

# MAT02025 - Amostragem 1

Experimentos, estudos observacionais e levantamentos

Rodrigo Citton P. dos Reis  
[citton.padilha@ufrgs.br](mailto:citton.padilha@ufrgs.br)

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DEPARTAMENTO DE ESTATÍSTICA

Porto Alegre, 2023

## Delineamento de pesquisa

# Introdução

- ▶ A realização de pesquisas empíricas levanta questões relativas ao **plano de observação da realidade** e à **escolha de métodos** para processamento e análise de dados.
- ▶ Estudar essas questões significa buscar opções lógicas e operacionais para a **composição de delineamentos** que assegurem o valor científico das informações obtidas e das conclusões alcançadas.

# Introdução

- ▶ Entende-se por **delineamento** o **plano estratégico de observação da realidade** que orienta o detalhamento posterior dos métodos e técnicas necessários à execução da pesquisa.
- ▶ De acordo com a natureza lógica e estrutural desse plano, as pesquisas podem ser reconhecidas como **experimentos, quase experimentos** ou **estudos observacionais**.

## Experimentos

- ▶ **Experimento** é o tipo de pesquisa em que o investigador controla a ocorrência das variáveis independentes (causa-fator-tratamento) com a finalidade de observar seus efeitos sobre as variáveis consideradas dependentes (efeito-resposta-doença).
- ▶ A validade dos efeitos detectados é assegurada por meio da **alocação aleatória** das unidades experimentais aos níveis definidos para as variáveis independentes.
- ▶ O **princípio da aleatorização** (também chamado de **princípio da randomização**) representou uma significativa contribuição da estatística às ciências empíricas, indicando caminhos para o desenvolvimento de planos de observação resistentes às diferentes fontes de erro, que podem comprometer a validade dos resultados.

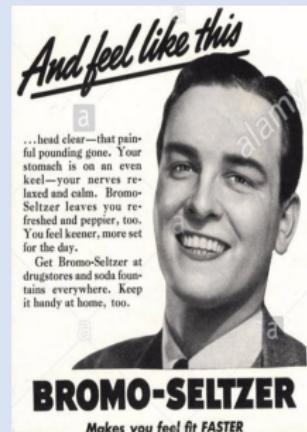
# Experimentos



*Do this*



... snap out of it the way doctors prove faster. Take Bromo-Seltzer! It does more for you. Not only relieves headache pain, but settles the stomach—calms the nerves—ALKALIZES! Soon you're smiling ...



# Experimentos

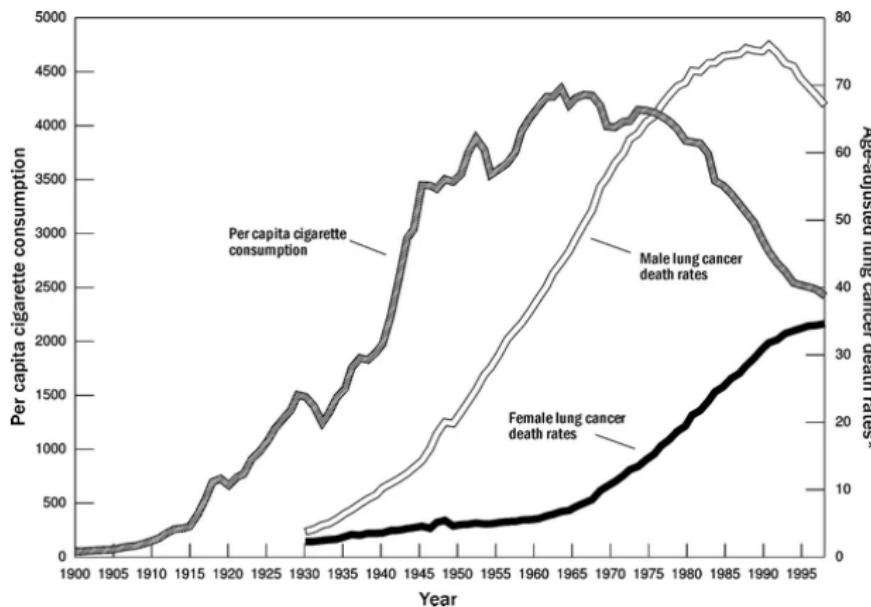
## Exemplo

- ▶ 12 pessoas sofrendo de dor cabeça.
- ▶ Algumas tomam aspirina ( $Z_i = 1$ ); outras não ( $Z_i = 0$ ).
  - ▶ Em um **delineamento experimental**, os grupos de **tratamento** e **controle** são aleatorizados. Por exemplo, cada indivíduo é alocado ao grupo de tratamento com  $\Pr(Z_i = 1) = 1/2$ .
- ▶ Uma hora depois, perguntamos para cada uma delas se a dor de cabeça sumiu ( $Y_i = 0$  se a dor não sumiu,  $Y_i = 1$  se a dor sumiu).

## Estudos observacionais

- ▶ Nas pesquisas em que a unidade de estudo é o ser humano (ou seres vivos), os investigadores, por questões éticas e práticas, não podem controlar a ocorrência da variável independente ou tratamento nem proceder à alocação aleatória das unidades observadas aos grupos de comparação, optando, assim, por delineamentos considerados **não experimentais (estudos observacionais)**.
  - ▶ Estudos de coorte;
  - ▶ Estudos do tipo caso-controle;
  - ▶ Estudos de cortes transversais ou levantamentos.

## Exemplo: “onde há fumaça ...”

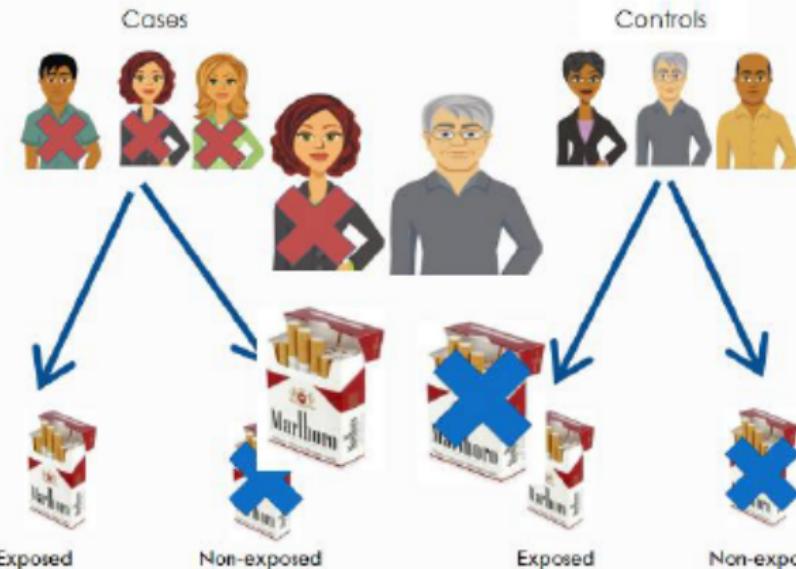


**Figura 1:** Uso de tabaco e mortalidade por câncer de pulmão, EUA, 1900-1997.  
Asterisco (\*) Por 100.000 e ajustado por idade para a população padrão dos EUA em 1970.

## Exemplo: “onde há fumaça ...”

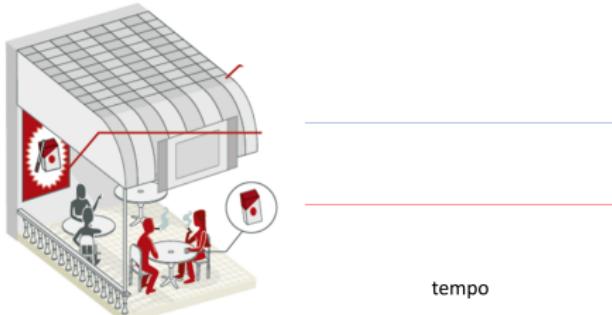
### Estudo caso-controle

#### Steps in a case-control study

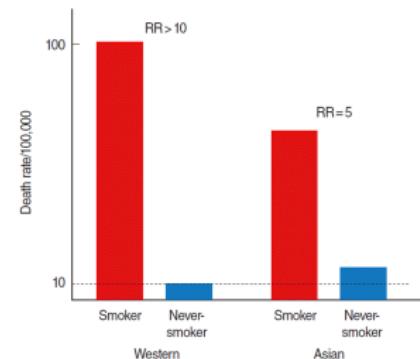


## Exemplo: “onde há fumaça ...”

### Estudo de coorte

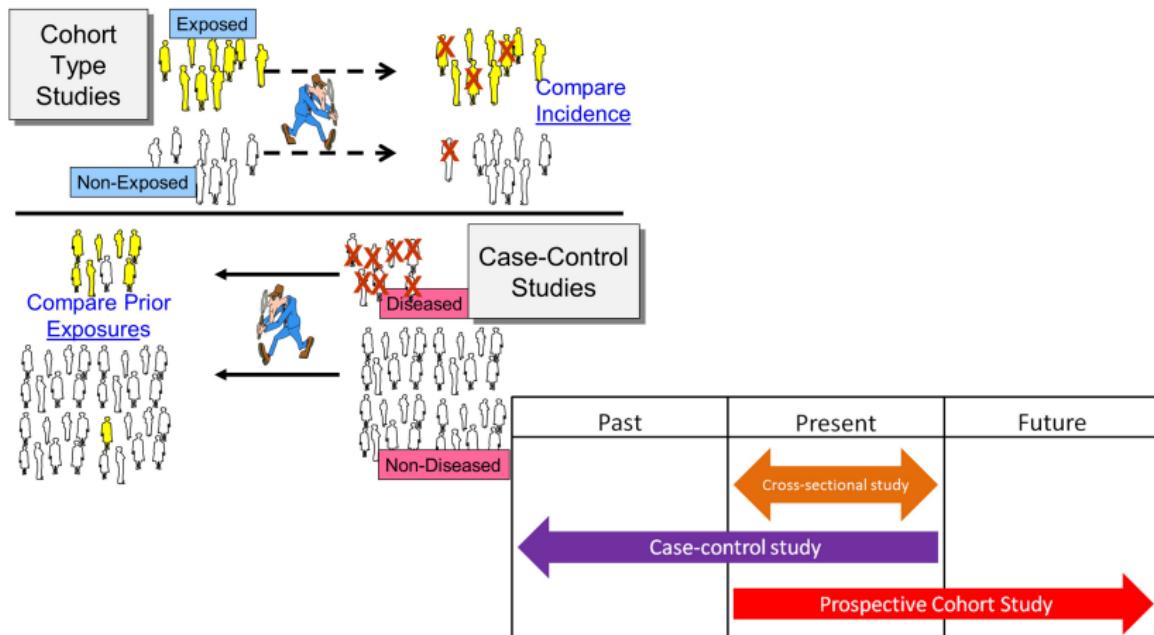


### Estudos de coorte



Fonte: Jung KJ, Jeon C, Jee SH. The effect of smoking on lung cancer: ethnic differences and the smoking paradox. *Epidemiol Health*. 2016;38:e2016060. doi: 10.4178/epih.e2016060

# Estudos observacionais



# Estudos observacionais

## Exemplo

- ▶ 12 pessoas sofrendo de dor cabeça.
- ▶ Algumas tomam aspirina ( $Z_i = 1$ ); outras não ( $Z_i = 0$ ).
  - ▶ Em um **delineamento observacional**, os grupos de **tratamento** e **controle** **não** são aleatorizados. Os indivíduos podem ser “alocados” ao grupo de tratamento devido a influência de covariáveis pré-tratamento ( $X_i$ ). Por exemplo, a **gravidade da dor de cabeça** pode influenciar na decisão de tomar ou não aspirina.
- ▶ Uma hora depois, perguntamos para cada uma delas se a dor de cabeça sumiu ( $Y_i = 0$  se a dor não sumiu,  $Y_i = 1$  se a dor sumiu).
  - ▶ Note que as covariáveis que afetam o tratamento, também podem influenciar o desfecho ( $Y$ ).

## Estudos observacionais

### Exemplo (cont.): delinamentos experimentais vs. observacionais

- ▶ Entre as pessoas que tomaram aspirina, a proporção de “dor de cabeça sumiu” foi 1/12; entre as que não tomaram aspirina, a proporção foi 4/12.
- ▶ **Qual a(s) conclusão(ões) podemos elaborar a respeito do tratamento “tomar aspirina”?**

# Levantamentos por amostragem

# Levantamentos por amostragem

- ▶ “Levantamento por amostragem” é a tradução ao português da expressão em língua inglesa *survey sampling* ou *sample surveys*.
- ▶ Com a finalidade de produzir “instantâneos” das realidades estudadas, os levantamentos por amostragem reúnem as seguintes características:
  1. Aplicam-se a **conjuntos reais e finitos** (na maioria dos casos), compostos de elementos, denominados **população de estudo**.
  2. Os elementos podem ser seres humanos, animais, árvores, fichas, prontuários, domicílios, áreas ou objetos.
  3. As características ou atributos são observados em cada elemento e, posteriormente, agregados por meio de medidas estatísticas chamadas **parâmetros** ou **valores populacionais**.
  4. Os dados são coletados em **amostras** das populações de estudo, e as medidas calculadas (estimativas) passam a ser a informação disponível para os valores populacionais desconhecidos.

# Levantamentos por amostragem

## Exemplo

- ▶ O **programa de controle da dengue no estado de São Paulo** exemplifica o uso do levantamento por amostragem como instrumento para conhecer o nível de infestação do *Aedes aegypti* em diversos municípios.
  - ▶ O conjunto de edificações localizadas em setores censitários constitui a **população de estudo**;
  - ▶ O **valor populacional desconhecido** é o número médio de recipientes com larvas do mosquito.
- ▶ Desse modo, são sorteadas amostras desse conjunto, registrando-se o número de recipientes existentes e infestados, para, a partir desses dados, ser calculado o **Índice de Breteau**.

## Levantamentos descritivos vs. analíticos

- ▶ Os levantamentos podem ter a finalidade **descritiva**, limitando-se a estimar frequências de elementos com determinada propriedade ou médias e variâncias de características quantitativas.
  - ▶ Ex: prevalência de diabetes na população brasileira de adultos (**proporção, frequência, erro padrão, intervalo de confiança**).
- ▶ Outros levantamentos, denominados **analíticos (ou investigações)**, definem grupos de comparação e, além de estimar, procuram detectar relações entre as características, buscando aumentar as explicações para o objeto pesquisado.
  - ▶ Ex: associação entre índice de massa corporal (IMC) e prevalência de diabetes (**razão de prevalências [RP] entre grupos de IMC, intervalo de confiança para RP**).
- ▶ Na prática, a maioria dos levantamentos cumpre as duas finalidades por meio de métodos de estimação e de técnicas de análise estatística multivariada.

# Levantamentos descritivos vs. analíticos

**CSP** CADERNOS DE SAÚDE PÚBLICA  
REPORTS IN PUBLIC HEALTH

## Evolução do diabetes mellitus no Brasil: dados de prevalência da Pesquisa Nacional de Saúde de 2013 e 2019

Table 2

Prevalence of diabetes mellitus in adults overall and by sex according to sociodemographic and clinical characteristics. Brazilian National Health Survey, 2013 and 2019.

Characteristic	2013				2019							
	Total	%	95%CI	Men	Women	Total	%	95%CI	Men	Women		
<b>Total</b>	6.2	5.9-6.6	5.3	4.8-5.8	7.0	6.5-7.5	7.7	7.4-8.0	6.9	6.5-7.4	8.4	8.0-8.8
Age (years)												
18-24	0.5	0.3-0.8	0.4	0.1-0.7	0.6	0.2-1.1	0.7	0.4-1.1	1.0	0.4-1.7	0.4	0.2-0.6
25-34	0.8	0.6-1.0	0.7	0.4-1.1	0.9	0.6-1.2	0.9	0.7-1.2	0.7	0.3-1.0	1.1	0.8-1.5
35-44	2.9	2.4-3.4	2.4	1.7-3.2	3.4	2.6-4.1	3.1	2.7-3.6	3.3	2.5-3.8	3.2	2.5-3.8
45-54	6.6	5.8-7.4	5.8	4.6-6.9	7.3	6.2-8.4	7.7	6.8-8.5	6.9	5.8-8.0	8.4	7.2-9.6
55-64	13.5	12.1-14.9	12.1	9.8-14.3	14.7	12.8-16.6	14.8	13.8-15.7	13.6	12.2-15.0	15.8	14.5-17.1
≥ 65	19.8	18.2-21.3	17.7	15.1-20.3	21.4	19.3-23.5	21.6	20.5-22.7	20.2	18.6-21.9	22.7	21.2-24.1

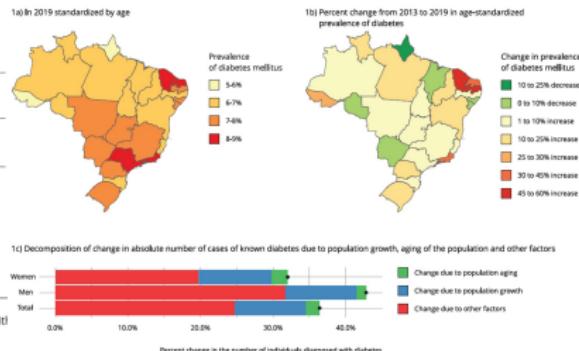
Table 3

Crude and adjusted \* prevalence ratio (PR) for diabetes in adults, overall and stratified by sex, in accordance with characteristics. Brazilian National Health Survey, 2019.

Characteristic	Total				Men				Women			
	Crude		Adjusted		Crude		Adjusted		Crude		Adjusted	
	PR	95%CI	PR	95%CI	PR	95%CI	PR	95%CI	PR	95%CI	PR	95%CI
Age (years)												
18-24	1.00		1.00		1.00		1.00		1.00		1.00	
25-34	1.25	0.74-2.13	1.30	0.77-2.22	0.64	0.29-1.44	0.65	0.29-1.44	2.77	1.54-4.98	2.95	1.64-5.31
35-44	4.34	2.69-7.00	4.41	2.74-7.10	3.01	1.59-5.68	3.03	1.61-5.70	7.74	4.56-13.16	7.92	4.66-13.49
45-54	10.64	6.73-16.84	10.41	6.58-16.46	6.69	3.69-12.12	6.68	3.68-12.13	20.51	12.30-34.22	19.66	11.08-32.71
55-64	20.36	12.84-32.27	19.45	12.26-30.88	13.07	7.08-24.12	12.90	6.96-23.88	38.56	23.50-63.25	35.65	21.67-58.66
≥ 65	29.82	18.95-46.92	27.20	17.25-42.90	19.47	10.65-35.59	19.09	10.37-35.12	55.44	33.82-90.87	47.35	28.75-77.98

Figure 1

Prevalence of diabetes in adults ages 18 and over. Brazilian National Health Survey, 2013 and 2019.



**Fonte:** Reis, R. C. P. dos, Duncan, B. B., Malta, D. C., Iser, B. P. M., & Schmidt, M. I. (2022). Evolution of diabetes in Brazil: prevalence data from the 2013 and 2019 Brazilian National Health Survey. *Cadernos de Saúde Pública*, 38(suppl 1). <https://doi.org/10.1590/0102-311x00149321>

## Vantagens do método das amostragens

## Uma pequena digressão

### Como obter uma boa amostragem e como tirar conclusões adequadas dos resultados?

Isso não importa, desde que o material do qual retiramos as amostras seja **homogêneo**, pois qualquer tipo de amostragem dá quase o mesmo resultado. Os diagnósticos de laboratório sobre nosso estado de saúde são feitos com apenas algumas gotas de sangue. Esse processo fundamenta-se na presunção de que o sangue em circulação está sempre bem misturado e que uma gota conta a mesma história que qualquer outra. Entretanto, quando o material está longe de ser homogêneo, **como acontece frequentemente**, o processo pelo qual se obtém a amostra se torna crítico e o estudo das técnicas que assegurem amostras dignas de confiança é importante.

## Custo reduzido

- ▶ Como os dados são obtidos apenas de uma pequena fração do agregado, as despesas são menores que se for empreendido um censo integral.
  - ▶ A Pesquisa Nacional de Saúde (PNS) foi uma pesquisa domiciliar nacional realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em parceria com o Ministério da Saúde. A PNS-2013 visitou 81.167 domicílios, e entrevistou 60.202 indivíduos com relação a percepção de saúde, violência e acidentes, estilos de vida e doenças crônicas. Peso, altura e pressão arterial foram medidos. (**Compare o tamanho da amostra de indivíduos em relação ao tamanho da população brasileira de adultos**)

## Maior rapidez

- ▶ Pelo mesmo motivo, os dados podem ser coletados e sintetizados mais rapidamente com uma amostragem, que com uma contagem completa.
  - ▶ Isso é um fator primordial, quando se necessita urgentemente das informações.

## Maior amplitude

- ▶ Em certos tipos de investigação, tem-se que utilizar pessoal perfeitamente treinado e equipamento altamente especializado<sup>1</sup>, cuja disponibilidade é limitada, para a obtenção dos dados.
  - ▶ O censo completo torna-se **impraticável**.
- ▶ Levantamentos por amostragem têm maior amplitude e flexibilidade.

---

<sup>1</sup>Veja o Manual de padronização e treinamento de Antropometria da PNS-2013:  
<https://www.pns.icict.fiocruz.br/wp-content/uploads/2021/02/Manual-de-Antropometria-e-de-Medida-de-Pressao-Arterial-PNS-2013.pdf>

## Maior exatidão

- ▶ Em virtude de se poder雇用 pessoal intensivamente treinado, e por se tornar exequível a supervisão mais cuidadosa do campo de trabalho e do processamento dos dados, dada a redução do volume de trabalho, uma amostragem pode, na realidade, proporcionar **resultados mais exatos** que a espécie de contagem integral que se consegue realizar.
  - ▶ Paradoxo? (“Fazer mais com menos”??)

# Exercício (ConVid)



ConVid Pesquisa de  
Comportamentos

## Pesquisa ConVid

A chegada da pandemia provocada pelo novo coronavírus no Brasil levou a uma série de iniciativas e de recomendações para a proteção das pessoas, que incluiu o isolamento social/quarentena. A presente pesquisa da Fundação Oswaldo Cruz, realizada em parceria com a Universidade Federal de Minas Gerais, e a Universidade Estadual de Campinas tem a finalidade de verificar como a pandemia afetou ou mudou a sua vida.

Os resultados da ConVid Pesquisa de Comportamentos poderão contribuir para uma melhor compreensão do impacto da quarentena sobre a saúde dos brasileiros para orientar as ações de saúde dirigidas a minimizar os efeitos adversos trazidos pelo isolamento social prolongado.

Leia mais em <https://convid.fiocruz.br/>

## Exercício (ConVid)

No exemplo da **pesquisa ConVid**, responda:

- ▶ Quem é a população de estudo?
- ▶ Os elementos são ...
- ▶ Quais características ou atributos são observados em cada elemento?
- ▶ Quais as possíveis quantidades populacionais podem ser estimadas a partir da pesquisa?
- ▶ Compartilhe e discuta com os colegas no Fórum Geral do Moodle.

## Para casa

1. Faça uma pequena busca sobre levantamentos amostrais realizados no Brasil (podem ser um levantamento realizado apenas na sua cidade, estado, ou ainda, de abrangência nacional). Escolha um levantamento do seu interesse.
2. Descreva o levantamento escolhido e apresente no Fórum Geral do Moodle.

## Próxima aula

- ▶ Principais fases de um levantamento por amostragem.

# Por hoje é só!

Bons estudos!

