

POSITION PAPER

Os Desafios da Reavaliação de Tecnologias de Saúde em Portugal

Pedro A. Laires, MSc^{*1}; Mónica Inês, MSc²; Miguel Gouveia, PhD³; Céu Mateus, PhD⁴; Luis Silva

Miguel, PhD⁵; Carlos Gouveia Pinto, PhD⁵

(em representação do Capítulo Português do ISPOR)

¹ Merck Sharp & Dohme - Outcomes Research / Centro Académico de Medicina de Lisboa / Centro de Investigação em Saúde Pública

² Pfizer – Outcomes & Evidence / Centro Académico de Medicina de Lisboa

³ CEA/UCP – Católica Lisbon School of Business and Economics

⁴ Lancaster University - Division of Health Research/Health Economics Group

⁵ CISEP – Centro de Investigação Sobre Economia Portuguesa/ISEG – Lisbon School of Economics & Management, Universidade de Lisboa

*** Correspondência:**

Pedro A. Laires

Rua Miguel Lupi, 20

1200-725 Lisboa

Telef. [+351 213925964](tel:+351213925964)

e-mail: isporportugal@gmail.com

SUMÁRIO

Introdução

O recém-criado Sistema Nacional de Avaliação de Tecnologias de Saúde (SiNATS) em Portugal permitirá que as tecnologias de saúde sejam objeto de reavaliação num contexto de prática clínica real. Este artigo (*position paper*) pretende elencar e descrever sumariamente os diversos desafios e limitações que poderão influenciar os pressupostos e resultados das reavaliações e por conseguinte o seu próprio propósito. O artigo reflete a posição do Capítulo Português da ISPOR (*International Society for Pharmacoeconomics and Outcomes Research*) – ISPOR Portugal sobre o tema.

Métodos

Um grupo de membros da ISPOR Portugal analisou a literatura disponível, nomeadamente a documentação de grupos de trabalho internacionais que anteriormente se debruçaram sobre matérias semelhantes e realizou uma análise crítica sobre os desafios da reavaliação aplicada ao contexto nacional. Foi ainda solicitada uma revisão final a todos os membros do Capítulo.

Resultados

Foram identificadas e descritas diversas limitações que potencialmente podem comprometer as reavaliações, designadamente aquelas relacionadas com a própria intervenção (i.e. a tecnologia de saúde), a população em estudo, a seleção do comparador e os resultados em Saúde a considerar neste tipo de análises. É ainda enfatizada a relevância das fontes de informação em que poderá assentar (pelo menos parcialmente) a investigação da efetividade comparativa, bem como as lacunas e limitações inerentes à mesma (i.e. vieses e confundimentos).

Conclusões

Um sistema de reavaliação das tecnologias de saúde deve ser alvo de uma análise *a priori* sobre as suas potencialidades e limitações. Neste artigo abordam-se estes tópicos tomando como referência os objetivos do SiNATS. No entanto, torna-se necessário dar continuidade a este trabalho, nomeadamente através da criação de grupos de trabalho que se debrucem mais detalhadamente sobre as matérias aqui mencionadas. Tal trabalho poderá mesmo ser um passo crucial para o sucesso

pleno da implementação de um sistema de reavaliação que sustente decisões de financiamento justas e eficientes das tecnologias de saúde em Portugal.

Introdução

Pelo decreto-lei nº97/2015, de 1 de Junho, foi criado o Sistema Nacional de Avaliação de Tecnologias de Saúde (SiNATS), que abrange todas as entidades públicas e privadas na área da saúde com o objetivo de obter ganhos em saúde, harmonizado com outros sistemas europeus que procuram atingir o mesmo objetivo. O SiNATS permitirá que as tecnologias de saúde sejam objeto de avaliação e reavaliação num contexto de prática clínica corrente. Tendo em conta alguns dos seus objetivos, designadamente a monitorização da utilização e efetividade das tecnologias de saúde, torna-se assim crucial que se desenvolvam os mecanismos apropriados para a medição dos ganhos clínicos, económicos e de qualidade de vida incrementais, e se identifiquem os desafios que esses objetivos colocam ao Sistema de Saúde.

A reavaliação das tecnologias de saúde compreende a realização periódica de novos estudos de avaliação (farmacoterapêutica e económica) de tecnologias que têm os mesmos fins terapêuticos de forma a demonstrar que os recursos públicos a elas afetados são utilizados de forma eficiente. As Orientações Metodológicas para Estudos de Avaliação Económica de Medicamentos (OM-EAEM)¹ em vigor podem ser tomadas como referência para a avaliação e reavaliação da generalidade das tecnologias de saúde. O presente documento aborda os pontos mais relevantes da reavaliação focando particularmente aqueles em que as OM-EAEM são menos específicas, nomeadamente a realização de estudos de efetividade cujas conclusões possam ser integradas nos processos de reavaliação das tecnologias.

Para que a integração destes dados na reavaliação farmacoterapêutica e económica seja efetuada de modo credível, e sustentada em evidência científica, é necessário ter em atenção a complexidade de estimar o efeito do tratamento com base em estudos observacionais. O desafio é calcular o impacto da tecnologia nos ganhos em saúde em contexto de prática clínica corrente, ou seja, partir de um efeito observado, e ajustá-lo em termos comparativos para outras características que alteram os ganhos em Saúde. O objetivo do ajustamento é obter-se de forma consistente e precisa o verdadeiro efeito da tecnologia, pois é esse que se pretende analisar na reavaliação.

Assim, a previsibilidade/periodicidade da reavaliação de cada tecnologia de saúde é um ponto importante a considerar para garantir a qualidade do processo de reavaliação, principalmente no caso de tecnologias cuja penetração no mercado possa não ser suficiente ou que necessitem de um período alargado de tempo de observação da sua utilização para se detetar o efeito verdadeiro nos ganhos em saúde. Exemplos deste fenómeno são tecnologias cujo resultado primário é a prevenção (p. ex. vacinação, prevenção de eventos cardiovasculares), podendo necessitar de um horizonte temporal de avaliação mais alargado, ou, em alternativa, recorrer a marcadores intermédios ou a estimativas obtidas através de modelização (eventualmente recorrendo ao modelo utilizado na avaliação inicial).

A reavaliação ocorrerá quando a tecnologia de saúde já está a ser utilizada correntemente, num ambiente eventualmente ocupado por múltiplos comparadores, num mercado sujeito a diversos factores passíveis de modificar a sua adoção (por exemplo, recomendações clínicas e opinião de peritos). Contudo, não tem sido comum a quantificação da efetividade relativa (comparada) nem a sua reavaliação terapêutica e/ou económica. Um ponto adicional é que o processo de reavaliação permite detetar ou confirmar a obsolescência tecnológica, principalmente se forem efetuadas comparações entre várias tecnologias.

Desafios das Reavaliações

Eficácia vs. Efetividade

É comum afirmar-se que a eficácia das tecnologias de saúde, em particular dos medicamentos, é superior à sua efetividade (*efficacy-effectiveness gap*)². Isto porque no contexto de ensaios clínicos, cujo propósito é essencialmente regulamentar e de aferição o mais precisa possível da relação benefício-risco, a variabilidade da resposta terapêutica é deliberadamente reduzida através da predefinição estreita das condições de tratamento (ambiente da prestação dos cuidados) e de critérios de elegibilidade restritos. A preocupação com a validade interna na determinação da eficácia sobrepõe-se, portanto, ao interesse em generalizar os resultados para uma população dita de “mundo real”. Posteriormente, a evidência torna-se de facto insuficiente, nomeadamente num contexto de financiamento de tecnologias de saúde durante a sua comercialização. Os cenários e a população alvo descrita no resumo das características do medicamento (RCM) são nesta fase de pós-comercialização mais abrangentes e a variabilidade da resposta terapêutica aumenta, geralmente afetando de forma negativa a efetividade. Agrava este cenário a possível concomitância de outras condições clínicas (e.g. comorbilidades), que podem influenciar os resultados, e a falta de adesão e persistência no tratamento, algo que também pode suceder, embora em menor dimensão, nos ensaios clínicos. Verifica-se assim, por um lado, uma deterioração nos benefícios da utilização do produto (menor resposta terapêutica) e, por outro, um aumento da susceptibilidade à ocorrência de efeitos adversos, cuja gravidade pode inclusivamente não ser acompanhada por uma assistência que se equipare àquela observada em contexto de investigação clínica, piorando em última instância os resultados em saúde no seu todo. No entanto, em situações menos frequentes, o oposto também poderá ocorrer (i.e. efetividade superior à eficácia).

Apesar deste desfasamento entre eficácia e efetividade, as decisões de financiamento são actualmente tomadas considerando apenas dados de eficácia que tipicamente permitem estabelecer um risco relativo de evolução da doença *versus* determinado comparador, não necessariamente o mais adequado em sede de avaliação de tecnologias de saúde (ATS ou HTA - *Health Technology Assessment* - tal como designado na literatura anglo-saxónica). E se é verdade que a utilização de metodologias como a análise segundo a intenção de tratar, a modelização de efetividade a longo prazo e os ensaios clínicos pragmáticos,³ bem como algumas iniciativas transnacionais de considerável dimensão (e.g. *GetReal Consortium* da *Innovative Medicines Initiative*)⁴, têm permitido explorar vias precoces no desenvolvimento dos medicamentos visando colmatar este desfasamento entre a eficácia e a efetividade, também é verdade que algumas destas metodologias podem não atender plenamente às necessidades num contexto de ATS. Para além do mais, algumas questões éticas têm sido identificadas.⁵ Por exemplo, a possibilidade de se obterem piores resultados clínicos na sequência da aleatorização que ocorre nos ensaios clínicos pragmáticos comparativamente ao que se poderia obter na sequência de uma decisão médica informada, consciente e melhor ajustada a cada doente (i.e. a afetação aleatória de uma determinada tecnologia de saúde pode constituir um maior risco *versus* a prática médica corrente).^{6,7}

Adicionalmente, é necessário ter padrões adequados de comparação do desempenho de uma dada tecnologia no mercado. A comparação da efetividade deve ser realizada com a alternativa de referência encontrada na prática clínica corrente (i.e. a estratégia terapêutica mais comum) a qual, naturalmente, muda ao longo do tempo. Poderá também ser importante realizar estudos comparativos com outras opções terapêuticas relevantes que tenham expressão no mercado. Em qualquer caso, é fundamental que exista acordo prévio entre os envolvidos relativamente à seleção de indicadores válidos para a medição da efetividade.

Assim, a medição de efetividade comparativa numa fase de comercialização torna-se crucial e deve ser simultaneamente planeada e flexível para, designadamente, permitir um financiamento adequado e fundamentado no contexto de um mercado dinâmico em que, designadamente, aparecem novas opções terapêuticas; normas e recomendações clínicas que se actualizam periodicamente; perfis epidemiológicos que não são forçosamente constantes; e quotas de mercado voláteis.

Efetividade Comparativa

A investigação da efetividade comparativa tem como objetivo comparar a efetividade e a segurança de alternativas para a prevenção, diagnóstico e tratamento das patologias⁸. É considerada fundamental para a tomada de decisões em saúde, designadamente no que ao financiamento dos medicamentos diz respeito⁹. Tem progredido de forma notória e sido alvo de um investimento considerável por parte dos decisores políticos, como o atesta o recente estímulo avultado concedido nos Estados Unidos da América à investigação nesta área, como se pode constatar no *Patient Protection and Affordable Care Act*,¹⁰ e no trabalho de organizações com ela relacionadas, designadamente a *AHRQ - Agency for Healthcare Research and Quality*¹¹ e o *PCORI - Patient-Centered Outcomes Research Institute*.¹²

Concretamente, o objetivo último da avaliação da efetividade comparativa é medir com precisão os ganhos clínicos incrementais de uma determinada tecnologia de saúde face a outra(s) considerada(s) relevante(s) para a indicação terapêutica em questão. Porém, antecipam-se diversos desafios, nomeadamente de ordem metodológica, que devem ser devidamente identificados de forma a mitigar o risco de uma reavaliação pós-comercialização indevida ou incorreta. Seguidamente apresentam-se alguns desses desafios.

Técnicas de Medição da Efetividade

Apesar de não ser consensual, os estudos observacionais bem desenhados podem ter uma validade que se aproxima dos ensaios clínicos aleatorizados^{13,14}. Para que tal seja possível é fundamental que alguns princípios metodológicos sejam reconhecidos e aplicados escrupulosamente, em particular o controlo dos vieses, o ajustamento aos factores de confundimento e a utilização de métodos estatísticos apropriados (e.g. os *propensity scores*, as variáveis instrumentais e a análise multivariável). Existem já diversas recomendações emanadas de organizações internacionais no sentido de alcançar patamares comuns de qualidade na investigação epidemiológica de natureza observacional. Destaca-se o STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*)¹⁵ e os grupos de trabalho específicos da ISPOR (*International Society for Pharmacoeconomics and Outcomes Research*).^{16,17,18}

Desenho do Estudo

Apesar de existirem muitos tipos de estudos epidemiológicos e econométricos, a escolha do desenho do estudo deve ser guiada pela pergunta científica e pela sua exequibilidade. É essencial que a concetualização e a implementação do estudo não sejam toldadas por questões de natureza pragmática ou política que comprometam a sua validade e fiabilidade. Alguns grupos de trabalho da ISPOR dedicam-se à análise dos desenhos dos estudos e respectivas virtudes e fragilidades de acordo com as necessidades da ATS.^{16,17,18}

A População em Estudo

Na avaliação da efetividade o ponto de partida é a questão científica, que deve geralmente emanar de 4 vectores principais, usualmente referidos pela sigla PICO – População, Intervenção, Comparador e Resultados (*Outcomes*).

A definição clara e rigorosa da população em estudo, não colocando em causa a representatividade da amostra (validade externa), é crucial como ponto de partida de qualquer estudo de efetividade comparativa. Associada a esta definição surge a caracterização do ambiente da investigação. É necessário identificar se o estudo deve ser implementado ao nível da comunidade, no ambulatório ou em ambiente hospitalar. Decidir também se são visadas as unidades de saúde do sector privado, para além do público. Os critérios de elegibilidade devem ser amplos e inclusivos, mas sem abdicar do rigor na definição da população em estudo bem como na diferenciação entre grupos de comparação, tendo em conta o viés de classificação.

A população em estudo pode reflectir a indicação terapêutica estrita, mas também pode obedecer a eventuais regras de financiamento mais específicas. Não faz sentido recolher dados de efetividade numa população que, apesar de estar abrangida pela indicação terapêutica do medicamento, não está em condições de beneficiar do financiamento do medicamento em causa pelo SNS. Pode ocorrer o caso de determinados doentes serem tratados dentro das indicações terapêuticas aprovadas, mas não serem recrutados para o estudo pós-comercialização por não obedecerem aos critérios de financiamento, assumindo que a investigação se situa no âmbito da reavaliação na fase de comercialização como aqui se pretende discutir.

Adicionalmente, o advento dos registos e dos modelos de partilha de risco baseados no desempenho (*performance-based agreements*) pode influenciar a seleção da população em estudo e consequentemente a análise da efetividade, na medida em que poderá haver uma maior pressão sobre a decisão médica (e.g. efeito Hawthorne que diz respeito à tendência dos indivíduos mudarem

o seu comportamento porque são alvo de atenção especial). Antecipando o novo modelo de ATS (i.e. SiNATS) pode assim haver o risco da reavaliação influenciar a prática clínica simplesmente porque está a ser observada/monitorizada.

Casos Prevalentes e Incidentes

Caso a população em estudo inclua tanto novos casos (incidentes) como prevalentes, isso pode gerar algumas questões metodológicas que devem ser discutidas. É esperado, por exemplo, que utilizadores prevalentes de um determinado medicamento apresentem uma adesão e persistência distintas daquelas obtidas em doentes *naïves*. A adesão e a persistência influenciam, por um lado, a efetividade e, por outro lado, resultam de inúmeros factores (e.g. maior nível educacional, melhor resposta à terapêutica, maior preocupação com a saúde e menos polimedicação e comorbilidades) que podem em si mesmo alterar os resultados de efetividade. Por isso a escolha da proporção de incidentes e prevalentes na população em estudo é importante tanto para a intervenção como para a comparação. Por exemplo, caso se considerem novos casos para a intervenção em estudo e casos prevalentes para a alternativa de comparação, esta última poderá ser favorecida pelo simples facto de incluir doentes mais persistentes e respondedores à terapêutica e também mais propensos a conseguir melhores resultados em saúde (situação semelhante ao viés de sobrevivência que afeta os estudos epidemiológicos de doença crónica). Existem métodos que visam especificamente esta questão (e.g. *incident user design*).^{19,20}

A Intervenção

Como já foi mencionado, a intervenção em estudo (e.g. novo medicamento participado) deve respeitar, não só os seus termos de utilização com base no RCM como também as regras de financiamento acordadas entre o titular de Autorização de Introdução no Mercado (AIM) e o

INFARMED. A investigação deve ter em conta o faseamento do acesso da população à tecnologia (*acesso e penetração no mercado*), que, por sua vez, depende de factores, como lacunas terapêuticas, recomendações clínicas, políticas comerciais, estratégias de *marketing* e a própria procura e expectativa dos doentes (que por sua vez pode ser influenciada por acções de informação e consciencialização sobre a doença e respectivos tratamentos e ainda pela própria partilha de experiências e opiniões, algo cada vez mais frequente e facilitado pelas redes sociais). Uma dada intervenção com rápido e amplo acesso no mercado terá certamente um efeito distinto daquela que tem menor acesso. Tal não se deverá ao aumento *per se* da dimensão da amostra, mas antes à ampliação da variabilidade da resposta terapêutica, tal como atrás discutido.

O modo de utilização (posologia), a duração do tratamento, a persistência e a adesão à terapêutica, bem como a possível interação com outros fármacos devem ser factores a ter em consideração na definição e avaliação da intervenção.

Comparadores

Tal como atrás referido, é necessário seleccionar para a investigação da efetividade comparativa a alternativa de referência na prática clínica corrente. Esta seleção poderá não ser linear e pode mesmo ser necessário incluir mais do que um comparador, mas principalmente poderá a comparação correta diferir daquela realizada nos ensaios clínicos. Apesar de tal poder também suceder durante a avaliação inicial, num contexto de reavaliação poderá conduzir ao confronto entre dados de efetividade obtidos na fase pós-comercialização para um comparador e dados de eficácia obtidos previamente com outro comparador (que até pode ser o placebo). A subsequente revisão dos moldes de financiamento pode assim tornar-se mais complexa.

Importa discutir se é sempre adequada a comparação com a alternativa terapêutica que correspondia à prática clínica mais relevante no momento da avaliação inicial. Em alguns casos, esta

alternativa poderá ter-se tornado irrelevante. Caso haja um comparador distinto daquele previamente utilizado e se pretenda actualizar os rácios incrementais a partir dos novos dados de efetividade, então ter-se-á que modelizar com esse novo comparador.

A inclusão dos genéricos nesta como noutras comparações pode ocorrer, mas diferenciar-se-á pelo preço. A diminuição do preço do comparador pode fazer com que uma tecnologia sob reavaliação deixe de ser custo-efetiva. Tal poderá fazer com que os doentes que iniciem o tratamento depois da avaliação inicial tenham acesso a um medicamento mais efetivo que aqueles que necessitem de tratamento num período após a reavaliação. Em termos de avaliação económica tal poderá estar correto, mas é preciso que tais situações sejam claramente identificadas e que as decisões de acesso sejam aceites pela sociedade.

Por fim, importa realçar o efeito de *crossover* (i.e. doentes que deixam a intervenção que lhes foi atribuída e passam a ser tratados com a outra intervenção em estudo), cujo impacto na modelização deve ser considerado e devidamente corrigido na modelização. No limite, a correção inapropriada de *crossover* pode causar decisões de preço e financiamento desajustadas face ao real benefício de um medicamento relativamente aos seus comparadores.²¹

Resultados em Saúde

A avaliação dos resultados em saúde deve ser compreensiva, incluindo-se não apenas os indicadores clínicos, como também económicos e aqueles relacionados com a qualidade de vida. Deve também ter-se em consideração a estreita combinação e interligação entre estes diferentes tipos de resultados (*ECHO model: Economic, Clinical and Humanistic Outcomes*),^{22,23,24} ao longo do tempo. Na ATS deve ser feita uma distinção clara entre resultados finais (i.e. *hard clinical endpoints*) e intermédios, como sejam os marcadores substitutos (*surrogate markers*). Esta distinção é crucial dada a sua implicação no desenho e no horizonte temporal do estudo. A distinção tem também

implicações no ambiente da investigação. Por exemplo, caso se pretenda obter resultados finais, como, por exemplo, o impacto sobre os internamentos, para um determinado medicamento utilizado em ambulatório onde os marcadores intermédios da efetividade são recolhidos, então já será necessário envolver na investigação várias dimensões do sistema de saúde. A duração do estudo deve ter em conta não só o tipo de resultados que se pretende, como a sua exequibilidade e a finalidade da investigação. Num contexto de reavaliação contínua de medicamentos e de um mercado dinâmico, com entradas de alternativas terapêuticas e novas práticas clínicas, poderá ser impraticável esperar pela determinação dos resultados finais que podem demorar uma década ou mais a apurar. O compromisso deve ser discutido proactivamente pelos proponentes e os diretamente envolvidos na investigação.

Em todo o caso, dado que a maioria dos estudos de avaliação económica de novas tecnologias na doença crónica utiliza horizontes temporais de fim de vida, haverá certamente um desfasamento de informação entre os ganhos estimados de longo prazo obtidos por modelização e os possíveis de obter em situações de prática clínica corrente, num horizonte temporal inferior. Assim, será difícil comparar tanto os ganhos incrementais de saúde como os rácios incrementais de custo-efetividade entre a avaliação inicial e a efetuada pós-comercialização sem recorrer novamente à modelização, o que em certo sentido desvia da pretensão em reavaliar as tecnologias de saúde com base em evidência estritamente observacional. Contudo, poderá não restar outra opção senão obter os valores para marcadores intermédios através de estudos observacionais e calcular os resultados de custo-efetividade para horizontes temporais de fim de vida com base em modelos.

Os resultados finais que se pretendem obter também estão na base da definição da frequência e dos *timings* de recolha de informação, bem como o cálculo de dimensão da amostra com as devidas implicações em termos de plano de recrutamento e participação de doentes, de centros e de investigadores.

Se possível, a perspectiva a adotar deve ser a da sociedade, tal como previsto nas OMEAEM. Preconizam-se que os custos indiretos relacionados com a perda de produtividade do trabalhador sejam ainda relatados mas separadamente. Ainda em relação aos custos, é necessário considerar o paradoxo da sobrevivência, em que ganhos incrementais de sobrevivência podem traduzir-se em maior despesa para a sociedade ou para o SNS.

Os resultados sobre a qualidade vida relacionada com a saúde devem ser considerados na investigação de efetividade comparativa, nomeadamente pela introdução de ferramentas de aferição apropriadas nos protocolos de investigação (e.g. questionários de *patient-reported outcomes* - PRO). Relativamente aos efeitos adversos dos novos medicamentos (que também podem ser avaliados no momento das revisões contratuais) convém salientar a menor tendência para registar efeitos adversos já conhecidos e associados a alternativas terapêuticas já utilizadas na prática clínica.²⁵

Consistência dos Resultados

Os resultados obtidos em qualquer investigação devem ser sempre confrontados com a literatura disponível. Em particular, devem ser encontradas e debatidas as razões de eventuais discrepâncias face a resultados de investigações semelhantes.

A avaliação da consistência dos resultados pode incidir sobre a direcção do efeito (e.g. superior, inferior ou similar efetividade) ou sobre a sua magnitude (e.g. grau de superioridade ou inferioridade).

A consistência dos resultados pode efetivamente fortalecer as conclusões e sedimentar a evidência que apoia a decisão política, no entanto é importante reconhecer a presença de viés de publicação na literatura disponível e garantir que esse potencial enviesamento não “contamine” a qualidade da evidência disponível, bem como as decisões que daí decorram. Este viés consiste no facto dos autores de textos científicos serem mais propensos a apresentar (e os editores de revistas mais interessados

em publicar) estudos com dados "positivos" nomeadamente aqueles em que os resultados rejeitam a não superioridade da efetividade de um novo medicamento. Assim, é importante considerar a consistência dos resultados, porém sem esquecer que a ciência publicada pode ser uma fracção do conhecimento passível de ter sofrido o filtro deste viés de publicação. Sobretudo é importante ter uma abordagem crítica, integradora e o mais imparcial possível. É isso que se deseja para sustentar uma decisão baseada na evidência.

Normas e Recomendações Clínicas

Um dos efeitos da entrada de um novo medicamento na prática clínica é a actualização das recomendações clínicas. Esta alteração poderá extravasar a simples introdução desse medicamento no algoritmo terapêutico. Por exemplo, pode levar à criação de mais linhas de tratamento; pode modificá-las (e.g. um medicamento já existente que estava em 1ª linha passar para linhas posteriores); pode “alavancar o protagonismo” de outros fármacos já existentes; e pode ainda levar à adoção de novos critérios de seleção dos doentes (e.g. inclusão de outras características na escolha terapêutica, como sejam as comorbilidades).

A introdução de um novo medicamento no mercado pode também gerar maior consciencialização para o cumprimento das orientações clínicas pelos médicos e aumentar as expectativas dos doentes.

Em bom rigor, o impacto da entrada de uma nova tecnologia no mercado não se cinge às características das próprias tecnologias, mas antes a toda a envolvente do Sistema de Saúde, a qual está em permanente mudança. Todos estes efeitos, nomeadamente o impacto nas recomendações e prática clínicas são realmente difíceis de confinar a um protocolo de investigação de efetividade.

Efeitos Colaterais (*spillover effects*)

A introdução de uma nova tecnologia pode alterar de tal forma o padrão da prestação de cuidados que o efeito da sua utilização ultrapassa a efetividade terapêutica direta. Por exemplo, a introdução de um novo medicamento pode gerar uma série de efeitos positivos colaterais, como sejam o aumento na investigação na área terapêutica respectiva, um maior interesse generalizado e efetiva mudança na atitude e prática clínica (por exemplo, testes necessários para a decisão de iniciar ou continuar um determinado medicamento podem em si mesmo auxiliar no controlo da doença *per se*). O próprio marketing farmacêutico pode potenciar a consciencialização da doença, a sua identificação e o seu tratamento precoces.

Adicionalmente, um novo medicamento pode ter efeitos benéficos noutras áreas terapêuticas que por vezes são dificilmente mensuráveis (*cross-disease spillovers*). Por exemplo, um psicofármaco poderá ter efeitos positivos na mortalidade e hospitalizações por causas cardiometabólicas simplesmente por terem permitido aos doentes melhores cuidados em saúde. Um exemplo de efeito colateral mais negativo pode ocorrer com a introdução de uma tecnologia muito efetiva que poderá conduzir a uma decisão de desvio dos recursos, em que um grupo de doentes beneficia, mas que outro poderá ser prejudicado. Por fim, há um outro efeito colateral associado ao ganho de sobrevivência e riscos concorrentes. Por exemplo, um dado medicamento pode reduzir a mortalidade cardiovascular enquanto aumenta indiretamente a mortalidade oncológica, unicamente por expor os indivíduos a outras doenças associadas ao envelhecimento, como é o caso das neoplasias.

Todos estes potenciais efeitos colaterais de um novo medicamento são praticamente impossíveis de quantificar, mas não devem ser ignorados quando se faz a reavaliação de uma tecnologia dado o seu impacto na quantidade e na qualidade da vida da população.

Fontes de Informação – validade; utilização de dados individuais; transparência; e disponibilidade

Uma das plataformas mais importantes para desenvolver estudos de efetividade comparativa são os registos de doentes. Um registo é um conjunto ordenado de documentos/informações sobre os doentes que partilham algumas características ao longo do tempo e deve recolher informação sobre os resultados bem (pré-)definidos com interesse para a análise (*ISPOR Taxonomy of Patients Registries*).²⁶

Os registos podem ser focados na doença (e.g. registo oncológico) ou em determinadas tecnologias de saúde (e.g. medicamentos biológicos). Podem seguir coortes de casos prevalentes ou de casos incidentes. Os registos permitem realizar comparações entre alternativas terapêuticas de uma forma sistemática e uniformizada porque promovem a padronização do recrutamento e da avaliação. São muito importantes para a condução dos estudos de avaliação das tecnologias. Porém, apresentam alguns riscos que devem ser acautelados (ver adiante), nomeadamente a possibilidade da amostra não ser representativa da população alvo (validade externa), a existência de lacunas nos dados e vieses de seleção e informação (validade interna). A participação assimétrica de centros tanto no recrutamento de doentes como no registo da informação pode distorcer os dados constantes do registo. Por exemplo, se um centro de investigação é mais activo na inclusão de dados em registos que a média nacional, então a sua prática clínica e respectivos resultados em saúde estarão “sobre-representados”. Esta questão será tanto mais grave quanto menor for o número de centros. Com a avaliação da efetividade na fase de comercialização pretende-se obter uma visão abrangente da realidade nacional e do desempenho das tecnologias de saúde nesse contexto. É por isso que os registos de informação devem evoluir não só na sua validade interna como na sua abrangência. A participação dos médicos é considerada o factor mais limitativo do sucesso de um registo sendo geralmente necessário o desenho de incentivos apropriados para o ultrapassar.

Os registos podem também ser utilizados para alicerçar programas de partilha de risco, mas é importante que as limitações mencionadas não sejam superiores aos potenciais benefícios da sua utilização.

Deve ser respeitado o princípio de que os métodos devem estar ao serviço das questões científicas que se pretendem estudar e não o inverso. No entanto, a falta de informação poderá contrariar este princípio. Por exemplo, se um registo não estiver preparado para acolher informação sobre um determinado indicador, necessário à avaliação de efetividade comparativa, isto poderá influenciar a decisão sobre a questão científica a ser colocada. Pode até suceder que áreas terapêuticas em que os registos estejam mais desenvolvidos sejam aquelas onde primeiro irá incidir a avaliação da efetividade comparativa. É importante que não seja a disponibilização dos meios para realizar a reavaliação (neste caso, os registos) que prevaleça sobre outros factores (e.g. os impactos epidemiológico e económico) no que diz respeito à definição estratégica de ATS em Portugal.

É também importante que a investigação não seja ela própria um entrave à prescrição e consequentemente no acesso à inovação. Por exemplo, se a prescrição de determinado medicamento estiver condicionada à inclusão de dados num registo isso poderá constituir um entrave à decisão médica pelo simples facto da prescrição necessitar de um maior esforço para a sua concretização. Este fenómeno poderá causar uma distorção na utilização do medicamento (por exemplo, reservar a sua utilização somente para os casos mais graves ou sem alternativas) e subsequentemente na própria investigação de efetividade.

Por fim, é fundamental explicitar regras sobre acesso e disponibilização dos dados obtidos nos registos, tendo como objetivo último maximizar a transparência dos processos de reavaliação.

Naturalmente, a proteção dos dados individuais terá que ser sempre salvaguardada.^{1*}

* Neste artigo não foram abordados especificamente os estudos de eficácia e segurança pós-autorização (PAES – Post Authorization Efficacy Study e PASS - Post-Authorisation Safety Study)²⁷ que apresentam um enquadramento regulamentar próprio e distinto da reavaliação para financiamento das tecnologias de saúde.

Limitações dos Estudos de Reavaliação

Confundimento

O confundimento é talvez a maior limitação dos estudos observacionais. Ocorre quando um factor está associado à exposição (e.g. tratamento) e simultaneamente ao resultado, distorcendo desse modo a relação de causalidade se não for controlado. A aleatorização dos ensaios clínicos teoricamente permite contornar este problema através do equilíbrio “forçado” dos possíveis factores confundidores nos diferentes braços de tratamento. Porém, nos estudos observacionais tal não é possível e, portanto, é essencial que se identifiquem de antemão estes potenciais factores, bem como se apliquem as ferramentas disponíveis para os controlar o mais possível (e.g. estratificação, emparelhamento, análise estatística multivariável).¹⁶ É importante ter em consideração que na presença de variáveis dependentes do tempo ou inter-temporais os métodos padrão de controlo de confundimento poderão ser desadequados. Esta questão do confundimento inter-temporal é particularmente importante no âmbito das análises de efetividade comparativa (modelos proporcionais de Cox e outros métodos²⁷ poderão auxiliar a investigação a lidar com esta limitação). Contudo, já existem muitos instrumentos para controlar estas dificuldades técnicas e que permitem a melhoria significativa da inferência causal, alguns deles bem estabelecidos e já referidos anteriormente. Existe um documento elaborado por uma *task force* da ISPOR em que são apresentados os meios para lidar com o confundimento.²⁸

Confundimento pela Indicação / Severidade

O confundimento pela indicação está associado a uma boa prática clínica²⁹ na medida em que resulta da aplicação correta dos instrumentos terapêuticos disponíveis. No entanto “contamina” a investigação da efetividade e segurança comparativas na medida em que pode induzir, tal como

sucedeu em diversos exemplos do passado, à falsa conclusão de que um determinado fármaco, ou qualquer outra tecnologia em saúde, está associado a piores resultados. Deste modo, um medicamento que tenha valor terapêutico acrescentado face às alternativas disponíveis pode ser utilizado com maior frequência em doentes mais graves e vulneráveis, que, por inerência da sua condição clínica, têm pior prognóstico, mesmo se melhor tratados, induzindo assim uma subestimação nos resultados. É necessário ter em consideração este tipo de confundimento no delineamento da investigação e utilizar o melhor desenho do estudo e a metodologia mais adequada (e.g. *propensity scores*³⁰).

Confundimento Residual

Diversos confundidores podem não ser identificados e/ou os seus efeitos não serem quantificados, quer porque o planeamento e desenho do estudo não os contemplaram, quer porque as fontes de informação disponíveis (por exemplo, estudo retrospectivo com utilização de registos, que recolhem um número limitado de variáveis excluindo outras que poderão actuar como confundidoras em determinadas associações estudadas) não possibilitam a recolha e o tratamento dessa informação. Este “confundimento escondido ou residual” pode evidentemente causar uma distorção substancial nas ilações extraídas pela investigação e como tal deverá ser alvo de reflexão e, se possível, correção (e.g. *case-crossover design*).¹⁸

Vieses

Existem inúmeras fontes de erros sistemáticos nos estudos observacionais e, em particular, nos que têm como objetivo a determinação da efetividade comparativa.³¹ Destacam-se a seguir alguns tipos de vieses.

Viés de Seleção

Este erro sistemático ocorre quando alguns doentes podem ter maior, ou menor, probabilidade de serem selecionados para um determinado grupo em estudo. O viés de seleção é um dos vieses mais relevantes e frequentes em epidemiologia e pode surgir por variados motivos, como por exemplo a auto-seleção dos doentes (viés do voluntário),^{32,33} o viés de Berkson,³⁴ o viés do trabalhador saudável,³⁵ e o viés de incidência-prevalência.^{32,37}

Para além de afetar a validade interna, o viés de seleção é um desafio para a própria validade externa. Por exemplo, o viés do voluntário pode levar ao predomínio de participantes de estratos educacionais e socioeconómicos diferentes na amostra do que o verificado na população devido à percepção da importância da investigação³⁶ causando um problema de representatividade da amostra.

Viés de Classificação

Outra fonte de imprecisão nos estudos epidemiológicos é o viés de classificação, que tanto pode incidir sobre a exposição (por exemplo, a identificação de exposição a um dado medicamento quando esta não está a ocorrer e vice-versa) e/ou sobre o resultado (por exemplo, classificação incorreta ou heterógena do resultado clínico em estudo). A utilização de critérios pré-definidos e padronizados na definição de caso nos protocolos de investigação é essencial para evitar a má classificação de exposição/resultados. Neste sentido, a “contaminação” entre os grupos de comparação também deve ser evidentemente limitada. Esta questão torna-se particularmente susceptível com desenhos retrospectivos pelo que a utilização de registos de doença para eventual medição de efetividade deve

ter em linha de conta esta limitação, bem como apresentar medidas que a contornem (e.g. comissão clínica independente para confirmação de casos).

Viés de Atrito

Este viés aplica-se aos estudos de coorte e é uma das suas principais fragilidades. A perda de participantes para seguimento ou seguimento incompleto pode causar diferenças relevantes entre os grupos em estudo. Medidas de minimização das perdas durante o seguimento (por exemplo, acompanhamento, contactos periódicos e intervenções de motivação à participação na investigação) devem ser consideradas. As técnicas de imputação de dados omissos também devem ser tidas em consideração, se tal se justificar.

Outros Vieses

Muitos outros vieses podem ocorrer^{31,37} e devem ser, dentro do possível, identificados, controlados e corrigidos. Os protocolos de investigação de efetividade comparativa devem ter uma secção onde esta questão possa ser abordada detalhadamente (avaliação do risco de enviesamento),³⁸ bem como a eventual adoção de métodos de correção de vieses, tais como *propensity scores*, emparelhamento, correção de persistência e variáveis instrumentais.³⁹

Conclusões e Perspetivas de Trabalho Futuro

Este artigo teve como propósito principal elencar os principais desafios que se põem ao desenvolvimento de estudos de reavaliação das tecnologias de saúde em Portugal no âmbito do SiNATS. Primeiramente foram identificados problemas de natureza mais conceptual que podem e

devem ser premunidos e, se possível, resolvidos de início como, por exemplo, a escolha do(s) comparador(es), o horizonte temporal e os resultados em saúde a incluir na análise. Seguidamente, foram identificadas potenciais limitações metodológicas, nomeadamente aquelas que comprometem a validade dos estudos observacionais (que estão na base da reavaliação) e que devem ser devidamente analisadas e tratadas de modo a que os dados em que os estudos se baseiam tenham a qualidade desejável. Nesta análise pressupôs-se que o objetivo fundamental dos estudos é garantir que o financiamento das tecnologias de saúde em Portugal se traduza em ganhos em saúde pelo menos semelhantes aos identificados na avaliação inicial.

No futuro dever-se-ão envidar todos os esforços no sentido de ir mais além agregando consensos sobre este tema para que se consolidem recomendações objetivas que sejam claras, transparentes e, sem comprometer o rigor, pragmáticas e flexíveis, superando os mencionados desafios e limitações. Estas recomendações poderão ser produzidas por grupos de trabalho especializados, à semelhança do que tem sido feito pela ISPOR.

A dimensão deste tema bem como a abrangência dos métodos e das possíveis soluções são demasiado vastas para que possam ser tratadas por um único grupo de trabalho. Este artigo pretendeu contribuir também para alertar para esta necessidade. Um sistema dimensionado na devida proporção dos desafios a que pretende responder possibilitará certamente a tomada de decisões mais justas e eficientes.

Agradecimentos

Os autores agradecem as contribuições de Isabel Monteiro, José Aranda da Silva, Nuno Silvério e Vanessa Jacinto, membros da ISPOR Portugal.

Referências

- 1 INFARMED. Orientações Metodológicas para Estudos de Avaliação Económica de Medicamentos. Nov 1998. Despacho nº 19064/99; 9 Sep 1999.
- 2 Eichler HG, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, Rowland M, Schneider CK, Bloechl-Daum B. Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov*. 2011 Jul 1;10(7):495-506.
- 3 Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol*. 2009 May;62(5):499-505.
- 4 <http://www.imi-getreal.eu>
- 5 Kalkman S, van Thiel GJ, Grobbee DE, van Delden JJ. Pragmatic randomized trials in drug development pose new ethical questions: a systematic review. *Drug Discov Today*. 2015 Jul;20(7):856-62.
- 6 Appelbaum PS, Roth LH, Lidz CW, Benson P, Winslade W. False hopes and best data: consent to research and the therapeutic misconception. *Hastings Cent Rep*. 1987 Apr;17(2):20-4.
- 7 Taylor KM, Margolese RG, Soskolne CL. Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer. *N. Engl. J. Med*. 1984 310, 1363–1367
- 8 Patient Centered Outcomes Research Institute (PCORI). Developing PCORI's National Priorities for Research and First Research Agenda. Available at: <http://www.pcori.org/assets/Program-Development-Committee-Report.pdf>
- 9 Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003 Sep 24;290(12):1624-32.
- 10 Patient Protection and Affordable Care Act, 42 U.S.C. § 18001 et seq. (2010).
- 11 <http://www.ahrq.gov/>
- 12 <http://www.pcori.org/>
- 13 Horwitz RI, Viscoli CM, Clemens JD, Sadock RT. Developing improved observational methods for evaluating therapeutic effectiveness. *Am J Med*. 1990 Nov;89(5):630-8.
- 14 Concato J, Lawler EV, Lew RA, Gaziano JM, Aslan M, Huang GD. Observational methods in comparative effectiveness research. *Am J Med*. 2010 Dec;123(12 Suppl 1):e16-23.
- 15 Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology*. 2007 Nov;18(6):805-35.
- 16 www.ispor.org/
- 17 Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report--Part III. *Value Health*. 2009 Nov-Dec;12(8):1062-73.
- 18 Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report--Part II. *Value Health*. 2009 Nov-Dec;12(8):1053-61.

-
- 19 Rothman KJ. *Epidemiology. An Introduction*. New York: Oxford University Press, 2002
- 20 Eric S. Johnson, Barbara A. Bartman, Becky A. Briesacher, Neil S. Fleming, Tobias Gerhard, Cynthia J. Kornegay, Parivash Nourjah, Brian Sauer, Glen T. Schumock, Art Sedrakyan, Til Stürmer, Suzanne L. West and Sebastian Schneeweiss. The incident user design in comparative effectiveness research. *Pharmacoepidemiology and drug safety* 2013; 22: 1–6
- 21 Jönsson L, Sandin R, Ekman M, Ramsberg J, Charbonneau C, Huang X, Jönsson B, Weinstein MC, Drummond M. Analyzing overall survival in randomized controlled trials with crossover and implications for economic evaluation. *Value Health*. 2014 Sep;17(6):707-13.
- 22 CM Kozma, CE Reeder, RM Schulz. Economic, clinical, and humanistic outcomes: A planning model for pharmaco-economic research. *Clin Ther*. 1993;15:1121–32.
- 23 Gunter MJ. The role of the ECHO model in outcomes research and clinical practice improvement. *Am J Manag Care*. 1999 Apr;5(4 Suppl):S217-24
- 24 Reeder CE. Overview of pharmaco-economics and pharmaceutical outcomes evaluations. *Am J Health Syst Pharm*. 1995 Oct 1;52(19 Suppl 4):S5-8.
- 25 Hazell L, Shakir SA. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf*. 2006;29(5):385-96.
- 26 Polygenis D, ed. *ISPOR Taxonomy of Patient Registries: Classification, Characteristics and Terms*. Lawrenceville, NJ; 2013.
- 27 Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med*. 2013 Apr 30;32(9):1584-618.
- 28 <http://www.ispor.org/TaskForces/documents/RDPartIII.pdf>
- 29 Walker AM. Confounding by indication. *Epidemiology*. 1996;7:335–336
- 30 Rosenbaum P, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
- 31 Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health* 2004;58:635–641.
- 32 Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979 32(1–2): 51–63.
- 33 Martinson BC, Crain AL, Sherwood NE, Hayes MG, Pronk NP, O'Connor PJ. Population reach and recruitment bias in a maintenance RCT in physically active older adults. *J Phys Act Health*. 2010 Jan;7(1):127-35.
- 34 Westreich D. Berkson's bias, selection bias, and missing data *Epidemiology*. 2012 Jan; 23(1): 159–164.
- 35 McMichael AJ, Spirtas R, Kupper LL. An epidemiologic study of mortality within a cohort of rubber workers, 1964-72. *J Occup Med* 1974; 16: 458-464
- 36 Robert Rosenthal and Ralph L. Rosnow (1975), *The Volunteer Subject*, which reviewed and summarized the copious research up until that time. John Wiley & Sons, New York.
- 37 Suissa S. Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol*. 2008 Aug 1;168(3):329-35.
- 38 Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters M, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, Treadwell JR. *Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care*

Interventions. Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008-2012 Mar 08. Available at: www.effectivehealthcare.ahrq.gov/

39 Gillespie IA, Floege J, Gioni I, Drüeke TB, de Francisco AL, Anker SD, Kubo Y, Wheeler DC, Froissart M; on behalf the ARO Steering Committee collaborators. Propensity score matching and persistence correction to reduce bias in comparative effectiveness: the effect of cinacalcet use on all-cause mortality. *Pharmacoepidemiol Drug Saf.* 2015 Jul;24(7):738-47.