# 4 Artigo

(...)

Resultados

Um total de 1.172 admissões foram consideradas elegíveis. Dos 268 prontuários selecionados, apenas 219 puderam ser acessados pelos pesquisadores. Desses, 20 foram excluídos: nove por tempo de internação hospitalar inferior a 24 horas, nove devido ao período de internação não corresponder ao período de inclusão no estudo e dois por se referirem a prontuários de doadores cadáveres para transplante de órgãos sólidos. Uma rodada extra de randomização selecionou 12 admissões adicionais, totalizando 211 prontuários (Figura 1). A indisponibilidade de prontuários para pesquisa foi justificada pelas características de armazenamento remoto de grandes prontuários e sua utilização para outras finalidades, como atendimento médico, auditoria ou faturamento.

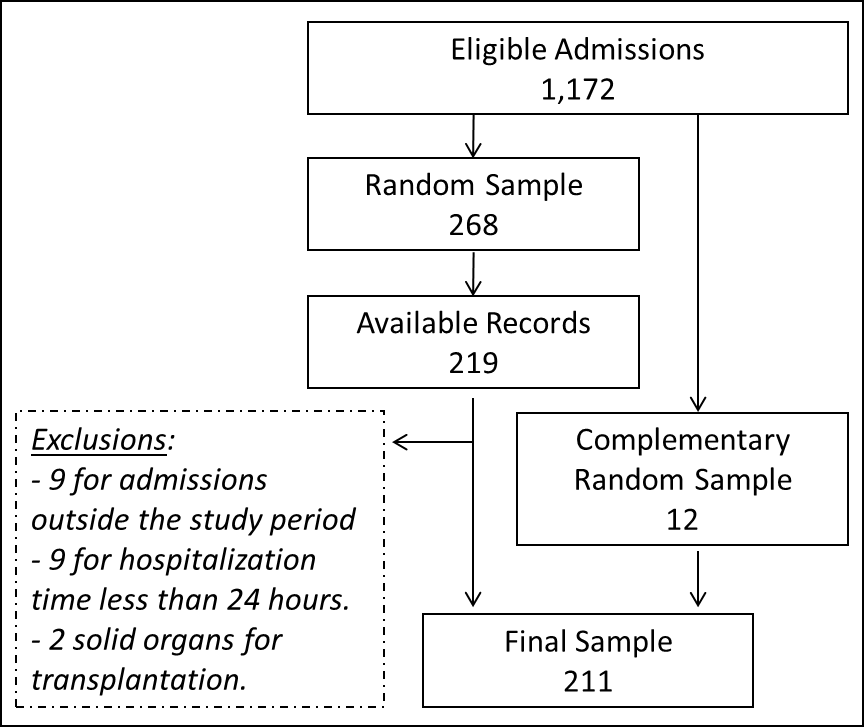


Figura 1 – Processo de seleção da amostra

Entre as 211 internações da amostra final, 134 (63,%%) eram de mulheres, 63 (29,9%) eram de pacientes com 60 anos ou mais, 172 (81,5%) ocorreram em caráter de urgência e 47 (22,3%) eram de pacientes com Índice de Comorbidade de Charlson maior ou igual a 3, ou seja, considerados alto ou muito alto. Quanto ao tipo de internação, 90 (42,7%) foram por motivos cirúrgicos, 76 (36,0%) clínicos e 45 (21,3%) obstétricos. O tempo médio de duração das internações incluídas nas amostras foi de 12,2 dias (Desvio padrão XXX), o que somou um total de 2.571 pacientes-dia. Devido às características metodológicas de corte transversal estabelecidas pelo padrão referência, apenas os dias de internação que estavam dentro do período da pesquisa foram contados, o que resultou em 1.393 pacientes-dia. Das 211 internações, 150 tiveram mais de 75% do tempo de hospitalização incluídos no estudo, o que representa 717 pacientes-dia, 23 tiveram 50-74% (358 pacientes-dia), 24 tiveram 25-49% (263 pacientes-dia) e 14 tiveram 4-24% (55 pacientes-dia).

Foram identificados 627 rastreadores pelo método padrão referência, dos quais 274 (43,7%) foram positivos para a ocorrência de eventos adversos. Os rastreadores mais frequentes foram os de cuidado quase em tempo real coletados por meio de entrevistas com os profissionais da assistência durante as visitas diárias aos setores de cuidado (n=357), seguidos daqueles referentes às solicitações de autorização para início de antibióticos (n=85) e aos resultados de exames laboratoriais alterados (n=49). A frequência de ocorrência e de positividade dos rastreadores do método padrão referência por cada tipo de fonte está descrita na tabela 1.

Tabela 1 - Rastreadores do método referência: frequência por tipo de fonte

|  |  |  |
| --- | --- | --- |
| **Rastreador por tipo de fonte** | **Total** | **Positivos para eventos adversos (%)** |
| Entrevista semiestruturada com profissionais | 357 | 165 (46,2) |
| Autorização para início de antibióticos | 85 | 29 (34,1) |
| Resultado de exames laboratoriais | 49 | 11 (22,5) |
| Avaliação de notas cirúrgicas | 34 | 9 (26,5) |
| Revisão diária das prescrições | 33 | 15 (45,5) |
| Achados em prontuário durante coleta ou análise de outros dados | 21 | 21 (100) |
| Relatórios de notificação voluntária | 21 | 11 (52,4) |
| Reinternação em até 30 dias da alta | 11 | 3 (27,3) |
| Avaliação de notas obstétricas | 7 | 2 (28,6) |
| Relatório sobre infecções relacionadas à assistência | 6 | 6 (100) |
| Relatórios da Agência Transfusional | 2 | 2 (100) |
| Avaliação da causa dos óbitos | 1 | 0 (0) |
| **Total** | 627 | 274 (43,7) |

Durante o período do estudo, os dois métodos combinados identificaram 233 eventos adversos em 92 internações, o que resultou em taxas globais de 43,6% das internações com pelo menos um evento adverso e 167 eventos adversos/1.000 pacientes-dia. Individualmente, o método padrão referência identificou 176 eventos adversos em 67 internações, com taxas de 31,8% das internações com pelo menos um evento adverso e 126 eventos adversos/1.000 pacientes-dia, e o *GTT* identificou 129 eventos adversos em 76 internações, com taxas de 36,0% das internações com pelo menos um evento adverso e 92 eventos adversos/1.000 pacientes-dia. Considerando-se a exclusão dos EAs que ocorreram fora do período da pesquisa, a porcentagem de internações com EAs pode estar subestimada.

Quase 70 % de todos os eventos adversos identificados foram classificados como E quanto à gravidade do dano (n=160) e nenhum evento foi classificado como I, ou seja, resultou em morte do paciente (Tabela 2). Em números absolutos, o *GTT* identificou mais eventos classificados como F-I que o método padrão referência, que por sua vez foi capaz de identificar quase o dobro de eventos classificados como E em comparação ao *GTT.* A notificação voluntária, que foi incluída como um dos rastreadores utilizados no método padrão referência, encontrou nove eventos adversos (oito classificados como E e um como H), o que corresponde a 3,9% do total de eventos identificados), sendo que apenas 3 deles foram identificados também pelo *GTT* (um H e dois E)*.*

Tabela 2 - Eventos adversos por classificação quanto à gravidade do dano e por método de identificação

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Eventos adversos por gravidade\*** | **Todos os métodos** | **Padrão Referência** | ***GTT-IHI*** | ***P\*\**** |
| E | 160 | 124 | 64 |  |
| F | 55 | 42 | 49 |  |
| G | 8 | 3 | 8 |  |
| H | 10 | 7 | 8 |  |
| I | 0 | 0 | 0 |  |
| F-I | 73 | 52 | 65 |  |

\*Classificação da gravidade segundo índice *NCC MERP adaptado pelo GTT-IHI:* categoria E - dano temporário ao paciente com necessidade de intervenção; categoria F - dano temporário ao paciente com necessidade de iniciar ou prolongar hospitalização; categoria G -dano permanente ao paciente; categoria H - intervenção necessária para manutenção da vida; categoria I - Morte do paciente.

\*\* Teste Qui-quadrado

Quanto à natureza, os tipos de eventos adversos mais frequentemente encontrados foram aqueles relacionados a acesso venoso periférico (n=67), eventos medicamentosos (n=54), cirúrgicos/anestésicos (n=30) e infecções (n=23) (Tabela 3). O *GTT* não identificou nenhum dos 16 eventos relacionados a cateteres gástricos/entéricos e apenas 2 dos 67 eventos relacionados a acesso venoso periférico. Entre os 176 eventos detectados pelo padrão referência, 103 foram identificados exclusivamente por rastreadores de cuidado que em tempo real colhidos de entrevistas com os profissionais, sendo 59 relacionados a acesso venoso periférico e 14 relacionados a cateteres gástricos/entéricos. Apesar dos revisores médicos do padrão referência terem encontrado registros em prontuário de 82 dos 103 eventos detectados a partir de relatos, o *GTT* identificou apenas 18 deles.

Tabela 3 - Eventos adversos por natureza por método de identificação e classificação quanto a gravidade do dano

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Eventos adversos por natureza** |  | **Todos os métodos** | |  | **Padrão Referência** | |  | ***GTT-IHI*** | |
|  | **Total** | **F-I** |  | **Total** | **F-I** |  | **Total** | **F-I** |
| Acessos venosos periféricos |  | 67 (28,8) | 0 (0) |  | 66 (37,5) | 0 (0) |  | 2 (1,6) | 0 (0) |
| Medicamentosos |  | 54 (23,2) | 13 (17,8) |  | 29 (16,5) | 10 (19,2) |  | 46 (35,7) | 12 (18,5) |
| Cirúrgicos/anestésicos |  | 30 (12,9) | 21 (28,8) |  | 19 (10,8) | 13 (25,0) |  | 27 (20,9) | 21 (32,3) |
| Infecções |  | 23 (9,9) | 14 (19,2) |  | 19 (10,8) | 13 (25,0) |  | 22 (17,1) | 13 (20,0) |
| Atraso na assistência |  | 16 (6,9) | 16 (21,9) |  | 12 (6,8) | 12 (23,1) |  | 12 (9,3) | 12 (18,5) |
| Cateteres gástricos/entéricos |  | 16 (6,9) | 0 (0) |  | 16 (9,1) | 0 (0) |  | 0 (0) | 0 (0) |
| Transfusão de hemoderivados |  | 6 (2,6) | 1 (1,4) |  | 2 (1,1) | 0 (0) |  | 6 (4,7) | 1 (1,5) |
| Flebite |  | 4 (1,7) | 0 (0) |  | 3 (1,7) | 0 (0) |  | 2 (1,6) | 0 (0) |
| Lesão por pressão |  | 3 (1,3) | 0 (0) |  | 1 (0,6) | 0 (0) |  | 3 (2,3) | 0 (0) |
| Radioterapia |  | 3 (1,3) | 3 (4,1) |  | 1 (0,6) | 1 (1,9) |  | 3 (2,3) | 3 (4,6) |
| Via área |  | 2 (0,9) | 2 (2,7) |  | 2 (1,1) | 2 (3,9) |  | 0 (0) | 0 (0) |
| Acessos vasculares centrais |  | 2 (0,9) | 0 (0) |  | 2 (1,1) | 0 (0) |  | 0 (0) | 0 (0) |
| Sondagem vesical |  | 2 (0,9) | 2 (2,7) |  | 0 (0) | 0 (0) |  | 2 (1,6) | 2 (3,1) |
| Assistência obstétrica |  | 1 (0,4) | 1 (1,4) |  | 1 (0,6) | 1 (1,9) |  | 1 (0,8) | 1 (1,5) |
| Queda |  | 1 (0,4) | 0 (0) |  | 1 (0,6) | 0 (0) |  | 1 (0,8) | 0 (0) |
| Transplante de células |  | 1 (0,4) | 0 (0) |  | 1 (0,6) | 0 (0) |  | 1 (0,8) | 0 (0) |
| Terapia dialítica |  | 1 (0,4) | 0 (0) |  | 0 (0) | 0 (0) |  | 1 (0,8) | 0 (0) |
| Lesão de pele por lesão mecânica |  | 1 (0,4) | 0 (0) |  | 1 (0,6) | 0 (0) |  | 0 (0) | 0 (0) |
| **Total** |  | 233 | 73 |  | 176 | 52 |  | 129 | 65 |

Considerando-se o total de eventos adversos identificados no estudo, o *GTT* foi capaz de encontrar 55,4% de todos os eventos adversos e 89,0% dos eventos classificados como F-I quanto a gravidade do dano. Os dados sobre a validade do *GTT* estão na tabela 4 (ou figura 2).

Tabela 4 – Sensibilidade, especificidade e acurácia do *GTT-IHI* para identificação de todos os eventos adversos e para identificação de eventos adversos classificados como “F” a “I” quanto a gravidade do dano

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Internação com eventos adversos** | | |  | **Eventos adversos individualmente** | | |
|  |  | **Total de eventos adversos** |  | **Eventos adversos F-I** |  | **Total de eventos adversos** |  | **Eventos adversos F-I** |
| **Sensibilidade\*** |  | 0,76 (XX – XX) |  | 0,90 (XX – XX) |  | 0,41 (XX – XX) |  | 0,85 (XX – XX) |
| **Especificidade\*** |  | 0,83 (XX – XX) |  | 0,90 (XX – XX) |  | 0,68 (XX – XX) |  | 0,68 (XX – XX) |
| **Acurácia\*** |  | 0,81 (XX – XX) |  | 0,90 (XX – XX) |  | 0,54 (XX – XX) |  | 0,54 (XX – XX) |

\*Valores (IC 95%)

Acurácia do *GTT* para identificação de eventos adversos de qualquer gravidade de dano quanto a internações

|  |  |  |  |
| --- | --- | --- | --- |
| ***GTT*** | **Padrão Referência** | | |
| **Positivo** | **Negativo** | **Total** |
| **Positivo** | 51 | 25 | 76 |
| **Negativo** | 16 | 119 | 135 |
| **Total** | 67 | 144 | 211 |
| **Sensibilidade = 0,76 (XX – XX)** | | | |
| **Especificidade = 0,83 (XX – XX)** | | | |
| **Acurácia = 0,81 (XX – XX)** | | | |

Acurácia do *GTT* para identificação de eventos adversos de classificação F-I de quanto a internações

|  |  |  |  |
| --- | --- | --- | --- |
| ***GTT*** | **Padrão Referência** | | |
| **Positivo** | **Negativo** | **Total** |
| **Positivo** | 37 | 17 | 54 |
| **Negativo** | 4 | 153 | 157 |
| **Total** | 41 | 170 | 211 |
| **Sensibilidade = 0,90 (XX – XX)** | | | |
| **Especificidade = 0,90 (XX – XX)** | | | |
| **Acurácia = 0,90 (XX – XX)** | | | |

Acurácia do *GTT* para identificação de eventos adversos de qualquer gravidade de dano quanto aos eventos individuais

|  |  |  |  |
| --- | --- | --- | --- |
| ***GTT*** | **Padrão Referência** | | |
| **Positivo** | **Negativo** | **Total** |
| **Positivo** | 72 | 57 | 129 |
| **Negativo** | 104 | 119 | 223 |
| **Total** | 176 | 176 | 352\* |
| **Sensibilidade = 0,41 (XX – XX)** | | | |
| **Especificidade = 0,68 (XX – XX)** | | | |
| **Acurácia = 0,54 (XX – XX)** | | | |

\*352 = 223 eventos adversos identificados em 92 internações + 119 internações sem eventos adversos identificados

Acurácia do *GTT* para identificação de eventos adversos de classificação F-I de gravidade de dano quanto aos eventos individuais

|  |  |  |  |
| --- | --- | --- | --- |
| ***GTT*** | **Padrão Referência** | | |
| **Positivo** | **Negativo** | **Total** |
| **Positivo** | 44 | 21 | 65 |
| **Negativo** | 8 | 153 | 161 |
| **Total** | 52 | 174 | 226\* |
| **Sensibilidade = 0,85 (XX – XX)** | | | |
| **Especificidade = 0,88 (XX – XX)** | | | |
| **Acurácia = 0,87 (XX – XX)** | | | |

\*226 = 73 eventos adversos F-I identificados em 58 internações + 153 internações sem eventos adversos F-I identificados

Discussão

Statement of principal findings

O objetivo desse estudo foi avaliar a validade do *GTT* como método para identificação de EAs. Apesar de amplamente utilizado por instituições individuais e em alguns contextos como ferramentas nacionais em programas de segurança do paciente, não há dados claros sobre sua acurácia (Hanskamp-Sebregts et al, 2016; Hibbert et al, 2016; Klein et al, 2020). Os resultados encontrados neste protocolo de pesquisa, ao compará-lo ao método padrão referência, mostram que o *GTT* apresenta valores de sensibilidade e especificidade satisfatórios, especialmente para EAs de maior gravidade, classificados como F-I pelo índice de *NCC MERP* adaptado pelo *IHI.*

* Interpretation within the context of the wider literature

Somando-se todos os eventos encontrados no estudo, independentemente do método de identificação, as taxas de EAs foram superiores aos dados já publicados (Classen et al, 2011, Hanskamp-Sebregts et al, 2016; Hibbert et al, 2016; Klein et al, 2020). Esse fato já era esperado, uma vez que estudos mostram que diferentes métodos identificam diferentes EAs e que a combinação de estratégias permite a detecção de um maior número de eventos (Naessens et al, 2009; Classen et al, 2011; Kennerly et al, 2014; Rutberg et al, 2014). Por outro lado, as taxas obtidas pelo *GTT* se aproximam daquelas descritas por Griffin e Resar (2009), mas são maiores que a de alguns estudos (Pleassen et al, 2021; Najjar et al, 2013; Kennerly et al, 2014; Kurutkan et al, 2015).

Além de questões relacionadas a diferenças no perfil de complexidade dos pacientes ou na qualidade do cuidado oferecido, outros elementos de ordem metodológica impactam nas taxas de EAs obtidas. Um deles é a ausência de consenso no uso de definições e classificações em segurança do paciente, em especial no julgamento do que seria considerado “dano leve” ou “dano temporário que necessita intervenção”, aqui classificados como E (Griffin e Resar, 2009; Schildmeijer et al, 2013; Mattsson et al, 2013). O entendimento por parte de alguns revisores de que alguns danos poderiam ser “aceitáveis” ou mesmo “inevitáveis”, implica no não reconhecimento de parte desses danos como um EA em saúde (Ref). Neste estudo, os revisores de ambos os métodos foram orientados a considerar EA qualquer incidente que causou dano do ponto de vista do paciente, exigindo qualquer monitoramento ou intervenção, mesmo que mínima, independente do potencial de evitabilidade (IOM, 2001; Griffin e Resar, 2009).

Para os métodos que se baseiam na revisão de prontuários, um fator adicional é o fato de que os eventos de menor dano são menos frequentemente registrados e, assim como o a notificação voluntária, a maior detecção de EAs leves pode significar apenas melhores registros e não pior segurança (Shojania e Marang-van de Mheen, 2020). No caso do *GTT,* além do impacto da qualidade dos registros, a limitação do tempo de revisão e a recomendação de que o prontuário não deve ser lido “from front page to back page” são fatores que favorecem sua aplicabilidade, mas implicam em reconhecida perda de sensibilidade para eventos com danos mais leves (Griffin e Resar, 2009; Mattsson et al, 2013). Os achados deste estudo reafirmam essa limitação do método ao mostrar que o *GTT* não detectou 96 do total de 160 EAs classificados na categoria E de gravidade do dano identificados. É importante ressaltar que desses 96 EAs não identificados pelo *GTT,* os revisores médicos do método padrão referência encontraram registro em prontuário de 85 deles, sendo que em 71 os registros estavam em notas da enfermagem, porção do prontuário não priorizada no processo de revisão pelos revisores primários que utilizam o *GTT* e frequentemente não acessados por completo nas revisões de internações mais prolongadas (Griffin e Resar, 2009; Sharek et al, 2011; Schildmeijer et al, 2013).

Quanto à natureza e considerando-se os dois métodos de identificação, os eventos relacionados a acessos venosos periféricos foram os mais frequentes, totalizando 30% do total, todos eles com dano E. Analisando-se apenas os eventos identificados pelo *GTT*, assim como em outros estudos (Classen et al, 2011; Naessens et al, 2009; Landrigan et al, 2010; Schildmeijer et al, 2012; Kennerly et al, 2014; Rutberg et al, 2014), os eventos mais frequentemente encontrados foram os medicamentosos, cirúrgicos/anestésicos e as infecções. A falha do *GTT* em detectar a maioria de eventos relacionados ao cuidado de enfermagem, entre eles os relacionados a acesso venoso periférico e a cateteres gástricos/entéricos, corrobora os achados de Schildmeijer sobre a impressão relatada por revisores experientes de que o *GTT* é voltado principalmente para os danos ligados às ações realizadas pelos médicos (Schildmeijer et al, 2013).

Os resultados desse estudo reforçam o papel de destaque do *GTT* como ferramenta válida para identificação de EAs, principalmente para os eventos de maior gravidade, classificados como F-I (Griffin e Resar, 2009; Mattsson et al, 2013). Embora a alta sensibilidade do *GTT* em comparação a outros métodos já tenha sido demonstrada (Naessens et al, 2009; Classen et al, 2011; Kennerly et al, 2014; Rutberg et al, 2014), poucos estudos apresentaram dados de acurácia (Klein et al, 2020). Classen e colaboradores encontraram sensibilidade de 94,9% para detectar pacientes com pelo menos um EA e especificidade de 100% para detectar pacientes sem eventos ao comparar os resultados do *GTT* com os achados de uma revisão retrospectiva detalhada de dados de prontuário e de dados administrativos realizada por um grupo de especialistas (Classen et al, 2011). Nesse estudo, os pesquisadores buscaram agregar fontes de dados para alcançar um padrão referência que se aproximasse ao máximo a um método padrão ouro. Entretanto, a busca se restringiu a revisão retrospectiva de documentos e dados e alguns EAs não registrados podem ter sido perdidos, o que pode ter gerado resultados de sensibilidade e especificidade superestimados. Além disso, o estudo apresenta os dados de validade apenas para o desfecho internações com ou sem EAs, não sendo possível analisar a capacidade de identificação individual de cada EA.

* Strengths and limitations

Um dos pontos fortes do estudo foi a elaboração de um método padrão referência altamente sensível para a avaliação da validade do *GTT*. Ele combinou diferentes fontes reconhecidas para obtenção de informações sobre a ocorrência de dano relacionado ao cuidado em saúde, incluindo dados administrativos e clínicos agrupados, além de relatos de profissionais da assistência direita, coletados quase em tempo real (Ref). Apesar de reconhecermos que a observação direta fornece mais informações sobre segurança do paciente em tempo real (Ref.), por questões práticas optamos por substituí-la pela coleta de rastreadores em entrevistas aos profissionais da assistência ao final de cada turno de plantão. Os rastreadores identificados por meio dessa fonte levaram a confirmação de 127 dos 176 EAs do padrão referência (72,2%), sendo que eles foram a única fonte para identificação de 103 (87 EAs com dano E, 10 com F, 2 com G e 4 com H).

O método padrão referência utilizado neste protocolo de pesquisa também apresenta limitações que interferiram na avaliação da validade do *GTT.* A mais importante se refere ao fato de que ele falhou em identificar 57 EAs em 25 internações. Apesar de corresponderem a EAs corretamente identificados pelo *GTT*, essas situações foram consideradas como “falsos positivos” para o cálculo da validade. Isso significa que os dados de sensibilidade, especificidade e acurácia global do *GTT* podem ser melhores do que os apresentados. Além disso, a decisão final sobre o dano ser atribuído ao cuidado em saúde e não a evolução da doença ficou a cargo de um revisor médico e está sujeita a vieses de interpretação e de registro. Tentamos reduzir a subjetividade desse julgamento acrescentando a escala de confiança sobre a existência de um EA e as perguntas para guiar o raciocínio do médico revisor.

É importante destacar também que esse estudo foi desenvolvido em um hospital de referência para pacientes de cuidado terciário e quaternário no estado de Minas Gerais, incluindo transplantes e oncohematologia. Trata-se de um hospital universitário, que tem se esforçado nos últimos anos para alcançar metas de qualidade internacionais, tanto no cuidado, quanto no registro em prontuário. Essas características, tanto relacionadas ao perfil de gravidade dos pacientes, quanto ao modelo de cuidado da instituição, podem ter interferido nos resultados obtidos, inclusive com impacto na avaliação da validade da ferramenta.

Outra limitação do estudo diz respeito a divergências na taxonomia em segurança do paciente e falta de consenso em definições e classificações. Essa limitação, descrita também por outros pesquisadores (Naessens et al, 2009; Mull et al, 2015; Shojania e Marang-van de Mheen, 2020), dificulta a comparação de dados entre estudos e interfere no aprendizado global em segurança do paciente. Destaca-se, por exemplo, a ausência de um sistema de classificação de EAs por natureza que seja simples e de fácil interpretação. Considerando-se os objetivos deste estudo, optamos por utilizar uma classificação própria que permitisse a compreensão dos EAs mais frequentemente identificados ou perdidos pelo *GTT.*

Apesar de terem sido empregados revisores primários com pouca experiência na aplicação do *GTT,* não consideramos isso como um ponto fraco*.* Estudos demonstraram que treinamento e experiência dos revisores aumentam a confiabilidade entre eles (Adler et al, 2008; Naessens et al, 2010; Schildmeijer et al, 2012), o que pode impactar também na validade da ferramenta. Contudo, a confiabilidade interexaminador dessa dupla de revisores primários constituída por acadêmicos de medicina, algo não descrito até o momento, foi descrita anteriormente com resultados de confiabilidade substancial para identificação de internações com EAs entre um acadêmico e outro e em relação a uma dupla de enfermeiros experientes e moderada em relação ao revisor médico (Moraes et al, 2021), o que é comparável a outros estudos que avaliaram a performance de revisores experientes (Sharek et al, 2011; Schildmaijer et al, 2021; Mattsson et al 2013).

Por fim, embora o *GTT* tenha se mostrado válidopara identificação de EAs, em especial os de maior gravidade, outras propriedades do método que não foram abordadas neste estudo precisam ser consideradas para que o *GTT* possa ser visto como uma estratégia adequada de medição em melhoria. Alguns estudiosos já demonstraram preocupação quanto ao impacto que a pequena amostra quinzenal (Ref) e que a moderada confiabilidade intra e interexaminadores obtida na maioria dos estudos que avaliaram reprodutibilidade (Ref) poderia ter nas conclusões das análises da evolução das taxas de EAs obtidas ao longo do tempo em programas de melhoria da segurança do paciente.

* Implications for policy, practice and research

Se por um lado os nossos resultados reforçam a validade do *GTT* para identificação de EAs mais graves, por outro eles ressaltam a sua fragilidade para a identificação de EAs que geram menor dano ao paciente. Esses eventos, embora causem pontualmente um menor impacto ao paciente e aos sistemas de saúde, podem se tornar extremamente relevantes pela frequência que acontecem. Neste estudo, por exemplo, os eventos relacionados a acessos venosos periféricos e a cateteres gástricos/entéricos somados correspondem a 35,6% de todos os EAs identificados e, portanto, devem ser vistos como indicadores de cuidado de má qualidade e alvo de ações de melhoria, o que inclui estratégias de medição adequadas.

Apesar de não ser nosso objeto de estudo, a análise dos EAs identificados pelas diferentes fontes de rastreadores no método padrão referência permite inferências sobre possíveis métodos complementares para identificação de determinados grupos de EAs que são geralmente perdidos pelo *GTT.* Um exemplo é a busca por rastreadores quase em tempo real por meio de entrevista com os profissionais que prestam assistência direta ao paciente ao final de cada turno de plantão para identificação de EAs relacionados aos cuidados de enfermagem. Esse método pode ser uma alternativa à observação direta, que, apesar de muito sensível, é um método caro e pouco factível para a maior parte das instituições, com possíveis ganhos em cultura de segurança pelo maior envolvimento dos profissionais da assistência direta (Shojania e Marang-van de Mheen, 2020). Contudo, mais estudos são necessários sobre o potencial desse método identificar diferentes tipos de EAs, estratégias de amostragem, possíveis vieses e custos envolvidos.

Analisando-se os EAs classificados como F-I quanto a gravidade do dano, o *GTT* não detectou apenas 8 dos 73 identificados, sendo 4 classificados como “Atraso na assistência”. Esses EAs se referem ao aumento da permanência hospitalar devido a espera para agendamento e realização de exames ou procedimentos indicados no tratamento. Isso pode ser justificado pela ausência de um trigger específico para esse tipo de EA, que é mais comum em serviços públicos de saúde de países com menos recursos. Situações como essa poderiam ser contornadas pela inclusão de *triggers* na perspectiva de cada contexto local, como já descrito por outros serviços de saúde (Ref).

Apesar de demandar recursos de tempo e profissionais treinados e experientes para revisão manual dos prontuários, o que implica em custos, o *GTT* é o método, entre os mais sensíveis, que requer menores esforços e que apresenta maior potencial de sustentabilidade a longo prazo (Naessens et al, 2010). Adler e colaboradores estimaram que o custo para a implementação do *GTT* em um sistema de saúde nos Estados Unidos foi de 710 dólares para o treinamento inicial da equipe de revisores e de 611 dólares ao mês por hospital para manutenção das atividades de revisão (Adler et al, 2008). Nossa experiência em empregar acadêmicos de medicina de anos finais do curso como revisores primários mostrou-se uma alternativa de menor custo, com bons resultados de confiabilidade quando comparados a enfermeiros experientes (Moraes et al, 2021). Entretanto, a detecção adequada da incidência de EAs corresponde apenas a uma pequena parcela das atividades para melhoria da qualidade dos cuidados. É fundamental que os programas de segurança do paciente criem estratégias para transformar os dados em aprendizado e melhorar os cuidados e os resultados para os pacientes (Klein et al, 2020). Qualquer avaliação de custo-efetividade deve considerar não só os custos de implementação da ferramenta, mas também os benefícios das atividades em melhoria da qualidade associadas (Mull et al, 2015).

* Conclusions

O *GTT* mostrou ser um método válido para identificação de EAs em pacientes adultos internados. Contudo, ele não foi capaz de identificar todos os EAs, deixando de detectar principalmente aqueles classificados como dano E e os relacionados a alguns cuidados de enfermagem. Por isso, não recomendamos que o *GTT* seja utilizado como método único de identificação de EAs. Outras ferramentas de medição devem ser associadas a ele para alcançar resultados representativos do perfil de segurança e qualidade nos cuidados em saúde oferecidos e, assim, orientar as melhores estratégias para melhoria.

As limitações inerentes aos diferentes métodos de identificação de EAs desqualifica métricas abrangentes de segurança do paciente, como as taxas gerais de EAs geradas individualmente por cada um deles. Medidas personalizadas para danos específicos ao paciente, que valorizam a compreensão da ocorrência de EAs por natureza, tem sido propostas para detectar o progresso ao longo do tempo e diferenças na segurança do cuidado oferecido entre as instituições (Shojania e Marang-van de Mheen, 2020). Nesse sentido, desde que sejam seguidas definições e classificações padronizadas em segurança do paciente, o *GTT* pode representar uma das fontes de dados para a medida de eventos específicos de interesse.