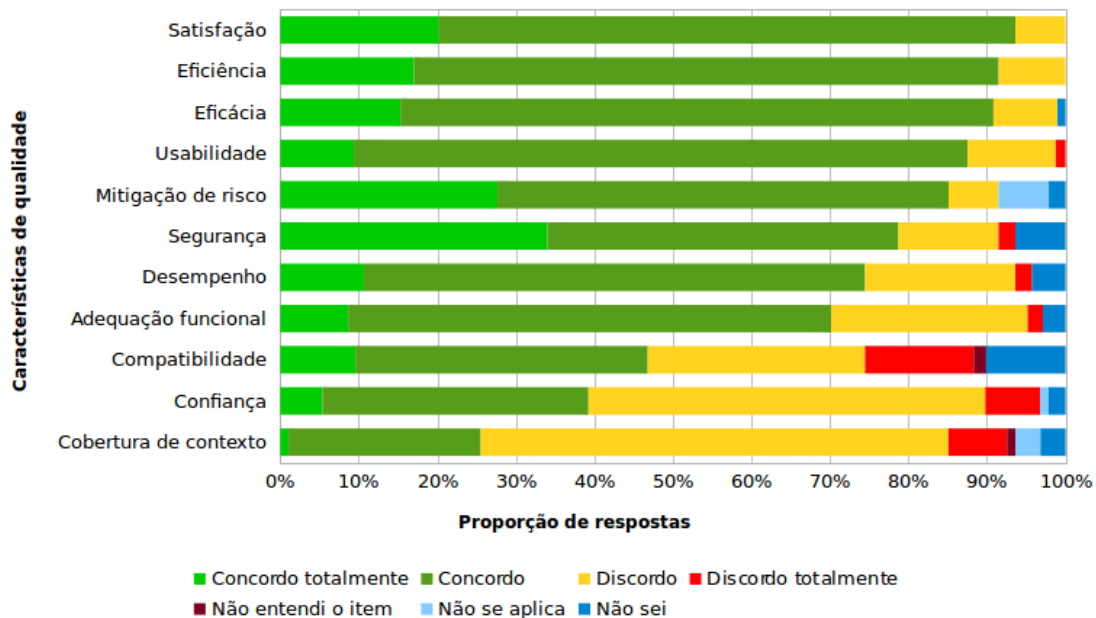
usuários do sistema DATATOX de diversos turnos de trabalho de várias localidades responderam o questionário; no total, foram obtidas 49 respostas, sendo 65% de atendentes e 35% de médicos. Desse número, 55% atestam terem usado o sistema por mais de um ano, 31% por mais de seis meses e apenas 14% por menos de 6 meses. A seção a seguir resume os dados coletados e os apresenta utilizando-se de estatística descritiva.

### 3. Resultados e Análise de Dados

Os dados coletados durante a execução do estudo de caso foram analisados usando duas métricas para cada característica de qualidade do modelo AdeQUATE, e uma terceira métrica para cada resposta dos participantes. As três métricas são definidas e analisadas na sequência.

**Métrica #1** Distribuição da soma das opções de resposta de cada item do questionário relacionado à qualidade de software.

O gráfico da Figura 2 mostra a distribuição de cada opção de resposta agrupada pelas características de qualidade do modelo AdeQUATE. Pôde-se observar que na maioria das características, as opções de resposta “Concordo totalmente” ou “Concordo” (barras verde e verde claro) correspondem a mais de 50% das respostas, com algumas exceções: “Compatibilidade”, “Confiança” e “Cobertura de contexto”. Como a opção “Não entendi o item” representa no máximo 2% das respostas, considera-se que todas as características foram bem compreendidas pelos participantes do estudo, atestando a validade do mesmo.



**Figura 2. Distribuição das opções de resposta no questionário entre as características.**

Para analisar mais profundamente o grau de qualidade de cada característica percebido pelos usuários participantes, foi calculada uma pontuação para cada característica

<sup>2</sup><https://www.limesurvey.org/>

usando (1). Na equação,  $\tilde{R}_i$  corresponde à mediana das respostas do questionário para cada característica  $n$ , por participante  $i$ , enquanto  $W(\tilde{R}_i)$  corresponde a uma função de atribuição de peso cujo domínio é o conjunto de opções de resposta e o contradomínio um peso associado: 3 para “Concordo totalmente”, 2 para “Concordo”, 1 para “Discordo” e 0 para “Discordo totalmente” (as demais opções não são avaliadas). O último componente da equação garante que cada  $S_n$  varie de 0 (grau mais pobre em qualidade – isto é, se todos os participantes respondessem “Discordo totalmente”) a 100 (grau de qualidade excepcional – isto é, se todos os participantes respondessem “Concordo totalmente”).

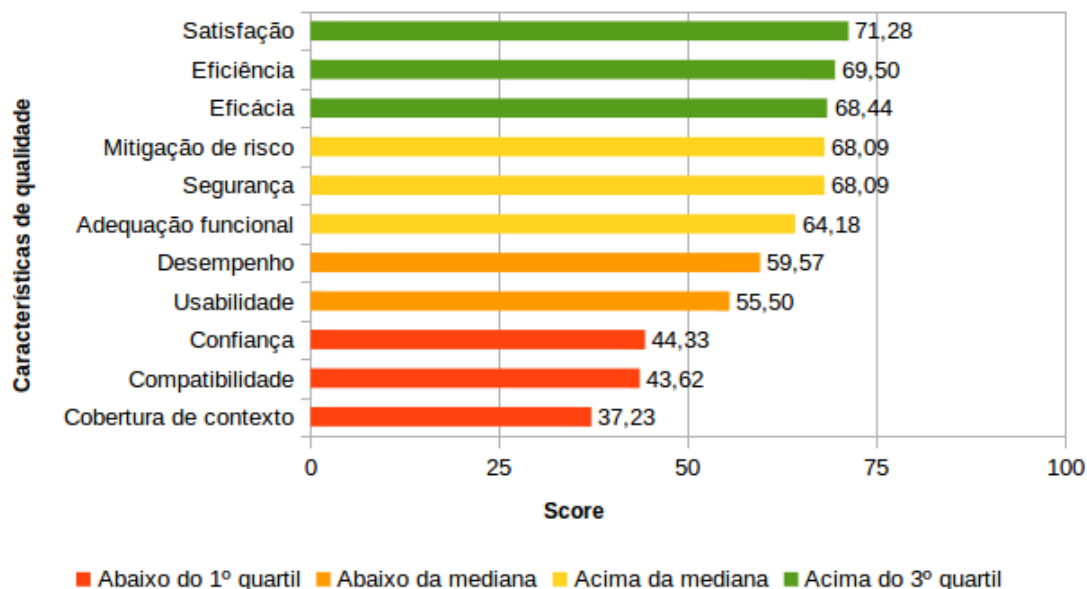
Esse sistema de pontuação é adaptado da pontuação definida pelo *System Usability Scale* (SUS) [Brooke 1996], por estar relacionado à qualidade de software e por ser uma ferramenta muito bem conceituada na indústria [Brooke 2013].

$$S_n = \left( \sum_{i=1}^{49} W(\tilde{R}_i) \right) * \frac{100}{147} \quad (1)$$

Usando (1), a segunda métrica é definida como segue:

**Métrica #2** A mediana das pontuações para cada item do questionário em sua característica.

Os dados obtidos que ilustram a análise da métrica #2 são exibidos no gráfico da Figura 3.



**Figura 3. Pontuação calculada para cada característica.**

Diferenças entre o gráfico da Figura 2 e o gráfico da Figura 3 mostram a influência do peso aplicado às opções de resposta no resultado final. Enquanto a Figura 2 mostra “Eficiência” e “Eficácia” como tendo aproximadamente o mesmo grau de qualidade, a Figura 3 mostra que a percepção dos usuários finais do grau de qualidade de “Eficiência” é maior, pois obteve mais respostas com “Concordo totalmente”. Essa diferença é ainda



mais evidente no caso das características “Confiança” e “Compatibilidade”. “Compatibilidade” obteve mais respostas com a opção “Concordo totalmente”, mas também obteve mais “Discordo totalmente” – mostrando uma certa divergência de opiniões entre os usuários; como resultado, “Confiança” pôde ser percebida como tendo um grau de qualidade maior que “Compatibilidade”. Por outro lado, outras características não foram afetadas pela mudança de métrica: “Satisfação” e “Cobertura de contexto” ainda são percebidas como tendo o grau de qualidade mais alto e mais baixo, respectivamente.

Todas as características obtiveram pontuações dentro de duas distâncias interquartis (IQR), ou seja, os resultados não se afastam muito da mediana. De forma complementar, não foram identificados *outliers* para a métrica #2.

### Métrica #3 Pontuação de qualidade por participante do questionário.

Enquanto as métricas #1 e #2 mostram no geral a percepção dos usuários com relação a cada característica de qualidade, a métrica #3 mostra a percepção de qualidade como um todo por participante, ilustrada pela Figura 4. Utilizando-se das referências usadas na escala SUS [Bangor et al. 2009], podemos ver que nenhum participante tem a percepção de que o sistema DATATOX possui um grau pobre de qualidade, e que a maioria das pontuações mostra uma percepção voltada para um grau de qualidade aceitável; contudo, apenas 16% dos participantes consideram o sistema como tendo um bom grau de qualidade.

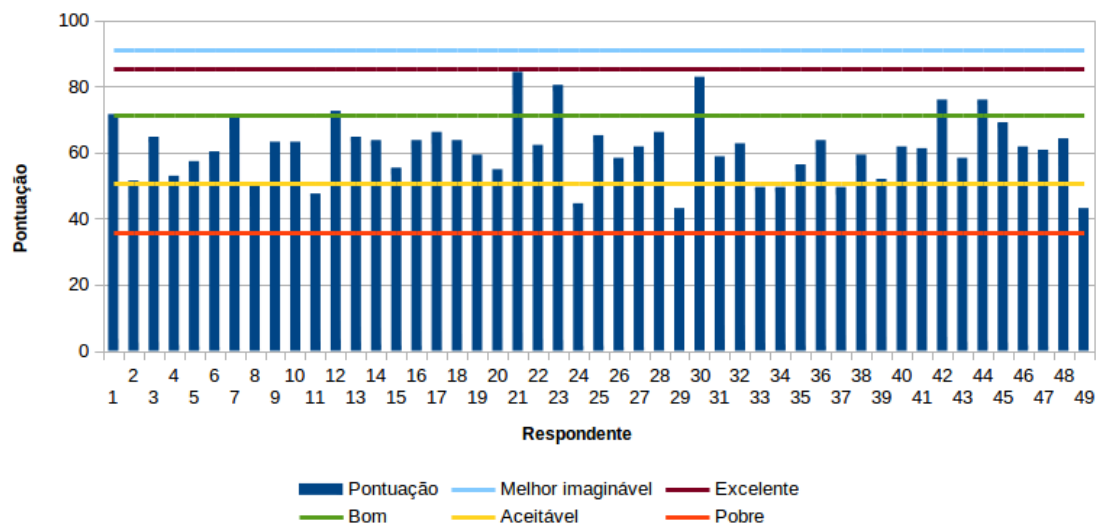


Figura 4. Pontuação calculada para cada característica.

A seção seguinte discute e contextualiza os dados analisados, ilustrando as possíveis razões para as percepções de baixo grau de qualidade em algumas características.

## 4. Discussão

Os dados obtidos na execução do estudo de caso mostram que, no geral, o sistema DATATOX foi percebido como tendo um grau aceitável de qualidade, como mostra a

Figura 4. As divisões em quartis exibidas na Figura 3 mostram como os usuários percebem o sistema relacionado a suas características de qualidade: “Satisfação”, “Eficiência” e “Eficácia” têm os graus de qualidade mais altos, seguidos por “Mitigação de risco”, “Segurança” e “Adequação funcional”. “Desempenho” e “Usabilidade” foram percebidos com um grau de qualidade moderadamente menor.

O baixo grau de qualidade percebido para “Cobertura de contexto” é dado pelo baixo grau de flexibilidade do sistema: o mesmo só pode ser acessado usando-se o navegador Mozilla Firefox, apenas em computadores *desktop* ou *laptop* e não pode ser acessado pelos requisitantes. Além disso, os usuários finais consideraram que o conjunto de funcionalidades não cobre todo o contexto de seus trabalhos, isto é, há funcionalidades desejadas pelos usuários que o sistema ainda não disponibiliza.

Como o DATATOX é um sistema *web* de teletoxicologia, ele está suscetível a flutuações na qualidade da conexão com a Internet. Tais instabilidades contribuem para um baixo grau de tolerância a falhas e disponibilidade, o que é percebido nos gráficos pela característica “Confiança”. O sistema também não é integrado com nenhuma base de dados toxicológicos, forçando seus usuários a depender de bases de dados externas, contribuindo para um baixo grau de “Interoperabilidade”.

#### 4.1. Ameaças à validade

Para aumentar a validade interna e evitar um viés de seleção, durante o planejamento do estudo de caso foi definido que seriam convidados usuários do sistema DATATOX de todas as unidades disponíveis, de todos os turnos de trabalho, pessoalmente ou via telefone. Pode-se assim garantir uma maior uniformidade dos dados, uma vez que cada CIT tem suas particularidades e flutuações na qualidade da conexão de Internet (o que pode gerar visões de qualidade diferentes). Para minimizar o viés de projeto foi adotado o modelo AdeQUATE no lugar de um questionário construído *ad-hoc*. Embora esse modelo e suas escalas de avaliação ainda não tenham sido validados por um estudo maior, assumiu-se que a derivação sistemática dos itens baseados em normas de qualidade internacionais como a ISO/IEC 25010 reduz ameaças à validade da construção do estudo de caso.

Durante este trabalho, o foco se manteve em um sistema específico de teletoxicologia – o sistema DATATOX. Logo, os resultados deste estudo de caso têm uma validade externa limitada. Contudo oferece valiosas informações para trabalhos futuros ligados à validação do modelo AdeQUATE.

## 5. Conclusão

Sistemas de controle toxicológico auxiliam toxicologistas a gerenciar casos de intoxicações por vários agentes, manipulando informações sobre o caso, informações sobre o paciente e acompanhando os casos quando necessário. Alguns sistemas auxiliam ainda mais os toxicologistas, oferecendo bases de dados toxicológicos especializadas. Nesse contexto, qualidade de software é um tópico delicado porque não há um padrão específico para guiar o desenvolvimento de tais sistemas, tampouco um modelo de avaliação de qualidade. Enquanto a norma ISO/IEC 25010 apresenta características e subcaracterísticas que permitem a engenheiros de *software* avaliar qualidade geral, as especificações do domínio de atenção à saúde são ignoradas.



Neste trabalho foi analisada a qualidade de *software* do sistema DATATOX – um sistema de teletoxicologia usado por vários CITs no Brasil – através de um estudo de caso que adota o modelo AdEQUATE. A análise dos dados mostra que o sistema foi percebido como tendo um bom grau de qualidade. Seus pontos fortes estão relacionados à satisfação, eficiência e eficácia, mostrando um sistema que satisfaz seus usuários e permite-os executar tarefas com acuracidade e completude, gastando poucos recursos. Por outro lado, o sistema DATATOX foi considerado como tendo um baixo grau de confiança, compatibilidade e cobertura de contexto, significando que o sistema algumas vezes não realiza algumas funções, não troca informações com outros sistemas e não é flexível o suficiente para cobrir todos os contextos de uso.

Adicionalmente, este estudo mostra a viabilidade de se avaliar qualidade de *software* usando a perspectiva dos usuários finais de sistemas de informação toxicológica adotando o questionário sistematicamente derivado do modelo AdEQUATE. Através dessa avaliação a ABRACIT tem a compreensão da qualidade atual do sistema, podendo investir em otimizações específicas para garantir um serviço de qualidade ainda melhor.

Após uma extensa e profunda validação do modelo AdEQUATE, um trabalho futuro pretende rever a percepção dos usuários da qualidade do sistema usando o modelo validado, e possivelmente com mais participantes, compará-lo com outros sistemas de teletoxicologia. Nesse ínterim, futuros trabalhos na área de banco de dados e inteligência artificial podem aumentar a qualidade do sistema, integrando o mesmo a uma base de dados toxicológicos (aumentando assim sua compatibilidade).

## Agradecimentos

Este trabalho teve o apoio da Secretaria de Estado da Saúde de Santa Catarina - SES/SC, do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq e da Fundação de Amparo à Pesquisa de Santa Catarina - FAPESC.

Os autores também gostariam de agradecer a todos os usuários do sistema DATATOX que dedicaram seu tempo e esforço para responder o questionário.

## Referências

- Alves, J. M., Wangenheim, C., Lacerda, T., Savaris, A., and Wangenheim, A. (2015). Adequate software quality evaluation model v1.0. Technical report, Instituto Nacional para Convergência Digital – INCoD.
- Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123.
- Bateman, D. N., Good, A. M., Laing, W., and Kelly, C. (2002). Toxbase: poisons information on the internet. *Emergency medicine journal*, 19(1):31–34.
- Behkamal, B., Kahani, M., and Akbari, M. K. (2009). Customizing iso 9126 quality model for evaluation of b2b applications. *Information and software technology*, 51(3):599–609.
- Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7.
- Brooke, J. (2013). Sus a retrospective. *Journal of Usability Studies*, 8(2):29–40.

- CIT/RS, C. (2015). Estatísticas dinâmicas. Date last accessed 2015-12-11.
- CIT/SC, C. (2015). Estatísticas anuais. Date last accessed 2015-12-11.
- International Standard Organization (2003). ISO/IEC 9126-2:2003 Software engineering – Product quality – Part 2: External metrics. Technical report.
- Mowry, J. et al. (2015). 2014 annual report of the american association of poison control centers' national poison data system (npds): 32nd annual report. *Clinical toxicology*, 53(10).
- Perry, L. et al. (2014). National toxicovigilance for pesticide exposures resulting in health care contact—an example from the uk's national poisons information service. *Clinical toxicology*, 52(5):549–555.
- Poison Control Centers (2015). Date last accessed 2015-12-14.
- Sauro, J. and Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2215–2224. ACM.
- Seifert, S. M. et al. (2013). An analysis of energy-drink toxicity in the national poison data system. *Clinical toxicology*, 51(7):566–574.
- Skolnik, A. (2013). Telemedicine and toxicology: back to the future? *Journal of Medical Toxicology*, 9(3):217–219.
- Tormey, W. and Moore, T. (2013). Poisonings and clinical toxicology: a template for ireland. *Irish journal of medical science*, 182(1):17–23.
- World Health Organization et al. (1997). Guidelines for poison control.