

PSS HEALTH: COMO CALCULAR TAMANHO DE AMOSTRA PARA TESTES DE COMPARAÇÃO DE MÉDIAS DE DOIS GRUPOS

PSS HEALTH: HOW TO CALCULATE SAMPLE SIZES FOR MEAN COMPARISON TESTS BETWEEN TWO GROUPS

Vânia Naomi Hirakata¹ , Stela Maris de Jezus Castro^{1,2} ,
Aline Castello Branco Mancuso¹ , Suzi Alves Camey^{1,2} ,
Vanessa Bielefeldt Leotti^{1,2} , Rogério Boff Borges¹ 

Clin Biomed Res. 2022;42(2):198-209

1 Unidade de Bioestatística, Grupo de Pesquisa e Pós-graduação (GPPG), Hospital de Clínicas de Porto Alegre (HCPA). Porto Alegre, RS, Brasil.

2 Departamento de Estatística, Instituto de Matemática e Estatística (IME), Universidade Federal do Rio Grande do Sul (UFRGS). Porto Alegre, RS, Brasil.

Autor correspondente:

Vânia Naomi Hirakata -
l-bioestatistica@hcpa.edu.br,
Unidade de Bioestatística, Grupo de Pesquisa e Pós-graduação (GPPG), Hospital de Clínicas de Porto Alegre (HCPA) Rua Ramiro Barcelos, 2350. 90035-007, Porto Alegre, RS, Brasil.

RESUMO

Dando continuidade à série de artigos que pretendem orientar o usuário na utilização da ferramenta PSS Health para o planejamento de uma pesquisa, esta edição apresenta um passo a passo de como realizar o cálculo e de quais informações são necessárias para comparar médias: de dois grupos dependentes ou independentes, de dois grupos independentes com duas medidas repetidas (deltas), e com duas ou mais medidas repetidas. Todos os exemplos também são ilustrados e disponibilizados em vídeos no canal da Unidade de Bioestatística.

Palavras-chaves: *Tamanho de amostra; Estudos longitudinais; Análise por pareamento*

ABSTRACT

Following the series of articles aiming to guide users in using the PSS Health tool for research planning, this issue presents a step-by-step guide on how to calculate and what information is needed to compare means between 2 dependent or independent groups, 2 independent groups with 2 repeated measures (deltas), and 2 independent groups with 2 or more repeated measures. All examples are accompanied by figures and available in video on the Biostatistics Unit's channel.

Keywords: *Sample size; Longitudinal study; Paired comparisons*

INTRODUÇÃO

Dando seguimento à série de artigos sobre a ferramenta PSS Health (*Power and Sample Size for Health Researchers*), apresentada por Borges et al.¹, este artigo mostra como calcular tamanho de amostra para testes de comparação de médias entre 2 grupos independentes ou relacionados (dependentes). Serão abordados exemplos de estudos transversais, com uma medida única, e de estudos longitudinais, com mais de uma medida da mesma unidade amostral. A partir de exemplos, será detalhado quais são as informações necessárias para cada situação, onde ou como obtê-las e o passo a passo da realização do cálculo na ferramenta, também ilustrado e disponibilizado em vídeo-aulas.

Cabe salientar, que existem outras situações de cálculos de tamanhos de amostra para comparação de médias de dois grupos que não serão abordadas no presente artigo, tais como para ensaios clínicos randomizados de equivalência, não-inferioridade e superioridade.

O PSS Health pode ser utilizado diretamente pelo computador através do software livre R², com o pacote PSS.Health³, ou on-line no site da Unidade de Bioestatística (sites.google.com/hcpa.edu.br/bioestatistica) ou diretamente no endereço hcpa-unidade-bioestatistica.shinyapps.io/PSS_Health. Neste artigo foi utilizada a versão online do aplicativo (acessado em outubro de 2021) e a versão 0.3.1 do pacote.

COMPARAÇÃO DE MÉDIAS DE DOIS GRUPOS INDEPENDENTES

Na área da saúde, existem estudos com objetivo de comparar grupos que representam populações distintas, isto é, populações que não têm relação entre si. Na estatística, tratamos estas amostras como grupos independentes. No caso específico de dois grupos, os mesmos são representados por uma variável dicotômica como, por exemplo, a presença ou não de uma intervenção (em ensaios clínicos randomizados), a presença ou não de uma doença (em estudos de caso-controle) e a presença ou não de uma condição ou exposição (em estudos longitudinais ou transversais).

Quando queremos extrapolar o resultado de uma amostra para uma população, estas comparações são feitas com base nos testes de hipóteses. Para definir qual teste escolher, deve-se primeiramente conhecer a natureza do desfecho em estudo, que pode ser quantitativo ou qualitativo. No caso específico de desfecho quantitativo, é comum ser realizado o teste t de Student para amostras independentes, por exemplo. Já nos estudos longitudinais, onde a medida é observada mais de uma vez para cada grupo, podemos utilizar abordagem de modelos lineares mistos ou equações de estimações generalizadas, por exemplo.

COMPARAÇÃO DE MÉDIAS DE DOIS GRUPOS DEPENDENTES

Outro objetivo também comumente encontrado na área da saúde é o de comparar dois momentos distintos de uma mesma população. Nestes casos, na estatística, tratamos estas amostras como grupos dependentes (pareados ou, ainda, relacionados), pois a unidade amostral de um grupo é a mesma do outro grupo, porém em um momento diferente. Em casos mais específicos, esta mesma abordagem também pode ser encontrada em comparações de grupos de unidades diferentes, porém altamente correlacionadas como, por exemplo, estudo com irmãos gêmeos.

Da mesma forma, quando queremos extrapolar o resultado de uma amostra para uma população, as comparações devem ser feitas com base em testes de hipóteses. Para comparação de dois grupos dependentes (ou dois momentos) é comum ser utilizado o teste t de Student para amostras dependentes.

Considerando a comparação de dois grupos independentes ou dependentes, o cálculo do tamanho de amostra, assim como do poder, é baseado no teste de hipóteses com o mesmo objetivo. Em ambos os casos, quando se trata de testes bilaterais, a

hipótese nula é de que não exista diferença entre as respostas médias de cada grupo e a alternativa é a de que existe diferença. Conceitos sobre testes de hipóteses e seus elementos podem ser revisados em publicação prévia⁴.

USO DO PSS HEALTH PARA CALCULAR O TAMANHO DE AMOSTRA

Para calcular o tamanho de amostra para testes de comparação de médias de dois grupos independentes e dependentes, no PSS Health, utiliza-se o menu “MÉDIAS” (Figura 1).

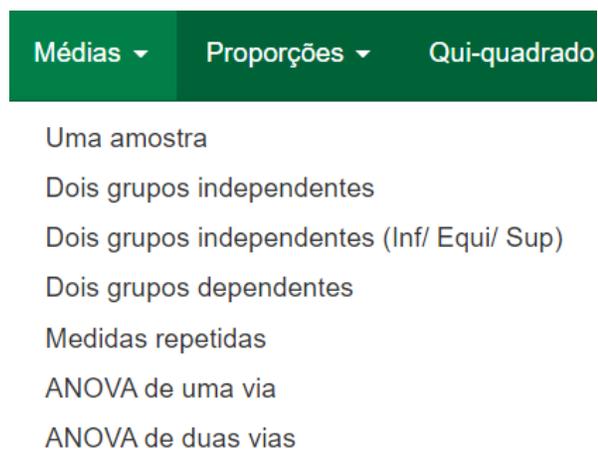


Figura 1 – Menu inicial do PSS Health para estimativas de cálculos envolvendo médias.

Para comparação de duas médias independentes

Quando o objetivo do estudo for a comparação de médias entre dois grupos independentes em um estudo transversal, o cálculo do tamanho da amostra é baseado no teste t de Student para duas amostras independentes⁵. O Quadro 1 apresenta as informações necessárias para o cálculo de tamanho de amostra para comparar as médias de uma variável quantitativa em dois grupos independentes e, assim como na ferramenta PSS Health, denotaremos a variável em estudo como variável Y. É importante salientar que as informações sobre a diferença mínima a ser detectada (esperada) e o tamanho de efeito (d de Cohen) devem ter relevância clínica. Estas informações podem estar disponíveis na literatura, em artigos que tenham evidenciado uma diferença em grupos semelhantes aos que se deseja estudar, mas também podem ser determinados pelo pesquisador. Maiores detalhes na seção “Qual a diferença entre significância estatística e relevância clínica?”⁶.

Quadro 1 – Informações necessárias para o cálculo de tamanho de amostra para comparação de médias de dois grupos independentes e as fontes para obtenção.

Informação	Fonte	
	Literatura	Pesquisador
Informações sobre a variável em estudo		
Tamanho do efeito (d de Cohen) ; OU	✓	✓
Diferença do Y a ser detectada em u.m. e	✓	✓
Desvio padrão esperado (comum aos grupos); OU	✓	
Desvio padrão de cada grupo	✓	
Informações sobre a amostragem e o teste de hipótese		
Balanceamento	✓	✓
Efeito do plano amostral	✓	✓
Nível de significância		✓
Poder		✓
Tipo de Teste		✓
Perdas e recusas		✓

No artigo anterior, Castro et al.⁷ mostraram como realizar o cálculo para uma pesquisa cujo objetivo era estimar a média de pressão arterial sistólica de crianças. Suponha que agora o objetivo seja comparar os níveis médios de pressão arterial sistólica de crianças em idade pré-escolar (de 2 a 6 anos) com e sem obesidade atendidas em unidades básicas de saúde de um município. Os pesquisadores desejam fazer um cálculo de tamanho de amostra capaz de detectar uma diferença entre as médias da pressão arterial sistólica que seja clinicamente relevante. Para exemplificar, utilizaremos informações do artigo de Garcia et al.⁸

Para calcular o tamanho de amostra para testar duas médias no PSS Health, após clicar no menu “MÉDIAS” escolhe-se a opção “DOIS GRUPOS INDEPENDENTES” (Figura 1).

Nesta opção são apresentadas três abas: duas abas com as opções de cálculo de tamanho de amostra para testar ou estimar uma diferença de médias e uma aba com a opção para cálculo de poder do teste de hipóteses. Iremos mostrar a utilização das duas primeiras. Não apresentaremos a aba de poder, no entanto, as informações necessárias são as mesmas, com a ressalva de que, ao invés do poder, deve ser informado o tamanho da amostra do estudo.

Cálculo de tamanho de amostra para TESTAR a diferença de médias

A aba “TESTAR” mostra os campos para colocar as informações requeridas para o cálculo de tamanho de amostra para testar a diferença entre médias (Figura 2). Para entender o significado de cada um dos campos basta clicar na interrogação apresentada ao lado do campo de interesse.

Conforme ilustrado na Figura 2, a primeira informação apresentada são as hipóteses a serem testadas, de

igualdade das médias (hipótese nula) e de diferença (hipótese alternativa). Logo abaixo, antes dos campos para informações do cálculo, pode-se clicar em “Mudar nomes” para nomear o desfecho (até então chamado de Y), a unidade de medida e os grupos em estudo. Estes serão substituídos nos campos solicitados e no texto final que descreve o cálculo realizado, conforme poderá ser observado na Figura 3.

Para calcular o tamanho da amostra, deve-se primeiramente decidir se ele será realizado a partir do desvio padrão ou do tamanho de efeito d de Cohen⁹. Neste exemplo, utilizaremos um desvio padrão obtido da literatura, mas é possível calcular o desvio padrão e/ou o d de Cohen no menu “Outras ferramentas”, no próprio PSS Health.

Fazendo o cálculo do tamanho da amostra a partir do desvio padrão, a primeira informação solicitada é a diferença que se deseja detectar entre os dois grupos. A partir dos dados da tabela 1 do artigo de Garcia et al.⁸, podemos calcular que essa diferença é de 5,5 mmHg (102,5-97,0). Note que neste exemplo os desvios padrão dos dois grupos foram iguais e, por isso, será utilizado o valor 11,0, mas nem sempre isto ocorre. Quando forem diferentes, deve-se primeiramente calcular o desvio padrão combinado (clicando na opção “Calcular o desvio padrão combinado”) e então informar no campo “Desvio padrão esperado”. Nas situações em que o desvio padrão não está apresentado, é possível obtê-lo de outras medidas de variabilidade⁷.

As informações seguintes são o balanceamento e o efeito do plano amostral. O balanceamento é a razão entre os tamanhos de amostra de dois grupos. Aqui, o balanceamento utilizado é de um, pois deseja-se grupos de igual tamanho (ou seja, um participante no grupo A para cada participante no grupo B). O efeito do plano amostral, representa o quanto o plano amostral proposto é mais ou menos eficiente,

em termos de variabilidade da estimativa, do que a amostragem aleatória simples. É utilizado para medir o efeito de um plano amostral sobre a variância de um estimador e, se o seu valor for igual a 1, então o plano amostral proposto é a amostragem aleatória simples ou é considerado tão eficiente quanto esta. Em nosso exemplo, manteremos com o valor de 1.

A informação do poder de amostra desejado só deve ser alterada caso o pesquisador deseje um poder diferente de 80% (*default* da ferramenta). Da mesma forma, o nível de significância, cujo *default* é de 5%, também pode ser alterado. Já o tipo de teste deve ser escolhido conforme a hipótese alternativa: o teste pode ser bilateral, unilateral superior ou unilateral inferior. Maiores informações em Hiraikata et al.⁴

Com estes campos preenchidos, já é possível observar o tamanho amostral calculado: 197 (40 obesos e 157 não obesos), conforme ilustrado na Figura 3, e o texto sugerido para descrição do cálculo. No entanto, ainda há a opção de incluir um percentual previsto para perdas ou recusas durante a

pesquisa, a partir do qual o tamanho final da amostra será ajustado, a fim de contemplar este percentual. Este ajuste pode ser observado na última frase do texto sugerido. Ainda sobre a sugestão de texto, é de extrema importância lembrar de alterar o nome do autor da referência, apresentado genericamente como “Fulano (1900)”, pois não há como adaptar esta informação direto na ferramenta.

Assim como apresentado em Castro et al.⁷, é possível, também, construir cenários de tamanhos de amostra. Os cenários podem ser visualizados através de um gráfico (Figura 4) e/ou através de uma tabela (Figura 5), sendo possível o download de ambos. Pode-se variar a diferença entre as médias (eixo vertical) para diferentes níveis de poder (linhas). Este exemplo de como utilizar o PSS Health para calcular o tamanho de amostra para comparar duas médias também pode ser visualizado no vídeo “PSS para comparação de médias em grupos independentes”¹⁰.

The screenshot shows the 'TESTAR' tab of the PSS Health tool. It contains the following fields and values:

- Hipóteses a serem testadas:**
 - $H_0 : \mu_{obesos} = \mu_{n\grave{o}obesos}$
 - $H_1 : \mu_{obesos} \neq \mu_{n\grave{o}obesos}$
- Mudar nomes:**
 - Calcular usando o d de Cohen
- Diferença do pressão arterial sistólica a ser detectada em mm Hg (Média do obesos) - (Média do grupo não obesos):** 5,5
- Desvio padrão esperado:** 11
- Calcular o desvio padrão combinado:** (checkbox)
- Balanceamento (obesos:não obesos):** 0,25
- Efeito do plano amostral:** 1
- Poder (%):** 80
- Nível de significância (%):** 5
- Tipo de teste de acordo com hipótese alternativa:** Bilateral
- Perdas/ Recusa (%):** 10

Figura 2 – Aba TESTAR para diferença de médias entre dois grupos independentes.

Tamanho amostral calculado: 197 (40 no obesos e 157 no não obesos)

Sugestão de texto:

Foi calculado o tamanho de amostra para detectar diferenças na média de **pressão arterial sistólica** entre os grupos obesos e não obesos, tendo uma diferença de **5.5 mmHg** como relevante para o estudo. Para isso foi utilizado a ferramenta PSS Health versão on-line (citação abaixo). Considerando poder de **80%**, nível de significância de **5%** e desvio padrão de **11 mmHg** (dados de Fulano (1900)), chegou-se ao tamanho de amostra total de **197** sujeitos, sendo **40** no obesos e **157** no não obesos. Acrescentando **10%** para possíveis perdas e recusas o tamanho de amostra deverá ser **220** (45 no obesos e 175 no não obesos).

Figura 3 – Sugestão de texto para descrição do cálculo de tamanho de amostra para testar a diferença de médias entre dois grupos independentes.

Construção de cenários

Utilize os argumentos abaixo para construir diferentes cenários. Você pode especificar valores de poder e uma sequência da diferença a ser detectada. Demais informações serão recuperadas do painel lateral.

Digite valores de poder para fazer o gráfico

80, 90, 95

Defina a sequência de valores para a diferença a ser detectada:

Mínimo Máximo Intervalo

5

6,5

0,5

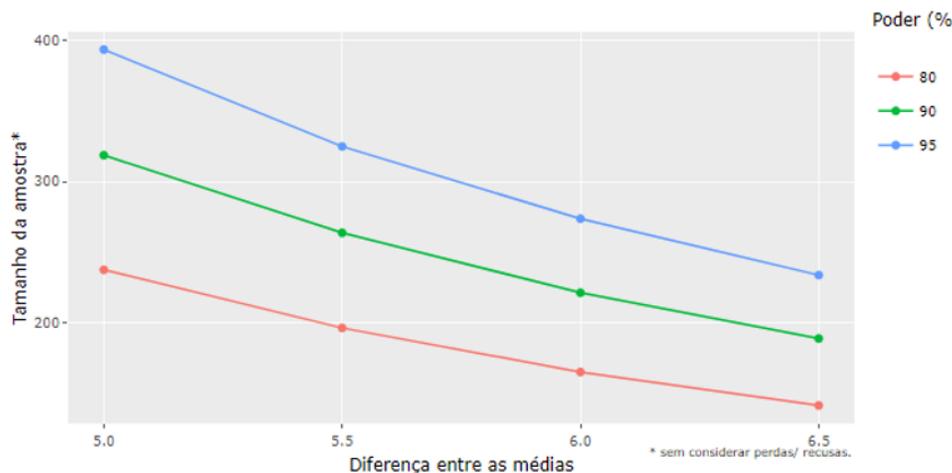


Figura 4 – Gráfico para diferentes cenários no cálculo de tamanho de amostra para comparação de médias de dois grupos independentes.

Diferença entre as médias	n total	n Tratamento	n Controle	Desvio padrão	Balanceamento	Poder (%)	Nível de significância (%)	Hipótese alternativa	Efeito do delineamento
5	237.5	47.5	190	11	0.25	80	5	2	1
5.5	196.25	39.25	157	11	0.25	80	5	2	1
6	165	33	132	11	0.25	80	5	2	1
6.5	141.25	28.25	113	11	0.25	80	5	2	1
5	318.75	63.75	255	11	0.25	90	5	2	1
5.5	263.75	52.75	211	11	0.25	90	5	2	1
6	221.25	44.25	177	11	0.25	90	5	2	1
6.5	188.75	37.75	151	11	0.25	90	5	2	1
5	393.75	78.75	315	11	0.25	95	5	2	1
5.5	325	65	260	11	0.25	95	5	2	1

Showing 1 to 10 of 12 entries

Previous 1 2 Next

Figura 5 – Tabela para diferentes cenários no cálculo de tamanho de amostra para comparação de médias de dois grupos independentes.

Cálculo de tamanho de amostra para estimar a diferença de médias

Na aba “ESTIMAR”, do menu de comparação de dois grupos independentes, é possível realizar o cálculo do tamanho da amostra necessário para, como o nome da aba já diz, estimar a diferença entre as médias de dois grupos. Como pode ser verificado na Figura 6, outros campos são solicitados e uma das diferenças em relação à aba “TESTAR” é o campo “Margem de erro ou semi-amplitude” que, por definição, representa a metade da largura total do intervalo de confiança. Ou seja, quanto menor seu valor maior será a precisão da estimativa,

porém o tamanho de amostra necessário também será maior.

Agora, suponha que o pesquisador deseje calcular o tamanho de amostra necessário para que o intervalo de confiança da diferença entre as médias tenha no máximo 4 mmHg de amplitude, ou seja, margem de erro igual a 2 mmHg, valor arbitrário atribuído pelo pesquisador. Assim, após descrever o nome e a unidade de medida do desfecho em estudo, digita-se no campo “Margem de erro ou semi-amplitude” o valor 2 e em “desvio padrão esperado” o valor 11, referenciado no exemplo anterior. O resultado do cálculo e o texto de descrição estão apresentados na Figura 7.

Testar
Estimar
Poder

Descreva o nome do desfecho ?

Descreva a unidade de medida de pressão arterial sistólica ?

Margem de erro ou semi-amplitude ?

Desvio padrão esperado ?

Nível de confiança (%) ?

Perdas/ Recusa (%) ?

Figura 6 – Aba ESTIMAR para diferença de médias entre dois grupos independentes.

Tamanho amostral calculado: 468 (234 em cada grupo)

Sugestão de texto:

O cálculo do tamanho de amostra foi realizado por meio da ferramenta PSS Health versão on-line (citação abaixo), para estimar a diferença entre as médias de **pressão arterial sistólica** nos grupos A e B, com margem de erro de **2 mm Hg**, nível de confiança de **95%** e desvio padrão esperado de **11 mm Hg** como referido em Fulano (1900), chegou-se ao tamanho de amostra de **468 sujeitos** (sendo **234** em cada grupo). Acrescentando **10%** para possíveis perdas e recusas o tamanho de amostra deverá ser **520**.

Figura 7 – Sugestão de texto para descrição do cálculo de tamanho de amostra para estimar a média da diferença entre dois grupos independentes.

Este exemplo de como utilizar o PSS Health para calcular o tamanho de amostra para estimar a diferença entre duas médias também pode ser visualizado no vídeo “PSS para comparação de médias em grupos independentes”¹⁰.

Para comparação de duas médias dependentes

Nesta seção vamos abordar estudos que avaliam dois grupos que sejam emparelhados por algum aspecto (por exemplo: mão direita e esquerda, irmãos

gêmeos) ou até mesmo um grupo único que seja avaliado em dois momentos (por exemplo: um grupo de pacientes que serão avaliados antes e depois de um procedimento). Nesses casos, é recomendada a utilização de testes para dados dependentes ou emparelhados. Aqui, a hipótese nula é de que a média das diferenças entre os pares é igual a zero. As informações necessárias para o cálculo de tamanho de amostra nesta situação, assim como a fonte de obtenção encontram-se no Quadro 2.

Quadro 2 – Informações necessárias para o cálculo de tamanho de amostra para comparação de duas médias emparelhadas ou dependentes e a fonte de obtenção.

Informação	Fonte	
	Literatura	Pesquisador
Informações sobre a variável em estudo		
Média das diferenças a ser detectada (ou esperada)	✓	✓
Desvio padrão da diferença	✓	
Informações sobre a amostragem e o teste de hipótese		
Nível de significância		✓
Poder		✓
Perdas/recusa		✓
Tipo de Teste de acordo com hipótese alternativa		✓

Para exemplificar, suponha uma pesquisa em que se deseja verificar se a pressão arterial sistólica de 24 horas, medida pela monitorização ambulatorial da pressão arterial (MAPA), se altera após a realização de exercício aeróbico em intensidade moderada (carga de 60% de sua capacidade máxima) em pacientes hipertensos. Para exemplificar, serão extraídas informações do estudo de Moreira et al.¹¹.

Para o cálculo do tamanho desta amostra, será utilizada a opção “Dois Grupos Dependentes” do menu “Médias”, ilustrado na Figura 1, aba “TESTAR” (Figura 8).

Primeiramente informaremos os parâmetros esperados para a variável em estudo. No campo “média das diferenças a ser detectada” deve ser informada a média das diferenças entre as medidas inicial e final que se deseja encontrar. Recomenda-se que essa diferença tenha relevância clínica. No exemplo em questão, será utilizada a diferença de 5 mmHg. A seguir, deve ser informado o desvio padrão das diferenças entre os momentos. Caso a informação não seja fornecida pelo artigo e apenas o desvio padrão de cada momento é apresentado, é possível calcular o desvio padrão da diferença

clitando em “OBTER O DESVIO PADRÃO DA DIFERENÇA ENTRE GRUPOS PAREADOS”. Neste caso, também é necessário informar a correlação esperada entre os valores inicial e final, que, para o exemplo em questão, será utilizado como sendo de 0,5, por ser um valor intermediário entre zero e um. Cabe ressaltar que, quanto menor for o valor da correlação, maior será o tamanho da amostra, e vice-versa. O pesquisador pode utilizar diferentes valores de correlação e escolher aquele que produz um tamanho de amostra factível, sendo que esta escolha deve constar nas publicações onde for relatado o cálculo do tamanho da amostra. Corre-se o risco da correlação ser menor do que a escolhida e, neste caso, haver perda de poder ou aumento do nível de significância. Os valores de desvio padrão, necessários para o cálculo, podem ser encontrados na tabela 3 do referido artigo: 13,3

mmHg no momento basal e 12,9 mmHg após o exercício. Assim, o desvio padrão da diferença entre os momentos será de 13,1 mmHg.

Os próximos campos a serem preenchidos são referentes ao teste de hipótese em questão. Assim como no exemplo para comparação de duas médias independentes, será considerado 5% de significância e 80% de poder para um teste bilateral.

Por fim, também há a opção de incluir um percentual previsto para perdas ou recusas durante a pesquisa. Neste exemplo usamos a previsão de 10% de perdas/recusas, assim o tamanho final da amostra será ajustado de 56 pares para 63, a fim de contemplar este percentual. A Figura 8 ilustra o cálculo para este exemplo.

Este exemplo também pode ser visualizado no vídeo “PSS para comparação de médias em grupos dependentes”¹².

Comparação entre duas médias de grupos dependentes

Nesta técnica deseja-se calcular o tamanho amostral de um estudo cujo objetivo é comparar se as médias de dois grupos dependentes, ou seja, dois grupos relacionados ou comumente dito pareados, diferem ou não em relação à resposta de interesse, com um certo nível de significância e poder.

Testar
Poder

Hipóteses a serem testadas

$H_0 : \mu_D = 0$ vs $H_1 : \mu_D \neq 0$

Média das diferenças a ser detectada

Desvio padrão da diferença

Obter o desvio padrão da diferença entre grupos pareados

Nível de significância (%)

Poder (%)

Perdas/ Recusa (%)

Tipo de teste de acordo com hipótese alternativa

Tamanho amostral calculado: 56 pares

Sugestão de texto:

Foi calculado o tamanho de amostra para detectar que a média das diferenças de Y entre os grupos 1 e 2 é diferente de zero, tendo uma média das diferenças de 5 u.m. como relevante para o estudo. Para isso foi utilizado a ferramenta PSS Health versão on-line (citação abaixo). Considerando poder de 80%, nível de significância de 5% e desvio padrão da diferença esperado de 13.1 u.m. (dados de Fulano (1900)), chegou-se ao tamanho total de 56 pares. Acrescentando 10% para possíveis perdas e recusas o tamanho de amostra deverá ser 63 pares.

Sugestões de citação:

ABNT: BORGES, Rogério Boff et al. Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde. *Clinical & Biomedical Research*, [S.l.], v. 40, n. 4, apr. 2021. ISSN 2357-9730. Available at <https://seer.ufrgs.br/hcpa/article/view/109542>. Date accessed: 19 jul. 2021.

APA: Borges, R., Mancuso, A., Camey, S., Leotti, V., Hirakata, V., Azambuja, G., & Castro, S. (2021). Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde. *Clinical & Biomedical Research*, 40(4). Retrieved from https://seer.ufrgs.br/hcpa/article/view/109542

BibTex: @article{PSSHealth, author = {Rogério Borges and Aline Mancuso and Suzi Camey and Vanessa Leotti and Vânia Hirakata and Guilherme Azambuja and Stela Castro}, title = {Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde.}, journal = {Clinical & Biomedical Research}, volume = {40}, number = {4}, year = {2021}, keywords = {tamanho de amostra, poder do teste, estimação de parâmetros, comparação de grupos, R}, issn = {2357-9730}, url = {https://seer.ufrgs.br/hcpa/article/view/109542}}

Comando R utilizado:

```
stats::power.t.test(n = NULL, delta = abs(5), sd = 13.1, type = 'paired', sig.level = 0.05, power = 0.8, alternative = 'two.sided')
```

Figura 8 – Janela do PSS Health com os campos a serem preenchidos ao ser solicitado o cálculo de tamanho de amostra para comparação de duas médias dependentes.

Para comparação de dois grupos independentes com duas medidas repetidas (deltas)

Outro tipo de estudo da área da saúde acompanha grupos independentes de sujeitos em mais de um momento no tempo. Por exemplo, na avaliação basal e após tratamento em um grupo de homens e outro

de mulheres, o objetivo poderá ser a comparação da variação média da variável de interesse, isto é, da diferença entre a medida após o tratamento e a medida basal, entre os homens e as mulheres. Assim, o cálculo do tamanho da amostra deve ser realizado considerando a variável diferença entre as duas avaliações, usualmente conhecida como

delta (Δ). Em termos de um teste de hipótese bilateral, a hipótese nula é de que a média do Δ dos dois grupos são iguais e a hipótese alternativa é de que elas sejam diferentes.

O cálculo de tamanho da amostra será muito similar ao exemplo anterior, pois, em suma, também haverá uma média para cada grupo. Porém, neste caso, devemos considerar a correlação existente entre a primeira e a segunda avaliação. Além disso, as informações sobre a diferença a ser detectada (esperada) devem ser pensadas em termos de Δ e, assim como visto no exemplo anterior, ter relevância

clínica. Geralmente estão disponíveis na literatura, em artigos que tenham evidenciado uma diferença em grupos semelhantes aos que se deseja estudar, mas também podem ser definidos pelo pesquisador.

Logo, para o então cálculo de tamanho de amostra para comparar a média do delta (Δ) de uma variável quantitativa medida em dois momentos entre dois grupos independentes, a ferramenta PSS Health apresenta a opção “Delta de dois grupos independentes”, ainda no menu Médias, que também pode ser visualizado na Figura 1. O Quadro 3 apresenta as informações necessárias para este cálculo.

Quadro 3 – Informações necessárias para o cálculo de tamanho de amostra para comparação de médias de dois deltas de grupos independentes e respectiva fonte de obtenção.

Informação	Fonte	
	Literatura	Pesquisador
Informações sobre a variável em estudo		
Diferença esperada/desejada entre os deltas OU mudança média (delta) de Y ao longo do tempo do grupo	✓	✓
Desvio padrão de Y no início do estudo	✓	
Desvio padrão de Y no final do estudo	✓	
Correlação das medidas de Y (entre início e final) OU desvio padrão da diferença de Y	✓	✓
Informações sobre a amostragem e o teste de hipótese		
Balanceamento		✓
Poder		✓
Nível de significância		✓
Perdas/recusas		✓

Assim, diferente do exemplo anterior, cujo objetivo era comparar médias de dois grupos independentes, agora o objetivo é comparar as diferenças ocorridas em dois momentos (Δ) de dois grupos independentes. Para exemplificar, será realizado o cálculo de tamanho de amostra mínimo necessário para detectar uma diferença, clinicamente relevante, na variação da pressão arterial sistólica (PAS) avaliada antes e após exercícios físicos, entre dois grupos de pacientes hipertensos que receberão diferentes cargas de exercícios (G1 e G2). Considere as informações apresentadas na tabela 2 do artigo de Moreira et al.¹¹, em que a média inicial da PAS no grupo G1 foi de 137,2 (DP = 14,9 mmHg) e do G2 foi de 144,4 (DP = 13,3 mmHg) e a média final do G1 foi de 135,2 (DP = 12,7 mmHg) e do G2 foi de 138,6 (DP = 12,9 mmHg).

Conforme ilustrado no quadro à esquerda da Figura 9, a primeira informação solicitada pela ferramenta é se serão utilizados os valores dos deltas de cada grupo para o cálculo. Deve-se assinalar a

opção se a informação a ser inserida for a média dos deltas de cada grupo ou manter desmarcada para informar apenas a diferença esperada/desejada entre os deltas. Como no exemplo em questão não é informado o valor do delta das medições, e sim as médias e desvios observados em cada momento, utilizaremos uma diferença esperada de 5 mmHg, pois é o valor que os próprios autores do trabalho utilizaram como sendo uma diferença clinicamente relevante. Seguindo, a ferramenta solicita o desvio padrão esperado no início e no final do estudo. Caso este valor não esteja disponível na literatura, é possível calculá-lo a partir da combinação de dois desvios padrão. Para tanto, na própria ferramenta, utiliza-se a opção “DESVIO PADRÃO COMBINADO” no menu “OUTRAS FERRAMENTAS”, assinalando a opção “ENTRAR COM OS VALORES DO DESVIO PADRÃO”. Os valores resultantes para o exemplo utilizado foram 14,1 (pré) e 12,8 (pós).

Outra informação necessária para o cálculo é a correlação entre as medidas pré e pós. Quando não houver referência na literatura sobre este valor, deve-se usar a estratégia descrita para a comparação de médias de dois grupos dependentes. É possível calcular a correlação se as informações dos desvios padrão pré, pós e

do delta intra-grupo estiverem disponíveis. Neste exemplo, utilizaremos novamente o valor de 0,5. Na Figura 9, observa-se os parâmetros utilizados e o resultado encontrado, e no vídeo “PSS para comparação de médias de grupos com 2 medidas repetidas através do delta”¹³, a demonstração do exemplo apresentado.

Comparação da mudança média ao longo do tempo de grupos independentes

Nesta técnica deseja-se calcular o tamanho amostral de um estudo cujo objetivo é comparar se a mudança média em um intervalo de tempo (Δ) de dois grupos independentes diferem.

Hipóteses a serem testadas

$H_0 : \bar{\Delta}_{Tratamento} = \bar{\Delta}_{Controle}$

$H_1 : \bar{\Delta}_{Tratamento} \neq \bar{\Delta}_{Controle}$

Mudar nomes

Utilizar os valores dos deltas de cada grupo

Diferença esperada/ desejada entre os deltas (em mm Hg)

5

Desvio padrão esperado de pressão arterial sistólica no

Início do estudo	Final do estudo
14,1	12,8

Usar o desvio padrão da diferença

Correlação das medidas de pressão arterial sistólica (início e fim)

0,5

Balaceamento (Tratamento:Controle)

1

Poder (%)

80

Nível de significância (%)

5

Perdas/ Recusa (%)

10

Tamanho amostral calculado: 230 (115 no Tratamento e 115 no Controle)

Sugestão de texto:

Foi calculado o tamanho de amostra para detectar diferenças, na mudança média de **pressão arterial sistólica**, em um intervalo de tempo **T** entre os grupos Tratamento e o Controle, por meio da ferramenta PSS Health versão on-line (citação abaixo). Considerando uma diferença esperada/ desejada de **5 mm Hg**, poder de **80%**, nível de significância de **5%**, correlação entre as medidas de **0,5**, desvios padrão de **14,1 e 12,8 mm Hg** no início e no final do estudo, respectivamente (dados de Fulano (1900)), chegou-se ao tamanho total da amostra de **230** sujeitos, sendo **115** sujeitos em cada grupo. Acrescentando **10%** para possíveis perdas e recusas o tamanho de amostra deverá ser **256**.

Sugestões de citação:

ABNT: BORGES, Rogério Boff et al. Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde. *Clinical & Biomedical Research*, [S.l.], v. 40, n. 4, apr. 2021. ISSN 2357-9730. Available at: <https://seer.ufrgs.br/hcpa/article/view/109542>. Date accessed: 10 dez. 2021.

APA: Borges, R., Mancuso, A., Camey, S., Leotti, V., Hirakata, V., Azambuja, G., & Castro, S. (2021). Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde. *Clinical & Biomedical Research*, 40(4). Retrieved from https://seer.ufrgs.br/hcpa/article/view/109542

BibTex: @article{PSSHealth, author = {Rogério Borges and Aline Mancuso and Suzi Camey and Vanessa Leotti and Vânia Hirakata and Guilherme Azambuja and Steia Castro}, title = {Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde}, volume = {40}, number = {4}, year = {2021}, keywords = {tamanho de amostra, poder do teste, estimação de parâmetros, comparação de grupos, R}, issn = {2357-9730}, url = {https://seer.ufrgs.br/hcpa/article/view/109542}}

Comando R utilizado:

```
epIR::epI.sscomp(control = 0, treat = 5, sigma = 13.497036711812, n = NA, power = 80/100, conf.level = 1 - 5/100, r = 1, design = 1, sided.test = 2)
```

*** Sempre procure um profissional de estatística para orientações no planejamento do estudo.**

Figura 9 – Comparação da mudança média ao longo do tempo de grupos independentes.

Para comparação de dois grupos independentes com duas ou mais medidas repetidas

Diferente do exemplo anterior em que se tinha apenas dois momentos, portanto, era possível o cálculo do delta, muitas vezes, o objetivo do estudo envolve comparar o comportamento de uma variável ao longo de diversos momentos entre dois grupos independentes. Por exemplo, um ensaio clínico com dois grupos, intervenção e controle, onde se quer avaliar se a variável que mede o efeito da intervenção se comporta da mesma forma que no grupo controle em dois ou mais momentos de avaliação. Um dos

métodos estatísticos que atende a este objetivo, dentre outros, é o modelo misto no contexto de medidas repetidas¹⁴ no tempo para testar os efeitos principais dos fatores grupo e tempo, e o efeito da interação entre eles. A ferramenta realiza o cálculo do tamanho de amostra baseada em um modelo misto onde o tempo é tratado como um fator (a variável tempo é assumida como categórica) e, então, faz-se necessário definir a diferença média mínima a ser detectada entre os dois grupos independentes no último momento da coleta de dados. Para a realização deste cálculo, as informações necessárias constam no Quadro 4.

Quadro 4 – Informações necessárias para o cálculo de tamanho de amostra para comparação de 2 ou mais medidas repetidas entre dois grupos independentes.

Informação	Fonte	
	Literatura	Pesquisador
Informações sobre a variável em estudo		
Número de momentos a serem avaliados		✓
Diferença do Y a ser detectada no último tempo de avaliação: média do tratamento – média do grupo controle	✓	✓
Desvio padrão esperado de Y no grupo tratamento	✓	
Desvio padrão esperado de Y no grupo controle	✓	
Tipo de matriz de correlação e parâmetros	✓	✓
Retenção esperada em cada momento e em cada grupo	✓	✓
Informações sobre a amostragem e o teste de hipótese		
Balanceamento		✓
Poder		✓
Nível de significância		✓
Perdas/recusas		✓

Suponha que se deseje realizar uma pesquisa para comparar o comportamento da pressão arterial sistólica durante 12 meses, com avaliação basal em 3, 6, 9 e 12 meses, em pacientes com diabetes que utilizam uma determinada medicação adicional à metformina (grupo A) comparada com um grupo de pacientes que utiliza apenas metformina (grupo B). Para exemplificar, serão utilizadas as informações retiradas do artigo de Derosa et al.¹⁵.

Para o cálculo do tamanho desta amostra, será utilizada a opção “Medidas repetidas” do menu “Médias”, ilustrado na Figura 1, aba “TESTAR”. Na janela aberta, o número de momentos a serem avaliados é definido pelo pesquisador de acordo com a necessidade, incluindo a medida basal, para o exemplo será digitado 5 (tempos basal, 3, 6, 9 e 12 meses). A seguir, deve-se informar a diferença média mínima na pressão arterial sistólica entre os grupos que se deseja encontrar no último momento de avaliação, lembrando que a mesma deve ter relevância clínica. Assim como anteriormente, será considerada uma diferença de 5 mmHg aos 12 meses de avaliação. Os desvios padrão esperados dos grupos A e B serão obtidos dos momentos basais de coleta (por serem os maiores encontrados em todo o período estudado, retirados da tabela 3 do artigo de Derosa¹⁵, lembrando que quanto maior a variabilidade, maior o tamanho da amostra), cujos valores são 7,4 e 6,1 mmHg. Quando um deles não for encontrado (ou o desvio padrão da intervenção ou do controle) pode se usar o desvio padrão encontrado nos dois campos.

Assim como na comparação de dois grupos avaliados em dois momentos, as medidas do mesmo sujeito são correlacionadas, logo, faz-se necessário informar as correlações entre estas medidas. Isto é feito através de uma matriz de correlações entre os momentos, a qual vai indicar o quanto o pesquisador espera que a medida do momento 1 esteja correlacionada com as medidas dos momentos 2, 3, 4 e 5, assim como a medida do momento 2 com as medidas dos momentos 3, 4 e 5, e assim sucessivamente. A ferramenta oferece três alternativas para as estruturas de correlação:

- Componente permutável: a correlação entre as medidas é sempre a mesma em todos os momentos;
- AR(1) – Auto-Regressiva de 1ª Ordem: a correlação entre medidas de momentos adjacentes é mais forte e vai enfraquecendo de acordo com o aumento da distância entre as mesmas;
- Não estruturada: as correlações entre as medidas de diferentes momentos podem assumir qualquer valor.

Esta não é uma informação que esteja disponível nos artigos em geral, portanto, sugere-se fazer o cálculo utilizando Componente permutável com 0,5 em cada grupo.

Na informação sobre a retenção esperada, deve ser incluído o percentual de pacientes que se espera que continuem no estudo em todos os momentos de avaliação. Esta informação deve ser pensada com cautela pelo pesquisador, e depende de várias condições da pesquisa (adesão/vínculo, gravidade). Por exemplo,

se os pacientes forem muito graves em sua linha de base, isto pode gerar um grupo com muita mortalidade, a retenção diminuirá ao longo do tempo. No exemplo, será considerada a retenção obtida no artigo de referência, que foram de 97,7% (84/86), 97,7%, 96,5% (83/86) e 94,2% (81/86) no Grupo A, e 98,8% (84/85), 97,6% (83/85), 96,5% (82/85) e 96,5% no Grupo B. As demais informações (balanceamento, nível de significância e poder) podem seguir as mesmas instruções dos modelos prévios. Este exemplo está ilustrado no vídeo “PSS para Comparação de Médias em grupos independentes com 2 ou mais medidas repetidas”¹⁶.

CONSIDERAÇÕES FINAIS

O PSS Health é uma ferramenta estatística, voltada aos pesquisadores da área da saúde, de fácil

acesso e utilização, com potencial para o ensino e a aprendizagem de quem deseja definir o tamanho amostral de um estudo ou calcular o poder de um teste de hipótese. No entanto, uma questão importante a ser definida no início do planejamento de um estudo, antes da realização do cálculo para tamanho amostral, é o seu objetivo principal. Ele será a questão norteadora não apenas para o delineamento e a metodologia, mas também para a análise estatística a ser empregada e, conseqüentemente, o tipo de cálculo amostral a ser realizado. Como pode ser observado nos exemplos apresentados, uma pequena mudança no objetivo ou no delineamento influencia diretamente no cálculo do tamanho amostral, mesmo se o desfecho permanecer o mesmo. Assim, ressalta-se a importância da correta realização do cálculo, sendo aconselhável uma consultoria estatística no caso de dúvidas.

REFERÊNCIAS

- Borges RB, Mancuso ACB, Camey SA, Leotti VB, Hirakata VN, Azambuja GS, et al. Power and Sample Size for Health Researchers: uma ferramenta para cálculo de tamanho amostral e poder do teste voltado a pesquisadores da área da saúde. Clin Biomed Res [Internet]. 2020 [citado 21 de maio de 2021];40(4). Disponível em: <https://seer.ufrgs.br/hcpa/article/view/109542>
- R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Disponível em: <https://www.R-project.org/>
- Borges R, Azambuja G, Mancuso A, Leotti V, Hirakata V, Camey S, et al. PSS.Health: Power and Sample Size for Health Researchers [Internet]. 2020 [citado 13 de novembro de 2020]. Disponível em: <https://CRAN.R-project.org/package=PSS.Health>
- Hirakata VN, Mancuso ACB, Castro SM de J. Teste de Hipóteses: Perguntas que você sempre quis fazer, mas nunca teve coragem. Clin Biomed Res [Internet]. 20 de agosto de 2019 [citado 29 de outubro de 2020];39(2). Disponível em: <https://seer.ufrgs.br/hcpa/article/view/93649>
- Hulley SB, Browner WS, Cummings SR, Grady D, Newman TB. Delineando a pesquisa clínica: uma abordagem epidemiológica [Internet]. ARTMED; 2008. Disponível em: <http://books.google.com.br/books?id=bshuPgAACAAJ>
- Castro SM de J, Mancuso ACB, Leotti VB, Hirakata VN, Camey SA. Bioestatística e Epidemiologia: Perguntas que você sempre quis fazer, mas nunca teve coragem. Clin Biomed Res Vol 39 No 3 2019 [Internet]. 2019; Disponível em: <https://seer.ufrgs.br/hcpa/article/view/96394>
- Castro SM de J, Branco AC, Camey SA, Leotti VB, Hirakata VN, Borges RB. PSS Health : como calcular tamanho de amostra para estimar média, proporção e correlação. Clin Biomed Res [Internet]. 5 de outubro de 2021 [citado 3 de novembro de 2021];41(3). Disponível em: <https://seer.ufrgs.br/hcpa/article/view/112466>
- Garcia FD, Terra AF, Queiroz AM, Correia CA, Ramos PS, Ferreira QT, et al. Evaluation of risk factors associated with increased blood pressure in children. J Pediatr (Rio J). 15 de janeiro de 2004;80(1):29–34.
- Lindenau JD-R, Guimarães LSP. Calculando o tamanho de efeito no SPSS. 2012 [Internet]. 2012;32. Disponível em: <http://seer.ufrgs.br/hcpa/article/view/33160>
- PSS para Comparação de Médias em grupos independentes [Internet]. [citado 4 de novembro de 2021]. (PSS Health). Disponível em: <https://www.youtube.com/watch?v=LpwwqjPkJk>
- Moreira WD, Fuchs FD, Ribeiro JP, Appel LJ. The Effects of Two Aerobic Training Intensities on Ambulatory Blood Pressure in Hypertensive Patients. J Clin Epidemiol. julho de 1999;52(7):637–42.
- PSS para Comparação de Médias em grupos dependentes [Internet]. [citado 4 de novembro de 2021]. (PSS Health). Disponível em: <https://www.youtube.com/watch?v=mp8qbyUsqV0>
- PSS para Comparação de Médias em grupos independentes com 2 medidas repetidas através do delta [Internet]. [citado 4 de novembro de 2021]. (PSS Health). Disponível em: <https://www.youtube.com/watch?v=c7zzQxgeaS4>
- Brown H, Prescott R. Applied Mixed Models in Medicine [Internet]. Wiley; 2014. (Statistics in Practice). Disponível em: <https://books.google.com.br/books?id=lrPbBQAAQBAJ>
- Derosa G, Cicero AFG, Franzetti IG, Querci F, Carbone A, Ciccarelli L, et al. Effects of exenatide and metformin in combination on some adipocytokine levels: a comparison with metformin monotherapy. Can J Physiol Pharmacol. setembro de 2013;91(9):724–32.
- PSS para Comparação de Médias em grupos independentes com 2 ou mais medidas repetidas [Internet]. [citado 4 de novembro de 2021]. (PSS Health). Disponível em: https://www.youtube.com/watch?v=oErZbk_cpWY

Recebido: 23 dez, 2021

Aceito: 21 mar, 2022