Aug 19, 2016 Clinical epidemiology study ~Actual procedures~



開藝 天保9年 (1838年)

Hirohide Yokokawa, M.D., Ph.D. Department of General Medicine, Juntendo University School of Medicine



>Research question

Steps of data analysis

Presentation of analyzed data

Implications of obtained results

What is a research question?

 The research question sets out what you hope to learn about the topic.

 This question, together with your approach, will guide and structure the choice of data to be collected and analyzed.

http://www.socscidiss.bham.ac.uk/r esearch-question.html http://twp.duke.edu/uploads/media_items/resea h-questions.original.pdf



Guidelines highlight some of the features of good questions

- Relevant
- Manageable in terms of research and in terms of your own academic abilities
- Substantial and with original dimensions
- Consistent with the requirements of the assessment
- Clear and simple
- Interesting



- Relevant: Arising from issues raised in literature and/or practice, the question will be of academic and intellectual interest.
- Manageable: You must be able to access your sources of data (be they documents or people), and to give a full and nuanced answer to your question.

Substantial and original: The question should showcase your imaginative abilities, however far it may be couched in existing literature.

• Fit for assessment: Remember, you must satisfy the learning outcomes of your course. Your question must be open to assessment, as well as interesting.

Clear and simple: A clear and simple research question will become more complex as your research progresses.

Interesting: Make your question interesting, but try to avoid questions which are convenient or flashy.



Descriptive study

 Checking distribution and characteristics of the participants: Knowing the distribution and characteristics leads to adequate further analytic analysis.

• Checking errors: Data errors must be corrected before main analysis.

Analytic study

- Estimating associations: To know associations between outcome and factors.
- Exploring associated factors: To explore/identify factors associated with the outcome.

An example of a research question



The association between waist circumference and lifestyle related disorders has been reported already.



However, there are few studies examining the association between actual visceral fat mass and lifestylerelated disorders.

Automated bioelectrical impedance analysis (BIA)









A research question

To clarify the association between visceral fat mass and lifestyle-related disorders, and estimate an appropriate cutoff value for visceral fat mass that associates with an increased risk of developing lifestyle-related disorders.

What is the first step to analyze the clinical data?



www.spanease.com

Establish the inclusion and exclusion criteria

 Checking distribution and characteristics of the participants

Checking errors

Inclusion and exclusion criteria

Inclusion criteria; Completed voluntary medical check-ups, and had their amounts of visceral fat measured using a BIA automated body composition analyzer.

Exclusion criteria; Use of medications for hypertension, dyslipidemia, or diabetes mellitus, and/or past history of cardiovascular disease or cerebrovascular disease.

How to present distribution and characteristics?

- The first Table (Table 1) usually shows distribution and characteristics.
- As representative values, mean (standard deviation) for parametric data, median (minimum, maximum) for non-parametric data, percentage for categorical data are usually used.

Let's calculate mean (standard deviation) for the items below.

Age, BMI, Wait circumference (WC), Visceral fat mass (VFM), Lipid profiles (LDL-C, HDL-C, TG), Blood pressures, HbA1c.

ile <u>E</u> dit	<u>V</u> iew <u>D</u> ata	Txansform	Analyze <u>G</u> aphs <u>U</u> tilities	Add-00	5	<u>W</u> indow <u>H</u> e	lp							
2 6		2	Reports Descriptive Statistics	ſ		Frequencies		42	1		Ø			
	Name	Туре	Compare Means			Descriptives		lis <mark>s</mark> ing	Columns		Alig			
4	Height	String	General Linear Model						6		Left			
5	Weight	Numeric	Generalized Linear Models			Explore)	12	=	Right			
6	BMI	Numeric	Mixed Models			<u>C</u> rosstabs		•	12	=	Right			
7	VisceralFat	Numeric	Correlate			TURF Analysis	s e		12	=	Right			
8	VFMQartiles	Numeric			1	<u>₩</u> <u>R</u> atio P-P Plots		•	12	=	Right			
9	VFM≥1.8kg	Numeric						•	12	=	Right			
10	WaistCircu	Numeric	L <u>o</u> gimear			0-0 Plots	e		12	-	Right			
11	TC	Numeric	Classify				TNOTE	•	12		Right			
12	HDLC	Numeric	Dimension Reduction			None	None	•	12	畫	Right			
13	TG	Numeric	Sc <u>a</u> le Nonparametric Tests Forecasting Survival			None	None	•	12	1	Right			
14	LDLC	Numeric		<u>N</u> onparametric Tests Forecas <u>t</u> ing	<u>N</u> onparametric Tests	Nonparametric Tests	<u>N</u> onparametric Tests	Nonparametric Tests		None None)	12	=	Right
15	Dyslipidemia	Numeric			Forecasting None None		•	12	畫	Right				
16	FBS	Numeric		Survival	None None		12	疅	Right					
17	HbA1c	Numeric	M <u>u</u> ltiple Response	•		None Non)	12	≣ Righ	Right			
18	Imparedgluc	Numeric	🚰 Missing Value Anal <u>y</u> sis			None	None		12	=	Right			
19	SBP	Numeric	Multiple Imputation	•		None Non		•	12	=	Right			
20	DBP	Numeric	Bimulation		None None		None None		12	=	Right			
21	ElevatedBP	Numeric							12	≣	Right			
22	VisceralFat	Numeric	ROC Curve			None	None		12	=	Right			
23	Totalnumber	Numeric	IPM SPSS Amon		ŧ	None	None)	12	=	Right			
24			IBM SPSS Amos											

ansform	<u>A</u> nalyze	<u>G</u> raphs	Utilities Add-ons	<u>W</u> indow	<u>H</u> elp					
	1	Ē 📥 E			- 4	\$		AB6		
Туре	Widt	n Decime	le Lahal	\/alues	Micci		Align	Measure	Role	
ng	3 🖫 🛙	requencies					📰 Left	💑 Nominal	🔪 Input	
neric	12			Variable(s):			🗮 Right	i Scale	🔪 Input	
neric	12 🔒	V1	-	Age		Statistics	🗮 Right	💑 Nominal	🔪 Input	
ıg	6	Age (10 year	s incr	🔗 BMI		Charts	들 Left	💑 Nominal	🔪 Input	
neric	11 💏	Height		🔗 Visceral Fa	at Mass (VF	<u>F</u> ormat	■ Right	i Scale	🔪 Input	
neric	11	Weight VEM (Optile)				Style	■ Right	🛷 Scale	🔪 Input	
neric	1	VFM (21.8 kg) [VFM				I Right	Scale	🔪 Input	
ieric	1.	Waist Circur	nferen			🔚 Frequencie	es: Statistics	(3))	×
eric	11	TC	_			- Percentile V	aluae		entral Let Jency-	
eric	12	HDLC					alues			
neric	12 🗹	oisplay freque	ency tables			Quartites	in for 40		Medion	
neric	12	ſ	OK Paste F	Reset Cance	I Help	Cut poin	IS IOF. 10	equal group	Mada	
neric	12	_				Percenti	e(s):		Mode Sum	
neric	12	0		None	None	Add		Ľ	<u>s</u> um	
neric	12	0		None	None	Chang	e			
neric	12	0		None	None	Remov	e			
neric	12	1		None	None					
neric	12	0	Impared glucos	. None	None	(3)			Values are group	n midnoints
neric	12	0		None	None				, va <u>i</u> ues are grou	y mapointo
neric	12	0		None	None				Distribution	
neric	12	0	Elevated BP	None	None	Std. devi	ationi Minimu	im 🛛	Ske <u>w</u> ness	
neric	12	1	Visceral Fat M	None	None	Variance	Maximu	um 🛛	<u>K</u> urtosis	
neric	12	0	Total number of	. None	None	Ra <u>n</u> ge	📃 S. <u>E</u> . me	ean		
						-	Cont	tinue	Help	

Let's calculate percentage for the items below.

High blood pressure Dyslipidemia Impaired glucose tolerance

ッ <u>ト1</u>] - IPMA	DCC Statistics Data Editor					
kans form	<u>Analyze</u> <u>G</u> rapis <u>U</u> tilities	Add- <u>o</u> n	s <u>W</u> indow	<u>H</u> elp		
	Rogenta		A State			
2	Descriptive Statistics	- F.	123 Frequencie	S	2	
Тур	Compare Means			1	ssing	Colum
ring	General Linear Model	•		е		3
imeric	Generalized Linear Models		and Explore	e		12
Imeric	Mixed Models	•	Crosstabs.	·· •		12
ring	Correlate		TURF Analy	/sis		6
Imeric	Regression	•	<u> R</u> atio	e		12
imeric	Loglinear		剂 P-P Plots	e		12
Imeric	Classify		🛃 Q-Q Plots			12
imeric	Dimension Reduction	•	INDIE	None		12
imeric	Scale		None	None		12
imeric	Nonnarametric Tests		None	None		12
imeric	Forecasting		None	None		12
imeric	Sunvival		None	None		12
imeric	<u>Surviva</u> Multiple Response	- K	None	None		12
imeric	Muluple Response	, r	None	None		12
imeric	Missing Value Analysis		None	None		12
imeric	Multiple Imputation		None	None		12
imeric	📅 Simulation		None	None		12
imeric	Quality Control	•	None	None		12
imeric	ROC Cur <u>v</u> e		None	None		12
imeric	IBM SPSS Amos		None	None		12
imeric	Liorare	м D1	None	None		12



Let's make a table for basic characteristics.

Table . Basic characteristics (N=422)

Number (%) or Mean (Standard deviation) 56.0 (8.8)

Age (years) Anthropometric measurements Body mass index (BMI) (kg/m²) Waist circumference (WC) (cm) Visceral fat mass (kg) Blood pressure-related factors Systolic blood pressure (mmHg) Diastolic blood pressure (mmHg) High blood pressure Lipid-related items High-density lipoprotein cholesterol (HDL-C) (mg/dL)Low-density lipoprotein cholesterol (LDL-C) (mg/dL)Triglycerides (TG) (mg/dL) Dyslipidemia Glucose-related items Hemoglobin A1c (HbA1c) (%) Impaired glucose tolerance



How to present analytic study data?



www.brandwithin.com

Let's conduct Logistic regression analysis

 Bivariate regression analysis; adjusting for age (10-year increase)

Transform Analyze Gaphs Utilities Add-ons Window Help Regent Descriptive Statistics Descriptive Statistics Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Type Compare Means Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics String General Linear Model Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Mixed Models Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Mixed Models Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Mixed Models Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Logdineer Linear Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Dimension Reduction Image: Column and the statistics Image: Column and the statistics Image: Column and the statistics Numeric Numeric <t< th=""></t<>
Noperiod Noperiod Type Descriptive Statistics Compare Means Values String General Linear Model Numeric Generalized Linear Models Numeric Mixed Models Numeric Mixed Models Numeric Regression Numeric Automatic Linear Modeling Numeric Classify Numeric Dimension Reduction String Curve Estimation Particle Particle Squaroo Numeric Dimension Reduction Scale Binary Logistic
Type Descriptive Statistics Type Compare Means String General Linear Model Numeric Generalized Linear Models Numeric Mixed Models Numeric Classify Dimension Reduction Curve Estimation Partial Loogt Oquaroo Umeric Mises Teste Numeric Mises Teste
Type Compare Means Values Missing Columistic String General Linear Model None None 3 Numeric Generalized Linear Models None None 12 Numeric Mixed Models None None 12 String Correlate None None 12 Numeric Mixed Models None None 6 Numeric Correlate None None 12 Numeric Legeinoor Linear 2 Numeric Classify Linear 2 Numeric Dimension Reduction Curve Estimation 2 Numeric None Scale Binary Logistic 2
String General Linear Model None None 3 Numeric Generalized Linear Models None None 12 Numeric Mixed Models None None 12 String Generalized Linear Models None None 12 Numeric Generalized Linear Models None None 12 Numeric Generalized Linear Models None None 12 Numeric Generalized None None 12 Numeric Classify Linear Linear 2 Numeric Dimension Reduction Generalized Generation 2 Numeric None Scale Binary Logistic
Numeric Generalized Linear Models None None 12 Numeric Mixed Models None None 12 String Correlete None None 6 Numeric Regression Leglinear Linear 12 Numeric Classify Linear Linear 12 Numeric Dimension Reduction Curve Estimation 12 Numeric Scale Sinary Logistic Binary Logistic
Numeric Mixed Models None None 12 String Correlate None None 6 Numeric Regression Linear 12 Numeric Classify Linear 2 Numeric Dimension Reduction Curve Estimation 2 Numeric Scale Sinary Logistic 8
String Correlate None None 6 Numeric Regression Automatic Linear Modeling 12 Numeric Lassify Linear 2 Numeric Dimension Reduction Image: Correlate 12 Numeric Scale Scale Image: Correlate 12 Numeric Scale Image: Correlate Image: Correlate 12 Numeric Scale Image: Correlate Image: Correlate 12 Numeric None None 12 12 Numeric Scale Image: Correlate Image: Correlate 12 Numeric None None None 12 Numeric None Image: Correlate Image: Correlate 12 Numeric None None 12 12 Numeric None None
Numeric Regression Automatic Linear Modeling Numeric Linear Numeric Dimension Reduction Numeric Scale Numeric Scale Numeric Binary Logistic
Numeric Loglinoor Linear Numeric Dimension Reduction Curve Estimation Numeric Scale Numeric Scale
Numeric Classify Image: Classif
Numeric Dimension Reduction Numeric Scale Numeric Scale Numeric Binary Logistic
Numeric Scale Binary Logistic
Numeric State State Binary Logistic
Vumeric Forecasting P
Vumeric Survival prohit
Numeric Multiple Response Imm Problem 2
Numeric Moniperion Nonlinear
Numeric Multiple Imputation Mail Weight Estimation 2
Numeric III Simulation III 2-Stage Least Squares 2
Vumeric Quality Control None None 12
Numeric ROC Curve None None 12
Numeric IBM SPSS Amos None 12
None None 12



None None	Dependent: Block 1 of 1 Previous <u>C</u> ovariates:	Categorical Qptions Style
HDLC G LDLC Dyslipidemia FBS HbAtc Impared glucose tol SBP DBP Visceral Fat Mass (k Total number of CK	VFMQartiles(Cat) Age10yearsincrese(Cat) <u>Age10yearsincrese(Cat)</u> <u>Method: Enter</u> <u>Selection Variable:</u> <u>Pate Reset Cancel Help</u>	Logistic Regression: Options Statistics and Plots Classification plots Logistic Regression: Options Correlations of estimates Hosmer-Lemeshow goodnees-of-fit Iteration history Casewise listing of residual Statistics Outliers outside 2 std. dev. Al cases Display O At each step O At last step
ber or None	kight is Norr	Probability for Stepwise Entry: 0.05 Removal: 0.10 Conserve memory for complex analyses or large datasets ✓ Include constant in model Continue C ncel Help

Variables in the Equation

								95% C.I.fo	or EXP(B)
		В	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1ª	Age10yearsincrese			3.992	3	.262			
	Age10yearsincrese(1)	.160	.576	.078	1	.781	1.174	.380	3.631
	Age10yearsincrese(2)	.310	.522	.352	1	.553	1.363	.490	3.789
	Age10vearsincrese(3)	648	529	1 501	1	221	1 911	678	5 386
(VFMQartiles			11.363	3	.010			
	VFMQartiles(1)	.066	.311	.045	1	.833	1.068	.581	1.964
	VFMQartiles(2)	.612	.301	4.139	1	.042	1.844	1.023	3.326
. L	VFMQartiles(3)	.804	.307	6.862	1	.009	2.233	1.224	4.075
	Constant	-1.260	.543	5.398	ï	.020	.284		

Odds ratio

a. Variable(s) entered on step 1: Age10yearsincrese, VFMQartiles.

Table . Odds Ratios for Lifestyle-Related DisordersAccording to Visceral Fat Mass (Logistic RegressionAnalysis)

		Bivariate ^{a)}				Ν	Aultivariate			
					Model 1 ^d)		Model 2 ^{e)}		
Visceral Fat Mass (kg)	OR ^{b)}	95% CIc)	Р	OR ^b	95% CIc)	Р	OR ^{b)}	95% CIc)	Р	
Elevated blood pressure										
Q1≤1.2	Refere	ence		Refe	rence		Referen	ce		
1.2 <q2≤1.8< td=""><td>1.07</td><td>0.58-1.96</td><td>0.83</td><td></td><td></td><td></td><td></td><td></td><td></td></q2≤1.8<>	1.07	0.58-1.96	0.83							
1.8 <q3<u>≤2.5</q3<u>	1.84	1.02-3.33	0.04							
2.5 <q4< td=""><td>2.23</td><td>1.22-4.08</td><td>< 0.01</td><td></td><td></td><td></td><td></td><td></td><td></td></q4<>	2.23	1.22-4.08	< 0.01							
Dyslipidemia Q1≤1.2 1.2 <q2≤1.8< td=""><td colspan="2">Reference</td><td colspan="3">Reference</td><td colspan="3">Reference</td></q2≤1.8<>	Reference		Reference			Reference				
1.8 <q3<u>≤2.5</q3<u>										
2.5 <q4< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></q4<>										
Glucose tolerance impairment Q1≤1.2 1.2 <q2≤1.8< td=""><td>Refere</td><td>ence</td><td></td><td>Refe</td><td>rence</td><td></td><td>Referen</td><td>ce</td><td></td></q2≤1.8<>	Refere	ence		Refe	rence		Referen	ce		
1.8 <q3≤2.5< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></q3≤2.5<>										
2.5 <q4< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></q4<>										
							/			



et's interpret the results.

Which is the most appropriate cutoff of VFM for lifestyle related disorders? (1.2, 1.8, or 2.5 kg)

Let's make a Receiver operating characteristics curve (ROC curve) and estimate the appropriate cut off values of BMI and WC for visceral fat mass (≥1.8 kg)





Environ Health Prev Med DOI 10.1007/s12199-014-0411-9

REGULAR ARTICLE

Association between lifestyle-related disorders and visceral fat mass in Japanese males: a hospital based cross-sectional study

Hironobu Sanada · Hirohide Yokokawa · Junichi Yatabe · Scott M. Williams · Robin A. Felder · Pedro A. Jose · Seiichi Takenosita

Received: 2 July 2014/ Accepted: 3 September 2014 © The Japanese Society for Hygiene 2014

Environ Health Prev Med. 2014; 19:429-435