

ENSAIOS CLÍNICOS EM DERMATOLOGIA – PARTE II

Artur César^{1,2}, Alberto Mota^{1,2}, Filomena Azevedo¹

¹Serviço de Dermatologia e Venereologia/Department of Dermatology and Venereology, Centro Hospitalar São João EPE, Porto/Oporto, Portugal

²Faculdade de Medicina, Universidade do Porto/Medical School, Oporto University, Portugal

RESUMO – O crescente número dos ensaios clínicos na área dermatológica exige do dermatologista a capacidade de avaliar a qualidade, compreender, e sintetizar informações de diferentes fontes para decidir sobre as estratégias terapêuticas/diagnósticas. Este artigo de educação médica contínua aborda os princípios básicos, considerados relevantes e adaptados à área dermatológica, fundamentais na interpretação dos resultados dos ensaios clínicos. A parte II deste artigo aborda os diferentes métodos de análise de dados em ensaios clínicos.

PALAVRAS-CHAVE – Ensaaios clínicos; Interpretação estatística de dados; Desenho de estudo.

CLINICAL TRIALS IN DERMATOLOGY – PART II

ABSTRACT – The increasing number of clinical trials in dermatology requires from the dermatologist the ability to evaluate the quality, understand, and synthesize information from multiple sources in order to make decisions regarding the therapeutic/diagnostic strategies. This medical education article explains basic principles, relevant and adapted to the dermatology area, fundamental to an appropriate interpretation of outcomes of clinical trials. Part II of this article discusses the different analysis methods in clinical trials.

KEY-WORDS – Clinical trials; Data interpretation, statistical; Research design.

Conflitos de interesse: Os autores declaram não possuir conflitos de interesse.

No conflicts of interest.

Suporte financeiro: O presente trabalho não foi suportado por nenhum subsídio ou bolsa.

No sponsorship or scholarship granted.

Recebido/Received - Novembro/November 2014; Aceite/Accepted – Dezembro/December 2014

Correspondência:

Dr. Artur Jorge Fernandes César
Serviço de Dermatologia e Venereologia
Centro Hospitalar São João EPE
Alameda Prof. Hernâni Monteiro
4200-319 Porto, Portugal
Tel.: 00351225512117
Email: arturjfc@gmail.com

Educação Médica Contínua

1. INTRODUÇÃO

Numa época de incessante inovação e farta em publicações científicas, é fundamental compreender quais as qualidades e as fragilidades dos diferentes desenhos de estudo, bem como entender as diferentes formas de apresentação de resultados fundamentais à tomada informada de decisões.

Ao dermatologista é exigido constantes decisões quer no processo de diagnóstico, terapêutica quer no estabelecimento de um prognóstico para cada dermatose. Esta decisão pode ser baseada na experiência clínica, mas idealmente baseia-se em dados científicos. Por este motivo é fundamental possuir uma compreensão dos princípios estatísticos básicos para a interpretação dos resultados apresentados/obtidos em ensaios clínicos de forma a avaliar a qualidade de estudos individuais e a sintetizar a informação neles patente. Quando se avaliam ensaios clínicos devem ser considerados vários factores já que as diferenças no desenho de estudo (abordado na parte I deste artigo) bem como os métodos de análise dos dados e apresentação dos resultados têm impacto na validade e reprodutibilidade dos seus resultados.

A parte I deste artigo explorou os diferentes desenhos de estudo e abordou o seu impacto nos resultados e validade. A parte II tem por principais objectivos explicar os métodos utilizados na análise dos dados obtidos em ensaios clínicos, além de procurar discutir diferentes testes estatísticos e o seu impacto nas conclusões de cada estudo.

No final do artigo encontram-se as definições dos termos utilizados no texto (sublinhados).

2. AVALIAÇÃO DOS PARÂMETROS ENVOLVIDOS NO ESTUDO

A dificuldade em **estratificar a gravidade das dermatoses** reflete-se na falta de validação e, por vezes, de credibilidade de vários dos **índices e escalas de avaliação de gravidade clínica** comumente utilizados (PASI - *Psoriasis Area and Severity Index*, BSA - *Body Surface Area*, EASI - *Eczema Area and Severity Index*, SCORAD - *Severity Scoring for Atopic Dermatitis*, entre outros), os quais demonstram uma significativa variação intra e interobservador. Por exemplo, no caso da psoríase, o índice semi-quantitativo PASI, descrito pela primeira vez em 1978 por Fredriksson e Pettersson¹, é o mais frequentemente utilizado para aferir a gravidade desta dermatose, mas nunca se demonstrou se o peso relativo de cada um dos parâmetros usados

na fórmula reflete a verdadeira gravidade clínica². Por outro lado, a psoríase pode ser pouco extensa e, portanto, ter um PASI relativamente baixo, mas clinicamente grave. É o caso da psoríase palmoplantar, a qual, embora localizada, pode condicionar uma significativa incapacidade funcional. Por outro lado, esta dermatose é clinicamente multifacetada, apresentando para além da forma vulgar ou em placas, variantes eritrodérmica, pustulosa e artropática, entre outras, e esta circunstância deve estar claramente definida nos critérios de inclusão dos doentes, como já referido. Acresce que para uma mesma doença cutânea podem usar-se escalas de avaliação diversas (EASI e SCORAD para a dermatite atópica) ou variantes de uma mesma escala (PASI e PASI modificado, por exemplo), o que dificulta a análise comparativa entre diferentes ensaios. O recurso a parâmetros mais simples do que as escalas de gravidade clínica, tais como o número e a proporção de doentes que alcançaram remissões parciais ou totais, poderá ser mais informativo, sobretudo nos **ensaios clínicos de abordagem pragmática**, isto é dotados de análise baseada no **princípio da intenção de tratar** (ITT - *Intention To Treat*), nos quais os scores das escalas de gravidade clínica dos doentes que não completam o estudo não podem ser utilizadas (ver definições no final do texto).

De igual modo, a **avaliação de eficácia** deve preferir **critérios e metas com relevância clínica real e evitar o recurso a variáveis ou marcadores substitutivos**, uma vez que se pretende que estes ensaios sejam executados em condições mais próximas da prática clínica usual e é com base nos mesmos que, habitualmente, se decide se um determinado tratamento é licenciado para uso humano. Contudo, existem determinados critérios clínicos cuja obtenção é demorada, como é o caso da avaliação do impacto de uma determinada intervenção na incidência de neoplasias malignas cutâneas, o que justifica o recurso aos **marcadores substitutivos**, mais rápidos de avaliar. É o caso por exemplo de se usar a contagem de nevos melanocíticos e de queratoses actínicas em substituição, respetivamente do risco de melanoma maligno e de carcinoma espinocelular. De igual modo, variáveis como a adesão a medidas de evitação da exposição solar e a duração do internamento hospitalar, são presumidas substitutivas, respetivamente do nível de exposição à radiação ultravioleta e de complicações de infeções cutâneas, como as dermo-hipodermite agudas.

Como já abordado, numa doença crónica, em que a cura não é expectável a breve trecho, é muito

Educação Médica Contínua

importante definirem-se quais os critérios de **eficácia a curto, médio e a longo prazo**. A título exemplificativo, é frequente nos ensaios clínicos que avaliam os denominados fármacos biológicos na psoríase (proteínas de fusão e anticorpos monoclonais) aferirem a eficácia através da proporção de doentes que às 12-16 semanas, denominado período de indução, alcançam uma redução de pelo menos 75% no PASI inicial (PASI 75). Para além de se tratar de uma avaliação de curto-médio prazo, a simples avaliação de escalas de gravidade clínica em dois períodos distintos pode não traduzir corretamente a evolução clínica ao longo do tempo. Neste caso, a avaliação da **área sob a curva PASI-tempo** pode ser mais detalhadamente informativa sobre o carácter ondulante da evolução clínica, à semelhança do parâmetro farmacocinético **área sob a curva** previamente abordado.

Outras dimensões que têm sido introduzidas nos ensaios clínicos em Dermatologia nos últimos anos são a avaliação da qualidade de vida e os estudos farmacoeconómicos. Como já referido, o impacto das doenças cutâneas na qualidade de vida do doente e da sua família pode ser muito significativo, como é o caso da dermatite atópica, na qual se demonstrou que a degradação da qualidade de vida é sobreponível à de outras doenças crónicas, tais como a diabetes mellitus³. Na área dermatológica, os estudos de avaliação da qualidade de vida podem basear-se em inquéritos que são específicos de uma doença, como são os casos do EDI (*Eczema Disability Index*) e do mais recente QoLIAD (*Quality of Life Index for Atopic Dermatitis*), para o caso da dermatite atópica, ou vocacionados para as doenças cutâneas enquanto classe geral, como é o exemplo do DLQI (*Dermatology Life Quality Index*), o qual pode ser adaptado a crianças (*Children DLQI*). No entanto, pela própria definição de qualidade de vida da Organização Mundial de Saúde, isto é como a "perceção que um indivíduo tem do seu lugar existencial no contexto cultural e do sistema de valores em que vive, em função dos seus objectivos, metas alcançadas e das suas preocupações", a avaliação desta dimensão é sempre subjetiva e sofre de significativa variabilidade cultural e linguística, pelo que os questionários têm de ser validados para a população em estudo.

Entre outros aspetos, a avaliação da qualidade de vida abre portas à aferição de uma outra dimensão nos ensaios clínicos, que é a **avaliação da eficácia do tratamento realizada pelo próprio doente participante** (por exemplo POEM - *Patient-Oriented Eczema Measure*)⁴ e não apenas centrada no investigador.

Infelizmente, poucos são os estudos em Dermatologia que utilizam a avaliação da eficácia clínica realizada pelo próprio doente, o qual, por ser o alvo do tratamento em estudo, está teoricamente mais bem colocado para aferir do seu "interesse"⁵. Esta relutância em envolver o doente participante na avaliação da eficácia conta na sua génese com vários fatores, nomeadamente a impressão por parte do investigador de que avaliação feita pelo próprio doente enferma de maior subjetividade, receio de que o resultado dessa avaliação contrarie a obtida pelos investigadores e a necessidade de aumentar o tamanho da amostra e, portanto, do investimento económico, para se obter uma adequada potência estatística. É interessante constatar que algumas das mais recentes **revisões sistematizadas** do Grupo Dermatológico *Cochrane* têm definido como principais medidas de avaliação da eficácia, as avaliações feitas pelos doentes incluídos.

A pressão económica derivada dos preços habitualmente mais elevados dos fármacos mais recentes, tem motivado o desenvolvimento de **estudos farmacoeconómicos**, os quais pretendem demonstrar que os novos tratamentos em avaliação podem ser economicamente eficientes, apesar do preço direto elevado, ao reduzirem os **custos indiretos**, tais como internamentos, deslocações ao hospital e absentismo escolar e profissional e, ainda, poderem reduzir os **intangíveis**, tais como o sofrimento psicológico e a dor associada à doença⁶. Em Dermatologia podemos encontrar diferentes tipos de **estudos farmacoeconómicos (análises custo-benefício e custo-efetividade** na maioria dos casos; ver definições no final do texto) em várias doenças, destacando-se a dermatite atópica, acne, lesões pré-malignas e malignas, psoríase, infeções bacterianas complicadas da pele e dos tecidos moles e a onicomicose.

Outro aspeto a ter em conta é a **sazonalidade das doenças**, isto é, determinadas dermatoses como a acne, dermatite atópica e a dermatite seborreica podem responder de forma diferente ao tratamento, consoante a estação do ano⁷. Nestes casos, e para que os estudos sejam verdadeiramente representativos, deve-se preferir a colheita de dados ao longo de todo o ano.

3. ANÁLISE ESTATÍSTICA DOS RESULTADOS

A Conferência Internacional de Harmonização E9 datada de 1998, forneceu os princípios que devem nortear a análise estatística dos ensaios clínicos⁸. A

Educação Médica Contínua

denominada **análise baseada no princípio da intenção de tratar** (ITT - *Intention To Treat*), deve ser preferida e inclui a avaliação da eficácia e da segurança de todos os doentes admitidos no estudo, incluindo os que o abandonaram. Assim, o reporte da proporção de abandonos, e motivos de tais abandonos, não só aumenta a credibilidade das conclusões do estudo, como fornece uma ideia aproximada da probabilidade de um doente não completar o esquema terapêutico em avaliação, o que é importante não só para o clínico, mas também para os doentes⁹. Reportar dados referentes apenas aos sujeitos que completaram o estudo, deixando de parte os que o abandonaram por perda de eficácia ou pela emergência de efeitos adversos (conhecida como **análise baseada no princípio "as-treated"**) resultaria em resultados altamente enviesados, pelo que não é recomendada. Uma das grandes dificuldades com o método de análise baseada no **princípio ITT** é lidar com a informação em falta daqueles que abandonaram o estudo. Para tal existem dois métodos, o **"last observation carried forward"** e o **"non-responder imputation"**. No primeiro, os resultados da última avaliação feita antes do abandono do estudo são assumidos como sendo os resultados desse caso no final do tempo do estudo (carreadas, portanto, para o final do período de avaliação). No segundo método, os casos em que a aquisição de dados foi incompleta são simplesmente classificados como não respondedores ao tratamento. Qualquer um dos métodos introduz viéses, como é paradigmático um artigo onde foram analisados os resultados da eficácia do etanercept no tratamento de 131 doentes com psoríase segundo estes três métodos estatísticos: **"as treated"**, **"intention to treat – last observation carried forward"** e **"intention to treat – non-responder imputation"**¹⁰. Segundo o método **"as treated"**, o etanercept demonstrava um aumento progressivo de eficácia desde a semana 24 até à semana 224. Nesse mesmo período, e segundo o método de análise de dados **"intention to treat – last observation carried forward"**, a eficácia do etanercept mostrava-se praticamente constante ao longo do tempo. Já com a análise dos mesmos dados pelo método **"intention to treat – non-responder imputation"**, que se mostrava o mais conservador dos três, o etanercept exibia um declínio progressivo na eficácia a partir da semana 24.

Outro aspecto a ter em conta na análise estatística é que muitos dos dados avaliados em Dermatologia são **não paramétricos**, isto é, não seguem uma distribuição denominada Normal ou gaussiana, como é o caso da avaliação da gravidade de uma dermatose

com base em escalas, ou então não é possível determinar o tipo de distribuição devido ao subdimensionamento da amostra. Neste contexto o uso de métodos estatísticos paramétricos (como o *t Student*) é inadequado, devendo-se recorrer **aos testes não paramétricos** (como o Mann-Witney U), mas que são muito pouco utilizados.

Outra questão prende-se com a deificação do **nível de significância de 0,05 do valor de p** para validar ou excluir diferenças, mas o qual não é suficiente para a interpretação dos resultados (ver definição no final do texto)¹¹. De facto, este valor deve ter em consideração se o **tamanho da amostra** foi adequado para detectar pequenas diferenças e se as diferenças consideradas "estatisticamente" significativas, o são, também, sob o ponto de vista clínico. O exemplo de um estudo em que a amostra estava subdimensionada foi o conduzido por Katz et al em 1989, o qual incluiu 51 doentes com psoríase vulgar, cada um tratado simultaneamente com pomada de furoato de mometasona 0,1% e de hidrocortisona 1,0%, em lesões bilateralmente opostas e durante 6 semanas¹². Este estudo concluiu que a pomada de furoato de mometasona era mais eficaz que a de hidrocortisona no tratamento da psoríase, com um potencial atrofogénico similar entre os dois corticosteróides tópicos. Contudo, da análise do desenho do estudo resulta que o tamanho da amostra conferia potência estatística suficiente para detectar diferenças significativas na eficácia entre os dois compostos, mas era notoriamente insuficiente para detetar diferenças em relação aos efeitos laterais, pelo que não se podia concluir pela alegada similitude na indução de atrofia cutânea. Ainda a respeito do tamanho amostral, mais um exemplo paradigmático, é o estudo controlado e duplamente cego conduzido por Spruance et al em 1997, que demonstrou que o creme de penciclovir era mais eficaz que o placebo no encurtamento do tempo de cicatrização de lesões de herpes orolabial¹³. Este ensaio envolveu um elevado número de doentes (n=2209) para detetar uma diferença máxima de 15% entre o produto ativo e o placebo, o que é considerada clinicamente pouco relevante, isto é, neste exemplo a amostra foi claramente sobredimensionada para permitir detetar uma pequena diferença estatisticamente significativa, mas de duvidoso interesse clínico.

É possível através de estudos preliminares determinar qual o tamanho da amostra necessário para que uma determinada diferença, considerada clinicamente significativa, alcance significado estatístico ($p < 0,05$) com uma probabilidade de 80% (potência estatística de 80%), nomeadamente através da aplicação da

Educação Médica Contínua

fórmula de Lehr¹⁴. Por vezes, seria necessário um número elevado de participantes para alcançar este objetivo, o que tomaria o ensaio clínico inviável ou muito oneroso¹⁵. É o caso, por exemplo, da demonstração de que o método pouco invasivo de dosear o pré-peptido de colagénio III na avaliação da toxicidade cumulativa do metotrexato,¹⁶ um fármaco de primeira linha no tratamento sistémico da psoríase, poderia substituir a biópsia hepática (habitualmente recomendada após uma dose cumulativa de 1-1,5g)¹⁷. Estima-se que seria necessário avaliar mais de 6000 doentes sob metotrexato para se obter uma potência estatística de 80% num estudo comparativo.

Outra dimensão importante na análise dos dados é determinar o **intervalo de confiança de 95%** (IC95; ver definição no final do texto) para a diferença entre tratamentos e compará-lo com a diferença que se estabeleceu previamente como clinicamente significativa, ou que é permitida para não considerar os tratamentos como diferentes. Na realidade, esta análise fornece informações adicionais muito importantes que não são suficientemente transmitidas pela análise isolada do **nível de significância** (valor de p)¹⁵. Em pormenor, permite aferir se uma diferença estatisticamente significativa entre tratamentos, também o é sob o ponto de vista de significado clínico, no caso dos **estudos de superioridade**, ou se os tratamentos são clinicamente equivalentes no caso dos denominados **estudos de equivalência ou de não inferioridade** (ver definição no final do texto)¹⁸. No desenho do estudo será sempre necessário definir previamente a diferença considerada clinicamente significativa ou o desvio permitido entre tratamentos para os definir como equivalentes. Um tratamento só será considerado superior (clinicamente significativo ou não equivalente, consoante o objetivo) ao comparador se o limite inferior do IC95 da diferença entre os tratamentos for sempre superior ao limite superior da diferença clinicamente significativa ou da diferença permitida entre os tratamentos (Fig. 1).

Numa análise em que há **avaliação de múltiplas variáveis** (ex: idade, género, nível de exposição solar, etc.) **em relação à mesma questão** (ex: mortalidade por melanoma maligno) obtêm-se naturalmente múltiplos valores de p . Neste caso será mais prudente diminuir o **nível de significância** de 0,05 para 0,01, de modo a reduzir a probabilidade de se detetarem diferenças que não existem, isto é falsos positivos ou erro estatístico de tipo 1 (referido, igualmente, como erro alfa ou de 1ª espécie) ou, então recorrer à **correção de Bonferroni**, a qual entra em linha de conta

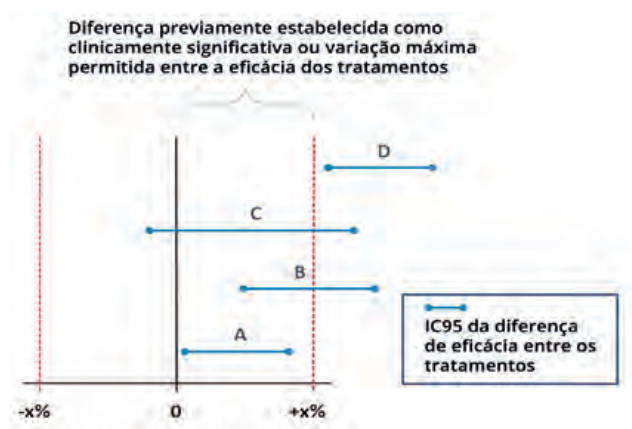


Fig 1 - Nos estudos comparativos A a D, as diferenças de eficácia entre os tratamentos foram estatisticamente significativas (**nível de significância** $p < 0,05$). Contudo, só a diferença do estudo D teve significado clínico ou um dos tratamentos deste estudo foi considerado superior ao comparador, em contraste com o estudo A (diferença sem significado clínico ou tratamentos equivalentes) e os B e C (inconclusivos).

com o número de p gerados (ver definição no final do texto). Por razões éticas, o **nível de significância** deve ser estabelecido logo no desenho do ensaio e não na sua conclusão. É possível que o recurso a outros métodos estatísticos, tais como a **análise de dados em microarrays** e o **método de Bayes**, possa vir a revolucionar o processamento de dados complexos e a eliminar erros comuns com os métodos tradicionais, ao não se basearem, nomeadamente, na hipótese zero ou nula (H_0).

4. OUTROS ASPETOS E CONSIDERAÇÕES

As diretrizes CONSORT (*Consolidation of Standards for Reporting Trials*, 2010) são atualmente a referência para a realização, avaliação e publicação de ensaios clínicos e foram adotadas por muitas revistas médicas, incluindo na área da Dermatologia¹⁹. Contudo, é de notar que devido à sua complexidade, alguns autores adaptaram outra metodologia que pretende avaliar de forma mais expedita a qualidade dos ensaios clínicos nos tratamentos tópicos da acne, nomeadamente²⁰.

Nos últimos anos tem crescido o número de **revisões sistematizadas**, incluindo as **metanálises**, na área dermatológica. Estes estudos de síntese, ao compilarem e analisarem estatisticamente, no caso das **metanálises**, os resultados de vários estudos clínicos, sejam positivos

Educação Médica Contínua

ou negativos, permitem uma avaliação mais exaustiva e credível sobre a reprodutibilidade e a consistência dos resultados em termos de eficácia e de segurança, aproximando a realidade experimental da clínica²¹.

5. CONCLUSÕES

O intuito da parte II deste artigo de Educação médica contínua foi fornecer ferramentas básicas necessárias na avaliação dos dados e informação sobre adequado processamento estatístico dos mesmos. Todos os exemplos acima expostos servem o propósito de demonstrar a elementar importância do exame exaustivo dos métodos estatísticos empregues perante a variabilidade da forma e conteúdo da informação científica reportada na literatura médica.

6. DEFINIÇÕES DE TERMOS USADOS NO ARTIGO

1. Estudo de abordagem pragmática – designação do ensaio que pretende avaliar se uma intervenção resulta na prática clínica e não apenas em determinadas condições ideais. A análise destes estudos é feita segundo o **princípio da intenção de tratar** (ITT – *Intention To Treat*). Em contraste, o **estudo explanatório** pretende avaliar a utilidade de uma intervenção em condições ideais ou determinar precisamente quais as condições ideais que tornam uma intervenção útil. A sua análise é feita numa abordagem *per protocol*.

2. Estudo de análise custo-benefício e custo-efetividade – estudo fármaco-económico que pretende avaliar se o benefício do tratamento justifica o seu custo, o que a verificar-se significa que o tratamento é custo-efetivo.

3. Nível de significância – probabilidade pré-definida abaixo da qual a hipótese nula H_0 (i.e., de que não há diferença entre os grupos comparados) do teste estatístico é rejeitada e, portanto, a diferença é considerada estatisticamente significativa. Habitualmente, é usada o valor de 0,05.

4. Intervalo e limites de confiança de 95% – em termos estatísticos, este intervalo de confiança significa que 95% (valor convencional) das observações de uma distribuição normal são delimitadas por 1,96 desvios-padrão de cada lado da média.

5. Estudo de equivalência – o objetivo primário deste tipo de comparação é demonstrar que dois tratamentos são equivalentes em relação a determinados parâmetros e é uma variação do estudo de não inferioridade. Por oposição, os estudos nos quais se pretende demonstrar diferenças, denominam-se de **estudos diferenciais** e são de dois tipos, de superioridade ou de inferioridade.

6. Correção de Bonferroni – ajustamento estatístico que é feito quando se interpretam múltiplos testes de significância que abordam a mesma questão básica. Neste caso, o significado estatístico só se atinge quando pelo menos um dos valores p é menor do que $0,05/k$, em que k representa o número de valores p calculados.

BIBLIOGRAFIA

1. Fredriksson T, Pettersson U. Severe psoriasis-oral therapy with a new retinoid. *Dermatologica*. 1978; 157:238-44.
2. Naldi L. Scoring and monitoring the severity of psoriasis. What is the preferred method? What is the ideal method? Is PASI passe? facts and controversies. *Clin Dermatol*. 2010; 28:67-72.
3. Lewis-Jones S. Quality of life and childhood atopic dermatitis: the misery of living with childhood eczema. *Int J Clin Pract*. 2006; 60:984-92.
4. Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol*. 2004; 140:1513-9.
5. Townshend AP, Chen CM, Williams HC. How prominent are patient-reported outcomes in clinical trials of dermatological treatments? *Br J Dermatol*. 2008; 159:1152-9.
6. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996; 276:1339-41.
7. Weiss SC, Rowell R, Krochmal L. Impact of seasonality on conducting clinical studies in dermatology. *Clin Dermatol*. 2008; 26:565-9.
8. U.S. Department of Health and Human Services FaDA, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). International Conference of Harmonisation of Pharmaceuticals for Human Use (ICH) E9.

Educação Médica Contínua

- Statistical principles for clinical trials. 1998. [consultado em] Disponível em: http://www.fda.gov/cder/guidance/ICH_E9-fnl.pdf.
9. Alosch M, Fritsch K, Soukup M, Wilkin J. Clinical trials and statistical analyses: what should dermatologists look for in a report? *Dermatol Therap.* 2009; 22:199-203.
 10. van Lumig PP, Driessen RJ, Kievit W, Boezeman JB, van de Kerkhof PC, de Jong EM. Results of three analytical approaches on long-term efficacy of etanercept for psoriasis in daily practice. *J Am Acad Dermatol.* 2013; 68:57-63.
 11. Blume J, Peipert JF. What your statistician never told you about P-values. *The J Am Assoc Gynecolc Laparosc.* 2003;10:439-44.
 12. Katz HI, Praver SE, Watson MJ, Scull TA, Peets EA. Mometasone furoate ointment 0.1% vs. hydrocortisone ointment 1.0% in psoriasis. Atrophogenic potential. *Int J Dermatol.* 1989; 28:342-4.
 13. Spruance SL, Rea TL, Thoming C, Tucker R, Saltzman R, Boon R. Penciclovir cream for the treatment of herpes simplex labialis. A randomized, multicenter, double-blind, placebo-controlled trial. Topical Penciclovir Collaborative Study Group. *JAMA.* 1997; 277:1374-9.
 14. Lehr R. Sixteen S-squared over D-squared: a relation for crude sample size estimates. *Stat Med.* 1992; 11:1099-102.
 15. Bhardwaj SS, Camacho F, Derrow A, Fleischer AB, Jr., Feldman SR. Statistical significance and clinical relevance: the importance of power in clinical trials in dermatology. *Arch Dermatol.* 2004;140:1520-3.
 16. Zachariae H, Sogaard H, Heickendorff L. Serum aminoterminal propeptide of type III procollagen. A non-invasive test for liver fibrogenesis in methotrexate-treated psoriatics. *Acta Dermato-Venereol.* 1989; 69:241-4.
 17. Roenigk HH, Jr., Auerbach R, Maibach H, Weinstein G, Lebwohl M. Methotrexate in psoriasis: consensus conference. *J Am Acad Dermatol.* 1998; 38:478-85.
 18. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA.* 2012; 308:2594-604.
 19. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol.* 2010; 63:e1-37.
 20. Revuz J, Moyses D, Poli F, Pawin H, Faure M, Chivot M, et al. A tool to evaluate rapidly the quality of clinical trials on topical acne treatment. *Journal of the European Academy of Dermatology and Venereology: J Eur Acad Dermatol Venereol.* 2008; 22:800-6.
 21. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ.* 2001; 323:42-6.

Educação Médica Contínua

VERIFIQUE O QUE APRENDEU

1. A análise de dados segundo o princípio *Intention-to-treat* (ITT) refere-se a qual dos seguintes?

- a) Inclusão no estudo de indivíduos que cumpram apenas os critérios de inclusão
- b) Todos os indivíduos incluídos no estudo devem ser incluídos na análise final dos dados independentemente se aderiram à intervenção ou não
- c) Apenas os indivíduos que completaram a intervenção até ao final do período de estudo são incluídos na análise final dos dados
- d) Todos os indivíduos devem concordar ser randomizados e possivelmente receber uma intervenção considerada placebo

2) Que tipo de estudo compara uma intervenção conhecida com uma intervenção com o objetivo de determinar se a nova intervenção não é menos eficaz que a já conhecida?

- a) Meta-análise
- b) Estudo de casos e controlos
- c) Estudo de não-inferioridade
- d) Ensaio clínico randomizado

3) Se o nível de significância de uma hipótese é 0.025, significa que a hipótese em análise:

- a) Deve ser rejeitada se $H_0=0.01$
- b) Não deve ser rejeitada se $H_0=0.02$
- c) Pode ou não ser rejeitada se $H_0=0.01$
- d) Pode ou não ser rejeitada se $H_0=0.02$