

How to enter data with Excel in health research



Vo Tuan Khoa



Epidemiological Research Training Course VI-2, 2015

Outline

- Introduction to Excel
- Prepare data with Excel
- Data enter with Excel

Introduction to Excel

- Microsoft Excel is a useful spreadsheet program
- Not designed to be a research data entry tool → commonly used because almost researcher knows how to use it basically
- Data files from Excel can be shared or imported directly into files formatted by most statistical software (SPSS, Stata, R, Minitab)
- Limitation: spreadsheets containing less than 256 variables (columns) and 65,536 records (rows).



Prepare data with Excel

- 1. Well-designed protocol (conceptual framework)
- 2. Questionnaire form (verify all variables if nessessary and clarify format)
- 3. Making data dictionaries
- 4. Making data tables

PROTOCOL Group of Hematology repeat Double platelet apheResis for donAtion to Compare the safety dUration: an interventionaL study in vietnAm (DRACULA study)

CONCEPTUAL FRAMEWORK



BERH VIỆN TRUYỀN MẦU HUYẾT HỌC	BỆNH VIỆN TRUYỀN MÁU HUYẾT HỌC Nghiên cứu BTH-TC2014 MSĐT nghiên cứu: :/::::::::::::::::::::::::::::::::::	Phụ lục 9 Ngày thu thập: □□ - □□ - 2015 Lần hiến: Đầu tiên Người thu thập: Số hồ sơ:
	THÔNG TIN CĂN BẢN B/	AN ĐẦU
1. Năm sinh: 2. Phái tính: 1. Nam 2. Nữ	19 🗆 🗆	
3. Số lần đã l	niến tiễu cầu:	
4. Trình độ hi 1. Biết đọ 2. Trung l 3. Trung l	ọc vấn: c biết viết hoặc tiểu học học cơ sở học phổ thông	

			-
μ	1001	1910	\sim
1	ruu	inc	0

				Рһџ lục 0								
BIH VIỆN TRUYỆN MÁU HUYẾT HỌC	BỆNH VIỆN TRUYỀN Nghiên cứu BTH-TC2 MSĐT nghiên cứu:	MÁU HUYÉT HỌC 2014 □/□□/□□□	ÁU HUYÉT HỌC Ngày thu thập: □ □ - □ 4 Lần hiến: □ □ /□□/□□□ Người thu thập: Số hồ sơ:									
PHIẾU THEO DÕI HIẾN TIỀU CẦU (Dành cho nhân viên thực hiện chiết tách tiểu cầu thu thập) I. Các chỉ số xét nghiệm trước và sau khi hoàn thành chiết tách tiếu cầu:												
	Chỉ số	Trước chiết tách	Sau chiết tách	Túi tiểu cầu								
1. Bạch	cầu (10 ^{3/} mm³)											
2. Hồng	cầu (10 ^{6/} mm³)											

			Ρhụ lục ố								
lll. Các phản ứng lâm sàng không mong muốn xảy ra trong quá trình chiết tách :											
A. Những phản ứng ở mức độ nhẹ:											
1.	Tê môi, tê tay, tê chân	1. Có	2. Không								
2.	Khó chịu	1. Có	2. Không								
3.	Chóng mặt	1. Có	2. Không								
4.	Nhức đầu	1. Có	2. Không								
5.	Vã mồ hôi	1. Có	2. Không								
6.	Xanh xao	1. Có	2. Không								
7.	Xỉu (bất tỉnh)	1. Có	2. Không								
B. Nh	iững phản ứng ở mức độ trun	g bình:									
1.	Buồn nôn	1. Có	2. Không								
2.	Nôn	1. Có	2. Không								

1. Đánh giá c	hung của bạn về hiến máu									
	1. Đánh giá chung của bạn về hiến máu									
<u>Hướng dẫn trả lời câu hỏi</u> : Hãy trả lời từng câu bằng cách khoanh tròn vào số tương ứng với câu trả lời. Không có câu trả lời nào đúng hay sai. Nếu chưa chắc chắn, vui lòng chọn câu trả lời mà bạn cho là phù hợp với bạn nhất.										
CÂU HỎI ĐÁNH GIÁ VỀ VIỆC HIẾN MÁU TÌNH NGUYỆN (Dành riêng cho người hiến tiểu cầu)										
BERH VIỆN TRUYỀN MÁU HUYẾT HỌC	BỆNH VIỆN TRUYỀN MÁU HUYẾT HỌC Nghiên cứu BTH-TC2014 MSĐT nghiên cứu: ::::::::::::::::::::::::::::::::::	Ngày thu thậ Lần đến: Đầu Số hồ sơ:	p: 🗆 – 🗆 I tiên	□ - 2015						

Data dictionaries and/or code book

- Data dictionary makes the column definitions explicit
- Data dictionary is a table of information about the database itself
 - -rows representing fields
 - column for field name, field type and field description

viết tắc	tên biến	loại biến	giá trị	ghi chú
subid	mã số	chuỗi	XXXXXX	
ho	họ và chữ lót	chuỗi		không dấu
ten	tên	chuỗi		không dấu
ngaync	ngày vào nghiên cứu	ngày	nn/tt/nnnn	
nhom	nhóm ngẫu nhiên	phân loại	2.nhóm 2 tuần	
			3.nhóm 3 tuần	
			4.nhóm 4 tuần	
ngaysinh	ngày sinh	liên tục	XXXX	1997-1945
phai	phái tính	nhị giá	1.nam	
			2.nữ	
hientc	số lần hiến tiểu cầu	liên tục		
hocvan	trình độ học vấn	phân loại	1.cấp l	
			2.cấp ll	
			3.cấp III	
			4.đại học/cao đẳn	g
vieclam	công việc làm	phân loại	1.trí óc	
			2.chân tay	
			3.hưu trí	
			4.không	

Variable names

- Most statistical programs allow long column headings or variable names
- Some rules for variable names
 - short enough to type quickly but long enough to be descriptive
 - English meaning
 - avoiding spaces and special characters (especially "dấu tiếng Việt")

Coded responses vs Free text

- Defining a variable should include specifing its range of allowed values
- Limiting responses to a ranged coded value > allowing free-text responses
- Set of response options to a question
 - exhaustive (all possible options are provided)
 - mutually exclusive (no two options are both correct)
- Consistent for coding yes/no (dichotomous) variables
- Consider "All that apply" questions

Data tables

- All computer databases have one or more data tables
 - rows = records or entities
 - columns = fields or attributes
- · Simplified data table
 - each row = an individual subject
 - each column = a subject-specific attribute (name, age, sex, predictor and outcome variable)
- Should assign a unique identification number (subject ID) to each study participant

X) • (¥ * 🖽 두	e. 74		and the second division of the second divisio	and the other division of	Sec. of			Dat	ta tables - Micr	osoft Excel	
Fil	e Ho	me Inser	t I	Page Layout	Formulas	Data Rev	iew	View	Add-Ir	15			
Norm	al Page Layout	Page Break Preview /orkbook Vie	Custo View ws	m Full s Screen	Ruler Gridlines Show	Formula Bar Headings	Zoom	100% Zoom	Zoom to Selectio	n Window	Arrange Freez All Panes	Byplit Hide Unhide	View Side by
	B14	-	(=	f_x									
1	A	В		С	D	E			F	G	Н	I	J
1									-	1.nam		1.cap II	1.tri oc
2						2.moi 2	tuan			2.nu		2.cap II	2.chan tay
3						3.moi 3	tuan					3. cap III	3.huu
4						4.moi 4	tuan				(so lan)	4.dh	4.khong
5	subid	ho		ten	ngaync	nhom		nam	sinh	phai	hientc	hocvan	vieclam
6													
7													
8													
9													

Data entry with Excel

- 1. Two in one (one reads code, one enters data)
 - Tab
 - Freeze panes tool
 - Data validation tool
- 2. Data form: each subject for each data entry form

Freeze panes

To view only and retain some top rows (such as row 1 contains your variable names) and some important column (such as column 1 contains study id) when scroll bars move

- Put your cursor in the cell that is simultaneously below the rows you want to freeze and to the right of the columns you want to freeze
- View / Freeze Panes: select Freeze Panes and undo by selecting Unfreeze Panes

MS Excel for Public Health

H6 Image: fill state of the state of	Split Hide Unhice	Freeze Panes v	Arrange All	New Windo G 1.nam 2.nu	Add-In Zoom to Selection	View 100% Zoom	Review r Review zoom E i 2 tuan	mulas Data Formula B lines I Headings Show D 2.mc	V Rul	Page Layor m Full s Screen <u>f</u> x C	isert L Cust View Views	me Ir Page Brea Preview /orkbook	Ho Page Layout H6 A	File formal
H6 f_x A B C D E F G 1 Image: Construction of the state	Split Hide Unhic	Freeze Panes •	Arrange ow All	ns New Windo G 1.nam	Add-In Zoom to Selection	View	Review Ir R Zoom	mulas Data Formula B Ilines I Headings Show	v Rul v Rul	Page Layor m Full s Screen fx C	isert I k View Views • B	me Ir Page Brez Preview Corkbook	Ho Page Layout H6 A	File
H6 fx A B C D E F G 1 1 1 1 1.nam 1.nam 2 2 2.moi 2 tuan 2.nu 2.nu 3 3 3.moi 3 tuan 4.moi 4 tuan (so 5 subid ho ten ngaync nhom namsinh phai hieu 6 7 2 1 <th1< th=""> 1 1 1<!--</th--><th>Split Hide</th><th>Freeze Panes *</th><th>Arrange ow All</th><th>ns Do New Windo</th><th>Add-In Zoom to Selection</th><th>View 100% Zoom</th><th>Review r Q Zoom</th><th>mulas Data er 🔽 Formula B llines 🗹 Headings Show</th><th>I Rul</th><th>Page Layor Im Full Is Screen</th><th>isert I uk Cust Views • (* B</th><th>me Ir Page Brea Preview Yorkbook</th><th>Ho Page Layout H6</th><th>File</th></th1<>	Split Hide	Freeze Panes *	Arrange ow All	ns Do New Windo	Add-In Zoom to Selection	View 100% Zoom	Review r Q Zoom	mulas Data er 🔽 Formula B llines 🗹 Headings Show	I Rul	Page Layor Im Full Is Screen	isert I uk Cust Views • (* B	me Ir Page Brea Preview Yorkbook	Ho Page Layout H6	File
H6 fx A B C D E F G 1 1 1 1.nam 1.nam 1.nam 2 2 2.moi 2 tuan 2.nu 2.nu 3 3 3.moi 3 tuan 4.moi 4 tuan (so 5 subid ho ten ngaync nhom namsinh phai hier 6 1 <t< th=""><th>Split Hide Unhic</th><th>Freeze Panes *</th><th>Arrange ow All</th><th>ns Do New N Windo</th><th>Add-In Zoom to Selection</th><th>View</th><th>r Q Zoom</th><th>mulas Data r 📝 Formula B llines 📝 Headings Show</th><th>v Rul V Rul</th><th>Page Layor m Full s Screen</th><th>isert I ik Cust Views Views</th><th>me Ir Page Brea Preview Yorkbook</th><th>Ho Page Layout H6</th><th>File</th></t<>	Split Hide Unhic	Freeze Panes *	Arrange ow All	ns Do New N Windo	Add-In Zoom to Selection	View	r Q Zoom	mulas Data r 📝 Formula B llines 📝 Headings Show	v Rul V Rul	Page Layor m Full s Screen	isert I ik Cust Views Views	me Ir Page Brea Preview Yorkbook	Ho Page Layout H6	File
H6 Image: Streen window Kriews Image: Streen window Kriews Image: Streen window Kriews H6 Image: Streen window Kriews H6 Image: Streen window Kriews Image: Streen window Kriews <td>Split Hide</td> <td>Freeze Panes *</td> <td>Arrange ow All</td> <td>ns D New N Windo</td> <td>Add-In Zoom to Selection</td> <td>View</td> <td>Review Ir Q Zoom</td> <td>mulas Data er 🕼 Formula B llines 🕼 Headings Show</td> <td>v Rul</td> <td>Page Layon</td> <td>isert I ok Cust Views</td> <td>me Ir Page Brea Preview Vorkbook</td> <td>Ho Page Layout</td> <td>File</td>	Split Hide	Freeze Panes *	Arrange ow All	ns D New N Windo	Add-In Zoom to Selection	View	Review Ir Q Zoom	mulas Data er 🕼 Formula B llines 🕼 Headings Show	v Rul	Page Layon	isert I ok Cust Views	me Ir Page Brea Preview Vorkbook	Ho Page Layout	File
H6 Image: file F G A B C D E F G 1 A B C D E F G 1 A B C D E F G 1 A B C D E F G 1 A B C D E F G 2 A A B C D E F G 3 A A B C D A A D 4 A A Moi 3 tuan A A S S 4 A A Moi 4 tuan (so S S 5 subid ho ten ngaync nhom namsinh phai hier 6 A A Ba Ruler Formula Bar Ruler K Add-Ins Imagee Imagee Imagee Imagee Imagee Imagee Imagee </td <td>Split Hide</td> <td>Freeze</td> <td>Arrange</td> <td>ns E New</td> <td>Add-In</td> <td>View</td> <td>Review Ir Q Zoom</td> <td>mulas Data er 📝 Formula B Ilines 📝 Headings</td> <td>Rul</td> <td>Page Layon</td> <td>isert I</td> <td>me Ir</td> <td>Ho Page</td> <td>File ormal</td>	Split Hide	Freeze	Arrange	ns E New	Add-In	View	Review Ir Q Zoom	mulas Data er 📝 Formula B Ilines 📝 Headings	Rul	Page Layon	isert I	me Ir	Ho Page	File ormal
H6 Image: state of s	Split			ns	Add-In	View	Review	mulas Data er 📝 Formula B	v Rul	Page Layo	isert I	me Ir	н	File
H6 fx A B C D E F G 1 1 1 1 1.nam 1.nam 1.nam 2 2 2.moi 2 tuan 2.nu 2.nu 3.moi 3 tuan 1.nam 3 3 3 3.moi 3 tuan (so 5 subid ho ten ngaync nhom namsinh phai hier 6 1				ns	Add-In	View	Review	mulas Data		Page Layo	isert	me Ir	Ho	File
H6 fx A B C D E F G 1									t Eo					
H6 Image: fx matrix	120												7	
H6 fx A B C D E F G 1								**************************************					6	
H6 Image: fx mark A B C D E F G 1 A B C D E F G 1 A B C D E F G 2 A B C D E F G 3 A B C D E F G 3 A A B C D E F G 4 A A B C D E F G Image: Comparison of the comparison o	ntc h	hier	hai	sinh p	nam		nhom	ngaync	n	t	ho	ubid	5 5	
H6 fx A B C D E F G 1 - - - 1.nam 2 - 2.moi 2 tuan 2.nu 3 - 3.moi 3 tuan -	lan) 4	(so				tuan	4.moi 4						4	
H6 fx A B C D E F G 1						tuan	3.moi 3						3	
H6 \bullet f_x A B C D E F G 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			2.nu	2		tuan	2.moi 2						2	
H6 \bullet f_x			nam	1			L	U	C	5		A	1	
H6 • fx		1	G	-	6		Е	D	C	2		Δ.	1	
WORDOOK VIEws 3110W 20011					20011			3110W	fr	- (-	TRUCOR	H6		
Layout Preview Views Screen Screen Selection Window All	Panes *	All	Window	Selection	7		adings	Chow	Screen	Views	Preview	Layout		
Normal Page Page Break Custom Full I Gridlines I Headings Zoom 100% Zoom to New Arrange	Freeze	Arrange	New	Zoom to	100%	Zoom	adings	Gridlines 🕅 He	Full	Custom	age Breal	Page P	Normal	
🔲 🔲 🗐 🐨 Ruler 🛛 Formula Bar 🝳 📑 🖬 🛁				Ó	7.	Q	mula Bar	Ruler 🔽 Fo						
File Home Insert Page Layout Formulas Data Review View Add-Ins				Add-Ins	View	iew	ita Revi	Formulas D	ge Layout	ert Pa	e In:	Hom	File	

Data validation tool

- To limit value in a column to a certain range or a set of values
- To prevent invalid value from being entered into a cell

Data validation tool

Highlight the column you want and then select Data / Data validation

er Existing Connection	Refresh All + @ Ed Connect	nnections $\begin{array}{c} \underline{A} \downarrow \\ \underline{A} \\ \underline{A} \\ \underline{A} \downarrow \\ \underline{A} \\ \underline{A} \downarrow \\ \underline{A} \downarrow \\ \underline{A} \\ \underline{A} \downarrow \\ \underline{A} \\ \underline{A} \downarrow \\ \underline{A} \\ $	Filter Sort & Filter	Clear Reapply Advanced	Text to Rer Columns Dup	→ Dat nove licates	a ion *	Consolidate What-If Analysis -	Group Ungr		
fx 1.	nam	E	E		н		L <u>i</u> rcle Llear	Data Validation			
L	U	L	1.	1.nam	1	1.cap II	1	Prevent invalid data from being entered into a cell. For example, you could reject			
		2.moi 2 tuan		2.nu		2.cap II	2				
	3.moi 3 tuan				3. cap II	13	invalid dates or numb than 1000.	pers greater			
		4.moi 4 tuan			(so lan)	4.dh	h 4	You can also force input to be			
ten	ngaync	nhom	namsinh	phai	hientc	hocvan	v	chosen from a dropd	own list of		
Van	1/7/2015	2	1972	1			2				
Thi	2/7/2015	3	1990					Press F1 for more	help.		

Data validation tool

In pop up menu: select type of data you want to enter and specify the range

Data Validation	Data Validation
Data Validation Settings Input Message Error Alert Validation criteria Allow: Any value Ignore blank Any value Ignore blank Decimal List Date Time Text length Custom	Data Validation Settings Input Message Error Alert Validation criteria Allow: Whole number Validation Input Message Error Alert Validation criteria Allow: Whole number Validation Input Message Error Alert Validation criteria Allow: Whole number Validation Validation criteria Apply these changes to all other cells with the same settings
<u>Clear All</u> OK Cancel	Clear All OK Cancel

Data validation tool

Now, we enter an invalid value, an error message pops up

А		В	С	D	E	F	G	Н	I	J
							1.nam		1.cap II	1.tri oc
					2.moi 2 tuan		2.nu		2.cap II	2.chan tay
					3.moi 3 tuan				3. cap III	3.huu
					4.moi 4 tuan			(so lan)	4.dh	4.khong
subid	ho		ten	ngaync	nhom	namsinh	phai	hientc	hocvan	vieclam
	1 Ng	uyen	Van	1/7/2015	2	1972	1		2	1
	2 Le		Thi	2/7/2015	3	1990	3			
					Microsoft Exc	el			×	
							ared is pot valid			
					- 3		ereu is not valiu.		to this call	
					A	user has restrict			to this cell.	-
						Retry	Cancel	Help		-
					W	as this informatio	on helpful?			-
										_

Note: not detect the incorrect value if it lies in the specified range

Data form tool

Create a form in excel, and when enter the form, data will be entered into the spread sheet

Data form tool

- File / Options
- In pop up menu: select Quick Access Toolbar, All Commands and then Form, Add and OK





Data form tool

- Highlight all colunms you want
- Click icon Form in corner of left top
- Then enter data in data form

А	В	С	D	E	F	G	Н		J	K	L	М
						1.nam		1.cap II	1.tri oc	1.co	1.nhieu	1.co
				2.moi 2 tuan		2.nu		2.cap II	2.chan tay	2.khong	2.it	2.khong
				3.moi 3 tuan				3. cap III	3.huu		3.khong	
				4.moi 4 tuan			(so lan)	4.dh	4.khong			(so huu)
subid	ho	ten	ngaync	nhom	namsinh	phai	hientc	hocvan	vieclam	thuocla	ruoubia	nha
					Sheet2		9	<u>×</u>)				
					subid:		New Ree	cord				
					h <u>o</u> :		New					
					ten:		Delet	e				
					ngaync:		Resto	re				
					nho <u>m</u> :		Find Pr	ev				
					n <u>a</u> msinh:							
					pha <u>i</u> :							
					hi <u>e</u> ntc:		Criteri	ia				
					hoc <u>v</u> an:		Close	2				
					vieclam:							
					th <u>u</u> ocla:							
					ruou <u>b</u> ia:							
					nha:		-					
▶ ► Sheet	1 Sheet2	Sheet3 🔶]/									
	A GINCLE A	oncorp A o	* /									

Referrences

- Alan C. Elliott, Linda S. Hynan, Joan S. Reisch, Janet P. Smith. 2006. Preparing Data for Analysis Using Microsoft Excel. *Journal of Investigative Medicine*. 54(6); 334-342.
- Microsoft Excel for data entry Health research.
 2010. Fernandez Hospital, Hyderabad (https://fernandezresearch.files.wordpress.com/2010/0 9/data-entry-with-excel.pdf)