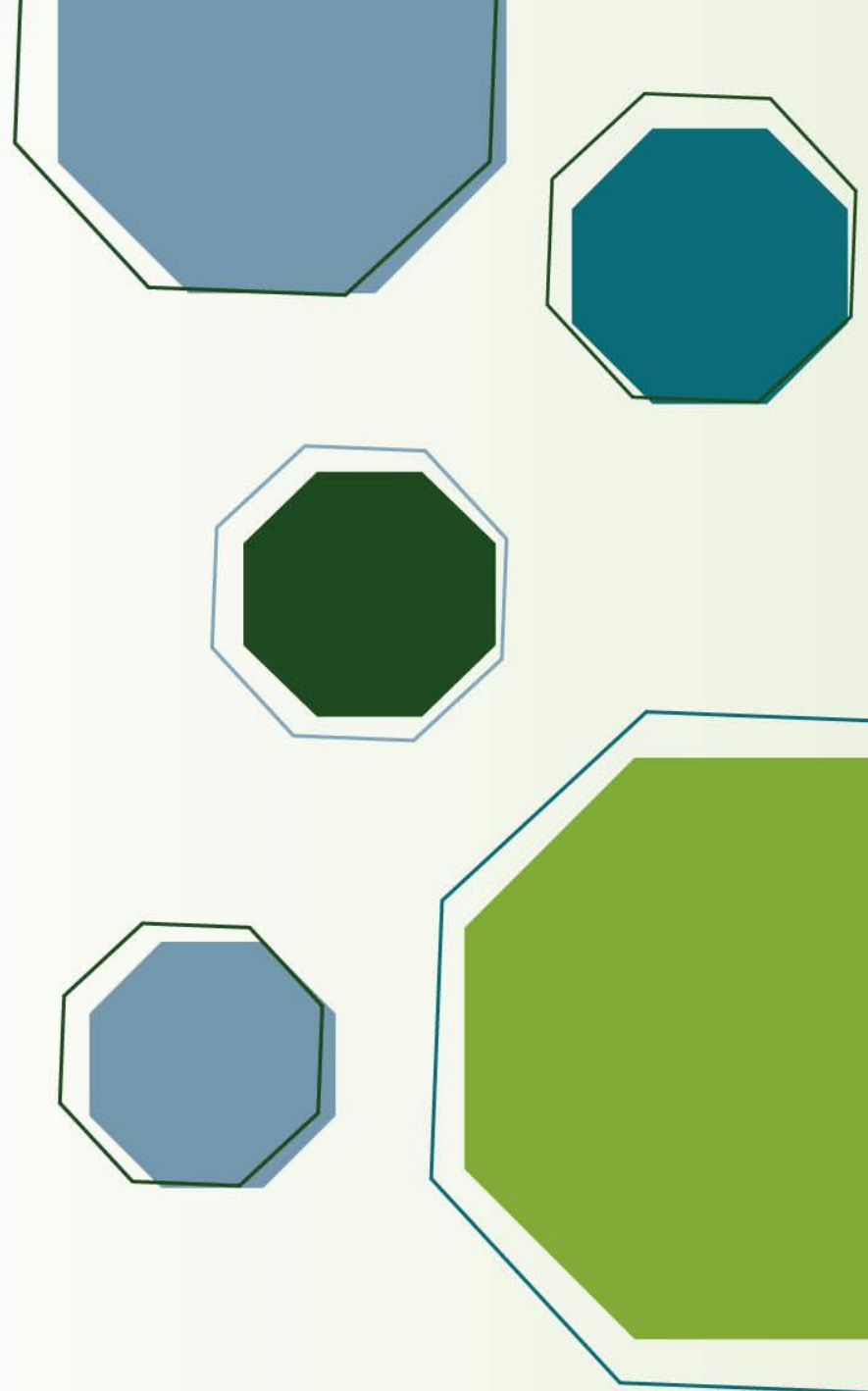




INSTITUTO DO
CÂNCER
DO ESTADO DE
SÃO PAULO
OCTAVIO FRIAS DE OLIVEIRA



Mestrado Profissional em Inovação e Avaliação de Tecnologias em Cancerologia

MCM5940 Métodos Estatísticos
Aplicados na Área de Saúde

Aula 2

Dra. Rossana V. Mendoza López

Construção de banco de dados

- **Coleta de dados**

Questionários e instrumentos validados

Definição de variáveis e tipos de resposta

Consistência do banco de dados

Dicionário de dados (*codebook*)

Exercício prático. Construção de banco de dados a partir de questionário *Survey*

Coleta de dados

- Questionário estruturado
- Instrumentos de coleta validados
- Exames clínicos ou laboratoriais

Questionários

- Identificação do caso/controle no banco de dados: identificação única, pode ser letra e número
 - Exemplo:
 - CA001 para o caso 1, CA002 para o caso 2, etc.
 - CO001 para o controle 1, CO002 para o controle 2, etc.
- Data da entrevista ou coleta
- Informações pessoais
 - Nome, data de nascimento, endereço, telefone, e-mail

Questionários

- Dados sócio demográficos
 - Sexo/gênero
 - Estado civil
 - Escolaridade

Questionários

- Estilo de vida (tipo de exposição e tempo de exposição)

Quais os fatores que podem estar relacionados com a doença? Como vamos a medir a exposição?

- Consumo de tabaco: idade de início, idade final, tipo de tabaco, frequência e quantidade, fumante atual? Ex-fumante?
- Consumo de álcool: idade de início, idade final, tipo de álcool (teor alcóolico diferente), frequência e quantidade, ainda bebe? Bebia no passado?

Questionários

- Consumo alimentar
 - Usar um questionário validado (QFA: Questionário de Frequência Alimentar; recordatório de 24 horas, etc.)
 - Usar um questionário não validado: tipos de alimentos, frequência e quantidade

Questionários

- Hábitos sexuais (maior cuidado com esta parte na hora de fazer a entrevista – pode gerar constrangimento para o entrevistado e/ou entrevistador)
- Perguntar de forma direta
- Utilizar linguagem simples

Questionários

- Exposição ocupacional
 - Tipo de exposição: substâncias as quais os indivíduos foram expostos
 - Tempo de exposição
 - Exposição passiva

Instrumentos

- Qualidade de vida
 - SF-36 (*Short form 36*): questionário validado em português-BR
 - WHOQOL-Bref
 - EQ-5D
- Específicos
 - HADS: *Hospital Anxiety and Depression Scale*
 - SPRINT: Estresse post-traumático

Versão Brasileira do Questionário de Qualidade de Vida -SF-36

Instrumentos

1- Em geral você diria que sua saúde é:

Excelente	Muito Boa	Boa	Ruim	Muito Ruim
1	2	3	4	5

2- Comparada há um ano atrás, como você se classificaria sua idade em geral, agora?

Muito Melhor	Um Pouco Melhor	Quase a Mesma	Um Pouco Pior	Muito Pior
1	2	3	4	5

3- Os seguintes itens são sobre atividades que você poderia fazer atualmente durante um dia comum. Devido à sua saúde, você teria dificuldade para fazer estas atividades? Neste caso, quando?

Atividades	Sim, dificulta muito	Sim, dificulta um pouco	Não, não dificulta de modo algum
a) Atividades Rigorosas, que exigem muito esforço, tais como correr, levantar objetos pesados, participar em esportes árduos.	1	2	3
b) Atividades moderadas, tais como mover uma mesa, passar aspirador de pó, jogar bola, varrer a casa.	1	2	3
c) Levantar ou carregar mantimentos	1	2	3
d) Subir vários lances de escada	1	2	3
e) Subir um lance de escada	1	2	3
f) Curvar-se, ajoelhar-se ou dobrar-se	1	2	3
g) Andar mais de 1 quilômetro	1	2	3
h) Andar vários quarteirões	1	2	3

MEDIDA DA ANSIEDADE E DA DEPRESSÃO EM PACIENTES NO PRÉ-OPERATÓRIO. ESTUDO COMPARATIVO

Instrumentos

Quadro I — Escala Hospitalar de Ansiedade e Depressão

Este questionário ajudará o seu médico a saber como você está se sentindo. Leia todas as frases. Marque com um "X" a resposta que melhor corresponder a como você tem se sentido na ÚLTIMA SEMANA. Não é preciso ficar pensando muito em cada questão. Neste questionário as respostas espontâneas têm mais valor do que aquelas em que se pensa muito. Marque apenas uma resposta para cada pergunta.

A (1) Eu me sinto tenso ou contraído:

- 3 () A maior parte do tempo
- 2 () Boa parte do tempo
- 1 () De vez em quando
- 0 () Nunca

D (2) Eu ainda sinto gosto pelas mesmas coisas de antes:

- 0 () Sim, do mesmo jeito que antes
- 1 () Não tanto quanto antes
- 2 () Só um pouco
- 3 () Já não sinto mais prazer em nada

A (3) Eu sinto uma espécie de medo, como se alguma coisa ruim fosse acontecer:

- 3 () Sim, e de um jeito muito forte
- 2 () Sim, mas não tão forte
- 1 () Um pouco, mas isso não me preocupa
- 0 () Não sinto nada disso

D (4) Dou risada e me divirto quando vejo coisas engraçadas:

- 0 () Do mesmo jeito que antes
- 1 () Atualmente um pouco menos
- 2 () Atualmente bem menos
- 3 () Não consigo mais

D (8) Eu estou lento para pensar e fazer as coisas:

- 3 () Quase sempre
- 2 () Muitas vezes
- 1 () De vez em quando
- 0 () Nunca

A (9) Eu tenho uma sensação ruim de medo, como um frio na barriga ou um aperto no estômago:

- 0 () Nunca
- 1 () De vez em quando
- 2 () Muitas vezes
- 3 () Quase sempre

D (10) Eu perdi o interesse em cuidar da minha aparência:

- 3 () Completamente
- 2 () Não estou mais me cuidando como deveria
- 1 () Talvez não tanto quanto antes
- 0 () Me cuido do mesmo jeito que antes

A (11) Eu me sinto inquieto, como se eu não pudesse ficar parado em lugar nenhum:

- 3 () Sim, demais
- 2 () Bastante
- 1 () Um pouco
- 0 () Não me sinto assim

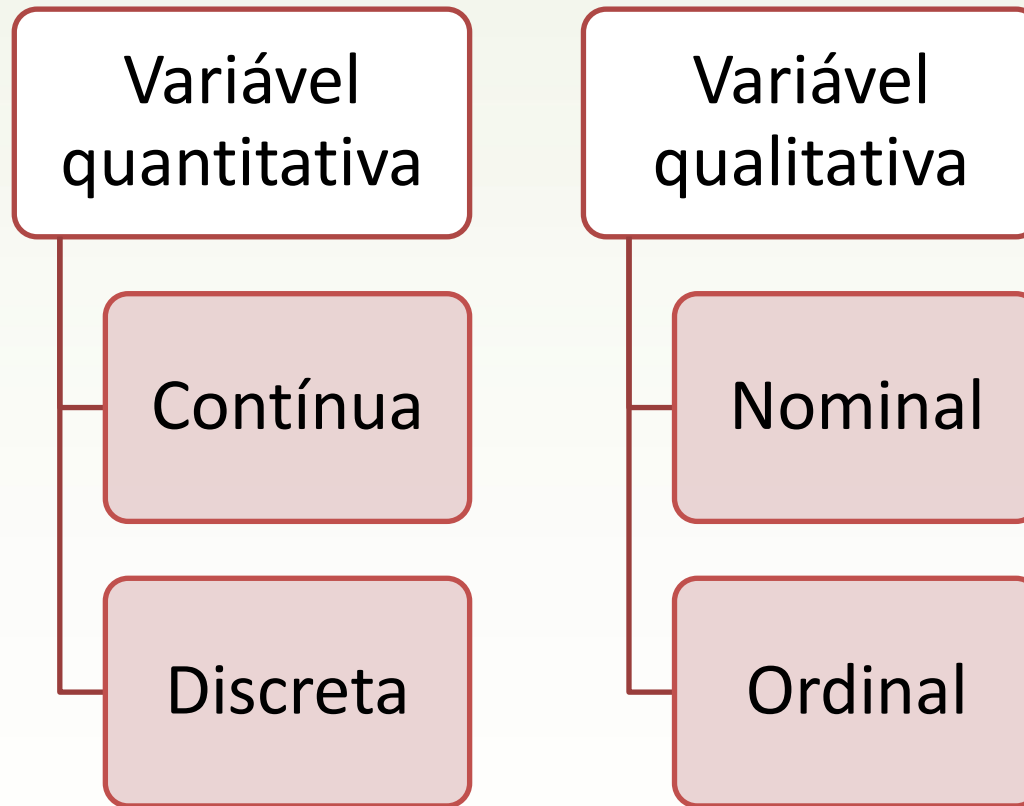
Pareamento

- Variáveis de confusão/ajuste
 - Sexo
 - Idade
 - Vizinhaça
- Variáveis utilizadas no pareamento não são consideradas na análise

Estrutura do banco de dados

- Planilha Excel
- RedCAP
- Outros sistemas de organização de dados

Identificar o tipo de variável



Existem diversos programas computacionais para elaboração de banco de dados. Certifique-se que o software que você está utilizando permite o armazenamento do tipo de dado que você quer analisar.

O programa Excel © é frequentemente utilizado pela facilidade na manipulação e armazenamento de dados, assim como na elaboração de fórmulas, gráficos e análises de dados. No entanto, quando os dados são armazenados em planilhas que serão transferidas para algum programa estatístico alguns cuidados são necessários.

Considerações

1. Se você tiver usado um formulário ou um questionário para realizar a pesquisa certifique-se de ter coletado todas as informações necessárias. Muitas vezes é melhor coletar mais informações e logo não utilizá-las do que ter poucas variáveis e esquecer-se de algumas que podem ser importantes.

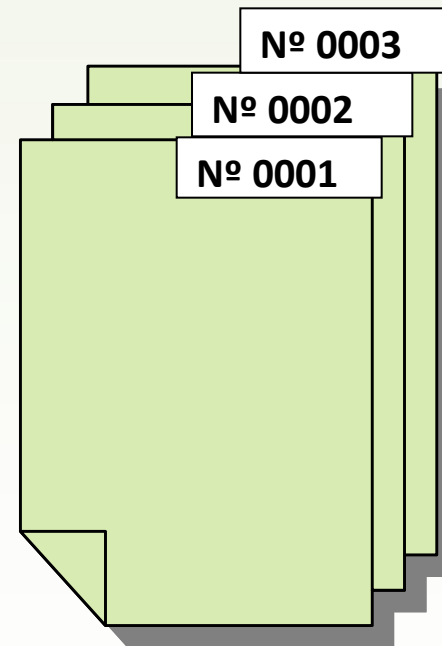
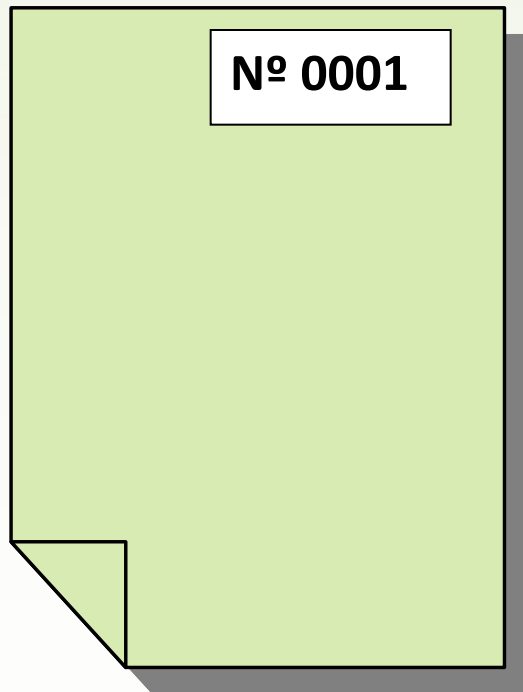
Considerações

2. Identifique todos os seus casos com um número de identificação único, de preferência um número sequencial: 1, 2, 3, ..., n.

Fica mais fácil identificar um caso por **idnum** (número de identificação) no caso de precisar atualizar essa informação.

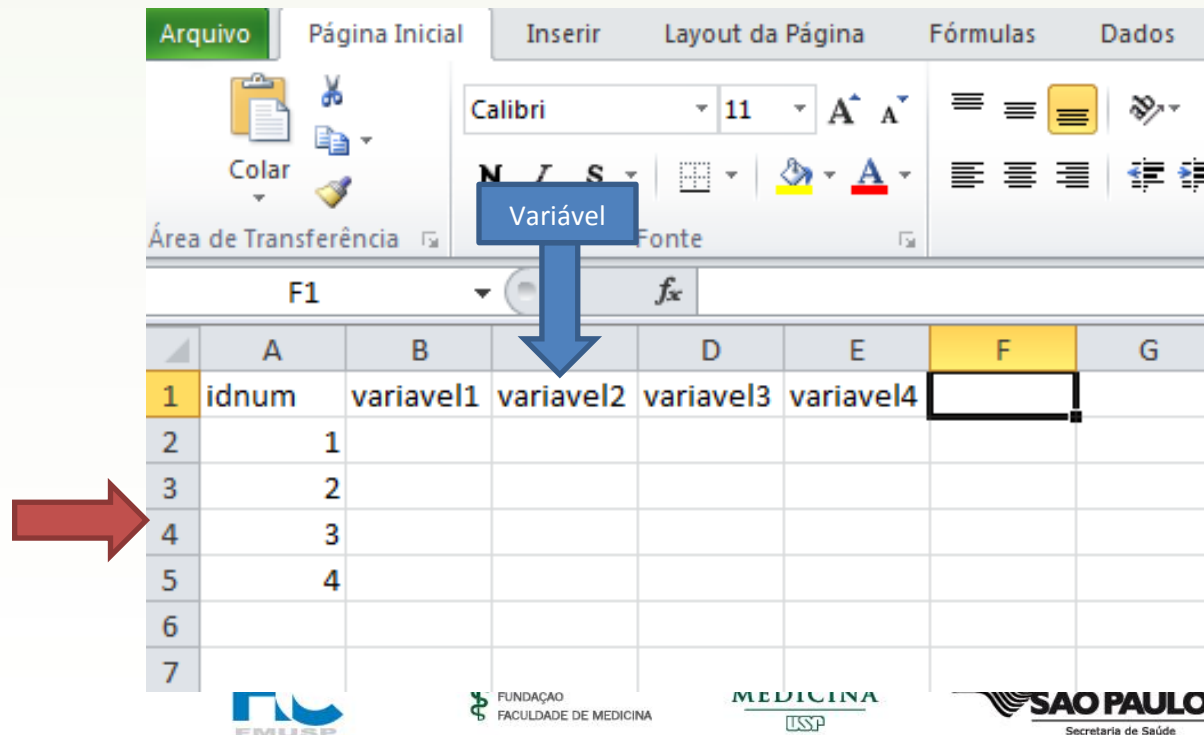
Não é recomendado utilizar as iniciais do nome do participante como identificação, duas pessoas podem ter as mesmas iniciais: Maria Aparecida da Silva (MAS) e Manoel Antônio Soares (MAS).

Identificação dos casos e controles



Considerações

3. Na sua planilha considere cada coluna como uma variável e cada linha como um registro.



The screenshot shows the Microsoft Excel interface with the following spreadsheet data:

	A	B	C	D	E	F	G
1	idnum	variavel1	variavel2	variavel3	variavel4		
2	1						
3	2						
4	3						
5	4						
6							
7							

A blue arrow labeled "Variável" points to the header row (row 1). A red arrow points to the first column (column A).

Considerações

4. Coloque nomes de variáveis curtos e de fácil reconhecimento, a descrição da variável deve ser incluída no “dicionário de variáveis”.

Errado

	A	B	C	D	E	F
1	idnum	idade	escolaridade	extravazamento extracapsular (presença de linfonodos)		
2		1	35 fundamental	sim		
3		2	42 superior	não		
4		3	29 ensino médio	sim		
5		4	99 não achei no prontuário	sim		
6		5 ??	incompleto	sim		
7						

Certo

No banco de dados

	A	B	C	D	E
1	idnum	idade	escolaridade	linfonodos	
2		1	35 fundamental	sim	
3		2	42 superior	não	
4		3	29 ensino médio	sim	
5		4	99 não achei no prontuário	sim	
6		5 ??	incompleto	sim	
7					

Considerações

4. Coloque nomes de variáveis curtos e de fácil reconhecimento, a descrição da variável deve ser incluída no “dicionário de variáveis”.

No “dicionário de variáveis”

	A	B	C	D
1	id	Variável	Descrição	
2		1 idnum	Identificação	
3		2 idade	Idade no diagnóstico	
4		3 escolaridade	Grau de escolaridade	
5		4 linfonodos	Extravazamento extracapsular (presença de linfonodos)	
6				

Considerações

7. Campos numéricos não devem ser misturados com letras, por exemplo, se a variável é idade não escrever nesse campo “ainda não foi coletado” ou colocar alguma outra observação.

Errado

	A	B	C	D	E
1	idnum	idade	escolaridade	peso	linfonodos
2	1	35	fundamental	65	sim
3	2	42	superior	59	não
4	3	29	ensino médio	60	sim
5	4	99	não achei no prontuário	não sei	sim
6	5	??	incompleto	??	sim
7					

Certo

	A	B	C	D	E
1	idnum	idade	escolaridade	peso	linfonodos
2	1	35	fundamental	65	sim
3	2	42	superior	59	não
4	3	29	ensino médio	60	sim
5	4	99	não achei no prontuário	999	sim
6	5	999	incompleto	999	sim
7					

Considerações

8. Para variáveis numéricas o código para “sem informação” dependerá das unidades da variável, aumentando em um dígito o valor de “99”. Por exemplo, na variável “idade” é melhor colocar “999” que “99”, pois é provável que um paciente tenha essa idade (veja figura acima).

9. Para as variáveis categóricas (por exemplo: sexo, estado civil, grau de escolaridade), se não tiver informação pode colocar o valor “99”.

Considerações

Errado

	A	B	C	D	E
1	idnum	idade	escolaridade	peso	linfonodos
2	1	35	fundamental	65	sim
3	2	42	superior	59	não
4	3	29	ensino médio	60	sim
5	4	99	não achei no prontuário	999	sim
6	5	999	incompleto	999	sim
7					

Certo

	A	B	C	D	E
1	idnum	idade	escolaridade	peso	linfonodos
2	1	35	fundamental	65	sim
3	2	42	superior	59	não
4	3	29	ensino médio	60	sim
5	4	99	99	999	sim
6	5	999	99	999	sim
7					

Considerações

10. Campos em formato data (por exemplo: dd/mm/aa) não devem ser misturados com letras ou números simples. Exemplo: se a variável é “data de diagnóstico”, e o pesquisador não tem essa data então melhor deixar em branco, não colocar “sem informação”.

Aula 2



Errado

D	E	F	G
peso	linfonodos	data_ult_info	
65	sim	31/01/2015	
59	não	20/04/2014 (óbito)	
60	sim	-	
999	sim	30/08/2016	
999	sim	perda de seguimento	

Certo

E	F	G	H
linfonodos	data_ult_info	observação	
sim	31/01/2015		
não	20/04/2014	óbito	
sim			
sim	30/08/2016		
sim		perda de seguimento	

Considerações

11. A utilização de “cores” na planilha Excel © pode ser útil para o pesquisador identificar alguma característica dos dados, no entanto, para as análises estatísticas não fazem diferença. Se você quer identificar alguma característica especial no seu banco de dados, por favor crie uma coluna com o nome da observação.

Considerações

Errado (cada cor significa um grupo)

	A	B	C	D	E	F	G
1	idnum	idade	escolaridade	peso	linfonodos	data_ult_info	observação
2	1	35	fundamental	65	sim	31/01/2015	
3	2	42	superior	59	não	20/04/2014	óbito
4	3	29	ensino médio	60	sim		
5	4	99		99	999	sim	30/08/2016
6	5	999		99	999	sim	perda de seguimento

Certo (variável "grupo")

	A	B	C	D	E	F	G	H
1	idnum	idade	escolaridade	peso	linfonodos	data_ult_info	observação	grupo
2	1	35	fundamental	65	sim	31/01/2015		1
3	2	42	superior	59	não	20/04/2014	óbito	1
4	3	29	ensino médio	60	sim			1
5	4	99		99	999	sim	30/08/2016	2
6	5	999		99	999	sim	perda de seguimento	2

Considerações

12. Evita criar uma planilha para cada conjunto de dados, por exemplo, não crie uma “planilha 1” para colocar os dados demográficos (sexo, idade, escolaridade, local de residência), e uma planilha 2 para os dados clínicos. Coloque tudo em uma única planilha, facilita as análises entre ambos os conjuntos de variáveis.

Considerações

13. De preferência crie códigos para as variáveis, isso facilita o tempo gasto na digitação e possíveis erros na hora de digitar categorias muito extensas.

Considerações

Errado

	A	B	C	D	E	F
1	idnum	idade	escolaridade	peso	linfonodos	data_ult_info
2	1	35	fundamental	65	sim	31/01/2015
3	2	42	superior	59	não	20/04/2014
4	3	29	ensino médio	60	sim	
5	4	99	Superior	999	sim	30/08/2016
6	5	999	Ensino fundamental	999	sim	

Certo

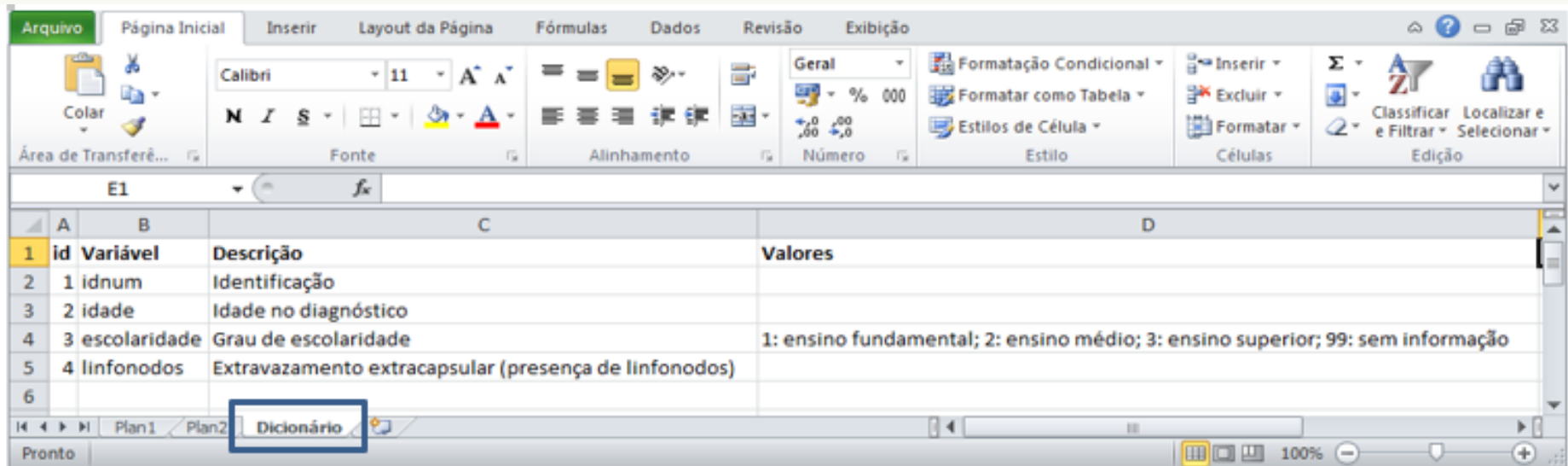
	A	B	C	D	E	F
1	idnum	idade	escolaridade	peso	linfonodos	data_ult_info
2	1	35	1	65	sim	31/01/2015
3	2	42	3	59	não	20/04/2014
4	3	29	2	60	sim	
5	4	99	3	999	sim	30/08/2016
6	5	999	1	999	sim	

Dicionário de dados

	A	B	C	D
1	id	Variável	Descrição	Valores
2	1	idnum	Identificação	
3	2	idade	Idade no diagnóstico	
4	3	escolaridade	Grau de escolaridade	1: ensino fundamental; 2: ensino médio; 3: ensino superior; 99: sem informação
5	4	linfonodos	Extravazamento extracapsular (presença de linfonodos)	

Considerações

14. Crie uma aba dentro do seu arquivo com o “dicionário das variáveis”, que contenha as informações das variáveis: nome, descrição e valores.



The screenshot shows the Microsoft Excel interface with the 'Dicionário' worksheet selected. The table contains the following data:

id	Variável	Descrição	Valores
1	idnum	Identificação	
2	idade	Idade no diagnóstico	
3	escolaridade	Grau de escolaridade	1: ensino fundamental; 2: ensino médio; 3: ensino superior; 99: sem informação
4	linfonodos	Extravazamento extracapsular (presença de linfonodos)	

Consistência dos dados

- Quais informações preciso verificar?
 - Idade x Doença
 - Sexo x Doença
 - Variáveis quantitativas (valores mínimo e máximo)
 - Fórmulas
 - Datas: verificar os tempos (se tem tempos negativos)

Consistência dos dados

- Informações de acordo com a pesquisa
- Informações clínicas
- Variáveis principais da pesquisa – **IMPORTANTE!**
- Coletei todas as informações necessárias?
- *Missing* vs. Não se aplica – **São diferentes!**

Aula 2



Exemplo



INTERCHANGE – Consorcio internacional de câncer de cabeça e pescoço

Estudo caso-controlado sobre fatores relacionados com câncer de cabeça e pescoço

Centros:

Brasil: São Paulo, Goiânia, Vitória

Argentina: IO Angel H. Roffo, Hospital Italiano (Buenos Aires)

Uruguai: Hospital de Clínicas (Montevidéu)

Maio, 2012

ID No. BR-|_|_|_|/|_|_|_|_|

ESTUDO INTERNACIONAL SOBRE CÂNCER DE BOCA, FARINGE E LARINGE

Questionário Inicial sobre Hábitos e Estilo de Vida

1. Número de identificação BR |_|_|_| |_|_|_|_|_|
 País Centro Número do indivíduo

O número de identificação é composto pela sigla BRI (Brasil – INTERCHANGE), e pelos códigos do centro e dos indivíduos.

O número de identificação do indivíduo é dado consecutivamente aos indivíduos incluídos no estudo, em cada centro, e não deve incluir o identificador do status de caso ou controle.

2. Orientações Gerais

- As colunas devem ser preenchidas com justificção à direita (exemplos: válido |_|_| 1| 2| ; inválido |1| 2|_|)
- Evitar os códigos “não sei/ignorado”; insistir para obter uma resposta, mesmo se for somente uma estimativa (resposta aproximada).
- Se você não obtiver uma resposta ou estimativa, as colunas deverão ser preenchidas com o número 9.
- Quando estiver escrito “por favor, especifique”, anotar a resposta na linha pontilhada.
- O número de identificação deve ser usado no questionário e nas amostras biológicas e deve ser escrito, juntamente com as iniciais do indivíduo em cada amostra biológica.

Aula 2



Banco de dados no SPSS

BD_PS_Interchange_Brazil_150518.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Analisar Gráficos Utilitários Extensões Janela Ajuda

1 : Drink_av 6 Visível: 151 de

	Idnum	Status	Center	Sex	Age	Race	Sch_Level	Smk_YN
1	120001	Caso	AC Camargo Cancer Center	Masculino	71	Amarelo (asiático)	Ensino médio completo	Former smoker
2	120002	Caso	AC Camargo Cancer Center	Masculino	47	Branco	Ensino médio completo	Current smoker
3	120003	Caso	AC Camargo Cancer Center	Feminino	54	Branco	Ensino médio completo	Never smoker
4	120004	Caso	AC Camargo Cancer Center	Masculino	72	Branco	Ensino superior completo	Former smoker
5	120005	Caso	AC Camargo Cancer Center	Masculino	69	Branco	Ensino superior completo	Former smoker
6	120006	Caso	AC Camargo Cancer Center	Masculino	27	Branco	Ensino médio completo	Former smoker
7	120007	Caso	AC Camargo Cancer Center	Masculino	56	Branco	Ensino médio completo	Former smoker
8	120008	Caso	AC Camargo Cancer Center	Feminino	85	Branco	Ensino superior completo	Never smoker
9	120009	Caso	AC Camargo Cancer Center	Masculino	47	Branco	Ensino fundamental completo	Current smoker
10	120010	Caso	AC Camargo Cancer Center	Masculino	74	Amarelo (asiático)	Ensino médio completo	Never smoker
11	120011	Caso	AC Camargo Cancer Center	Feminino	78	Branco	Ensino fundamental incompleto	Never smoker
12	120012	Caso	AC Camargo Cancer Center	Masculino	66	Branco	Ensino fundamental incompleto	Former smoker



Log In

 A. C. Camargo Cancer Center

Solicitação de acesso à Colaboradores para Estudos Multicêntricos / Pesquisadores Externos

Clique aqui para solicitar a criação de acesso à colaboradores de instituições afiliadas.

Acesso para Colaboradores / Pesquisadores Internos / Alunos vinculados ao Ensino do Hospital

Clique aqui se estiver na rede do hospital, para maiores informações de acesso e criação de bancos no REDCap.

Caso não lembre sua senha clique aqui para recuperar sua senha (somente para usuários dentro da rede do A.C. Camargo).

Please log in with your user name and password. If you are having trouble logging in, please contact [Informatica Medica - Ramal 2958](#).

Username:

mp.curadoj

Password:

<https://dados.accamargo.org.br/redcap/>

Aula 2



REDCap Home My Projects + New Project Help & FAQ Training Videos Send-it Messenger

Logged in as mp.curado My Profile Log out

Solicitação de acesso à Colaboradores para Estudos Multicêntricos / Pesquisadores Externos
Clique aqui para solicitar a criação de acesso à colaboradores de instituições afiliadas.

Acesso para Colaboradores / Pesquisadores Internos / Alunos vinculados ao Ensino do Hospital
Clique aqui se estiver na rede do hospital, para maiores informações de acesso e criação de bancos no REDCap.

Caso não lembre sua senha clique aqui para recuperar sua senha (somente para usuários dentro da rede do A.C. Camargo).

Listed below are the REDCap projects to which you currently have access. Click the project title to open the project.
[Read more](#) To review which users still have access to your projects, visit the [User Access Dashboard](#).

Dashboard pages:
- [User Access Dashboard](#)
- [Sponsor Dashboard \(8 users\)](#)

My Projects Organize Filter projects by title

Project Title	Records	Fields	Instrument	Type	Status
EPGC <small>(drafted changes are in review)</small>	1,540	1,762	16 forms		
MIDMAMA - Mida e Câncer de Mama <small>(drafted changes are in review)</small>	504	432	7 forms		

<https://dados.accamargo.org.br/redcap/>



Aula 2



Logged in as mp.curado | Log out

My Projects

Project Home or Project Setup

REDCap Messenger **!**

Project status: **Production**

Data Collection

Record Status Dashboard

Add / Edit Records

Hide data collection instruments

Identificacao Do Entrevistado

Estilo De Vida

Frequencia Alimentar

Dados Clinicos e Patologicos

TCLE - Epidemiologia

TCLE - Belem

TCLE - Ceara

Análise CTC

Dados Prontuario Controle 1

Dados Clínicos e Patológicos Epidemiologia

Endoscopia

Oncogenetica

Oncogenetica Variantes Painel

Dados Prontuários Casos

Etapa do tratamento

Ciclo de tratamento

Applications

Calendar

Data Exports, Reports, and Stats

Project Home Project Setup Other Functionality Project Revision History

Quick Tasks

- Codebook**: The Codebook is a human-readable, read-only version of the project's Data Dictionary and serves as a quick reference for viewing field attributes.
- Export data**: Export your data from REDCap to open or view in Excel or various stats packages.
- Create a report**: Build custom reports for quick views of your data, and export reports to Excel/CSV.
- Check data quality**: Build or execute data quality rules to find discrepancies and errors in your project data.
- User Rights**: Grant new users access to this project or modify user privileges for current users.
- Online Designer and Data Dictionary Upload**: Create new fields/questions on your data collection instruments or modify existing ones using the Online Designer or by uploading a Data Dictionary. Quick link: [Download the current Data Dictionary](#) OR [Download Data Dictionary with drafted changes](#)
- Copy this project**: Create an exact duplicate of this project, which copies over all data collection instruments, any surveys that exist, as well as the option to copy all users and reports to the new project.
- Data Access Groups**: Create groups of users to limit user access to certain records/responses, in which only users within a given Data Access Group can access records created by users within that group.

Project Dashboard

The tables below provide general dashboard information, such as a list of all users with access to this project, general project statistics, and upcoming calendar events (if any).

Current Users (37)	
User	Expires
amanda.ramos	never

Project Statistics	
Records in project	1,540
Most recent activity	19/06/2019 8:51am

[redcap gastrico](#)

Aula 2



Log In



Criação de Conta de Usuário

Se você ainda não possui uma conta de acesso ao REDCap-HCFMUSP, preencha o [Formulário de Cadastro de Usuários do REDCap-HCFMUSP](#)

Mover o Projeto para Produção

Já terminou o desenvolvimento do seu projeto no REDCap e pretende iniciar a coleta de dados? Os passos necessários são:

- 1) **Faça o cadastro do seu Projeto.**
- 2) Realize um teste de todos os formulários e funcionalidades do projeto.
- 3) Solicite que o seu projeto seja movido para produção e inicie a coleta de dados de forma segura.

Estudos Multicêntricos coordenados por Pesquisadores do HCFMUSP

Para solicitar a criação de contas de acesso para os colaboradores de centros coparticipantes preencha o [Cadastro de Projetos Multicêntricos - REDCap-HCFMUSP](#)

Utilização do "REDCap Mobile App" para coleta de dados offline

Para solicitar a autorização de uso do REDCap Mobile App (coleta de dados offline) preencha o [Cadastro de Utilização do REDCap Mobile App - HCFMUSP](#).

Please log in with your user name and password. If you are having trouble logging in, please contact [Núcleo Especializado em Tecnologia da Informação](#).

Username:

Password:

Log In

[Forgot your password?](#)

<https://redcap.hc.fm.usp.br/>



Aula 2



Alguns exemplos de bancos de dados

Exemplo 1 – Diferentes planilhas no mesmo arquivo, colunas sem nome, etc.

	S	T	U	V	W	X	Y	Z	AA	AD	AE	AF	AG	AH
1	Tacro	Evero	Siru	MMF	NUMERO DE BIOPSIA		CEC		CBC	cbc_data	Melanoma	Bowen	QUERATO/	campo can
2	0	0	0	2000	7				5	jan/14		X (fev/14)		
3	2	0	0	720	4	jan/12	X (jan12)		0			X (jan/15)		
4	0	0	0	0	2				0					
5	0	0	0	0	10	fv/14	XXXX		0					X
6	0	0	0	0	1									X
7	2	0	0	1440	2	mai/13	X							
8	2	0	0	1440	3				XX		X			
9	6	0	0	0	1									
10					0									
11					0									
12	3	0	0	1440	0									
13	10	0	0	720	2									
14	0	0	0	0	1									
15	0	0	0	0	2									
16	3	4,5	0	0	0									
17	2	0	0	1080	0									
18					13		XX		XXXXXXXX			X		
19	2	0	0	1440	4				XXXXXXXX			X		
20	3	0	0	1440	14				XXXXXXXX					
21	0	0	2	720			X							
22	0	0	0	0	10		XXXXXX		XX					
23	4	0	0	1440	2									
24	0	0	2	0	1									
25					1									
26	8	0	0	1440	1									

Aula 2



Alguns exemplos de bancos de dados

Exemplo 2 – Diferentes colunas com descrição das variáveis, sem dicionário de dados

	A	B	C	D	E	F	G	H	I	J	K	
1												
2	Questionário sobre o conhecimento da ação do álcool sobre feto quando ingerido pela mãe durante a gestação											
3												
4		Q1	Q1	Q2	Q3	Q4	Q5	Q6	Q6A	Q7	Q8	
5	ID	Q1CAT_PROF	Q1CAT_PROF_OUTRO	Q2TEMP_PROF	Q3SEXO	Q4ING_ALCOOL	Q5PREJUIZO	Q6QUAN_ACEITAVEL	Q6ACODIGO	Q7QUESCONSUMO	Q8EFADVERSOS	
6	1	3			2	2	1	1	nenhuma	1	1	
7	2	3			1	1	1	1	nenhuma	1	1	
8	3	3			2	2	1	1	um gole	0	1	
9	4	3			1	2	1	1	nenhuma	1	0	
10	5	3			1	2	1	1	nenhuma	1	1	
11	6	3			4	2	1	1	nenhuma	1	1	
12	7	3			2	2	1	1	nenhuma	1	1	
13	8	3			1	2	1	1	nenhuma	1	1	
14	9	3			2	2	1	0	nenhuma	1	1	
15	10	3			3	1	1	1		2	1	
16	11	3			1	2	1	0	nenhuma	1	1	
17	12	3			4	2	1	0	nenhuma	1	1	
18	13	3			1	2	1	0	nenhuma	1	1	
19	14	3			2	2	1	0	nenhuma	1	1	
20	15	3			1	2	1	0	nenhuma	1	0	
21	16	3			4	2	1	1	não sabe	2	1	

Aula 2



Alguns exemplos de bancos de dados
Exemplo 3 – Cada planilha é um paciente

8	Aline	17/01/1994	f	19/09/2001		
9	Aluan	04/08/1987	m	08/06/1998	01/07/1998	
10	Aman	17/06/1993	f	29/12/2008	05/01/2009	
11	Ana B	12/05/2004	f	24/08/2006		
12	Ana B	13/09/2005	f	28/08/2006		
13	Ana C	22/11/1990	f	11/02/1996		
14	Ana C	15/09/1998	f	01/06/2001		
15	Ana C	15/06/2007	f	18/05/2009		
16	Ana C	10/07/1980	f	27/02/1995		
17	Ana J	28/07/2011	f	22/05/2012		
18	Ana K	20/09/2010	f	03/09/2011		
19	Ande	08/08/2000	m	27/08/2001		
20	André	04/08/1989	m	04/10/2000		
21	André	10/11/1995	m	16/11/1998		
22	Anna	01/04/2003	f	07/07/2008		
23	Antôn	22/03/1981	m	04/10/1992		
24	Arian	09/10/1998	f	02/04/2001		
25	Arthu	25/02/2004	m	15/06/2011		
26	Arthu	05/05/1998	m	12/09/1999		

Sheet1

Agnes

Alana V

Albert

Alguns exemplos de bancos de dados Exemplo 4 – Dados em Word

QUEILITE ACTINICA

- 1- NOME :
- 2- GÊNERO:
 - a. MASCULINO X
 - b. FEMININO
- 3- IDADE: 62
- 4- FOTOTIPO:
 - a. I
 - b. II X
 - c. III
 - d. IV
 - e. V
 - f. VI
- 5- MATRÍCULA: 1201
- 6- HE:
 - a. ANTES :
 - i. DISPLASIA LEVE
 - ii. DISPLASIA MODERADA X
 - iii. DISPLASIA GRAVE
 - b. DEPOIS:
 - i. DISPLASIA LEVE
 - ii. DISPLASIA MODERADA X
 - iii. DISPLASIA GRAVE

Aula 2



Alguns exemplos de bancos de dados

Exemplo 5 – Dicionário e dados na mesma planilha

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Variáveis																			
2	IG=idade gestacional em semanas					Sup. Resp= suporte respiratório do Rn. Pode ser VMI= ventilação mecânica														
3	Peso (g)= peso em gramas					CPAP= pressão positiva de vias aéreas														
4	corticóide=uso de corticóide materno																			
5	RCIU= retardo de crescimento intra-uterino					ECO=exame ecocardiográfico para confirmação ou não de canal arterial com repercussão														
6	hiper=hipertensão arterial materna					canal= persistente (S); fechado (N) ou SEM INFORMAÇÃO (SI) quando Rn não tem ECO (não posso afirmar q														
7	inf.= infecção precoce do recém-nascido					indo= recebeu (S), não recebeu (N) indometacina para fechamento do canal.														
8	SDR= Síndrome de desconforto respiratório					DV indo= dias de vida em que recebeu indometacina (consta o dia de início do tratamento com indoemtacina.														
9	surf. Uso= uso de surfactante no RN					n ciclos= nº de ciclos de indometacina														
10	d. canal= diâmetro do canal arterial em mm.																			
11	FATORES PREDISPONENTES PARA PERSISTÊNCIA DO CANAL ARTERIAL																			
12	REG	data	tipo	Apgar		sexo	class.	IG	Peso	perda	CORTICÓI	RCIU	INF.	RM	SDR	surf.	nº	Sup.resp.		
13		nascimento	parto	1	5			(sem.)	(g)	peso(%)	uso		Hipert.	S/N		uso	DOSES	VMI	CPAP	
14	861477	01/12/2001	C	7	8	F	AIG	27	840	20,2	S	N	N	S	N	S	2	S	S	
15	861479	01/12/2001	C	5	8	M	AIG	27	720	19,4	S	N	N	N	N	S	2	S	N	
16	869448	30/01/2002	C	7	9	F	AIG	27	680	18,9	S	S	N	S	N	S	1	S	N	
17	870959	07/02/2002	N	5	9	M	AIG	29 5/7	920	22	N	S	N	S	N	S	1	S	N	
18	876623	14/03/2002	N	6	8	M	AIG	27	1060	30	S	N	N	S	N	N	0	S	N	
19	883407	25/04/2002	C	5	8	M	PIG	34 4/7	1490	6,7	N	S	N	N	N	N	0	N	N	
20	883711	28/04/2002	C	7	8	F	AIG	25 4/7	740	18,9	N	N	N	S	S	S	2	S	S	
21	885230	09/05/2002	C	7	8	F	PIG	31	770	19,2	S	S	N	N	N	S	0	S	S	
22	885228	10/05/2002	C	7	8	F	AIG	30 3/7	1145	8,2	S	N	N	N	N	S	2	S	S	
23	885229	10/05/2002	C	8	10	M	AIG	30 3/7	1340	4,4	S	N	N	N	N	S	2	S	S	
24	937045	12/05/2002	C	4	9	F	AIG	27	1080	8,3	S	S	N	N	N	S	1	S	N	
25	887337	21/05/2002	C	1	7	M	AIG	28 3/7	1090	2,9	N	N	N	S	N	S	1	S	S	
26	891135	17/06/2002	C	8	10	M	AIG	31 5/7	830	17,4	S	S	N	S	S	S	1	S	N	
27	891138	17/06/2002	C	4	7	M	PIG	31 5/7	795	14,4	S	S	N	S	S	S	2	S	N	
28	895309	15/07/2002	C	2	6	M	AIG	27	880	16,4	N	N	N	S	S	S	2	S	S	
29	895310	15/07/2002	C	3	8	F	AIG	27	880	16,4	N	N	N	S	N	S	2	S	N	
30	895311	15/07/2002	C	4	8	M	AIG	27	930		N	N	N	S	N	S	2	S	N	
31	896649	24/07/2002	C	8	10	M	AIG	28 5/7	1150	18,2	S	N	N	S	(19DIAS)	S	1	S	S	
32	897319	30/07/2002	C	8	10	M	AIG	29 6/7	1350	15,5	S	N	N	N	N	N	0	N	S	
33	898229	05/08/2002	N	9	10	F	AIG	31	1340	16,7	N	N	N	N	N	S	1	N	N	
34	911150	01/11/2002	C	1	6	M	AIG	28	825	13,9	S	N	N	S	S	S	2	S	N	

The impact of the time to start radiation therapy on overall survival in newly diagnosed glioblastoma

Vanessa Montes Santos¹ · Gustavo Nader Marta^{2,3} · Marcella Coelho Mesquita⁴ ·
Rossana Veronica Mendoza Lopez⁵ · Edla Renata Cavalcante⁴ · Olavo Feher^{4,6}

Table 2 Univariable and multivariable analysis for overall survival

	Univariable		Multivariable			
	HR (95% CI)	p Value	Model 1		Model 2	
	HR (95% CI)	p Value	HR (95% CI)	p Value	HR (95% CI)	p Value
Gender						
Male	1.00					
Female	1.06 (0.66–1.70)	0.810				
Age						
≤ 65 years	1.00		1.00		1.00	
> 65 years	2.53 (1.37–4.66)	0.003	2.50 (1.28–4.86)	0.007	2.62 (1.40–4.90)	0.003
Type of surgery						
Total	1.00					
Subtotal	1.21 (0.76–1.91)	0.424				
Biopsy	0.82 (0.25–2.68)	0.819				
PS ECOG						
0 or 1	1.00					
2 or 3	1.00 (0.55–1.83)	0.992				
Re-operation						
No	1.00					
Yes	1.40 (0.73–2.66)	0.308				
Interval between surgery and radiation (days)						
Up to 30	1.00		1.00			
31–60	0.60 (0.27–1.34)	0.215	0.83 (0.36–1.94)	0.667		
> 60	0.74 (0.37–1.46)	0.381	1.20 (0.48–2.13)	0.969		
Up to 60	1.00				1.00	
> 60	1.03 (0.64–1.66)	0.890			1.15 (0.71–1.87)	0.570

Exercício 2

Organizar o banco de dados “Survey”

Criar o dicionário de variáveis

Verificar a consistência dos dados

Enviar o exercício até **terça feira** dia **23 de agosto às 18h por e-mail**

rossana.veronica@hc.fm.usp.br

alessandro.campolina@hc.fm.usp.br

Obrigada



Rossana V. Mendoza López
Bioestatística e Epidemiologista
Centro de Investigação Translacional em Oncologia
ICESP
rossana.veronica@hc.fm.usp.br

