

CONSTRUÇÃO DE BASE DE DADOS

Paulo Nicola

Unidade de Epidemiologia

Instituto de Medicina Preventiva

Faculdade de Medicina de Lisboa

CONSTRUÇÃO DE BASE DE DADOS

■ Objectivos

■ Introduzir e praticar:

- a construção de uma Base de Dados relacional, para um estudo clínico
- a construção de um formulário para entrada de dados
- a consulta (*query*) da Base de Dados para monitorização.

a partir de um exemplo.

CONSTRUÇÃO DE BASE DE DADOS

- **Tipos de Bases de Dados em Investigação Clínica**
 - Combinações de processos em papel, folhas de Excel, entrada directa de dados para o programa de estatística
 - Base de Dados relacional num computador pessoal
 - Base de Dados relacional num computador cliente-servidor
 - Base de Dados em servidor da internet

CONSTRUÇÃO DE BASE DE DADOS

- **Tipos de Bases de Dados em Investigação Clínica**
 - Combinações de processos em papel, folhas de Excel, entrada directa de dados para o programa de estatística
 - **Base de Dados relacional num computador pessoal**
 - Base de Dados relacional num computador cliente-servidor
 - Base de Dados em servidor da internet

CONSTRUÇÃO DE BASE DE DADOS

- **Vantagens de uma Base de Dados informática**
 - Facilidade na entrada de dados
 - Por exemplo, uso de sistemas de classificação incorporados, listas de opções, etc.

CONSTRUÇÃO DE BASE DE DADOS

- **Vantagens de uma Base de Dados informática**
 - Facilidade na entrada de dados
 - Validação automática dos dados
 - A informação é inserida num formato uniforme
 - Facilidade na colheita de informação condicional (apenas necessária em certas condições)

CONSTRUÇÃO DE BASE DE DADOS

- **Vantagens de uma Base de Dados informática**
 - Facilidade na entrada de dados
 - Validação automática dos dados
 - Verificação automática de erros

CONSTRUÇÃO DE BASE DE DADOS

- **Vantagens de uma Base de Dados informática**
 - Facilidade na entrada de dados
 - Validação automática dos dados
 - Verificação automática de erros
 - (Re)formatação fácil da informação
 - Construção de condições de consulta, com produção de novas tabelas e novos formatos de dados

CONSTRUÇÃO DE BASE DE DADOS

- **Vantagens de uma Base de Dados informática**
 - Facilidade na entrada de dados
 - Validação automática dos dados
 - Verificação automática de erros
 - (Re)formatação fácil dos dados
 - A alternativa é... uma pilha enorme de formulários em papel! e +++ trabalho...

CONSTRUÇÃO DE BASE DE DADOS

■ Vantagens do Microsoft Access

- “User-friendly”: por exemplo – uso de assistentes (*wizards*), ligação com o Excel e o Word, interfaces para construir consultas (*queries*)
- Vem com o Microsoft Office
- Perfeitamente capaz para Bases de Dados de tamanho pequeno ou médio
- Grande suporte técnico e de formação
- Tem uma linguagem de programação associada (VisualBasic)

CONSTRUÇÃO DE BASE DE DADOS

■ Áreas do Access

Definição de relações
entre tabelas

The screenshot shows the Microsoft Access interface with the 'Datasheet' tab selected. The ribbon includes 'Home', 'Create', 'External Data', 'Database Tools', and 'Datasheet'. The 'Datasheet' ribbon has groups for 'Views', 'Fields & Columns', 'Data Type & Formatting', and 'Relationships'. The 'Relationships' group contains 'Relationships' and 'Object Dependencies' buttons. The 'All Access Objects' pane on the left shows a list of tables: Baby, Death, Doctor, Exam, IkpRace, and IkpSpecialty. The main window displays a datasheet for the 'Doctor' table with columns: DocID, DocFNam, DocMNM, DocNam, DocSex, DocDOB, SocialSec, and DocPhone. A query is applied to the 'Exam' table, showing columns: ExamID, Subject, ExDate, ExWgt, ExHgt, ExNPSco, and Add New Field. The data rows show exam results for various subjects and dates.

Tabelas: Definição de variáveis e armazenamento de dados

Consultas (*queries*): seleccionar e re-organizar dados

Formulários: para entrada de dados

Relatórios: para comunicação de resultados

CONSTRUÇÃO DE BASE DE DADOS

■ Construção de tabelas

Chave:

numeração automática

Nomes das variáveis: únicos

Tipo de dados

Descrição

Formato de introdução

Subtipo, tamanho, precisão

Texto para o formulário

Valor predefinido

Limites, identificação de erros

Texto de alerta

Pode não ser introduzido?

Field Name	Data Type	Description
ParticipID	AutoNumber	Identificador único do participante
PopulID	Number	Identificador único do elemento da população
ArtriteReuma	Yes/No	Tem artrite reumatóide?
InsufCard	Yes/No	Tem insuf cardíaca na data de diagnóstico de AR?
DataDiagn	Date/Time	Data de diagnóstico de Artrite Reumatóide
Hipertensao	Text	Tem hipertensao na data de diagnóstico de AR?
Displidemia	Memo	Tem displidemia na data de diagnóstico de AR?
Diabetes	Number	Tem Diabetes na data de diagnóstico de AR?
Peso	Date/Time	Peso (kgs) na data de diag de AR
Altura	Currency	Altura (cms) na data de diag de AR
Investid	AutoNumber	Identificador único do investigador que entrou

General	Lookup
Format	Short Date
Input Mask	00-00-0000;0/;
Caption	Data de Diagnóstico da Artrite Reumatóide
Default Value	
Validation Rule	<=Now()
Validation Text	Verifique a data de diagnóstico!
Required	Yes
Indexed	No
IME Mode	No Control
IME Sentence Mode	None

CONSTRUÇÃO DE BASE DE DADOS

- **Conselhos para a criação de tabelas**
 - Comece por fazer uma tabela por fonte de dados – expanda se necessário
 - Não se pode ter mais que 255 variáveis numa tabela
 - Cada registo numa tabela deve estar identificado de forma única por uma variável chave
 - As variáveis chave devem ligar as tabelas

CONSTRUÇÃO DE BASE DE DADOS

- **Conselhos para a criação de variáveis**
 - **cada variável deve ser atômica**
 - Diz respeito apenas a uma entidade
 - Exemplo:
 - NomePrim, NomeUlt / não Nome
 - Rua, Localidd, CodigoPos / não Morada
 - DNascDia, DNascMes, DNascAno / não DNasc
 - **cada variável deve ser única: evitar redundância**
 - **Não recolha o que pode calcular**
 - Exemplo
 - $\text{DataObs} - \text{DataNasc} = \text{Idade}$

CONSTRUÇÃO DE BASE DE DADOS

■ Criação de variáveis

■ Designação das variáveis

- Nome curto (até 8 caracteres)
- Sem espaços ou pontuação
(incluindo: _ / - . , : etc)
- Uso de capitulações no meio do nome
 - ParticipID, DataNasc, DataExam, NomeInic, NomeUltm, etc
- Usar o singular (Exame e não Exames)

CONSTRUÇÃO DE BASE DE DADOS

■ Criação de variáveis

■ Evite re-introduzir dados já existentes

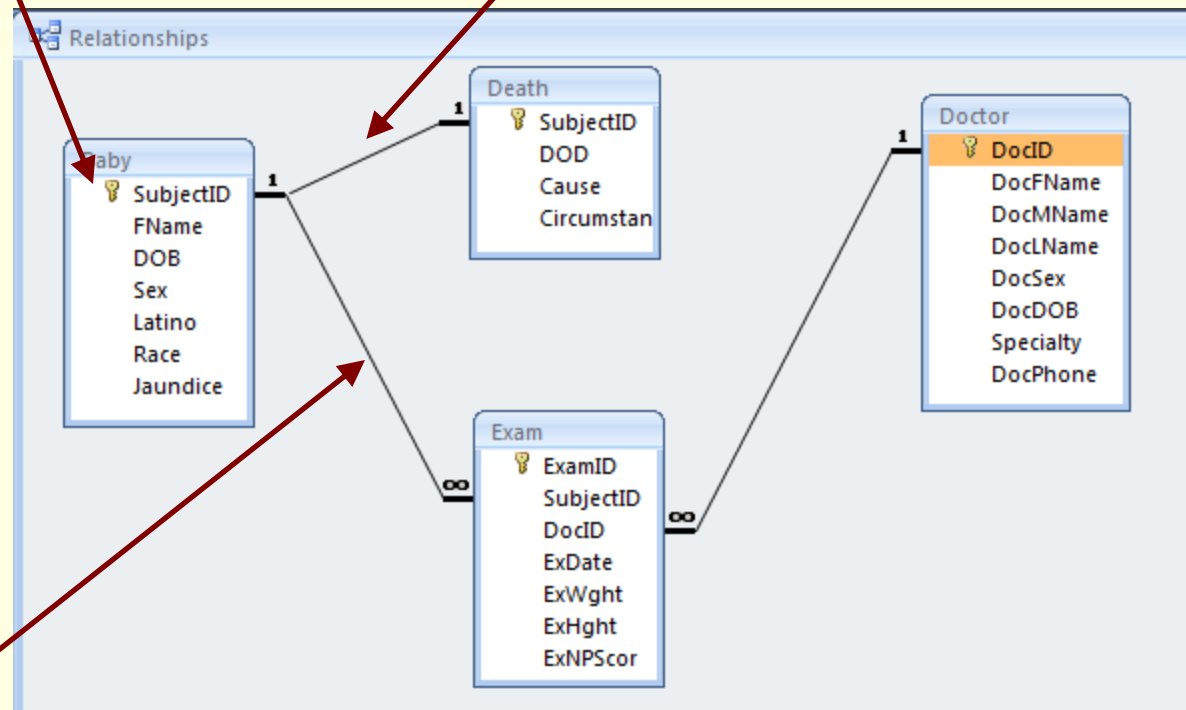
- Por exemplo, dados administrativos
- A maioria dos programas permite a exportação de dados em formatos simples (Comma Separated Values CSV, TXT, etc) que podem ser importados
- Exige “limpeza e verificação”, mas pode poupar tempo e “novos erros”

CONSTRUÇÃO DE BASE DE DADOS

■ Relações entre tabelas

Relação um-para-um:
cada registo de uma tabela tem
apenas um registo noutra tabela

Chave



Relação um-para-muitos:
um registo de uma tabela
pode ter vários registos
noutra tabela

CONSTRUÇÃO DE BASE DE DADOS

■ Construção de consultas

The image displays a Microsoft Access interface. At the top, two tables are shown: 'Baby' and 'Death'. The 'Baby' table has fields: SubjectID (primary key), FName, DOB, and Sex. The 'Death' table has fields: SubjectID (primary key), DOD, Cause, and Circumstances. A 1:1 relationship is established between the SubjectID fields of both tables. Below the tables is a query design grid. The first row is for the 'Baby' table, with 'SubjectID' selected. The second row is for the 'Death' table, with 'SubjectID' selected. The third row is for a calculated field named 'Died', with the formula: `Died: IIf(IsNull([Death].[SubjectID]);No;Yes)`. The fourth row is for the 'Criteria' row, with 'Yes' entered in the 'Died' column. The fifth row is for the 'or' row, which is empty. The sixth row is for the 'Show' row, with checkboxes checked for the 'Baby' and 'Died' columns. The seventh row is for the 'Sort' row, with 'Ascending' selected for the 'Baby' column. The eighth row is for the 'Table' row, with 'Baby' selected. The ninth row is for the 'Field' row, with 'Baby.*' selected. The query view below shows the resulting query with the same fields and criteria.

Field:	Baby.*	Died: IIf(IsNull([Death].[SubjectID]);No;Yes)	
Table:	Baby		
Sort:	Ascending		
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Criteria:		"Yes"	
or:			

Variáveis

Ordenação de dados

Cálculo de fórmulas (IMC)

Condições

CONSTRUÇÃO DE BASE DE DADOS

■ Construção de consultas

- As consultas organizam, ordenam, calculam e formatam os dados
- Existe uma linguagem própria (SQL) mas o Access tem uma ferramenta gráfica para a construção de consultas
- As consultas podem juntar dados de várias tabelas, se estas estiverem relacionadas

CONSTRUÇÃO DE BASE DE DADOS

■ Construção de formulário

Campos de entrada

The screenshot displays the Microsoft Access interface for a database named "[Participantes]". The main window shows a form with several data entry fields: "ParticipID" (value: 1), "ArtriteReum" (checkbox), "InsufCard" (checkbox, checked), "Diabetes" (checkbox), "Dislipidemia" (checkbox, checked), "Hipertensão" (checkbox), "Peso" (value: 24,5), "Altura" (value: 121), and "InvestID" (value: 4). A date field "Data de Diagnóstico da Artrite Reumatóide" contains the value "10/12/2006". A red arrow points from the text "Campos de entrada" to this date field.

An open "Caixa de texto: DataDiagn" properties window is shown on the right. It lists various properties for the text box, such as "Nome", "Origem do controlo", "Formatar", "Casas decimais", "Máscara de introdução", "Valor predefinido", "IME Hold", "Modo IME", "Modo de frase IME", "Regra de validação", "Texto de validação", "Texto da barra de estado", "Comportam. da tecla Enter", "Permitir correcção automática", "Visível", "Mostrar quando", "Vertical", "Quebra de linha asiática", "Activado", "Protegido", "Pesquisador de filtro", "Tabulação automática", "Paragem de tabulação", "Índice de tabulação", "Barras de deslocamento", "Ampliável", and "Redutível". A red arrow points from the text "Máscara de formatação" to the "Máscara de introdução" property, which is set to "00/00/0000;0;". Another red arrow points from the text "Alerta" to the "Regra de validação" property, which is set to "Data de Diagnóstico da Artrite Reumatóide". A third red arrow points from the text "Propriedades" to the "Nome" property, which is set to "DataDiagn".

A "Microsoft Office Access" warning dialog box is open in the foreground, displaying a yellow warning icon and the text "Verifique a data de diagnóstico!". It has "OK" and "Ajuda" buttons. A red arrow points from the text "Alerta" to this dialog box.

Máscara de formatação

Alerta

Propriedades

CONSTRUÇÃO DE BASE DE DADOS

■ Construção de formulários

- ***Use os formulários electrónicos o mais depressa possível***
 - Permite a detecção de erros enquanto o participante está disponível
 - Menos passos na introdução de dados equivale a menos erros
- ***Minimize o registo de dados em papel, mas se o tiver que fazer...***
 - Faça-o o mais perto da data de colecção de dados possível
 - De preferência pela mesma pessoa que fez a observação

CONSTRUÇÃO DE BASE DE DADOS

■ Algumas convenções no uso de controlos

- **Escolhas mutuamente exclusivas:**
usar botões de rádio
- Indica dados em falta (*missing data*)

Branco Europeu

Branco Não Europeu

Asiático

Negro

Melanésio

CONSTRUÇÃO DE BASE DE DADOS

■ Algumas convenções no uso de controlos

- Escolhas não mutuamente exclusivas: usar caixas de verificação (*check boxes*)
- **Cuidado:** Não indica dados em falta ou não verificados (isto é possível por programação)

Dispneia	<input checked="" type="checkbox"/>
Edema Membros Inf	<input type="checkbox"/>
Cardiomegalia	<input checked="" type="checkbox"/>
Edema Pulmonar	<input type="checkbox"/>
Engurgitamento Jugular	<input type="checkbox"/>

CONSTRUÇÃO DE BASE DE DADOS

■ Base de Dados relacional

- É uma colecção de tabelas tipo folha de cálculo (2 dimensões)
 - Linhas são registos
 - Colunas são variáveis
- Tabelas estão relacionadas entre si
 - Uma para uma
 - Uma para muitas
 - (muitas para muitas)

CONSTRUÇÃO DE BASE DE DADOS

■ Vantagens de uma Base de Dados relacional

- Elimina a redundância
 - Nunca é necessário introduzir a mesma informação mais que 1 vez (mesmo que seja a identificação do indivíduo)
- Garante a integridade dos dados
 - Toda a informação encontra-se “em ligação”
- Uso mais fácil de classificações complexas

CONSTRUÇÃO DE BASE DE DADOS

■ Cenário

- Estudo “*Impacto da Icterícia Neonatal no Desenvolvimento Neurológico*”
 - Estudo de coorte para determinar as sequelas neuropsiquiátricas após 5 anos dos recém-nascidos com icterícia neonatal

Adaptado do Curso “Data Management for Research”, 2005, UCSF

CONSTRUÇÃO DE BASE DE DADOS

■ Cenário

- Estudo “*Impacto da Icterícia Neonatal no Desenvolvimento Neurológico*”
 - **Estudo de coorte**
 - **Localização:** um único departamento
 - **População:** crianças com icterícia neonatal e um grupo aleatório de crianças sem icterícia
 - **Variável dependente:** presença ou ausência de icterícia neonatal
 - **Variável independente:** score neuropsiquiátrico (55 a 145 pontos) nas crianças com 5 anos

CONSTRUÇÃO DE BASE DE DADOS



Impacto da Icterícia Neonatal no
Desenvolvimento Neurológico

CADERNO DE RECOLHA DE DADOS

CRIANÇA

Nome: Filipe Monteiro

Data de Nascimento: 01/08/2004

Sexo: M F

Raça / etnia:

- Branco De Origem Europeia
- Branco De Origem Não-Europeia
- Negro
- Mulato
- Cigano
- Asiático (Sub-Continente Indiano)
- Asiático (Chinês, Japonês)
- Melanésico (Timor)
- Outro , País De Origem: _____

Data da observação:

15/06/2006

Nome do examinador:

Daniel Cardoso

Peso (Kgs): 24,9

Altura (cms): 112

Score: 104

CONSTRUÇÃO DE BASE DE DADOS



Impacto da Icterícia Neonatal no Desenvolvimento Neurológico

ID	Nome	DataNascim	Sexo	RacaEtnia	Data de Exa	NomeExaminador	Peso	Altura	SCORE
1	Filipe Monteiro	01-08-2004	<input checked="" type="checkbox"/>	Branco Europeu	19-05-2011	Daniel Cardoso	24,9	112	103

CADERNO DE RECOLHA DE DADOS

CRIANÇA

Nome: **Filipe Monteiro**

Data de Nascimento: **01/08/2004**

Sexo: M F

Raça / etnia:

- Branco De Origem Europeia
- Branco De Origem Não-Europeia
- Negro
- Mulato
- Cigano
- Asiático (Sub-Continente Indiano)
- Asiático (Chinês, Japonês)
- Melanésico (Timor)
- Outro , País De Origem: _____

Data da observação:

15/06/2006

Peso (Kgs): **24,9**

Nome do examinador:

Daniel Cardoso

Altura (cms): **112**

Score: **104**

Field Name	Data Type	
ID	AutoNumber	
Nome	Text	Primeiro e último nome
DataNascimento	Date/Time	dd/mm/aaaa
Sexo	Yes/No	Masculino=sim
RacaEtnia	Text	Lista de opções
DataExame	Date/Time	dd/mm/aaaa
NomeExaminador	Text	Primeiro e último nome
Peso	Number	Em kgs
Altura	Number	Em cms
SCORE	Number	Aos 5 anos, de 55 a 145

General	Lookup
Format	Short Date
Input Mask	00-00-0000;0;/
Caption	Data de Exame
Default Value	Now()
Validation Rule	<= Now()
Validation Text	Data superior à data actual!
Required	Yes
Indexed	No
IME Mode	No Control
IME Sentence Mode	None
Smart Tags	

CONSTRUÇÃO DE BASE DE DADOS

ID	Nome	DataNascim	Sexo	RacaEtnia	Data de Exai	NomeExaminador	Peso	Altura	SCORE
1	Filipe Monteiro	01-08-2004	<input checked="" type="checkbox"/>	Branco Europeu	19-05-2011	Daniel Cardoso	24,9	112	103

“Dicionário de Dados”

- Definição das variáveis
 - Nome
 - Tipo de dados
- Padrão da entrada
- Valores pré-definidos
- Regras de validação
- Texto de alerta
- Permissão de dados omissos

Field Name	Data Type	
ID	AutoNumber	
Nome	Text	Primeiro e último nome
DataNascimento	Date/Time	dd/mm/aaaa
Sexo	Yes/No	Masculino=sim
RacaEtnia	Text	Lista de opções
DataExame	Date/Time	dd/mm/aaaa
NomeExaminador	Text	Primeiro e último nome
Peso	Number	Em kgs
Altura	Number	Em cms
SCORE	Number	Aos 5 anos, de 55 a 145

General	Lookup
Format	Short Date
Input Mask	00-00-0000;0;/
Caption	Data de Exameção
Default Value	Now()
Validation Rule	<= Now()
Validation Text	Data superior à data actual!
Required	Yes
Indexed	No
IME Mode	No Control
IME Sentence Mode	None
Smart Tags	

CONSTRUÇÃO DE BASE DE DADOS

■ Tabela com uma observação por participante

ExamID	SubjectID	DocID	ExDate	ExWght	ExHght	ExNPScor	FName	Age
11	4184	5	23-02-2010	21,0	106	124	Jordan	4
12	4430	5	23-02-2010	23,3	105	135	Cecil	4
13	4670	5	27-02-2010	20,2	106	56	Amelia	4
14	5051	5	07-03-2010	19,3	111	145	William	5
15	5305	2	23-02-2010	17,5	116	66	Andrew	5
16	5310	5	24-02-2010	19,6	115	95	Eric	5
17	5592	5	08-03-2010	19,7	109	105	Dalia	5
18	5787	5	22-03-2010	20,4	113	122	Gordan	5
19	5926	5	12-03-2010	20,1	106	132	Kevin	5
20	5988	4	12-03-2010	16,3	121	78	Chace	5
21	6068	2	21-03-2010	16,2	106	86	Joseph	5
22	6075	5	22-03-2010	19,9	117	66	Nicholas	5
23	6178	4	22-03-2010	16,2	108	88	Mason	5
24	6308	1	23-03-2010	16,9	116	63	Bradford	5
25	6517	3	16-03-2010	15,7	110	96	George	5
26	6592	1	02-04-2010	14,4	111	80	Brian	5
27	6744	2	26-03-2010	16,3	113	85	Corry	5
28	6812	5	28-03-2010	16,5	111	58	Caroline	5
29	7049	5	28-03-2010	17,7	111	89	Caitlin	5
30	7305	2	25-03-2010	17,4	99	92	Eric	5
31	7399	2	05-04-2010	18,8	103	103	Isaac	5
32	7405	3	07-04-2010	15,4	109	93	Margaret	5
33	7435	5	11-04-2010	20,4	112	93	Eric	5
34	7513	5	03-04-2010	18,8	111	93	Roxanne	5

CONSTRUÇÃO DE BASE DE DADOS

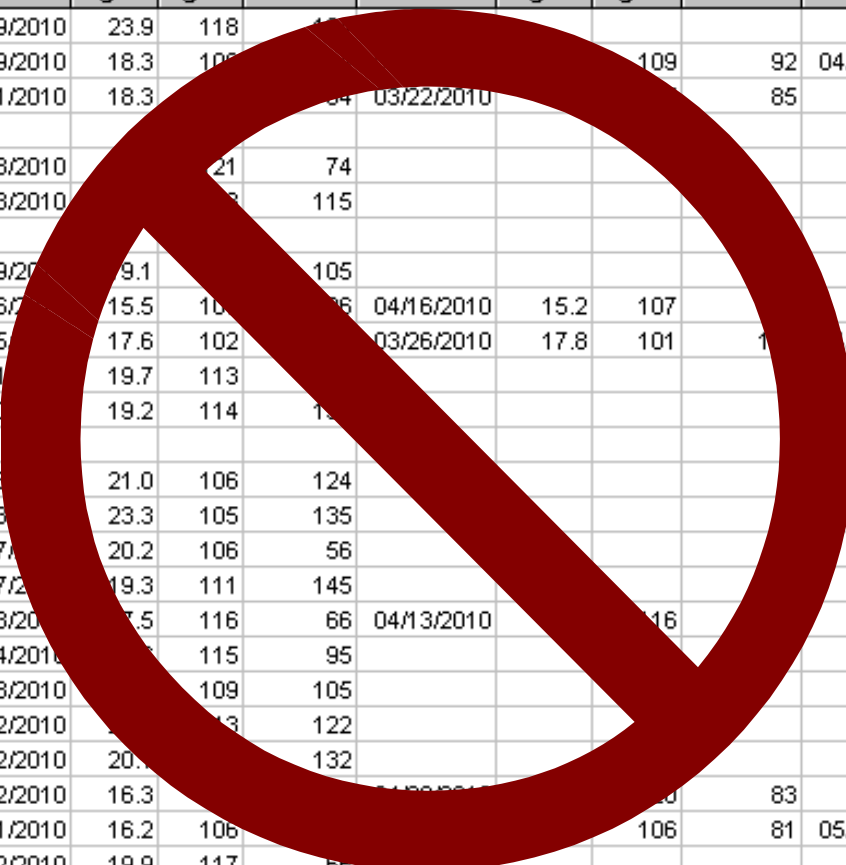
- Se alguns participantes tiverem mais de uma observação, o que fazer?

	Study	FName	Sex	Jaun	Date1	Wght1	Hght1	NPScor1	Date2	Wght2	Hght2	NPScor2	Date3	Wght3	Hght3	NPScor3
▶	2101	Robert	M	<input checked="" type="checkbox"/>	01/29/2010	23.9	118	104								
	2322	Helen	F	<input type="checkbox"/>	01/29/2010	18.3	109	94	03/19/2010	18.6	109	92	04/05/2010	19.1	110	88
	2376	Amy	F	<input checked="" type="checkbox"/>	02/01/2010	18.3	117	84	03/22/2010	18.5	117	85				3
	2390	Alejandr	M	<input type="checkbox"/>												
	2497	Isiah	M	<input type="checkbox"/>	02/18/2010	20.5	121	74								
	2569	Joshua	M	<input checked="" type="checkbox"/>	02/13/2010	24.8	113	115								
	2819	Ryan	M	<input type="checkbox"/>												
	3019	Morgan	F	<input type="checkbox"/>	02/09/2010	19.1	105	105								
	3031	Cody	M	<input type="checkbox"/>	02/26/2010	15.5	107	126	04/16/2010	15.2	107	132				
	3290	Alia	F	<input checked="" type="checkbox"/>	02/05/2010	17.6	102	136	03/26/2010	17.8	101	145	04/12/2010	18.0	102	125
	3374	Zachary	M	<input checked="" type="checkbox"/>	02/11/2010	19.7	113	3								
	3625	David	M	<input checked="" type="checkbox"/>	02/10/2010	19.2	114	134								
	3901	Jacksor	M	<input type="checkbox"/>												
	4184	Jordan	F	<input type="checkbox"/>	02/23/2010	21.0	106	124								
	4430	Cecil	M	<input type="checkbox"/>	02/23/2010	23.3	105	135								
	4670	Amelia	F	<input type="checkbox"/>	02/27/2010	20.2	106	56								
	5051	William	M	<input checked="" type="checkbox"/>	03/07/2010	19.3	111	145								
	5305	Andrew	M	<input checked="" type="checkbox"/>	02/23/2010	17.5	116	66	04/13/2010	17.4	116	65				
	5310	Eric	M	<input type="checkbox"/>	02/24/2010	19.6	115	95								
	5592	Dalisay	F	<input checked="" type="checkbox"/>	03/08/2010	19.7	109	105								
	5787	Gordan	M	<input type="checkbox"/>	03/22/2010	20.4	113	122								
	5926	Kevin	M	<input checked="" type="checkbox"/>	03/12/2010	20.1	106	132								
	5988	Chace	M	<input type="checkbox"/>	03/12/2010	16.3	121	78	04/30/2010	16.3	120	83				
	6068	Joseph	M	<input type="checkbox"/>	03/21/2010	16.2	106	86	05/09/2010	16.7	106	81	05/26/2010	16.7	106	83
	6075	Nicholas	M	<input type="checkbox"/>	03/22/2010	19.9	117	66								

Record: 1 of 405

CONSTRUÇÃO DE BASE DE DADOS

- Se alguns participantes tiverem mais de uma observação, o que fazer?

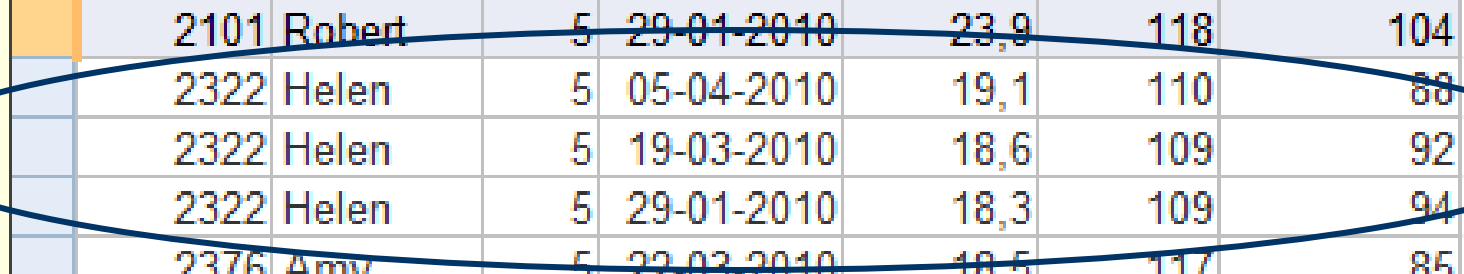


Study	FName	Sex	Jaun	Date1	Wght1	Hght1	NPScor1	Date2	Wght2	Hght2	NPScor2	Date3	Wght3	Hght3	NPScor3
2101	Robert	M	<input checked="" type="checkbox"/>	01/29/2010	23.9	118	105								
2322	Helen	F	<input type="checkbox"/>	01/29/2010	18.3	109	92	04/05/2010	19.1	110	88				
2376	Amy	F	<input checked="" type="checkbox"/>	02/01/2010	18.3	104	85	03/22/2010							3
2390	Alejandr	M	<input type="checkbox"/>												
2497	Isiah	M	<input type="checkbox"/>	02/18/2010		21	74								
2569	Joshua	M	<input checked="" type="checkbox"/>	02/13/2010			115								
2819	Ryan	M	<input type="checkbox"/>												
3019	Morgan	F	<input type="checkbox"/>	02/09/2010	19.1		105								
3031	Cody	M	<input type="checkbox"/>	02/26/2010	15.5	106	96	04/16/2010	15.2	107					
3290	Alia	F	<input checked="" type="checkbox"/>	02/05/2010	17.6	102	101	03/26/2010	17.8	101	101	01/12/2010	18.0	102	125
3374	Zachary	M	<input checked="" type="checkbox"/>	02/11/2010	19.7	113									
3625	David	M	<input checked="" type="checkbox"/>	02/10/2010	19.2	114	105								
3901	Jacksor	M	<input type="checkbox"/>												
4184	Jordan	F	<input type="checkbox"/>	02/23/2010	21.0	106	124								
4430	Cecil	M	<input type="checkbox"/>	02/23/2010	23.3	105	135								
4670	Amelia	F	<input type="checkbox"/>	02/27/2010	20.2	106	56								
5051	William	M	<input checked="" type="checkbox"/>	03/07/2010	19.3	111	145								
5305	Andrew	M	<input checked="" type="checkbox"/>	02/23/2010	17.5	116	66	04/13/2010		116					
5310	Eric	M	<input type="checkbox"/>	02/24/2010		115	95								
5592	Dalisay	F	<input checked="" type="checkbox"/>	03/08/2010		109	105								
5787	Gordan	M	<input type="checkbox"/>	03/22/2010		113	122								
5926	Kevin	M	<input checked="" type="checkbox"/>	03/12/2010	20.1		132								
5988	Chace	M	<input type="checkbox"/>	03/12/2010	16.3						83				
6068	Joseph	M	<input type="checkbox"/>	03/21/2010	16.2	106				106	81	05/26/2010	16.7	106	83
6075	Nicholas	M	<input type="checkbox"/>	03/22/2010	19.9	117	80								

Record: 1 of 405

CONSTRUÇÃO DE BASE DE DADOS

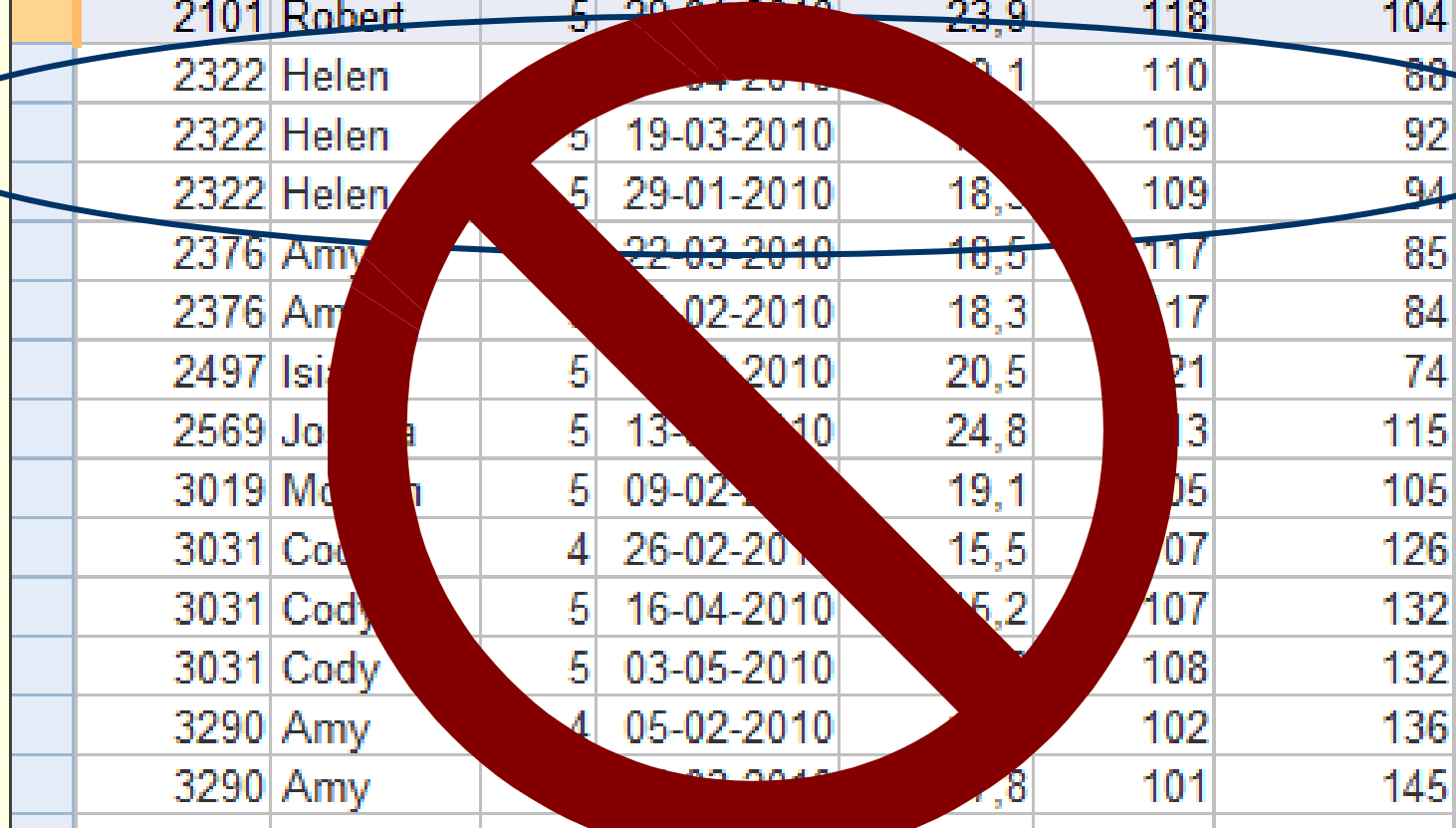
- Se alguns participantes tiverem mais de uma observação, o que fazer?



StudyI	FName	Ag	ExDate	ExWgl	ExHgl	ExNPSci
2101	Robert	5	29-01-2010	23,9	118	104
2322	Helen	5	05-04-2010	19,1	110	88
2322	Helen	5	19-03-2010	18,6	109	92
2322	Helen	5	29-01-2010	18,3	109	94
2376	Amy	5	22-03-2010	18,5	117	85
2376	Amy	5	01-02-2010	18,3	117	84
2497	Isiah	5	18-02-2010	20,5	121	74
2569	Joshua	5	13-02-2010	24,8	113	115
3019	Morgan	5	09-02-2010	19,1	105	105
3031	Cody	4	26-02-2010	15,5	107	126
3031	Cody	5	16-04-2010	15,2	107	132
3031	Cody	5	03-05-2010	15,6	108	132
3290	Amy	4	05-02-2010	17,6	102	136
3290	Amy	5	26-03-2010	17,8	101	145

CONSTRUÇÃO DE BASE DE DADOS

- Se alguns participantes tiverem mais de uma observação, o que fazer?



StudyID	FName	Age	ExDate	ExWgt	ExHgt	ExNPScore
2101	Robert	5	29-01-2010	23,9	118	104
2322	Helen	5	19-03-2010	19,1	110	88
2322	Helen	5	19-03-2010	19,1	109	92
2322	Helen	5	29-01-2010	18,5	109	94
2376	Amy	5	22-03-2010	18,5	117	85
2376	Amy	5	02-2010	18,3	117	84
2497	Isis	5	02-2010	20,5	121	74
2569	Joy	5	13-02-2010	24,8	113	115
3019	Molly	5	09-02-2010	19,1	105	105
3031	Cody	4	26-02-2010	15,5	107	126
3031	Cody	5	16-04-2010	15,2	107	132
3031	Cody	5	03-05-2010	15,2	108	132
3290	Amy	4	05-02-2010	17,8	102	136
3290	Amy	4	02-2010	17,8	101	145

CONSTRUÇÃO DE BASE DE DADOS

■ Normalização

Se algumas crianças têm mais que uma observação, “normaliza-se” os registo em duas tabelas

- Uma tabela para as crianças
- Uma tabela para as observações

CONSTRUÇÃO DE BASE DE DADOS

Normalização

Subject	FName	DOB	Se	Latir	Race	ExamI	Subject	DocID	ExDate	ExWgt	ExHgt	Ext
2101	Robert	08-01-2005	M	<input checked="" type="checkbox"/>								
2322	Helen	06-01-2005	F	<input type="checkbox"/>	Asian							
2376	Amy	13-01-2005	F	<input type="checkbox"/>	Asian							
2390	Alejandro	14-01-2005	M	<input checked="" type="checkbox"/>								
2497	Isiah	18-01-2005	M	<input type="checkbox"/>	Black or African Ar	709	2322	3	29-01-2010	18,3	109	
2569	Joshua	23-01-2005	M	<input checked="" type="checkbox"/>		706	2101	4	29-01-2010	23,9	118	
2819	Ryan	26-01-2005	M	<input type="checkbox"/>	White	711	2376	2	01-02-2010	18,3	117	
3019	Morgan	29-01-2005	F	<input type="checkbox"/>	White	8	3290	2	05-02-2010	17,6	102	
3031	Cody	15-03-2005	M	<input type="checkbox"/>	White	6	3019	4	09-02-2010	19,1	105	
3290	Amy	16-02-2005	F	<input type="checkbox"/>	White	10	3625	5	10-02-2010	19,2	114	
3374	Zachary	21-02-2005	M	<input type="checkbox"/>	Asian	20	5988	4	12-03-2010	16,3	121	
3625	David	22-02-2005	M	<input type="checkbox"/>	White	19	5926	5	12-03-2010	20,1	106	
3901	Jackson	28-02-2005	M	<input type="checkbox"/>	White	25	6517	3	16-03-2010	15,7	110	
4184	Jordan	02-03-2005	F	<input type="checkbox"/>	White	708	2322	2	19-03-2010	18,6	109	
4430	Cecil	02-03-2005	M	<input type="checkbox"/>	Asian	21	6068	2	21-03-2010	16,2	106	
4670	Amelia	06-03-2005	F	<input checked="" type="checkbox"/>	White	22	6075	5	22-03-2010	19,9	117	
						23	6178	4	22-03-2010	16,2	108	
						18	5787	5	22-03-2010	20,4	113	
						710	2376	4	22-03-2010	18,5	117	
						26	6592	1	02-04-2010	14,4	111	
						34	7513	5	03-04-2010	18,8	111	
						31	7399	2	05-04-2010	18,8	102	
						707	2322	1	05-04-2010	19,1	110	
						39	7737	5	06-04-2010	23,5	109	
						27	7405	3	07-04-2010	15,4	108	

CONSTRUÇÃO DE BASE DE DADOS

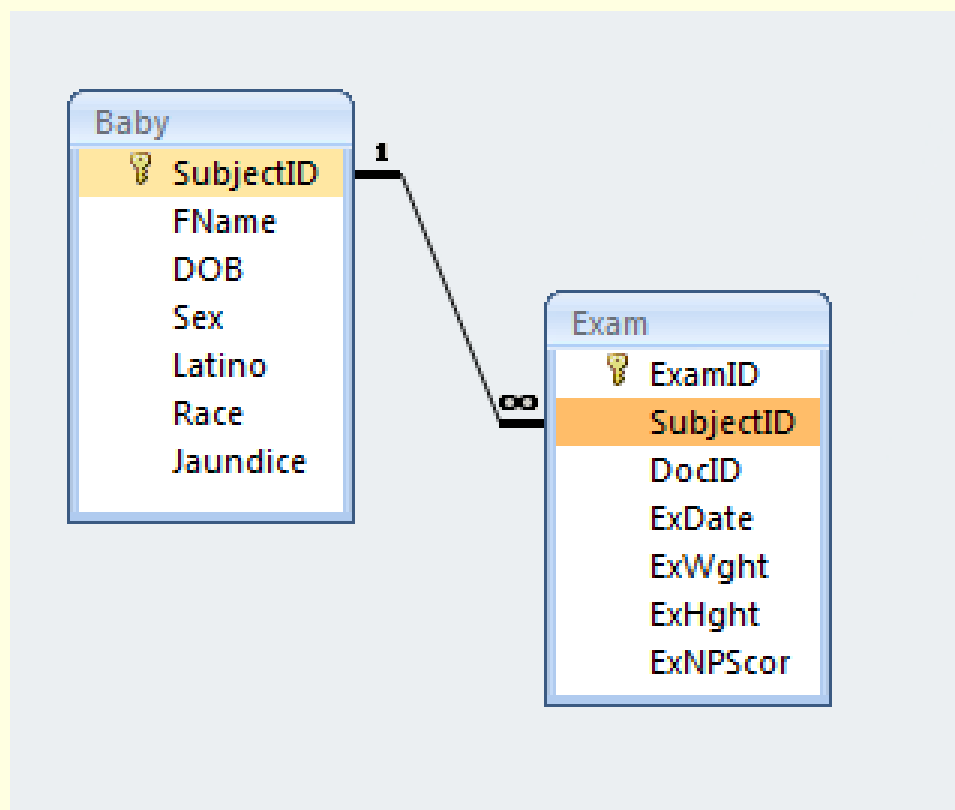
Normalização

Chaves!

Subject	Name	DOB	Sex	LatIn	Race	ExamID	Subject	DocID	ExDate	ExWgt	ExHgt	Ext
2101	Robert	08-01-2005	M	<input checked="" type="checkbox"/>								
2322	Helen	06-01-2005	F	<input type="checkbox"/>	Asian							
2376	Amy	13-01-2005	F	<input type="checkbox"/>	Asian							
2390	Alejandro	14-01-2005	M	<input checked="" type="checkbox"/>								
2497	Isiah	18-01-2005	M	<input type="checkbox"/>	Black or African Ar	709	2322	3	29-01-2010	18,3	109	
2569	Joshua	23-01-2005	M	<input checked="" type="checkbox"/>		706	2101	4	29-01-2010	23,9	118	
2819	Ryan	26-01-2005	M	<input type="checkbox"/>	White	711	2376	2	01-02-2010	18,3	117	
3019	Morgan	29-01-2005	F	<input type="checkbox"/>	White	8	3290	2	05-02-2010	17,6	102	
3031	Cody	15-03-2005	M	<input type="checkbox"/>	White	6	3019	4	09-02-2010	19,1	105	
3290	Amy	16-02-2005	F	<input type="checkbox"/>	White	10	3625	5	10-02-2010	19,2	114	
3374	Zachary	21-02-2005	M	<input type="checkbox"/>	Asian	20	5988	4	12-03-2010	16,3	121	
3625	David	22-02-2005	M	<input type="checkbox"/>	White	19	5926	5	12-03-2010	20,1	106	
3901	Jackson	28-02-2005	M	<input type="checkbox"/>	White	25	6517	3	16-03-2010	15,7	110	
4184	Jordan	02-03-2005	F	<input type="checkbox"/>	White	708	2322	2	19-03-2010	18,6	109	
4430	Cecil	02-03-2005	M	<input type="checkbox"/>	Asian	21	6068	2	21-03-2010	16,2	106	
4670	Amelia	06-03-2005	F	<input checked="" type="checkbox"/>	White	22	6075	5	22-03-2010	19,9	117	
						23	6178	4	22-03-2010	16,2	108	
						18	5787	5	22-03-2010	20,4	113	
						710	2376	4	22-03-2010	18,5	117	
						26	6592	1	02-04-2010	14,4	111	
						34	7513	5	03-04-2010	18,8	111	
						31	7399	2	05-04-2010	18,8	102	
						707	2322	1	05-04-2010	19,1	110	
						39	7737	5	06-04-2010	23,5	109	
						37	7405	3	07-04-2010	15,4	108	

CONSTRUÇÃO DE BASE DE DADOS

■ Normalização: relação um-para-muitos



CONSTRUÇÃO DE BASE DE DADOS

- **Normalização**
 - Mantém a integridade dos dados
 - Facilita a pesquisa de dados nas consultas e a re-organização das tabelas
- **Principal vantagem do uso do Access versus Excel ou outros sistemas tipo folhas de cálculos**

CONSTRUÇÃO DE BASE DE DADOS

■ Cenário

- Estudo “*Impacto da Icterícia Neonatal no Desenvolvimento Neurológico*”
 - A avaliação neuropsiquiátrica foi efectuada por diferentes médicos
 - Um dos objectivos é examinar se as características do médico (especialidade, idade, sexo) influenciam a avaliação neuropsiquiátrica
 - Um médico avalia várias crianças
 - Cada criança pode ter mais que um exame
 - Cada criança nunca é examinada pelo mesmo médico mais do que uma vez

CONSTRUÇÃO DE BASE DE DADOS

■ Dados dos observadores

	DocId	DocFNam	DocMNI	DocLNam	DocSe	DocDOB	Specialty	DocPhone
+	1	Charles	Everett	Koop	F	10/14/1916	Pediatric Surgery	(555) 631-2442
+	2	Antonia	Coello	Novello	F	8/23/1944	Pediatrics	(555) 502-5398
+	3	Minnie	Joycelyn	Elders	F	8/13/1933	Pediatrics	(555) 555-5555
+	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
+	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999

CONSTRUÇÃO DE BASE DE DADOS

- Como identificar os observadores em cada observação?

ExamI	Subject	Exam	ExDate	ExWgt	ExHgt	ExNPSco	Doct	DocFNam	DocMNI	DocLNam	DocSe	DocDOB	Specialty	DocPhone
709	2322	3	29-01-2010	18,3	109	94	3	Minnie	Joycelyn	Elders	F	8/13/1933	Pediatrics	(555) 555-5555
706	2101	4	29-01-2010	23,9	118	104	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
711	2376	2	01-02-2010	18,3	117	84	2	Antonia	Coello	Novello	F	8/23/1944	Pediatrics	(555) 502-5398
8	3290	2	05-02-2010	17,6	102	136	2	Antonia	Coello	Novello	F	8/23/1944	Pediatrics	(555) 502-5398
6	3019	4	09-02-2010	19,1	105	105	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
10	3625	5	10-02-2010	19,2	114	134	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
5	2569	4	13-02-2010	24,8	113	115	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
4	2497	4	18-02-2010	20,5	121	74	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
12	4430	5	23-02-2010	23,3	105	135	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
11	4184	5	23-02-2010	21,0	106	124	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
15	5305	2	23-02-2010	17,5	116	66	2	Antonia	Coello	Novello	F	8/23/1944	Pediatrics	(555) 502-5398
16	5310	5	24-02-2010	19,6	115	95	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
7	3031	1	26-02-2010	15,5	107	126	1	Charles	Everett	Koop	F	10/14/1916	Pediatric Surgery	(555) 631-2442
13	4670	5	27-02-2010	20,2	106	56	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
14	5051	5	07-03-2010	19,3	111	145	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
17	5592	5	08-03-2010	19,7	109	105	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
20	5988	4	12-03-2010	16,3	121	78	4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
19	5926	5	12-03-2010	20,1	106	132	5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999
25	6517	3	16-03-2010	15,7	110	96	3	Minnie	Joycelyn	Elders	F	8/13/1933	Pediatrics	(555) 555-5555

CONSTRUÇÃO DE BASE DE DADOS

- Como identificar os observadores em cada observação?

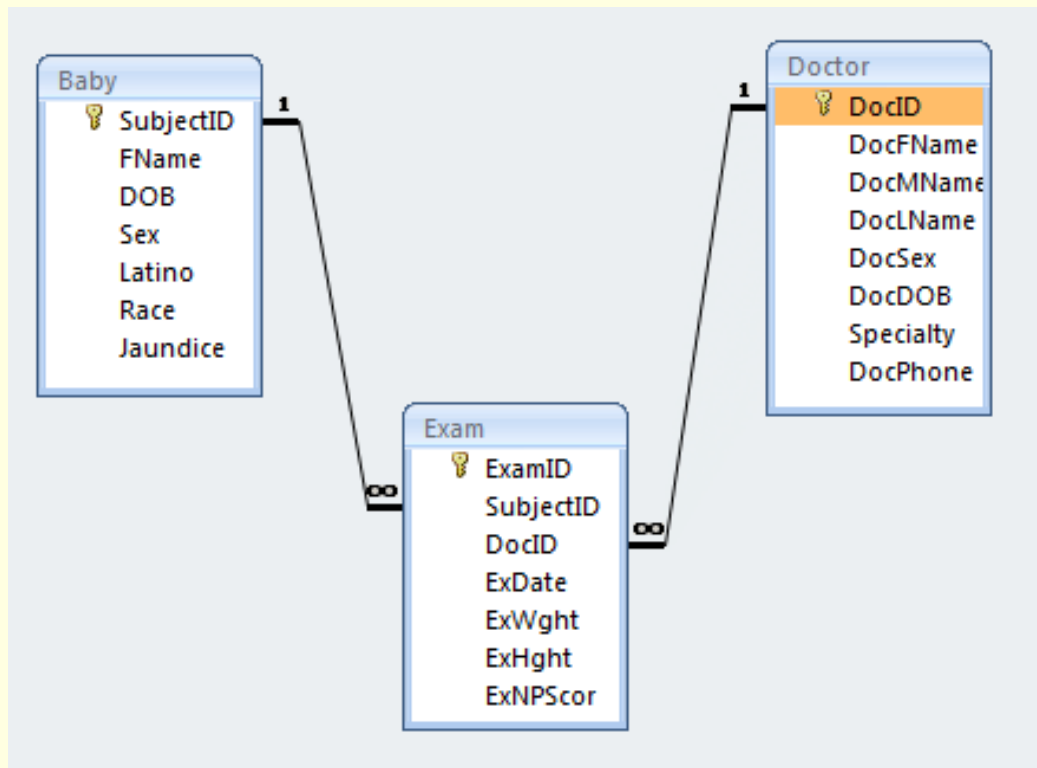
Chaves!

ExamID	SubjectID	DocID	ExDate	ExWgt	ExHgt	ExNPSco
705	2323	3	29-01-2010	18,3	109	94
706	2101	4	29-01-2010	23,9	118	104
711	2376	2	01-02-2010	18,3	117	84
8	3290	2	05-02-2010	17,6	102	130
6	3019	4	07-02-2010	19,1	105	105
10	3625	5	10-02-2010	19,2	114	134
5	2569	4	13-02-2010	24,8	113	75
4	2497	4	18-02-2010	20,5	121	74
12	4430	5	23-02-2010	23,3	105	135
11	4184	5	23-02-2010	21,0	107	124
15	5305	2	23-02-2010	17,5	116	66
16	5310	5	24-02-2010	19,6	115	95
7	3031	1	26-02-2010	15,5	107	126
13	4670	5				
14	5051	5				
17	5592	5				
20	5988	4				

DocID	DocFNam	DocMNI	DocLNam	DocSe	DocDOB	Specialty	DocPhone
1	Charles	Everett	Keop	F	10/14/1916	Pediatric Surgery	(555) 631-2442
2	Antonia	Coello	Novello	F	8/23/1944	Pediatrics	(555) 502-3398
3	Minnie	Joycelyn	Elders	F	8/13/1933	Pediatrics	(555) 555-5555
4	David		Satcher	M	3/2/1941	Family Practice	(555) 362-4360
5	Julius	B.	Richmond	M	9/26/1916	Pediatrics	(999) 999-9999

CONSTRUÇÃO DE BASE DE DADOS

- Como identificar os observadores em cada observação?



CONSTRUÇÃO DE BASE DE DADOS

- **A relação um-para-um**
 - **Alguns campos são específicos do participante, mas apenas são preenchidos em situações especiais**
 - **Exemplo: Registo de morte**

CONSTRUÇÃO DE BASE DE DADOS

Subject	FName	DOB	Se	Jaundic	DOD	Cause	Circumstances
2101	Robert	06-01-2005	M	<input checked="" type="checkbox"/>			
2322	Helen	06-01-2005	F	<input type="checkbox"/>			
2376	Amy	13-01-2005	F	<input checked="" type="checkbox"/>			
2390	Alejandro	14-01-2005	M	<input type="checkbox"/>			
2497	Isiah	18-01-2005	M	<input type="checkbox"/>			
2569	Joshua	23-01-2005	M	<input checked="" type="checkbox"/>			
2819	Ryan	26-01-2005	M	<input type="checkbox"/>			
3019	Morgan	29-01-2005	F	<input type="checkbox"/>			
3031	Cody	15-03-2005	M	<input type="checkbox"/>			
3290	Amy	16-02-2005	F	<input checked="" type="checkbox"/>			
3374	Zachary	21-02-2005	M	<input checked="" type="checkbox"/>	21-10-2008	Acute Myelogenous Leukemia	Diagnosed with AML in January
3625	David	22-02-2005	M	<input checked="" type="checkbox"/>			
3901	Jackson	28-02-2005	M	<input type="checkbox"/>			
4184	Jordan	02-03-2005	F	<input type="checkbox"/>			
4430	Cecil	02-03-2005	M	<input type="checkbox"/>			
4670	Amelia	06-03-2005	F	<input type="checkbox"/>			
5051	William	14-02-2005	M	<input checked="" type="checkbox"/>			
5305	Andrew	02-02-2005	M	<input checked="" type="checkbox"/>			
5310	Eric	03-02-2005	M	<input type="checkbox"/>			

**Os campos são específicos dos participantes, mas os dados existirão apenas para alguns participantes!
Manter estas colunas na tabela dos participantes leva a muitos campos vazios e desperdício de espaço.**

CONSTRUÇÃO DE BASE DE DADOS

- **A relação um-para-um**
 - **Alguns campos são específicos do participante, mas apenas são preenchidos em situações especiais**
 - Exemplo: Registo de morte
 - **A criação de uma tabela própria permite**
 - Eliminar campos vazios
 - Reduzir o espaço reservado (memória, tamanho do ficheiro)
 - Ajudar na pesquisa de dados

CONSTRUÇÃO DE BASE DE DADOS

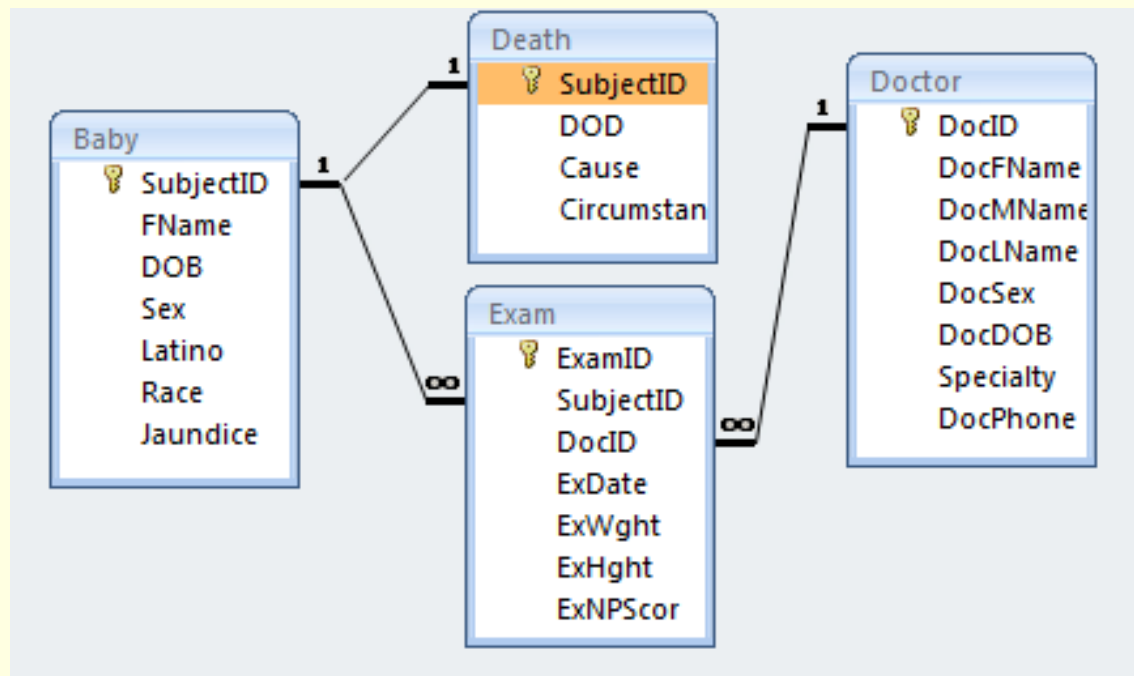
The image shows a database interface with two tables: 'Baby' and 'Death'. The 'Baby' table has columns: Subject, FName, DOB, and Sex. The 'Death' table has columns: Subjec, DOD, Cause, and Circumstance. A blue oval highlights the subject 3374 in both tables, and a blue arrow points from the 'Death' table to the 'Baby' table.

Subject	FName	DOB	Sex
2101	Robert	06-01-2005	M
2322	Helen	06-01-2005	F
2376	Amy	13-01-2005	F
2390	Alejandro	14-01-2005	M
2497	Isiah	18-01-2005	M
2569	Joshua	23-01-2005	M
2819	Ryan	26-01-2005	M
3019	Morgan	29-01-2005	F
3031	Cody	15-03-2005	M
3290	Amy	16-02-2005	F
3374	Zachary	21-02-2005	M
3625	David	22-02-2005	M
3901	Jackson	28-02-2005	M

Subjec	DOD	Cause	Circumstance
54283	01-01-2009	Meningococemia	When in the Course
47609	09-10-2007	Respiratory failure; Viral pneumonia	I say to you today, r
26425	01-01-2006	Multiple congenital anomalies	Four score and seve
3374	21-10-2008	Acute Myelogenous Leukemia	Diagnosed with AM
2101			

CONSTRUÇÃO DE BASE DE DADOS

■ Relação um-para-um



CONSTRUÇÃO DE BASE DE DADOS

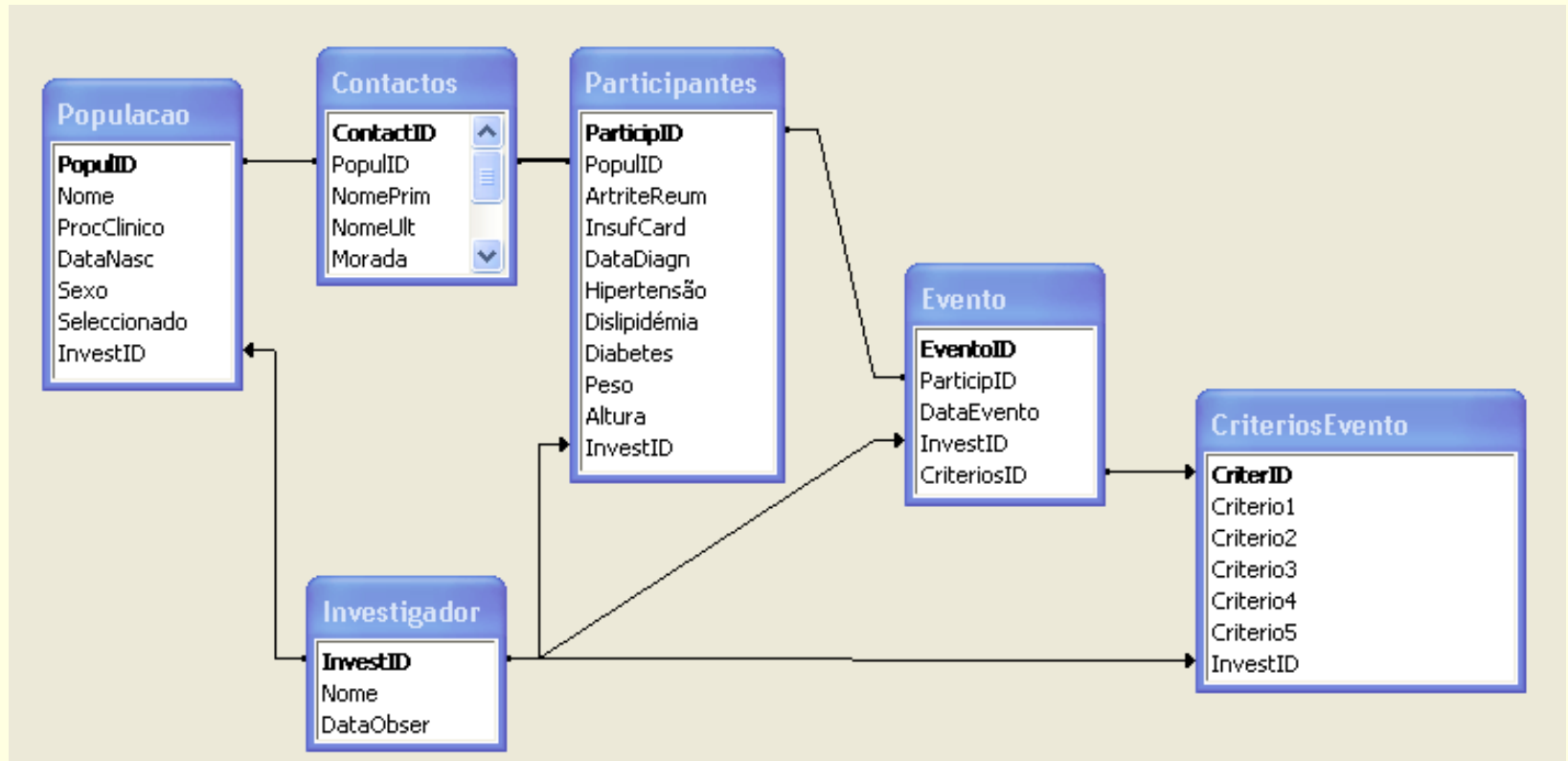
- **A relação um-para-um**

- **Concentrar dados confidenciais ou que serão excluídos da análise estatística**
 - Nome, morada, telefone, números identificadores, etc.
 - Dados administrativos
- **Cuidado:** garantir que a exclusão destes dados não “destrói” a integridade da Base de Dados!

CONSTRUÇÃO DE BASE DE DADOS

■ A relação um-para-um

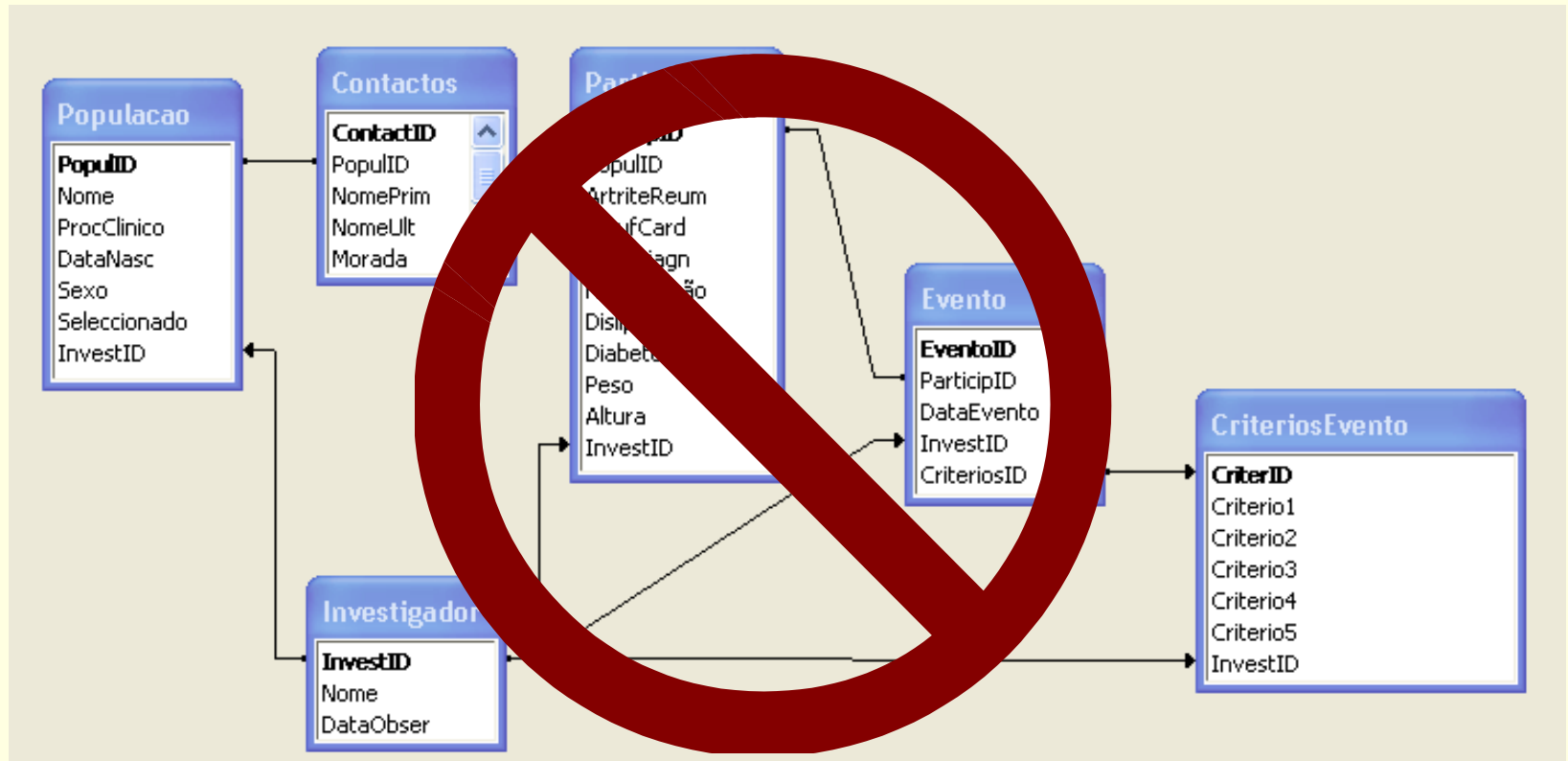
- **Cuidado:** garantir que a exclusão destes dados não “destrói” a integridade da Base de Dados!



CONSTRUÇÃO DE BASE DE DADOS

■ A relação um-para-um

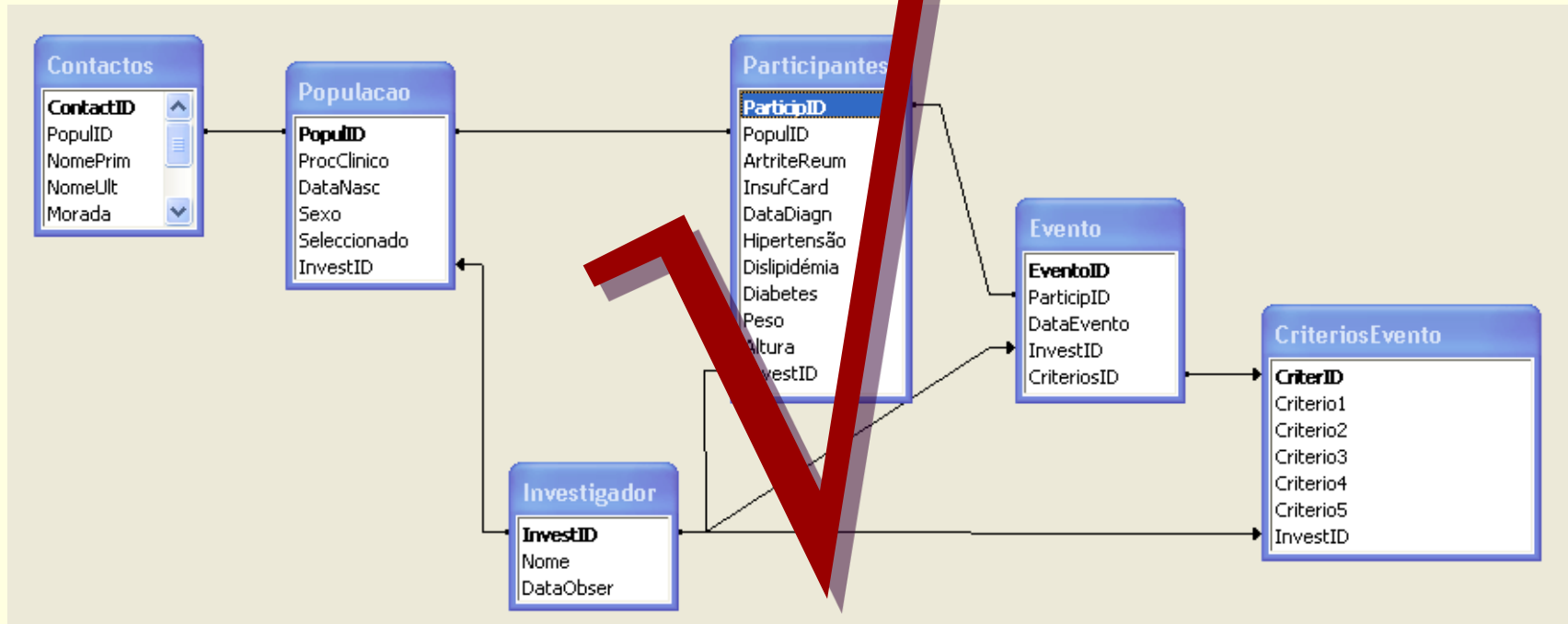
- **Cuidado:** garantir que a exclusão destes dados não “destrói” a integridade da Base de Dados!



CONSTRUÇÃO DE BASE DE DADOS

■ A relação um-para-um

- **Cuidado:** garantir que a exclusão destes dados não “destrói” a integridade da Base de Dados!



CONSTRUÇÃO DE BASE DE DADOS

■ A relação um-para-um

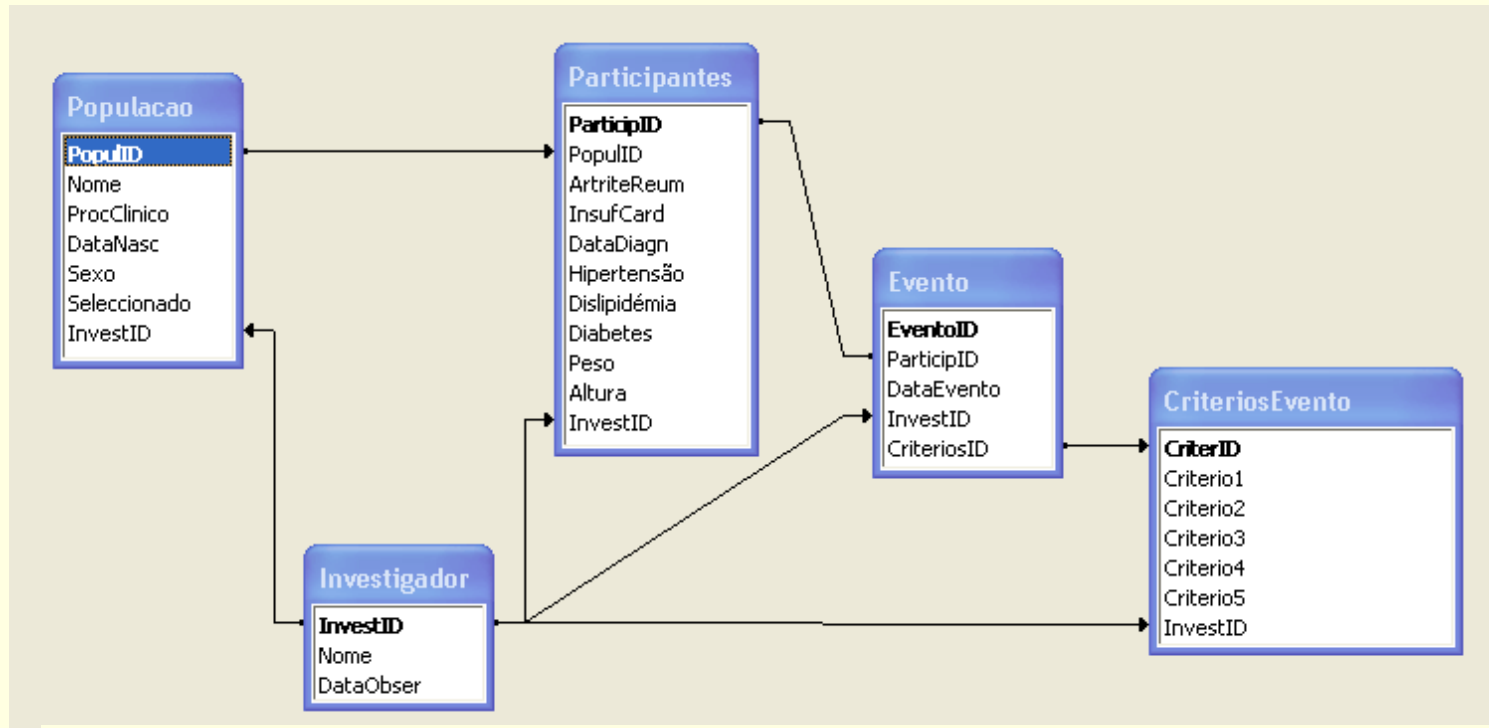
- Tabelas que provêm ou são **importadas de outros ficheiros electrónicos** (exemplo: diagnósticos, processos clínicos electrónicos, dados administrativos).
- Tabelas que provêm de fontes distintas e **podem entrar em conflito** se “misturadas”
 - Entrevista vs. questionário
 - Dados de familiares
 - Processo clínico de internamento vs. processo do serviço de urgência
 - etc

CONSTRUÇÃO DE BASE DE DADOS

- **Criar um modelo de dados (*Data model*)**
 - **Considerar a inclusão de todos os dados relativos ao estudo**
incluindo:
 - Fazer ligações (*links*) a outros ficheiros (exemplo, imagens)
 - Incluir dados administrativos, dos investigadores, agenda, etc.
 - Incluir dados de selecção dos participantes
 - etc., etc., etc.

CONSTRUÇÃO DE BASE DE DADOS

- Exemplo: estudo da associação entre a artrite reumatóide e a insuficiência cardíaca



Ligação entre a identificação do investigador e a colheita de dados nos vários campos

CONSTRUÇÃO DE BASE DE DADOS

- **Criar um modelo de dados (*Data model*)**
 - **Sinais de um mau modelo**
 - **Redundância** – repetição na inserção dos mesmos dados ou dados relativos (exemplo: data de nascimento e idade)
 - **Muitos campos vazios**
 - **Impossibilidade de definir subgrupos**

CONSTRUÇÃO DE BASE DE DADOS

- **A realização de backups no Access**
 - Faça uma cópia regular da Base de Dados
 - A corrupção do ficheiro “vai acontecer”!
 - Para um meio externo (não apenas no disco rígido!)
 - Verifique a adequação da cópia, abrindo-a

CONSTRUÇÃO DE BASE DE DADOS

- **Outras potencialidades avançadas**
 - Uso online
 - Uma base de dados em uso por vários “clientes”
 - Aplicações do VisualBasic no Access

CONSTRUÇÃO DE BASE DE DADOS

■ Questões?