



Instituto de Pesquisa clínica Evandro Chagas – Fundação Oswaldo Cruz  
Programa de pós-graduação de pesquisa clínica em doenças infecciosas - 2013.

“Computação: introdução a aplicativos de informática para auxílio à pesquisa”



EpiData - Roteiro de aula

Coordenador: Pedro Emmanuel A. A. do Brasil

Coordenador: Paula Mendes Luz



Referências:

[http://www.epidata.dk/php/downloadc.php?file=epidata\\_pdf.pdf](http://www.epidata.dk/php/downloadc.php?file=epidata_pdf.pdf)

<http://www.epidata.dk/php/downloadc.php?file=epitour.pdf>

<http://www.epidata.dk/php/downloadc.php?file=dmePIDATA.pdf>

<http://onkpat.ki.se/1?node=47234>

<http://www.pitt.edu/~Super1/lecture/lec7171/001.htm>

ÍNDICE

OBJETIVOS DO MÓDULO .....	3
RAZÕES PELAS QUAIS EPIDATA 3.1® FOI ESCOLHIDO PARA ESTE MÓDULO:.....	3
O PROGRESSO DOS DADOS.....	4
INSTALAÇÃO .....	5
OPÇÕES .....	6
ENTENDENDO A ESTRUTURA .....	6
DEFININDO OS DADOS (CRIANDO O ARQUIVO .QES).....	6
<i>Nome</i> .....	7
<i>Texto descritivo</i> .....	7
<i>Formato do campo</i> .....	8
<i>Exemplo de .QES</i> .....	8
CRIANDO A ENTRADA DE DADOS (CRIANDO O ARQUIVO .REC) .....	11
ADICIONANDO/REVISANDO OS CONTROLES (EDITANDO ARQUIVO .CHK) .....	11
CONSISTÊNCIA DE CONTROLES .....	23
ENTRANDO COM OS DADOS .....	23
<i>Dupla Digitação</i> .....	24
<i>Verificação de arquivos duplicados</i> .....	25
GERANDO A DOCUMENTAÇÃO DOS DADOS (“CODEBOOK” E RELATÓRIO SIMPLES) .....	25
EXPORTANDO DADOS PARA OUTROS FORMATOS .....	26
OUTROS RECURSOS .....	27
<i>Importar dados</i> .....	27
<i>Criar .QES a partir de .REC</i> .....	27
<i>Revisar arquivo de dados</i> .....	28
<i>Anexar/fundir dados</i> .....	28

## Objetivos do módulo

O curso é estritamente prático e o aluno deve manipular programa durante as oficinas. Os objetivos deste curso são: sensibilizar que erros na informatização de dados que já foram coletados em formulários de papel são comuns, e independentes da experiência do digitador, e que estes devem e podem ser evitados; sensibilizar sobre documentações avançadas sobre tratamento de dados em todas as etapas e como encontrá-las; apresentar o EpiData para que os alunos sejam capazes de construir e executar uma entrada de dados e habilitar o aluno procurar ajuda para o aplicativo solucionando dúvidas de tópicos não abordados neste texto.

### Razões pelas quais EpiData 3.1® foi escolhido para este módulo:

- Este aplicativo foi especialmente desenvolvido para investigações com funções que ajudam a processar os dados em cada etapa.
- Este aplicativo é muito amigável, muito simples, o que facilita o aprendizado principalmente para iniciantes.
- Aplicativo é de acesso livre
- Não necessita de computador poderoso para funcionar.
- É capaz de exportar dados para virtualmente todos os formatos acessíveis para a maioria dos aplicativos utilizados em análises sofisticadas.

EpiData 3.1 é um aplicativo que deve ser utilizado para a informatização de dados que já foram coletados e posteriormente serão submetidos a explorações, tabulações e análises (em outro aplicativo). O forte deste programa é a facilidade/simplicidade para a execução da **entrada de dados**. Este aplicativo consegue manipular dados simples como um único questionário, assim como situações mais complexas como questionários complexos/longos e campos condicionais, bem como bancos relacionados e possui uma variedade muito grande de funções e controles na entrada que dá muita maleabilidade a manipulação de dados nesta etapa. Não há limites definidos para quantidades de registros que podem ser armazenados no EpiData nem o tamanho de arquivo que consegue trabalhar (apesar do arquivo de controle ser limitado em 64Kb). O EpiData trabalha com arquivos de texto simples (editáveis pelo bloco de notas) e transforma esses no formulário, entrada de dados e nos controles de entrada. Uma cópia deste material está disponível em <http://blog.ipec.fiocruz.br/lapclin-chagas/>

EpiData 3.1 teve seu desenvolvimento interrompido em 2008 e uma nova versão baseada em arquivos XML iniciou com a intenção de substituir a versão 3.1. Essa iniciativa baseia-se no

princípio que o a manipulação dos dados deve ser protegida e hierarquizada de forma a atender requisitos de boas praticas clinicas de agências reguladoras. No entanto, a versão disponível no início de 2013 ainda era uma versão de teste e o próprio time de desenvolvimento recomendava não utilizar essa versão para trabalho real.

O EpiData 3.1 não é o único aplicativo capaz de executar a sua tarefa e a escolha entre eles depende das funcionalidades desejadas e empatia. Há o EpiInfo 6.04d (desenvolvimento interrompido e do qual o EpiData 3.1 é dissidente), o EpiInfo 7 (<http://wwwn.cdc.gov/epiinfo/>), CSPro (<http://www.csprousers.org/software/>) Acces®, Excel®, LibreOffice (<http://pt-br.libreoffice.org/>). Destes, o Acces e Excel não são de acesso livre. O Excel e o LibreOffice tem a desvantagem grande de serem planilhas de cálculo, com pouquíssimos recursos de entrada de dados (apesar de possível), e por isso permitem erros com muita frequência, como por exemplo o usuário trocar uma linha ou coluna durante uma digitação. A maioria dos aplicativos de análise de dados (como R, SPSS, SAS e STATA) permite edição/entrada de dados, mas em uma estrutura semelhante a uma planilha de cálculo – com todas as suas desvantagens.

É muito comum, por exemplo, em ensaios multicentricos, os dados serem digitados em um portal da internet com os formulários disponíveis “online” ou em um software desenvolvido pela equipe do próprio ensaio. Esses portais geralmente utilizam linguagem SQL ou MySQL que são linguagens de armazenamento de dados em servidores. Softwares de livre acesso que exemplificam esse grupo são RedCap (<http://catalyst.harvard.edu/services/redcap/>) e OpenClinica (<https://community.openclinica.com/>). Ainda, é cada vez mais freqüente a utilização de entrada de dados eletrônica diretamente em computadores, pulando a fase dos papeis. Em inquéritos de campo um software para computadores de mão é muito comum nessas situações, como o Pendragon®.

Outro software que está sendo utilizado com frequência é o Teleform®. Este gera formulários em papel e depois alimenta o banco de dados a partir de imagens de um scanner que digitalizou os formulários preenchidos.

Todos esses recursos têm as suas vantagens e desvantagens e ao mesmo tempo não modificam a essência do progresso dos dados e evitam (mais ou menos) os erros (humanos) durante esse processo.

## **O progresso dos dados.**

O processo usual dos dados, desde a sua forma bruta na fonte até geração de informações a partir de análises geralmente segue os seguintes passos – não necessariamente nesta ordem - (esses passos podem ser modificados de acordo com o tipo de delineamento da investigação e o grau de informatização do processo):

1. O entrevistador coleta os dados e completa o formulário.
2. O entrevistador corrige o formulário e corrige os erros, retorna a fonte caso necessário.
3. Um supervisor verifica os questionários, corrige erros e coleta novamente os dados de uma amostra das fontes.
4. Os dados são informatizados em um computador
5. Os dados são informatizados uma segunda vez.
6. Os dois arquivos gerados pela informatização são comparados, os erros de digitação são comparados e os formulários são consultados para correção dos erros.
7. Durante a informatização, ou após esta, os dados são conferidos para valores permitidos ou inconsistência de valores que são condicionados a outros campos (*e.g.* sexo = “masculino” ou “feminino”; se sexo = “masculino” então gestação = “não se aplica”)
8. Quando os dados estão limpos, haverá necessidade de editá-lo, recodificando e criando novas variáveis ou selecionando registros.
9. Os dados podem ser (e geralmente são) relacionados com dados de outras fontes como exames de laboratório.
10. Os dados então deverão ser exportados para um formato amigável para análise.

Muitas vezes, erros nestas etapas passam despercebidos e podem comprometer resultados. Há vários documentos que abordam questões teóricas sobre os erros de aferição, como estes devem ser considerados e estratégias que poderiam ser adotadas para evitá-los. Neste curso, serão abordadas somente as questões práticas relacionadas ao aplicativo sendo eventualmente comentadas questões teóricas.

## **Instalação**

O aplicativo pode ser encontrado em [www.epidata.dk](http://www.epidata.dk) em diversos idiomas, incluído o português do Brasil. Para instalá-lo em computadores com Windows, basta clicar duas vezes em cima do arquivo baixado do sitio e seguir as orientações da tela. Para os que usam computadores multi-usuário e não são administradores da máquina (com privilégios de instalação) colar o diretório do EpiData no c:\ a partir de outro computador ou pen-drive parece funcionar sem

problemas. Para usuários de Linux, não há orientações no manual para instalação, mas teoricamente é possível fazê-lo através do WINE.

## Opções

Podem ser controlados as cores dos campos e da tela, tipo e tamanho das fontes nos diferentes arquivos e outras opções, através das opções no menu “Arquivo/opções”. Modificando as opções jamais haverá danos dos dados mesmo que as opções tenham sido modificadas após os dados serem informatizados.

Via de regra as opções padrões do EpiData são bem satisfatórias. Uma única dica é fazer com que as fontes e seus tamanhos sejam os mesmos no .QES e no .REC para que não haja dificuldade de fazer os campos caberem dentro da tela do formulário.

## Entendendo a estrutura

O EpiData trabalha com uma estrutura que é composta por 3 arquivos principais. O arquivo .QES (questionário) é o arquivo texto que irá dizer ao EpiData como será o rosto da entrada de dados. O arquivo .CHK (“check”) que são os controles de entrada para este banco de dados. O arquivo .REC que é o banco de dados propriamente dito, ou seja, o arquivo que guarda os seus dados. Todos estes devem ter o mesmo nome (salvo a extensão) e estar no mesmo diretório para funcionar corretamente.

## Definindo os dados (criando o arquivo .QES)

Repare que o EpiData apresenta duas barras de ferramentas. A primeira, mais acima, é semelhante à de outros aplicativos com os itens “Arquivo”; “Controles”, “Entrada/Saída de dados”; “Documento”; “Janela” e “Ajuda”, a segunda apresenta numeração de 1 à 6 com as funções mais utilizadas no processo de informatização dos dados.

- Pressione o botão “1. Definir dados”.
- Um pequeno menu com duas opções aparecerá.
- Pressione a opção novo arquivo .QES.

Você perceberá que a barra numerada permanecerá, porém com algumas funções desativadas e a barra superior apresentará itens diferentes. Agora a barra de menus apresentará

funções apenas para a definição de dados para a criação do arquivo .QES. Aparecerá na tela um arquivo texto em branco com o nome “Sem título”.

- Pressione o menu “Arquivo\Gravar como” e salve o seu arquivo com um nome desejável no diretório onde armazenará os seus dados (e.g. “c:\banco” ou “c:\dados\pesquisaX”) [*dica: não utilize espaços, “ç”, acentuação ou símbolos na nomeação pois alguns aplicativos de análise podem não reconhecer estes símbolos*]

Para compreender como se deve confeccionar um arquivo .QES é necessário entender que (em situações ordinárias) cada linha do arquivo representará uma variável e esta deverá conter:

### **Nome**

Nome é o termo que identificará a variável dentro do banco de dados (o nome da coluna). Por padrão, o EpiData entende que a primeira palavra da linha no arquivo .QES será o termo designado como nome da variável. [*dica: mesmo que o nome da variável seja um número, coloque uma letra antes do número como “c” de campo, “v” de variável ou “p” de pergunta, pois o EpiData não compreende bem os nomes iniciados com números*]

Não é aconselhável nomear as variáveis com termos com mais de 8 ou 10 caracteres. Também não é aconselhável colocar símbolos nos nomes das variáveis, pois o EpiData pode acusar erros, principalmente se forem símbolos utilizados pelo .QES em definições de campo como “\_”, “#” ou “@”.

A minha sugestão é; em formulários pequenos – não mais do que 15 ou 20 campos, que dê pra enxergar todos ao mesmo tempo na mesma tela, e que não serão compartilhados entre vários colegas envolvidos no mesmo projeto – a utilização de abreviações nos nomes da variável (e.g. Data do início -> dataini) funciona bem para identificação de seu conteúdo; em formulários grandes ou naqueles que serão compartilhados entre vários participantes, a nomeação através de números funciona melhor (v1, v2, v3...) e nesta situação os textos descritivos (rótulos o etiquetas) devem ser bem elaborados/amadurecidos.

### **Texto descritivo**

O texto descritivo é o texto que será posteriormente identificado como o rótulo (ou etiqueta) da variável (não das respostas) no momento da exportação e pode ser a própria pergunta.

Geralmente a pergunta no questionário é muito longa para caber na etiqueta, e por isso seu resumo ou termos indicadores pode ser mais adequado.

O texto descritivo deve sempre estar antes dos campos de entrada. Caso estejam depois do campo, não serão utilizados nas etiquetas. As etiquetas não devem ser muito longas (mais que 4 palavras) pois serão exportadas e utilizadas nas gerações de tabelas e relatórios pelos aplicativos de análise que não possuem espaço ilimitado na saída de comandos.

Alguns pesquisadores gostam de colocar mais de um campo na mesma linha. Se for o caso, depois do primeiro campo, deverá haver o nome da segunda variável e .QES terá a seguinte aparência:

p79      POSOLOGIA ALTERADA #              p710 NECESSITOU TRATAMENTO ADICIONAL #

### Formato do campo

O campo é o espaço disponível para que haja entrada de dados na variável. Este poderá ser numérico (#), uma data (dd/mm/yy) ou alfa-numérica (\_\_) ou texto caixa alta <A >. Nos campos alfa-numéricos e texto com caixa alta, são semelhantes, no sentido de que permitem a entrada de letras ou números, ou combinações de letras e números e serão sempre armazenados/exportados como texto. A diferença é que, independente da CAPS LOCK nos campos de caixa alta todas as letras entrarão como maiúsculas. Reparem que este define que tipo de dado pode estar contido na variável.

Os tipos de campos a serem utilizados podem ser acessados no botão “Listagem de seleção de campos” (o segundo da direita para esquerda) na barra de botões ou pressionando Ctrl+Q. Os tipos mais utilizados e simples de campos são (explicação para outros campos pode ser encontrada no manual e no EpiTour [documento de referência]):

Tipo	Exemplo
Alfa-numérico	_____
Número ID	<IDNUM>
Numérico	##; ##.##
Texto caixa alta	<A >
Datas	<dd/mm/yyyy>

### Exemplo de .QES

Agora, entraremos com o texto que será uma cópia do formulário que foi utilizado para a coleta de dados para a pesquisa! É recomendável que se coloque um título no formulário para fácil identificação de qual arquivo está sendo utilizado. Podemos utilizar este simples exemplo que se segue: *[dica: quanto mais numéricos forem os campos, mais se utilizará o teclado numérico a direita do teclado e mais fácil fica a digitação dos dados]*

\*\*\*\*\*

Exemplo de formulário

\*\*\*\*\*

```
numero NUMERO IDENTIFICADOR: ##
aluno INICIAIS DO ALUNO: _____
sexo SEXO: _____
gesta GESTACOES: _____
datapre DATA DO PREENCHIMENTO: <dd/mm/yyyy>
datanas DATA DE NASCIMENTO: <dd/mm/yyyy>
idade IDADE AO PREENCHIMENTO: _____
estado ETADO CIVIL: _____
altura ALTURA ESTIMADA: _____
peso PESO ESTIMADO: _____
calcado NUMERO DO CALCADO DO ALUNO: _____
abo GRUPO SANGUINEO ABO: _____
rh GRUPO SANGUINEO RH: _____
curso CURSO DE POS-GRADUACAO: _____
graduacao CURSO DE GRADUACAO: _____
```

- Após a confecção do arquivo .QES como copia idêntica do formulário utilizado salve o arquivo com Ctrl+S ou menu “Arquivo/Gravar”.
- Agora, visualize como o formulário de entrada ficará através do botão “Visualizar formulário de dados” (o terceiro da direita para esquerda) ou Ctrl+T.

Repare agora que o arquivo .QES e a pré-visualização estão abertas e demonstradas através de abas no canto inferior esquerdo. Utilize F10 para fechar a aba selecionada ou o menu “Arquivo/fechar formulário”.

Alguns pesquisadores gostam de colocar no arquivo texto.QES, as opções de entrada de cada campo abaixo ou a direita da entrada (e.g. 1.Masculino e 2.Feminino). Essa racionalidade é interessante quando se utilizará o EpiData para imprimir os formulários que serão utilizados como CRF (*cilnical research forms*). Fica um pouco tosco por conta dos poucos recursos como editor de texto, mas é bem eficiente. Os limites e opções estarão presentes na entrada através do arquivo .CHK - será explicado mais adiante – e por isso para o digitador não há necessidade de colocar as opções no texto do formulário.

- Edite novamente o arquivo caso necessário e salve novamente.
- Antes de criar a entrada do banco de dados propriamente dita, o arquivo .REC, você pode alinhar os campos.
- Caso possua uma variável por linha no arquivo .QES, deixe o cursor na linha a qual deseja alinhar as restantes e pressione o menu “Editar/Alinhar campos”
- Por outro lado, se possuir mais de uma variável por linha, alinhe manualmente com espaços pois o comando alinhar campos não funciona bem nesta situação
- Visualize novamente a entrada com o comando Ctrl+T e caso satisfeito, salve novamente e feche ambas as abas.
- Por algum motivo desconhecido, o alinhamento pode colocar uma ou duas das linhas fora do lugar, essas serão necessariamente corrigidas com espaços inseridos a mão.

Agora o nosso arquivo de exemplo deverá se parecer com algo como este:

```
numero          NUMERO IDENTIFICADOR: ##
aluno           INICIAIS DO ALUNO: _____
sexo           SEXO: _____
gesta          GETACOES: _____
datapre        DATA DO PREENCHIMENTO: <dd/mm/yyyy>
datanas        DATA DE NASCIMENTO: <dd/mm/yyyy>
idade          IDADE AO PREENCHIMENTO: _____
estado         ETADO CIVIL: _____
altura         ALTURA ESTIMADA: _____
peso          PESO ESTIMADO: _____
calcado        NUMERO DO CALCADO DO ALUNO: _____
abo           GRUPO SANGUINEO ABO: _____
rh            GRUPO SANGUINEO RH: _____
curso          CURSO DE POS-GRADUACAO: _____
graduacao      CURSO DE GRADUACAO: _____
```

Se nós seguirmos a dica de colocar todas as variáveis possíveis como numéricas este pareceria com algo assim:

```
numero          NUMERO IDENTIFICADOR: ##
aluno           INICIAIS DO ALUNO: _____
sexo           SEXO: #
gesta          GETACOES: #
datapre        DATA DO PREENCHIMENTO: <dd/mm/yyyy>
datanas        DATA DE NASCIMENTO: <dd/mm/yyyy>
idade          IDADE AO PREENCHIMENTO: ###
estado         ETADO CIVIL: #
altura         ALTURA ESTIMADA: #.#
peso          PESO ESTIMADO: ###.#
calcado        NUMERO DO CALCADO DO ALUNO: ##
abo           GRUPO SANGUINEO ABO: #
rh            GRUPO SANGUINEO RH: #
curso          CURSO DE POS-GRADUACAO: #
graduacao      CURSO DE GRADUACAO: #
```

Posteriormente, para cada campo colocaremos códigos e valores permitidos a cada um deles.

**Nunca modifique os nomes das variáveis no .QES depois que já há registros no banco.**

### **Criando a entrada de dados (criando o arquivo .REC)**

Repare que até agora o arquivo criado foi apenas o texto que dará forma ao rosto da entrada (arquivo .QES) e ao banco de dados propriamente dito. O rosto da entrada de dados poderia ser visualizado pela função Ctrl+T durante a edição do .QES, mas somente visualizar, não podendo entrar os dados.

Agora, criando o arquivo .REC, a entrada de dados, o arquivo que armazenará os dados propriamente dito, será possível digitar e informatizar os dados. À semelhança do .QES, o arquivo .REC também é um arquivo texto editável pelo bloco de notas, porém com um dicionário que codificará o banco de dados.

- Pressione o botão 2 da barra numerada.
- Procure na janela que se abriu o diretório em que foi salvo o arquivo .QES desejado
- Selecione o arquivo .QES desejado a partir do qual será gerado o arquivo .REC
- Pressione OK

Repare que o EpiData criou um arquivo com o mesmo nome, no mesmo diretório porém com a extensão .REC. Pode-se modificar o nome e o diretório destino, mas não é aconselhável, salvo se for em casos de usuários mais experientes.

No EpiInfo 6.04d havia uma série de comandos possíveis para a entrada. Estes eram executados antes ou no acesso do arquivo .REC para que a entrada se comportasse de uma determinada maneira. No EpiData®, a maioria desses comandos foram incorporados nos controles, no arquivo .CHK, tornando a sua geração e acesso muito mais simples.

### **Adicionando/revisando os controles (editando arquivo .CHK)**

Este momento é o momento mais trabalhoso da confecção da entrada de dados e que exige maior atenção. Pois aqui é onde os controles de valores legais, etiquetas dos códigos dos valores, as perguntas condicionais, as fórmulas, os pulos, a auto-procura por registros já informatizados, backup automático e outros recursos mais complexos de controle de entrada são editados formando as regras de entradas. Funciona como um arquivo que discrimina comandos que devem ser

executados toda vez que dados serão inseridos em um determinado campo. Embora as regras ou controles de entrada sejam opcionais no EpiData®, estas são altamente recomendáveis, para minimizar os erros durante a entrada.

- Pressione o botão “3. Controles”
- Selecione o diretório e o arquivo .REC ao qual se aplicarão os controles.
- O EpiData criou um arquivo idêntico ao .QES e ao .REC, desta vez com a extensão .CHK.

Perceba que o formulário é idêntico ao .REC, no entanto não é possível a entrada de dados. Junto com tela do formulário aparece uma pequena janela com recursos rápidos para edição dos campos, a lista das variáveis e a opção de edição avançada em texto de cada variável. Os cinco aspectos básicos das regras pela janela de controle são:

1. As etiquetas dos valores – texto descritivo que explica e limita valores numéricos.
2. Restrição de entrada para valores previstos – “legal” e “range”.
3. Especificar seqüência de entrada – pulos e regras condicionais.
4. “Mustenter” – campo em que há necessidade de preenchimento.
5. “Repeat” – repete o mesmo valor do registro anterior.

Um exemplo genérico:

- Coloque o cursor no campo variável ou selecione a variável na janela de controles.
- No campo “Range, Legal” coloque os valores permitidos para esta variável (*e.g.* 1,2 [1 e 2] ou 1-9 [1 a 9] ou 1-4,9 [1 a 4 e 9])
- Selecione as etiquetas dos valores abrindo as opções do campo “Value label” ou pressione o sinal de “+” ao lado caso a etiqueta desejada ainda não tenha sido confeccionada.
- Pressionando o “+” uma segunda janela de texto abrirá e nesta você poderá editar as etiquetas e o nome da etiqueta (edição avançada explicada a seguir).
- O campo “Jump” condiciona um pulo para outro campo dependendo do valor inserido – caso o valor neste campo *sexo* seja 1 (masculino) pule a pergunta *gestação* – 1>p4
- Pode-se condicionar mais de um pulo – 1>p4, 9>p5

- As opções “Mustenter” e “Repeat” são sim ou não na caixa de diálogo.
- Para cada variável os mesmos passos devem ser seguidos.

Um exemplo genérico de um .CHK visualizado pelo editor de CHK ou bloco de notas.

```
DATATEND => NOME DA VARIÁVEL  
  RANGE 01/08/2004 INFINITY => PRIMEIRO COMANDO (2 ESPAÇOS)  
  REPEAT => SEGUNDO COMANDO  
END => FIM DE COMANDOS PARA VARIÁVEL
```

```
p17  
  COMMENT LEGAL USE rotulo_simnao  
END
```

O botão editar na caixa de controle permite a edição de controles mais complexos através de texto simples dentro do arquivo .CHK. A estrutura do arquivo deve seguir os seguintes aspectos (como no exemplo genérico acima): o nome da variável na primeira linha, a seqüência de comandos nas linhas abaixo com dois espaços da margem cada comando, e o termo “END” que marca o fim de comandos para esta variável na mesma margem.

Tomemos como exemplo o .REC gerado anteriormente.

A variável *numero* é uma variável numérica (que permite digitação apenas de números). Esta tem a intenção de identificar o registro de forma única. Ou seja, cada paciente/sujeito/formulário tem um número. Pode entrar qualquer número inteiro com 2 dígitos. O importante é que sujeitos não tenham números repetidos ou que o mesmo sujeito não seja digitado duas vezes. Para isso há dois recursos: (1) comando KEY e (2) o comando “AUTOSEARCH”.

O comando “KEY” deve ser utilizado em variáveis que identificam o registro e pode ser inserido em diversas variáveis simultaneamente. Este comando permite a indexação do campo para uma procura mais fácil caso deseje procurar um registro por esta variável posteriormente. Mas se este for para identificação de forma única, como o número de prontuário, número de registro ou algum outro campo numérico que possa ser identificador único deve ser completado como “KEY UNIQUE”. Este faz com que o EpiData não consiga entrar com dois valores iguais na mesma variável.

O comando “AUTOSEARCH” permite uma função semelhante ao “KEY”, porém mais amigável. O “AUTOSEARCH” executa uma procura por valores que foram inseridos na variável em registros anteriores durante a entrada. Caso encontre, ele pergunta se o digitador deseja editar o

registro anterior ou inserir um novo. Em alguns bancos de dados o número prontuário, por exemplo, identifica de forma única o sujeito, mas não o registro, pois o mesmo sujeito poderia entrar mais de uma vez no banco (não é o caso do nosso exemplo). Assim a extensão “AUTOSEARCH LIST” ajuda, pois exibe uma lista de todas as linhas que tem o mesmo prontuário. Este também pode ser executado com a coincidência de mais de um campo como, por exemplo, *numero* e *datacol* (e.g. “AUTOSEARCH LIST numero datacol”). Usar os comandos “AUTOSEARCH” e “KEY UNIQUE” ao mesmo tempo porque não faz muito sentido (principalmente se o mesmo sujeito entrar mais de uma vez no banco), mas usar “AUTOSEARCH” e “KEY” talvez facilite a procura dos registros. Para inserir estes comandos no nosso exemplo:

- Coloque o cursor no campo *numero*
- Pressione editar na pequena janela de edição do CHK
- Uma nova pequena janela se abrirá com o nome da variável e o termo END
- A dois espaços da margem coloque os comandos desejados. Ao final a janela deverá ter um texto semelhante ao que se segue...

```
NUMERO
  AUTOSEARCH NUMERO
  KEY
END
```

- Pressione a tecla “Aceitar e fechar” e a janela anterior ficará em evidência

O comando “MUSTENTER” tem uma funcionalidade semelhante aos comandos “KEY” e “AUTOSEARCH”. Funciona bem em campos chaves em que os dados discriminam as observações ou então para campos dos quais muitos outros são dependentes ou relacionados. Este comando obriga o digitador a entrar com dados no campo não permitindo que este fique em branco, somente seguindo para os campos subsequentes caso este seja preenchido. Assim, o nosso .CHK poderia também parece como:

```
NUMERO
  MUSTENTER
  AUTOSEARCH NUMERO
  KEY
END
```

Na variável *sexo* seria adequado restringir os valores possíveis para masculino ou feminino. Os comandos “COMMENT LEGAL” são utilizados com a especificação de uma etiqueta logo em seguida irá restringir os valores legais para a variável que está sendo editada.

No EpiData, deve-se utilizar o recurso de rótulos ou etiquetas genéricas. Estes permitem que os valores com as respectivas etiquetas sejam exportadas ao final da digitação e permitem que o mesmo rótulo possa ser utilizados em várias variáveis (*e.g.* rótulo 0-não 1-sim poderiam ser aplicados em diversas variáveis binárias).

- Na caixa de diálogo dos comandos CHK, ao lado das opções de “Value label” deve-se pressionar o botão “+”
- Na nova janela de texto que se abrirá criaremos os rótulos ou etiquetas.
- Os rótulos tem uma estrutura semelhante à estrutura de comandos de variáveis dentro do .CHK.

Nesta janela o texto que aparece deve ser algo parecido com:

```
Rotulo_sexo  
END
```

Vamos editar para ficar igual a:

```
Rotulo_sexo  
 1 MASCULINO  
 2 FEMININO  
END
```

- Para aceitar as modificações pressione “Aceitar e fechar”, lembre-se sempre disso.

Em seguida, devemos ter certeza de que na caixa de diálogo dos comandos CHK, no campo das opções de “Value label” a opção sexo deve estar aparecendo, indicando que o rótulo sexo está sendo aplicado para esta variável. Se pressionarmos o botão gravar e em seguida editar, veremos na janela de edição um texto semelhante a:

```
SEXO  
  COMMENT LEGAL USE Rotulo_sexo SHOW  
END
```

Assim fica fácil perceber que podemos utilizar o mesmo rótulo para diversas variáveis como, por exemplo, quando muitas têm a mesma resposta 1“sim” e 0“não”. Bastaria criar uma variável com o rótulo sim e não e na janela “value label” na caixa de diálogo selecionar este rótulo sempre que necessário. Isso significa que o comando USE está sendo aplicado em todos os Roteiro de aula – Módulo: Epidata® - dúvidas (em inglês): [epidata-list@lists.umanitoba.ca](mailto:epidata-list@lists.umanitoba.ca)

COMMENT LEGAL ao invés de especificarmos qual o comentário dos valores a serem utilizados. Ainda, o comando SHOW faz com que os rótulos apareçam durante a entrada independente da solicitação do digitador, bastando que o cursor chegue no campo com uma etiqueta.

Passemos para a variável *gesta* que se refere ao número de gestações. Repare que mesmo sendo numérica como a variável *sexo*, ela não apresenta categorias, pois é uma variável contínua. Como possui apenas um dígito, as possíveis respostas podem variar de 0 a 9. Mas quando o sexo é masculino, necessariamente o número de gestações deve ser 0. Então para evitar erros de digitação, podemos condicionar essa resposta à variável *sexo*.

Para isso, os comandos “IF” e “THEN” são bastante úteis, pois podem fazer qualquer tipo de condicionamento e podem ser associados a outros comandos e operadores. Pode-se, por exemplo, condicionar pulos como no “JUMP”, pode-se condicionar resultados de outras variáveis, condicionar formulas, condicionar valores aceitáveis a outras variáveis, e mensagens de ajuda.

- Com o cursor na variável *gesta*, pressione o botão editar:

No nosso exemplo:

```
GESTA
  AFTER ENTRY
    IF SEXO = 1 THEN
      LET GESTA = 0
      GOTO DATAPRE
    ENDIF
  END
END
```

- Pressione a tecla “Aceitar e fechar” e a janela anterior ficará em evidência

Repare que neste exemplo outros três comandos foram exibidos (1) “AFTER ENTRY” (2) LET e (3) GOTO.

Os comandos “BEFORE” e “AFTER” condicionam o momento do comando para antes ou depois do arquivo, do registro ou da entrada: “BEFORE FILE”, “BEFORE RECORD” ou “BEFORE ENTRY” respectivamente. No nosso exemplo, significa que independente do valor inserido, depois da entrada de algum valor em *gesta* algum comando será executado. Em seguida o “IF... THEN” que condicionou ao sexo ser masculino para que o comando seja executado. “GOTO” faz com que depois de executados os comandos de interesse o cursor siga para um determinado

campo; “LET” faz com uma ordem seja executada: como “faça” a idade ser este valor ou esta fórmula.

Então, o que fizemos foi: após a entrada de qualquer valor, se sexo for masculino, o valor de gesta deve ser modificado para 0 e em seguida o cursor deve ir para a variável *datapre* (aqui *datapre* é a variável seguinte, mas poderia ser qualquer uma, podendo funcionar como um pulo).

Repare que no exemplo acima haverá um pulo de um campo. Assumindo que o formulário preenchido possui essa variável o digitador pode inadvertidamente querer preencher o valor de *gesta* em *datapre* porque não atentou para o pulo. Como o digitador tem o hábito e dever de copiar os dados no formulário para banco de dados, logo os controles deveriam seguir esse lógica. Reformulando o exemplo acima, teríamos:

```
GESTA
  AFTER ENTRY
    IF (SEXO = 1) AND (GESTA <> 0) THEN
      HELP "Valor inseperado para numero de gestações em homes!"
      CLEAR GESTA
      GOTO GESTA
    ENDIF
  END
END
```

Aqui acrescentamos mais dois comandos não citados anteriormente. O comando “CLEAR” tem a função de apagar o que foi digitado no campo. O comando “HELP” faz com que a apareça na tela uma mensagem que deve ser discriminada entre aspas após o comando. Então, fizemos com que se após a entrada do valor de gesta, se o sexo for masculino e o número de gestações for diferente de 0, então haverá um aviso e o cursor retorna para gesta com o campo vazio.

Nas duas próximas variáveis, as datas serão digitadas. Na primeira, a data de coleta e a segunda a data de nascimento. Iremos, como no exemplo acima condicionar essas duas para que eventuais erros de digitação não criem inconsistências absurdas. A primeira data é a data de digitação e por isso não há como controlar muito. Assumindo-se que o curso está sendo ministrado em 2008, talvez seja razoável condicionar este campo a uma data posterior a janeiro de 2008. As datas se comportam como números com etiquetas. Isso quer dizer que uma data se refere ao 1 (e.g. 1 de janeiro de 1960). As outras datas são números que são contados adicionando-se ou subtraindo-se dias a esta data referência. Assim, 1 de janeiro de 1961 corresponderia ao número 365. Esta lógica adotada por todos os aplicativos permite contas com datas.

- Com o cursor no campo *datapre* pressione o botão editar
- Na nova janela insira um comando condicionando esta data a ser maior que janeiro de 2009.

Este deve ficar alguma coisa como se segue:

```
DATAPRE
  IF DATAPRE < "01/01/2008" THEN
    CLEAR DATAPRE
    GOTO DATAPRE
  ENDIF
END
```

- *Dica: as datas devem ser sempre expressas entre aspas dentro do .CHK.*

Podemos sofisticar um pouco este comando emitindo um aviso para o digitador entender porque a data digitada não está sendo aceita. Poderia parecer como se segue:

```
DATAPRE
  IF DATAPRE < "01/01/2008" THEN
    HELP "A VARIÁVEL DATA DE DIGITAÇÃO DEVE SER POSTERIOR A JANEIRO/2008"
    CLEAR DATAPRE
    GOTO DATAPRE
  ENDIF
END
```

- Pressione a tecla “Aceitar e fechar” e a janela anterior ficará em evidência.

No exemplo acima, o comando dado foi: se *datapre* for anterior à 1 de janeiro de 2008, então avise que a data deve estar errada, limpe a data digitada e permaneça com o cursor no campo *datapre*.

A data de nascimento também pode ser condicionada a partir de princípios assumidos. Por exemplo, se os voluntários a preencherem este formulário são alunos de pós-graduação, então todos devem ter curso superior e por isso provavelmente têm mais de 20 anos de idade. Assim vamos condicionar com esta regra.

- Coloque o cursor no campo *datanas* e pressione o botão editar
- Escreva uma regra para que essa data seja pelo menos 20 anos antes da data da digitação

Esse comando deve parecer com alguma coisa como se segue:

```
DATANAS
  AFTER ENTRY
  IF ROUND (INT ((DATAPRE-DATANAS)/365.25)) < 20 THEN
```

```
HELP "O ALUNO DEVE TER MAIS DE 20 ANOS. DIGITE AS DATAS NOVAMENTE"  
CLEAR DATAPRE  
CLEAR DATANAS  
GOTO DATAPRE  
ENDIF  
END  
END
```

- Pressione a tecla “Aceitar e fechar” e a janela anterior ficará em evidência

Aqui nós condicionamos com uma formula que há o cálculo de idade no momento do preenchimento do formulário. A idade é expressa em anos e as contas são feitas em dias, por isso os aplicativos têm que entender que depois da conta deve haver uma transformação para número (“integer”) e deve ser arredondado (“round”) depois de dividir pelo numero de dias em um ano (365.25).

Agora vamos fazer com que a idade seja preenchida automaticamente através da mesma fórmula utilizada anteriormente.

- Coloque o cursor no campo *idade* e edite o campo com uma formula para calcular a idade

```
IDADE  
BEFORE ENTRY  
LET IDADE = ROUND (INT ((DATAPRE-DATANAS) / 365.25) )  
END  
END
```

- Pressione a tecla “Aceitar e fechar” e a janela anterior ficará em evidência

Nos campos subseqüentes no formulário de exemplo não há muitos recursos complexos a serem adicionados. Construa e adicione os rótulos de cada campo como exercício.

Seguem algumas descrições e exemplos de comandos do .CHK que são muito úteis no controle de entrada de dados.

O comando “CONFIRM” faz com que haja necessidade de se pressionar a tecla “Enter” depois de entrar com os valores no campo para passar para o campo seguinte. O padrão utilizado pelo EpiData é passar para o campo seguinte quando todos os possíveis espaços forem preenchidos, assim agilizando a digitação. Quando o campo possui apenas um caractere, imediatamente o cursor segue ao campo seguinte, mas se o campo for um nome de 60 caracteres e o dado possui apenas 20, o EpiData vai ficar aguardando a tecla “Enter” ou completar 60 caracteres. O comando “CONFIRM” permite que a passagem para o campo seguinte seja igual em todos os campos e assim

tornando a digitação mais uniforme. Esse comando pode ser utilizado em cada variável para ser aplicado somente naquela variável, ou pode ser utilizado com o comando “BEFORE FILE” para se aplicar a todas as variáveis. Assim poderia ser como se segue:

```
BEFORE FILE
  CONFIRM
END
```

ou

```
SEXO
  CONFIRM
  COMMENT LEGAL Rotulo_sexo
  END
END
```

O comando “NOENTER” serve para colocar uma variável na entrada que não está no formulário original e por isso não deverá ser digitada. Este comando impede que qualquer dado seja digitado neste campo. Exemplificando com o campo idade que não existiria no formulário, mas que será calculado durante a entrada.

```
DATANAS
  AFTER ENTRY
  LET IDADE = INT(ROUND((DATAPRE-DATANAS)/365.25))
  END
END

IDADE
  NOENTER
END
```

Uma outra forma de fazer a mesma coisa é definir variáveis temporárias e condicionais durante a digitação. Algumas delas poderiam ser desprezadas após a mudança de registro. No entanto, as variáveis definidas durante a entrada são temporárias e não podem ser exportadas

O comando “TYPE COMMENT” faz com que os comentários discriminados para os valores legais permitidos em comandos “COMMENT LEGAL” ou que sejam atribuídos rótulos apareçam a direita do campo após os valores serem digitados. Útil para a verificação se o código digitado é mesmo o que deveria entrar durante a entrada. Pode ser utilizado somente em alguns campos ou em todos os campos combinado com o comando “BEFORE FILE”.

```
ANIMAL
  COMMENT LEGAL Rotulo_especie
  TYPE COMMENT
  END
```

```
BEFORE FILE
  TYPE COMMENT ALLFIELDS
  CONFIRM
END
```

A seguir seguem mais alguns exemplos genéricos de comandos .CHK

```
AFTER FILE
  HELP "NAO ESQUECA DE FAZER COPIA DE SEGURANCA. MANDE UMA COPIA DO
ARQUIVO REC PARA O SUPERVISOR POR EMAIL"
END
```

```
p17
  COMMENT LEGAL USE rotulo_simnao
  AFTER ENTRY
    IF P17 = 0 THEN
      GOTO p66e
    ENDIF
    IF DESENHO = 1 THEN
      GOTO co18
    ENDIF
  END
END
```

```
IFITIT2
  AFTER ENTRY
    IF IFITIT2 < 40 THEN
      LET IFIRES2="NEGATIVO"
      GOTO DIAG2
    ENDIF
    IF IFITIT2 = 40 THEN
      LET IFIRES2="INDETERMINADO"
      GOTO DIAG2
    ENDIF
    IF IFITIT2 > 40 THEN
      LET IFIRES2="POSITIVO"
      GOTO DIAG2
    ENDIF
  END
END
```

Repare que as formulas e condições exigem operadores. Não há necessidade de espaços entre o texto e os operadores. Os operadores comuns são: = (igual); > (maior); <(menor); <>(diferente), \*(multiplicar); / (dividir); + (somar) e – (subtrair).

- Pressione o botão Gravar com frequência e ao final para que o trabalho não seja perdido.
- Pressione o botão sair ou F10

EpiData também consegue executar programas do windows ou mostrar arquivos externos ao banco de dados. Suponhamos que estejamos armazenando dados de pacientes de que tem lesão

de pele. Pra cada formulário há uma foto. As fotos todas seriam armazenadas em um diretório e numeradas de acordo com o número do paciente.

```
FOTONUM
  DEFINE tmpName _____
  AFTER ENTRY
    LET tmpName="d:\pics\picture"+fotonum+".jpg"
    EXECUTE @tmpName WAIT
  END
END
```

O campo FOTONUM é um campo numérico. Se o usuário entra o número 3, então uma figura em .jpg é chamada em "d:\pics\picture3.jpg" e será executada pelo programa padrão para arquivos de .jpg. O commando WAIT indica ao EpiData a esperar até que o visualizador de fotos é fechado. @ indica que tmpName deve ser substituído pelo conteúdo da variável tmpName.

As fórmulas também podem ser aplicadas a textos quando conveniente, por exemplo:

ID	v1	v2	v3
---	---	---	---
1	BW	101	BW101
2	CW	102	CW102

etc..

```
V3
  BEFORE ENTRY
    IGNOREMISSING
    V3=V1+V2
  END
END
```

Aqui o IGNOREMISSING vai evitar um erro na hora de calcular a formula se umas das variáveis estiver em branco. Assim. Se uma de v1 ou v2 estiver em branco v3 será igual a que não está em branco.

Alguns iniciantes tem dificuldade de entender que o EpiData trabalha com três arquivos e que somente o .REC guarda os dados e não sabem o que mandar por correio eletrônico ou o que guardar como copia de segurança. Por isso, talvez seja interessante pedir pro EpiData pedir para fazer uma copia de segurança dos dados automaticamente. Para isso utiliza-se o comando BACKUP aninhado no comando AFTER FILE como se segue:

```
AFTER FILE
  BACKUP g:\ ZIP recrutamentos[pacientes_rastreados] [DATE]
END
```

Neste exemplo o EpiData fará: uma copia dos arquivos .REC, .QES, .CHK em um arquivo zip na partição G:\ com o nome “recrutamentos[pacientes\_rastreados][data da atualização].zip”. Repare que o destino não precisa ser o mesmo que os dados estão sendo digitados. Também é possível colocar uma senha para proteger o arquivo.

Uma lista completa de comandos do .CHK e operadores para fórmulas, alguns muito úteis e outros nem tanto, estão disponíveis no manual do EpiData nas páginas 33 à 56. Mais exemplos do .CHK podem ser encontrados no manual do EpiData e em outros documentos didáticos e oficiais disponíveis no sitio do programa.

Após salvar as suas modificações, antes de dar a confecção por encerrada, abra o seu .CHK pelo menu “Arquivo/Abriu” ou Ctrl+O e selecione arquivos do tipo chk. Agora você estará observando como o .CHK realmente é. Conhecer esta estrutura é interessante, pois comandos como BEFORE FILE ou AFTER RECORD são mais fáceis de editar nessa interface do que pela janela de controles.

### **Consistência de controles**

Quando se trabalha com um questionário muito grande, e muitos controles e condições estão presentes, comumente comentemos erros fazendo os controles executarem comandos que são divergentes. Antes de passar para entrada confira inconsistências dos controles.

- Pressione o menu “Documento/Consistência de controles”.
- Caso haja inconsistências, esta opção irá indicar se há um ou mais controles conflitantes.

### **Entrando com os dados**

Entrar com os dados parece a mais fácil das tarefas até aqui. Não se iluda!!! É a etapa onde os erros realmente acontecem com frequência. Manter a concentração nesta etapa é o mais difícil.

- Pressione o botão “4. Entrar dados”

A entrada de dados deve parecer familiar. Repare que agora os botões numerados não aparecem mais e a barra de menus é diferente. Este é o padrão do EpiData mas poderá ser modificado no menu “Arquivo/Opções”. A barra de menus apresenta agora somente funções

aplicáveis à entrada de dados. Durante a digitação o EpiData é mais amigável à utilização do teclado, Por isso, evite utilizar o “mouse”.

O menu “Ir para” tem algumas funções que são úteis na entrada de bancos grandes como: localizar campos (F4), localizar registro (Ctrl+G), localizar um valor em um campo dentre todos os registros (Ctrl+F), ir para um registro ou ir para um campo. Pode-se usar as teclas Ctrl+home para ir para o primeiro campo, Ctrl+end para ir para o ultimo campo, F7 (Ctrl+PgDown) para o registro anterior, F8 (Ctrl+PgUP) para o próximo registro, Ctrl+N para novo registro e F9 mostra as opções de entrada para o campo onde se localiza o cursor.

Repare que embaixo há uma barra de status que indica qual o numero do registro atual, quantos registros há no banco, setas de navegação entre os registros, e um “X” que apaga os registros (Shift+del). Perceba que apagar os registros não significa realmente apagar todo o registro do banco, mas somente marcar o registro para que este não seja contado e tabulado. Este permanece visível à navegação da entrada, porém será apagado quando houver exportação dos dados.

## **Dupla Digitação**

Depois de entrar com as centenas de formulários da sua investigação, agora você gostaria de minimizar os erros de digitação digitando todos ou uma parte uma segunda vez. Isso pode ser realizado de duas maneiras.

O EpiData permite uma re-entrada de dados para dupla digitação.

- Abra o menu “Ferramentas/Verificação de dupla digitação”.
- Escolha o diretório e o arquivo que já foi digitado e será conferido.
- Uma nova janela com algumas opções se abrirá
- Escolha o diretório em que o arquivo duplo será criado.
- Considere as opções “ignorar campos alfabéticos” (nesta opção o EpiData verificará somente os campos numéricos) e “combinar os registros por um campo identificador único” (este permite que digitação não necessariamente seja na mesma ordem e dependente do registro único)
- Pressione OK. EpiData criará um arquivo com mesmo nome no mesmo diretório adicionando “\_dbl.rec”.
- Pressione o botão “4. Entrar dados”
- Selecione o arquivo \*\_dbl.rec criado e um aviso de dupla digitação aparecerá.

- Digite normalmente... e toda vez que um campo for preenchido com um valor diferente do arquivo original o EpiData irá acusar, e solicitará uma das seguintes opções: manter o valor do arquivo original, manter o valor da nova entrada ou digitar novo valor.

### **Verificação de arquivos duplicados**

O EpiData pode verificar arquivos .REC (gerados do mesmo .QES) informatizados através de outro recurso, mas agora os dois arquivos já deveriam estar digitados (dupla digitação), como por exemplo o recrutamento.rec e o recrutamento2.rec.

- Vá ao menu “Documentos\Validar arquivos duplicados”
- Escolha o diretório onde os arquivos de interesse se encontram e selecione os arquivos
- Pressione OK
- Uma nova janela se abrirá e nesta deve-se selecionar o campo identificador único de cada registro e as opções do que deve ser ignorado
- Selecione as opções de interesse e diga qual o campo (variável) que identifica de forma única cada registro.
- Pressione OK

O EpiData gerará um relatório constando os registros e cada valor discordante de cada registro que poderá ser salvo em arquivo texto. Posteriormente, os formulários em papel originais podem ser consultados para as correções pertinentes.

### **Gerando a documentação dos dados (“codebook” e relatório simples)**

Esta opção serve para uma descrição de todas as possíveis informações dos dados contidos no banco. É possível visualizar os dados como em uma planilha, é possível gerar um mapa dos dados onde o banco é descrito (“Estrutura do arquivo”) com o rótulo/etiqueta do banco, número de variáveis, descrição das variáveis como nome, rótulo, tipo, controles e rótulo dos controles.

Há ainda, a opção de gerar relatório descrevendo o número de registros com as informações (e quantos em branco) com frequências sumárias de cada campo, amplitude, média e desvio padrão.

Esses recursos são úteis para descrição dos dados quando das análises sumárias e para trabalhar novamente os dados após longo período entre a confecção da entrada e da análise.

## **Exportando dados para outros formatos**

Usualmente trabalha-se com mais de um aplicativo devido a disponibilidade ou intimidade do pesquisador com recursos que são menos ou mais complexos dependendo do aplicativo (*e.g.* trabalho com SPSS e desejo fazer uma particular análise que o SAS possui mais recurso, deve-se transformar os dados do formato SPSS para SAS). Há programas que apenas transformam os dados de muitos formatos para outros como “StatTransfer” e facilitam esta operação. Alguns aplicativos, como o R, e versões recentes do SPSS apresentam recursos próprios de importação e exportação para muitos formatos. Quando não se possui essas ferramentas, a transformação dos dados para o formato desejado deve passar por algum formato em comum entre os dois aplicativos. Formatos comuns são arquivos textos tabulados ou separados por vírgulas, ou em formatos antigos como o DBase (.dbf), no entanto os formatos em comum podem variar (*e.g.* SPSS => Texto tabulado => SAS). A desvantagem de se utilizar esses formatos mais simples é que se perdem os dicionários dos bancos que contem os códigos e as etiquetas elaboradas pela entrada de dados.

Por isso o EpiData possui um recurso de exportação de dados para alguns formatos de aplicativos muito utilizados. Neste recurso aproveita-se tudo o que já foi realizado nos controles, etiquetas, anotações e estrutura.

- Pressione o botão “6. Exportar dados”
- Escolha o aplicativo a ser utilizado a partir de então
- Dependendo da escolha algumas outras opções surgirão, como as variáveis que deseja exportar, se os registros apagados devem ser realmente apagados, qual a versão do aplicativo a ser utilizado dentre outras.
- EpiData salvará arquivo com a extensão do aplicativo selecionado anteriormente no diretório selecionado.
- No caso de exportação para o STATA, o EpiData gera um banco já formado. No caso de SPSS e SAS, o EpiData gera um banco de dados em arquivo texto e um script de SPSS (ou SAS) com todos os comandos de importar e organizar os dados com as etiquetas e etc.

## Outros recursos

### Importar dados

O EpiData também é capaz de importar dados. No entanto, esse recurso é bem mais limitado do que o de exportação. O EpiData é capaz de importar bancos já com registros ou sem registros dos formatos Stata®, DBase e texto separado por vírgulas ou tabulações. Neste caso, a minha sugestão é: sempre transforme os seus dados para formato STATA, importe para dentro do EpiData®. Você perceberá que não só um .REC é criado mas um .CHK pronto pra usar também. No caso de outros formatos, o melhor seria utilizar o aplicativo original para exportar para o formato EpiData®/EpiInfo 6.04 (DOS) e então rever o .REC a partir de um .QES. Mas com esta estratégia, um .CHK deverá ser gerado manualmente depois. As outras opções podem ser utilizadas, mas não só os controles podem não serão incorporados automaticamente, como um QES. Deverá ser definido (ou criado) antes da importação – dá mais trabalho.

- Pressione o menu “Entrada\saída de dados\Importar”
- Escolha o formato desejado
- Dependendo do formato desejado, algumas opções devem ser estabelecidas como formato das datas, arquivo .QES a ser utilizado ou ainda os separadores.
- Escolha o nome e diretório do arquivo a ser salvo.

### Criar .QES a partir de .REC

Este recurso pode ser útil em ocasiões em que houve extravio dos arquivos .QES ou por que o arquivo foi criado por outra pessoa e agora deseja-se acrescentar mais registros. Também é útil quando houve exportação de outro aplicativo para o EpiData e agora se deseja “arrumar” a entrada para que a entrada de dados fique mais fácil e eficiente. Para que no momento da entrada o .REC fique com uma interface amigável a edição deve ser realizada no .QES e então atualizar o . REC a partir deste novo .QES. Nesta atualização, não há perda de dados desde que não haja modificação dos nomes das variáveis. **Nunca modifique os nomes das variáveis no .QES depois que já há registros no banco.** Se você quiser modificar os nomes das variáveis sem perder os dados nela contidos, utilize recurso do EpiData próprio para este fim.

- Pressione o menu “Ferramentas/Arquivo .QES de arquivo .REC”

- Selecione o diretório e o arquivo .REC a ser utilizado
- Selecione o diretório e nome do arquivo .QES a ser criado
- Pressione OK

### **Revisar arquivo de dados**

O arquivo de dados é o arquivo .REC. No exemplo anterior que se importa dados e se deseja atualizar algum campo ou parte do questionário, duas coisas devem estar em mente. Gerar um novo arquivo .REC com o mesmo nome significa apagar todos os dados e mudar o nome da variável significa apagar aquela variável do banco. Por isso toda atualização no formato do banco deve ser realizada no .QES e em seguida atualizada da seguinte forma:

- Abra o arquivo .QES correspondente ao banco que deseja atualizar
- Acrescente a variável que deseja ou edite o .QES **sem modificar nos nomes da variáveis**
- Salve o .QES
- Agora pressione o menu “Ferramentas\Revisar arquivo de dados”
- Ou pressione o botão 4 entrar dados e escolha o .REC correspondente
- No ultimo caso o EpiData irá perguntar se deseja atualizar os dados, pois irá perceber que o QES. é mais novo que o .REC

### **Anexar/fundir dados**

Com frequência mais de um digitador é envolvido no processo e de informatização dos dados contidos em formulários. Pode ocorrer do digitador 1 ser responsável pelos formulários de 1 a 200 e do digitador 2 ser responsável pelo 201 a 400. Ou então, o digitador 1 digita um banco com todos os dados coletados em campo e o digitador 2 digita todos os resultados dos laboratórios. A junção dos bancos pode ser realizada com o EpiData ou com os aplicativos de análise que fazem essas funções igualmente bem ou melhor como Stata, SPSS, R ou SAS.

- Pressione o menu “Entrada/Saída de dados\Anexar/fundir”
- Escolha os arquivos que serão utilizados como fonte
- Abrirá uma nova janela onde deverá ser escolhido o diretório e o nome do arquivo resultante

- Na mesma janela algumas opções devem ser especificadas para que o processo funcione. (as opções são intuitivas se for claro a compreensão da junção lado a lado ou abaixo)
- A opção “completar” deverá ser usada quando um banco é igual ao outro, mas com registros diferentes como no primeiro exemplo. O segundo banco será conectado embaixo do primeiro dando seqüência aos registros.
- A opção “associar”, diferente da anterior, juntará os bancos lado a lado e se aplicará no segundo exemplo em que os digitadores estão informatizando dados diferentes dos mesmos indivíduos. Esse processo depende de uma variável identificadora única presente em ambos os bancos.
- As outras opções serão fáceis de entender pelo desenho do aplicativo dados que estas acima já foram compreendidas

Boas pesquisas e que força esteja com vocês! ☺