Data management, dummy tables, figures

Planning for Data Analysis

Outline

- Data management
  - Types of measurement
  - Data code book
  - Organizing data: wide vs long form
  - Data cleaning & checking accuracy of the data
- Describing data
  - Study Flow Charts
  - Dummy tables
  - Figures: Kaplan Meier curve, longitudinal curve
- Sharing Experiences, Q & A

Types of Measurement

- What is being measured?
  - A construct or a concept
  - Examples: speed, muscle strength, ability to clearly see, ability of a tool to detect changes in physiology, complications, satisfactions in one's life
- How is it measured?
  - Speed: against how fast your walk vs. distance per a unit of time
  - Muscle strength: depending on the muscle? Hang grip strength test
  - Visual acuity: foot (20/20) vs. LogMAR vs. self-reported
  - Complications: number of complications vs. focus on one important complication vs. patient-reported complication
  - Life satisfaction: Yes/No vs. Life Satisfaction Questionnaire

Examples of Type of Measurements

Constructs	Proxy Measures
Health	Self-rated or physician-rated health status
Effectiveness of a Therapy	Number of patients cured
Obesity	Change in cloth size, Body Mass Index

Discrete vs Continuous Data

- Discrete variables have values that can assume only whole number.
  - Number of siblings a person has
  - Number of lesions on a right hand
- Continuous variables have any values within a defined range.
  - Weights and temperature

Binary, Nominal, Ordinal, Interval, and Ratio Data

- Binary - Only take two values
- Nominal (or Categorical) - Only categorize
- Ordinal/Ranked - Categorize and an order of values are meaningful
- Interval - Categorize, put in order, and have equal distances between values
- Ratio Data - Same as interval and has a meaningful zero

Examples

- Binary: Gender - female and male
- Nominal:
  - Marital Status - Single/married/separated/widowed/divorced
  - Regions of Thailand - North, South, Northeast, and Central
- Ordinal/Ranked
  - Smoking status - Never, Former, Current
  - Cancer stages - Stage I, II, III, IV

Examples

- Interval
  - Number of physicians working at KCMH
  - Number of children a woman has before 45 years of age
- Ratio - length or duration

De-identified Data

- EU *General Data Protection Regulation* (*GDPR*)
- HIPAA regulation (US): example of HIPAA identifiers
  - Name
  - Telephone numbers; Fax number; Email address; Social Security Number; Medical record number
  - Address (all geographic subdivisions smaller than state, including street address, city county, and zip code)
  - All elements (except years) of dates related to an individual (including birthdate, admission date, discharge date, date of death, and exact age if over 89)

De-identified Data

- One File with study ID linking patient’s name or other identifier
- Data file with study ID and other variables collected in the study
- Study ID is a link between two files

Data Codebook

- Table with variable names, description, and details categories
- List of symptoms and medication: dynamic
  - Example: phenotype of drug hypersensitivity

Suggested codes for binary, categorical variable

- Binary variable: 0, 1
- Categorical variable: 1, 2, 3, …
- Cautions:
  - One column contain one piece of information only
    - Want to capture pain (Y/N): has information on location of pain in parenthesis: (knee), (upper arm)
  - Underlying diseases: One column had three number: 1, 2, 3; 2, 5, 7
    - One column – one underlying disease
    - One column – for number of co-morbidity or number of underlying disease

Organization of data: wide vs long

- Measuring outcome one time vs repeated measures
- Wide: records of variables for multiple time points in one row
- Long: one row contains value of variable at one time point

Wide Format Example

Pre vs. Post Design

ID	Intervention	Pretest scores	Posttest scores
1	Txt	24	50
2	Txt	19	56
3	Txt	35	63
4	Txt	30	55

Repeated measure design over 12 months

ID	Intervention	QOL_3mo	QOL_6mo	QOL_12mo
1	Txt	54	78	85
2	Usual Care	23	55	60
3	Txt	30	45	60
4	Usual Care	40	70	65

Wide Format

- Wide format: cross-sectional designs, cohort design with no repeated data
- Needed this format for
  - One-way repeated measure ANOVA (SPSS)
  - Two-way repeated measure ANOVA (SPSS)
  - Paired t-test (STATA, SPSS)

Long Format

- For longitudinal analysis
- Examples:
  - Study of depression trajectory over time
  - Whether the satiety hormone level reduces over time after receiving an educational intervention on foods

Long Format Example

id	group	sex	Time (mo.)	Weight	% body fat	Complication (Y/N)	SF-36
1	0	1	0	5	15	0	50
1	0	1	3	6	17	1	54
1	0	1	6	8	18	0	63
1	0	1	12	15	25	0	60
2	1	0	0	4	10	1	45
2	1	0	3	6	12	1	59
2	1	0	6	7	15	0	70
2	1	0	12	9	18	0	75
3	1	1	0	8	20	0	60
3	1	1	3	11	21	0	45

Data format for Survival Analysis

Patient's ID	Time to event (months)	Event	Medication	Covariate 1
1	30	Died (1)	Placebo	Male
2	60	Alive (0)	Txt	Female
3	55	Loss to follow up (0)	Txt	Female
4	60	Alive (0)	Placebo	Male
5	15	Loss to follow up (0)	Placebo	Female

Study of cancer patients receiving chemotherapy vs placebo – follow-up for 60 months

Data cleaning and checking the accuracy of the data

- Randomly select 5-10% to check the data accuracy
- Outliers: asking yourself if this observed values are possible
- Dates: typos on years
  - Common era – 2019 vs Buddhist era – 2562
  - Format – dd/mm/yyyy vs. mm/dd/yyyy
  - Excel – change computer date, the format can change
- Missing: check missing data

Using REDCap for data collection

- REDCap: http://crc.md.chula.ac.th/redcap/
  - No Cost
  - HIPAA-compliant environment
  - Space limit: 200 MB
- Benefits:
  - Type of variables planned at the beginning of data collection
  - Set the limit or range of values: minimized typos
  - Can prints the forms for subjects/participants
  - Options for exporting data into Stata, SAS, SPSS, excel with variable labels

REDCap: Contacts

- Phanupong Phutrakool (Del)
  - Tel. +662-251-6702
  - Email address: phanupong.p@chulacrc.org

Using shared drives in project collaborations

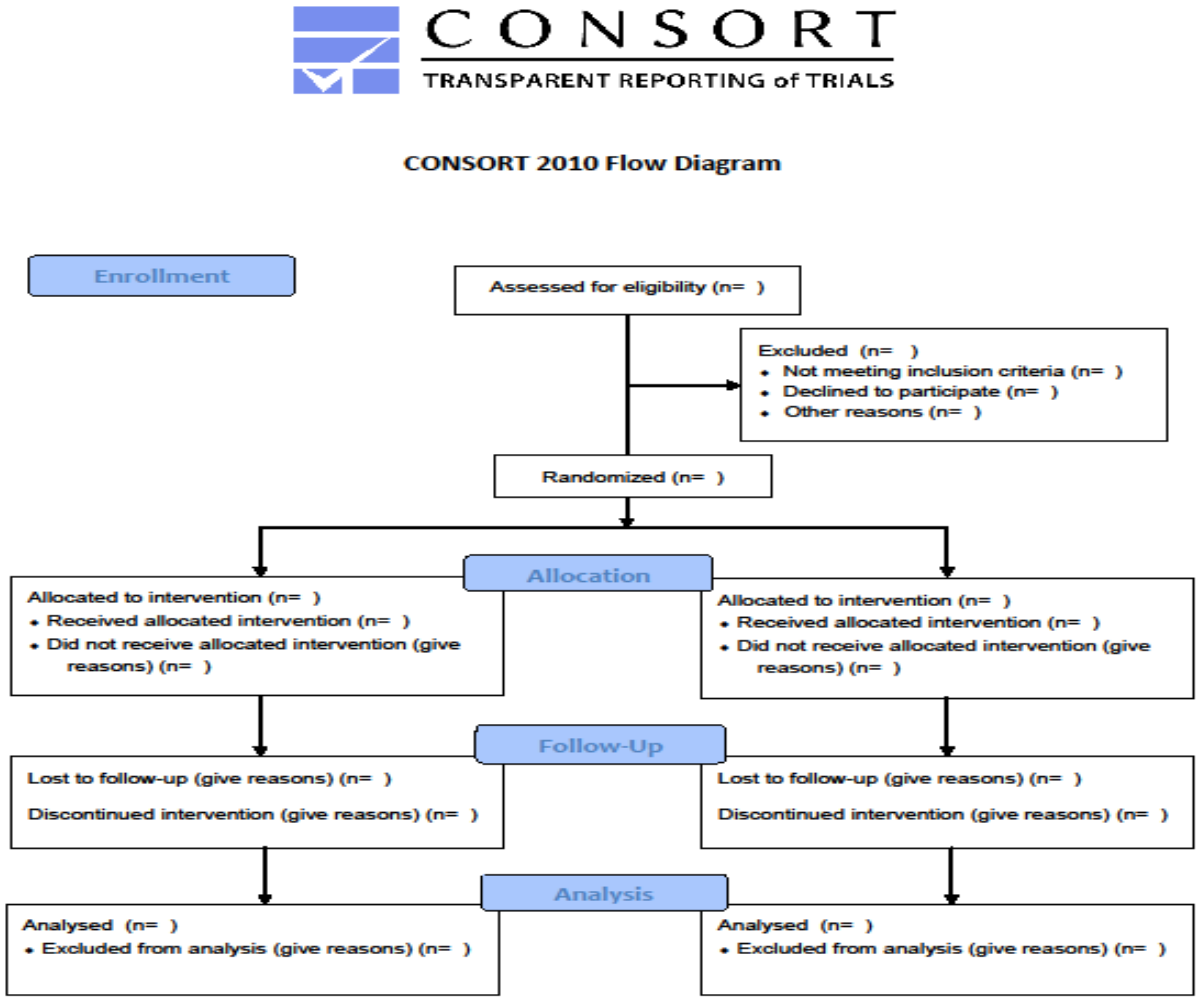
- OneDrive: 5 GB
- Dropbox: 2 GB
- Backup and Sync from Google (Google Drive): 15 GB including other Google account and photos
- Google Apps from Chulalongkorn University: no space limit
  - http://www.it.chula.ac.th/th/googleapps

Use Reporting Guideline in Data Management and Data Analysis

- Equator Network: http://www.equator-network.org/
- SPIRIT (randomized controlled trial, RCT): protocol guideline
  - https://www.equator-network.org/reporting-guidelines/spirit-2013-statement-defining-standard-protocol-items-for-clinical-trials/
- CONSORT: reporting guideline
  - Suggested Flow Diagram (RCT): http://www.equator-network.org/reporting-guidelines/consort/
  - Significant test for baseline values (RCT): http://www.consort-statement.org/checklists/view/32--consort-2010/510-baseline-data

Dummy Tables

- Table shells of manuscript being prepared
- Tables should be able to stand alone or self-explanatory
  - Readers should be able to read only the table and understand what the outcome variable(s) would be and what are being compared



RCT: Study Flow Chart

http://www.consort-statement.org/consort-statement/flow-diagram

RCT Table 1: Example 1

Table 1. Patient Demographics and Baseline Characteristics

Characteristic	Patients, No. (%)		
	350 µg of Dexamethasone DDS (n=57)	700 µg of Dexamethasone DDS (n=57)	Observation (n=57)
Age, mean (SD), y	63.8 (10.2)	63.8 (11.6)	62.9 (12.0)
Range	22-81	22-86	23-85
Sex			
Male	30 (52.6)	29 (50.9)	31 (54.4)
Female	27 (47.4)	28 (49.1)	26 (45.6)
Race			
White	41 (71.9)	43 (75.4)	41 (71.9)
Black	3 (5.3)	4 (7.0)	5 (8.8)
Hispanic	13 (22.8)	9 (15.8)	7 (12.3)
Asian	0 (0.0)	0 (0.0)	4 (7.0)
Native American	0 (0.0)	1 (1.8)	0 (0.0)
Duration of current macular edema, y			
<0.5	16 (28.1)	21 (36.8)	21 (36.8)
0.5-1.0	30 (52.6)	15 (26.3)	13 (22.8)
>1.0	11 (19.3)	21 (36.8)	23 (40.4)
Hemoglobin A_{1c}, median, %[a]	7.6	7.3	7.3
Prior cataract extraction	11 (19)	12 (21)	11 (19)
Prior photocoagulation	34 (60)	35 (61)	29 (51)
Baseline visual acuity, mean (SD), No. of letters	54.4 (9.96)	54.7 (11.00)	54.4 (11.88)
Mean baseline central retinal thickness, mean (SD), µm	446.5 (123.7)	428.3 (155.9)	417.5 (126.8)

Abbreviation: DDS, drug delivery system.
[a]To convert to proportion of total hemoglobin, multiply by 0.01.

Adapted from “Randomized Controlled Trial of an Intravitreous Dexamethasone Drug Delivery System in Patients With Diabetic Macular Edema” by Haller et al. (2010), retrieved from doi:10.1001/archophthalmol.2010.21

RCT Results: Example 1

Table 2. Other Efficacy Measures at Day 90[a]

	350 µg of Dexamethasone DDS	700 µg Dexamethasone DDS	Observation
Central retinal thickness[b]			
Patients, No.[c]	14	11	19
Change in macular thickness, mean (SD), µm	−42.57 (95.96)	−132.27 (160.86)	+30.21 (82.12)
P value (95% CI for difference) vs observation[d]	.07 (−151.3 to 5.8)	<.001 (−247.0 to −78.0)	
Fluorescein leakage			
Patients, No.[c]	54	55	56
≥2 levels of improvement, No. (%)	12 (22.2)	20 (36.4)	3 (5.4)
P value (95% CI for difference) vs observation[d]	.01 (4.3 to 29.4)	<.001 (17.0 to 45.0)	1 (1.8)
≥3 levels of improvement, No. (%)	7 (13.0)	15 (27.3)	
P value (95% CI) vs observation[d]	.03 (1.6 to 20.8)	<.001 (13.2 to 37.8)	

Abbreviations: CI, confidence interval; DDS, drug delivery system.
[a]The outcome measures were not assessed at day 180.
[b]Measured with optical coherence tomography at selected sites only.
[c]The number of patients in the analyses varies because these measurements were not available for all patients.
[d]*P* values for retinal thickness are based on pairwise contrast from a 1-way analysis of variance model at each visit with factor of treatment. *P* values for fluorescein leakage are based on Fisher exact test. All are comparisons with the observation group.

Adapted from “Randomized Controlled Trial of an Intravitreous Dexamethasone Drug Delivery System in Patients With Diabetic Macular Edema” by Haller et al. (2010), retrieved from doi:10.1001/archophthalmol.2010.21

RCT: Example 2 – Flow Diagram

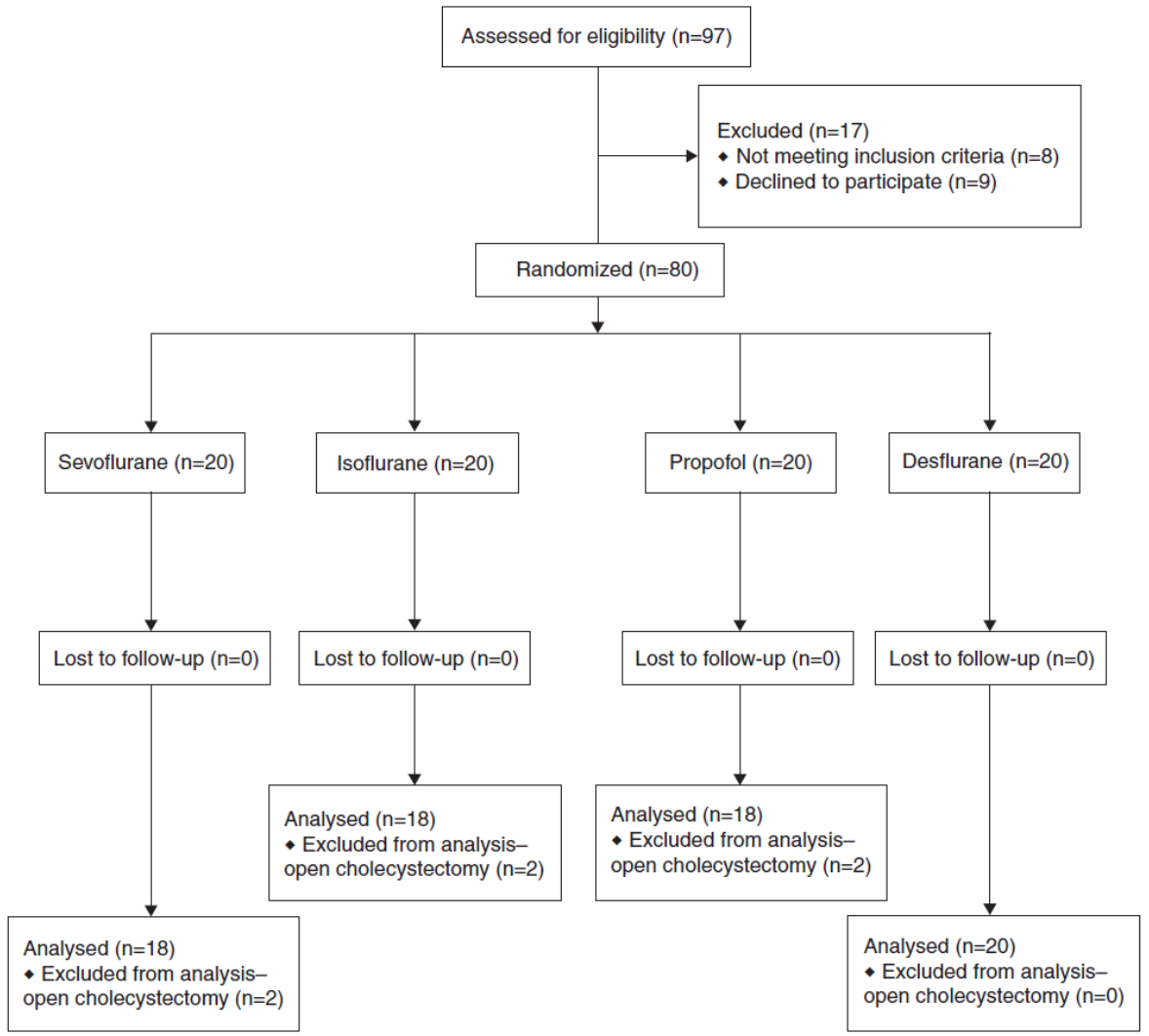


Figure 1 CONSORT flow diagram.

Adapted from “ARTICLERandomized, controlled trial comparing the effects of anesthesiawith propofol, isoflurane, desflurane and sevoflurane on painafter laparoscopic cholecystectomy” by Ortiz et al. (2014), retrieved from http://dx.doi.org/10.1016/j.bjane.2013.03.011

RCT: Example 2 - Table 1

Table 1 Patient demographics and surgical characteristics.

	PROP (n = 18)	ISO (n = 18)	DES (n = 20)	SEVO (n = 18)
Age	29(7)	34(12)	33(12)	34(14)
Weight (kg)	76(22)	80(16)	77(27)	74(16)
Height (in.)	62(2)	63(3)	63(3)	63(4)
Female	18(100)	16(89)	14(70)	15(83)
ASA class				
1	10(55)	5(28)	5(25)	7(39)
2	7(39)	13(72)	14(70)	10(55)
3	1(6)	0(0)	1(5)	1(6)
Diagnosis				
AC	11(61)	10(55)	12(60)	9(50)
BC	4(22)	5(28)	3(15)	8(44)
GP	3(17)	3(17)	5(25)	1(6)
Surgery time (min)	93(16)	102(45)	88(23)	86(28)
Anesthesia time (min)	148(19)	155(47)	142(24)	142(33)
Estimated blood loss (mL)	39(25)	47(54)	42(34)	37(28)
Nausea				
No	15(83)	13(72)	16(80)	16(89)
Yes	3(17)	5(28)	4(20)	2(11)

Continuous variables are presented as mean(SD) and categorical variables are presented as n(%).

Adapted from “ARTICLERandomized, controlled trial comparing the effects of anesthesiawith propofol, isoflurane, desflurane and sevoflurane on painafter laparoscopic cholecystectomy” by Ortiz et al. (2014), retrieved from http://dx.doi.org/10.1016/j.bjane.2013.03.011

RCT Results: Example 2

Table 2 Analgesic comparison.

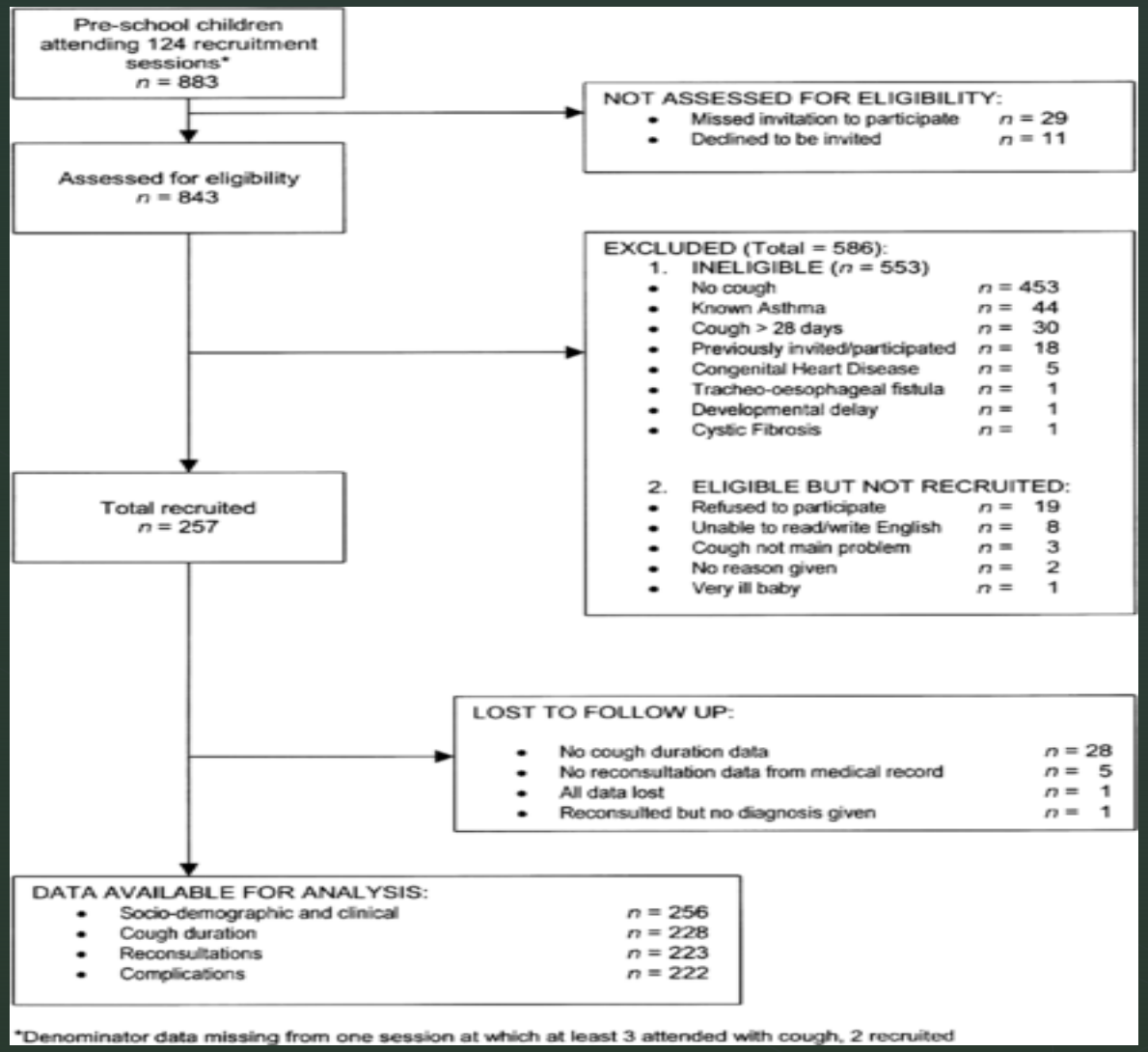
	PROP (n = 18)	ISO (n = 18)	DES (n = 20)	SEVO (n = 18)	p
Preop pain score (0–10)	1.3(2.4)	0.4(1.1)	1.7(2.1)	1.1(2.1)	0.28
Intraop fentanyl					
>250 mcg	6(33)	11(61)	8(40)	5(28)	
<250 mcg	12(67)	7(39)	12(60)	13(72)	0.21
Intraop morphine (mg)	6.1(4.3)	5.1(4.1)	3.6(4.0)	6.1(4.8)	0.24
24 h morphine (mg)	16(8)	15(11)	12(7)	13(8)	0.61
Hydrocodone/APAP (#)	1.9(1.8)	1.9(2.1)	2.2(1.6)	1.3(1.8)	0.53

Continuous variables are presented as mean(SD) and categorical variables are presented as n(%). p-values obtained by comparing summary measures across treatment groups using one-way ANOVA for continuously measured variables and Fisher's exact test for categorical variables.

Adapted from "ARTICLERandomized, controlled trial comparing the effects of anesthesiawith propofol, isoflurane, desflurane and sevoflurane on painafter laparoscopic cholecystectomy" by Ortiz et al. (2014), retrieved from http://dx.doi.org/10.1016/j.bjane.2013.03.011

Example of STROBE Flow Diagram

Flow diagram from Hay et al. [141].



Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLOS Medicine 4(10): e297. https://doi.org/10.1371/journal.pmed.0040297
https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0040297

PLOS | MEDICINE

Cross-Sectional Study: Table 1

Table 1: Demographic, distance and temporal relationships with malaria infection, *Plasmodium vivax* (Pv) and *Plasmodium falciparum* (Pf), in Gilgel-Gibe dam area, south-western Ethiopia, 2005

	Pv		Pf	
Variable	Rate	Crude OR (95% CI)	Rate	Crude OR (95% CI)
Age (years)				
<1 'control'	4/96 (4.2%)	1	3/96 (3.1%)	1
'at-risk'	13/190 (6.8%)	1.69 (0.52,5.52)	9/190 (4.7%)	1.54 (0.49,4.79)
1–4 'control'	18/396 (4.5%)	1	13/396 (3.3%)	1
'at-risk'	34/429 (7.9%)	1.81 (1.21,2.71)**	23/429 (5.4%)	1.67 (1.42,6.66)
5–9 'control'	12/282 (4.3%)	1	1/282 (0.4%)	1
'at-risk'	36/462 (7.8%)	1.9 (0.76,4.77)	27/462 (5.8%)	17.4 (1.22,249.24)**
Village/groups				
'control'	34/774 (4.4%)	1	17/774 (5.4%)	1
'at-risk'	83/1081 (7.7%)	1.81 (1.17,2.79)**	59/1081 (2.2%)	2.57 (1.01,6.57)**
Month				
October 'control'	15/253 (5.9%)	1	10/253 (4.0%)	1
'at-risk'	56/559 (10.0%)	1.76 (0.88,3.53)*	34/559 (6.1%)	1.57 (0.32,7.71)
November 'control'	13/260 (5.0%)	1	3/260 (1.2%)	1
'at-risk'	25/262 (9.5%)	2.00 (1.38,2.92)**	7/262 (2.7%)	2.35 (0.17,32.73)
December 'control'	6/261 (2.3%)	1	4/261 (1.5%)	1
'at-risk'	2/260 (0.8%)	0.33 (0.02,4.96)	18/260 (6.9%)	4.78 (1.03,22.23) **

* = significant at 0.1 level ** = significant at 0.05 level

Adapted from “Malaria and water resource development: the case of Gilgel-Gibe hydroelectric dam in Ethiopia” by Yewhalaw et al. (2009), retrieved from https://doi.org/10.1186/1475-2875-8-21

Cross-Sectional Study: Results

Table 2: Adjusted odds ratios (ORs) using a design-based logistic regression of malaria infection for *Plasmodium vivax* (Pv) and *Plasmodium falciparum* (Pf) by age, gender, month and village of residence in Gilgel-Gibe dam area, south-western Ethiopia, 2005.

Variable		Adjusted OR Pv	p-value	Adjusted OR Pf	p-value	Adjusted OR *Plasmodium* postivity	p-value
Village	'at risk'	1.63	0.015 **	2.40	0.085 *	1.97	0.013**
	'control'	1.00	--	1.00.	--	1.00	--
Month	October	1.00	--	1.00	--	1.00	--
	November	0.88	0.428	0.39	0.199	0.68	0.200
	December	0.18	0.096 *	0.89	0.828	0.41	0.062*
Age (yrs)	<1	1.00	--	1.00	--	1.00	--
	1–4	1.19	0.209	1.17	0.241	1.20	0.083*
	5–9	1.15	0.710	.94	0.890	1.08	0.837
Sex	Male	1.00	--	1.00	--	1.00	--
	Female	1.79	0.054	0.87	0.541	1.33	0.146

* = significant at 0.1 level ** = significant at 0.05 level

Adapted from “Malaria and water resource development: the case of Gilgel-Gibe hydroelectric dam in Ethiopia” by Yewhalaw et al. (2009), retrieved from https://doi.org/10.1186/1475-2875-8-21

Case Control Study: Table 1

Table 1 Background information for cases and controls

Data	All cases (n = 959)	Cases with Parkinson's disease (n = 767)	Controls (n = 1989)	All cases vs controls (p value)
Age, mean (SD)	69.9 (9.5)	69.8 (9.2)	69.8 (10.0)	
Sex, No (%)				
Male	537 (56)	426 (56)	1057 (53)	
Female	422 (44)	341 (44)	932 (47)	
Age left school (years), mean (SD)	14.5 (2.6)	14.5 (2.6)	14.4 (2.7)	0.20†
Currently working, No (%)	90 (9)	66 (9)	334 (17)	0.001‡
Friend/relative helped with responses, No (%)	287 (30)	227 (30)	214 (11)	0.001‡
Age at diagnosis (years), mean (SD)	62.4 (10.3)	61.6 (9.9)	n/a	
Interview quality assessment, No (%)*				
Implausible	4	2	9	0.001§
Poor/confused, but plausible	176 (19)	137 (18)	253 (13)	
Good	766 (81)	616 (82)	1685 (87)	

*Interview quality was categorised by the interviewer as implausible when the subject was unable to respond to basic questioning or provided data that were illogical or implausible. Where recall was occasionally poor or confused, but in the main sounded plausible, this was coded as poor/confused, but plausible. Where the subject provided good, precise responses and the work history was well described the interview quality was categorised as good.
†*t* Test; ‡χ^2 test; §χ^2 test combining "implausible" and "poor/confused, but plausible" categories.

Adapted from "Environgmental risk factors for Parkinson's disease and parkinsonism: the Geoparkinson Study" by Dick et al. (2007), retrieved from doi:10.1136/oem.2006.027003

Case Control Study: Table 3

Table 3 Adjusted results† (all cases vs controls)

	OR (95% CI)
Ever used tobacco containing product‡	0.50 (0.42 to 0.60)
Ever consumed beer, wine or spirits regularly	1.01 (0.83 to 1.23)
House with water supply from river or well‡§	1.18 (0.97 to 1.43)
Ever been knocked unconscious‡	1.57 (1.29 to 1.91)
Knocked unconscious:	
Once vs never	1.35 (1.09 to 1.68)
More than once¶ vs never	2.53 (1.78 to 3.59)
Ever had a general anaesthetic for an operation	0.81 (0.67 to 0.98)
Ever been treated by doctor after exposure to gas/smoke	0.99 (0.49 to 1.20)
Ever taken sleeping pills for >1 year	1.33 (1.07 to 1.65)
Ever taken medicines for anxiety for >1 year	1.95 (1.54 to 2.47)
Ever taken medicines for depression for >1 year	1.92 (1.49 to 2.49)
First-degree family history of Parkinson's disease‡	4.85 (3.43 to 6.86)
Any exposure to solvents	1.01 (0.84 to 1.23)
Any exposure to pesticides	1.29 (1.02 to 1.63)
Any exposure to iron	1.21 (0.87 to 1.44)
Any exposure to manganese	1.05 (0.81 to 1.37)
Any exposure to copper	1.00 (0.74 to 1.34)
Average annual intensity of exposure	
Solvents:	
Low exposure* vs no exposure	1.17 (0.92 to 1.50)
High exposure* vs no exposure	0.88 (0.69 to 1.12)
Pesticides:	
Low exposure* vs no exposure	1.13 (0.82 to 1.57)
High exposure* vs no exposure	1.41 (1.06 to 1.88)
Iron:	
Low exposure* vs no exposure	1.11 (0.79 to 1.56)
High exposure* vs no exposure	1.14 (0.82 to 1.59)
Manganese:	
Low exposure* vs no exposure	1.22 (0.86 to 1.73)
High exposure* vs no exposure	0.92 (0.64 to 1.32)
Copper:	
Low exposure* vs no exposure	1.05 (0.70 to 1.59)
High exposure* vs no exposure	0.94 (0.64 to 1.40)

*Cut-off point for low/high exposure taken to be median value of those exposed.
†Logistic regression adjusting for age, sex, country, ever used tobacco-containing product, ever knocked unconscious and first-degree family history of Parkinson's disease.
‡Odds ratios derived from a single logistic regression model with these factors as the only covariates.
§Excluding Malta water supply data.
¶Number of times knocked unconscious were once (n = 460), twice (n = 74), three times (n = 37), four times (n = 19), five times (n = 8), six times (n = 4), seven times (n = 1), 10 times (n = 4) and 20 times (n = 1).

Adapted from "Environgmental risk factors for Parkinson's disease and parkinsonism: the Geoparkinson Study" by Dick et al. (2007), retrieved from doi:10.1136/oem.2006.027003

Figure: Bar Chart

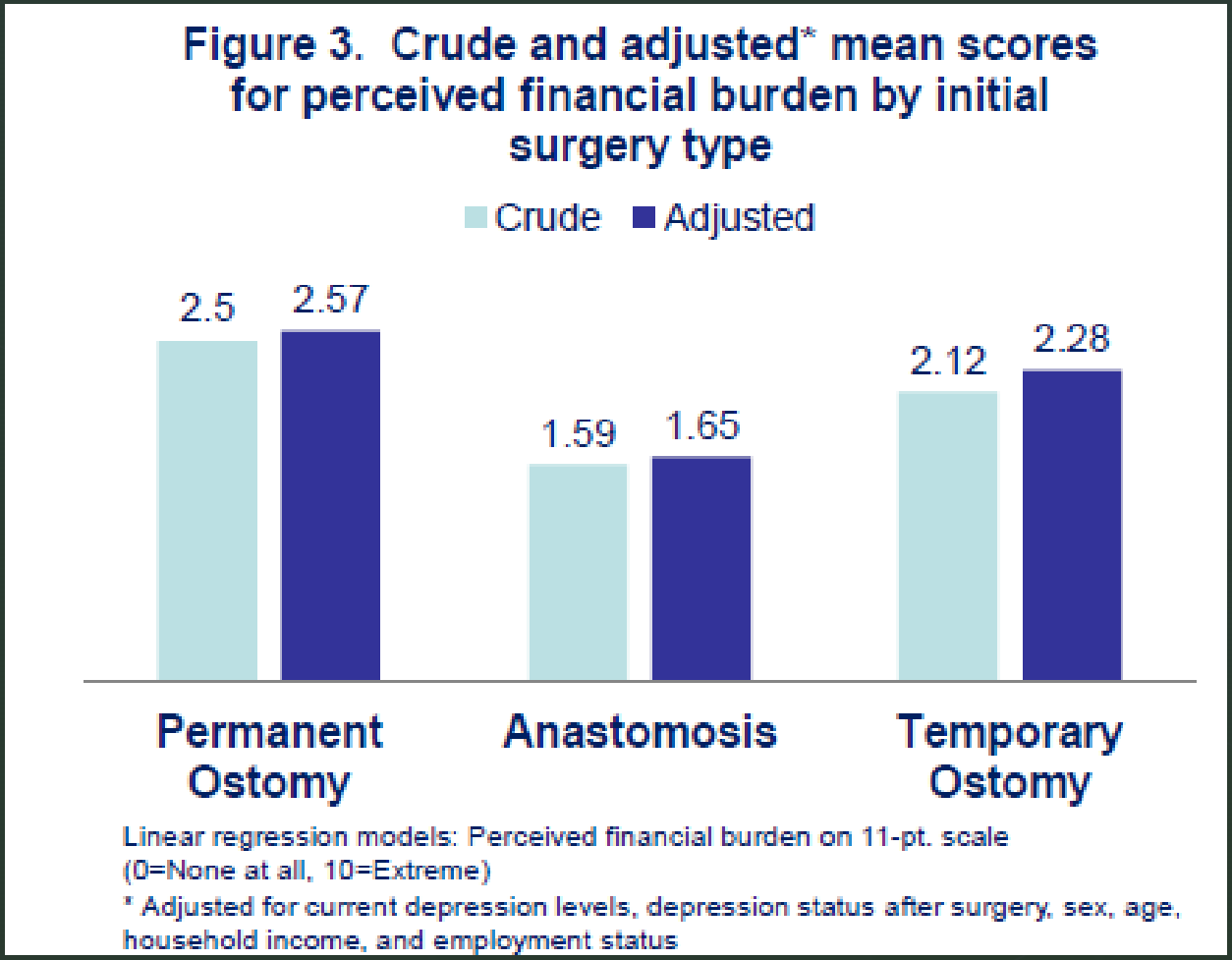
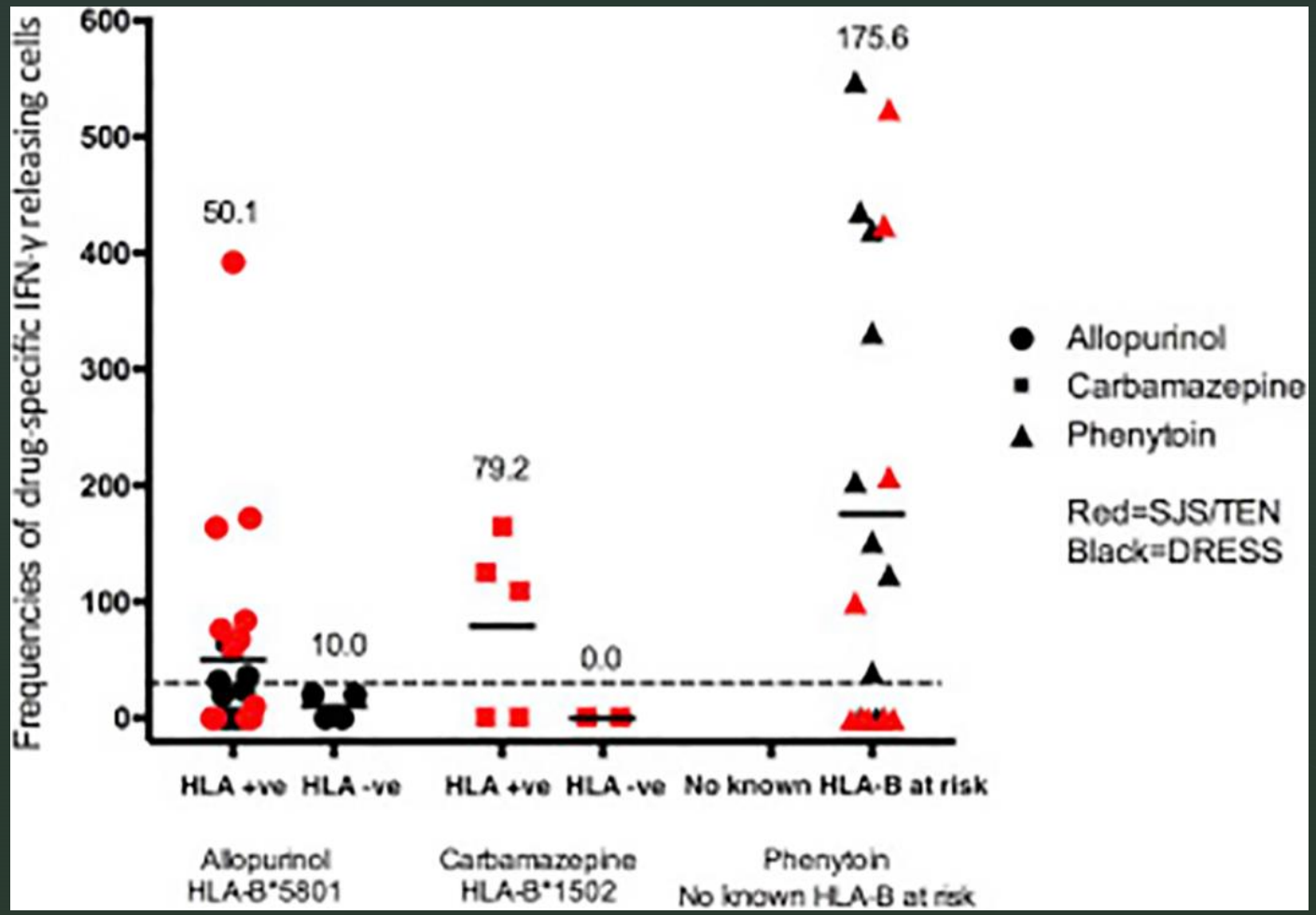


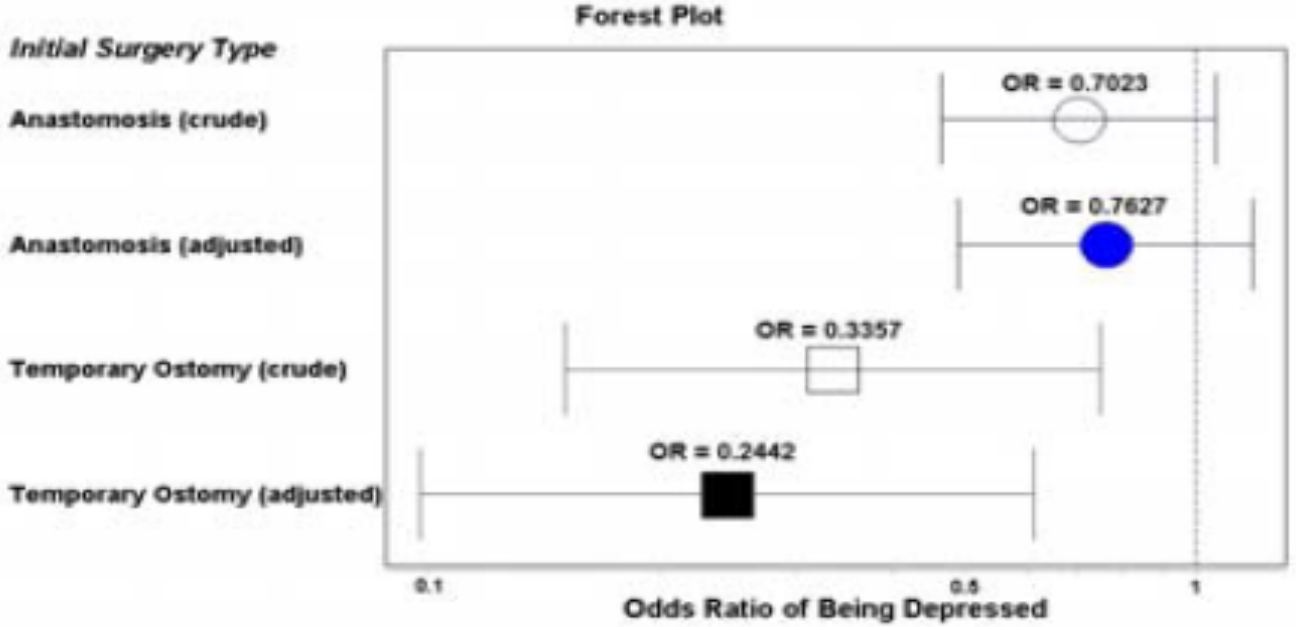
Figure: Dot Density Graph



Adapted from "*Analysis of HLA-B Allelic Variation and IFN-γ ELISpot Responses in Patients with Severe Cutaneous Adverse Reactions Associated with Drugs.*" by Klaewsongkram et al. (2019) retrieved from doi: 10.1016/j.jaip.2018.05.004.

Figure: Forest Plot

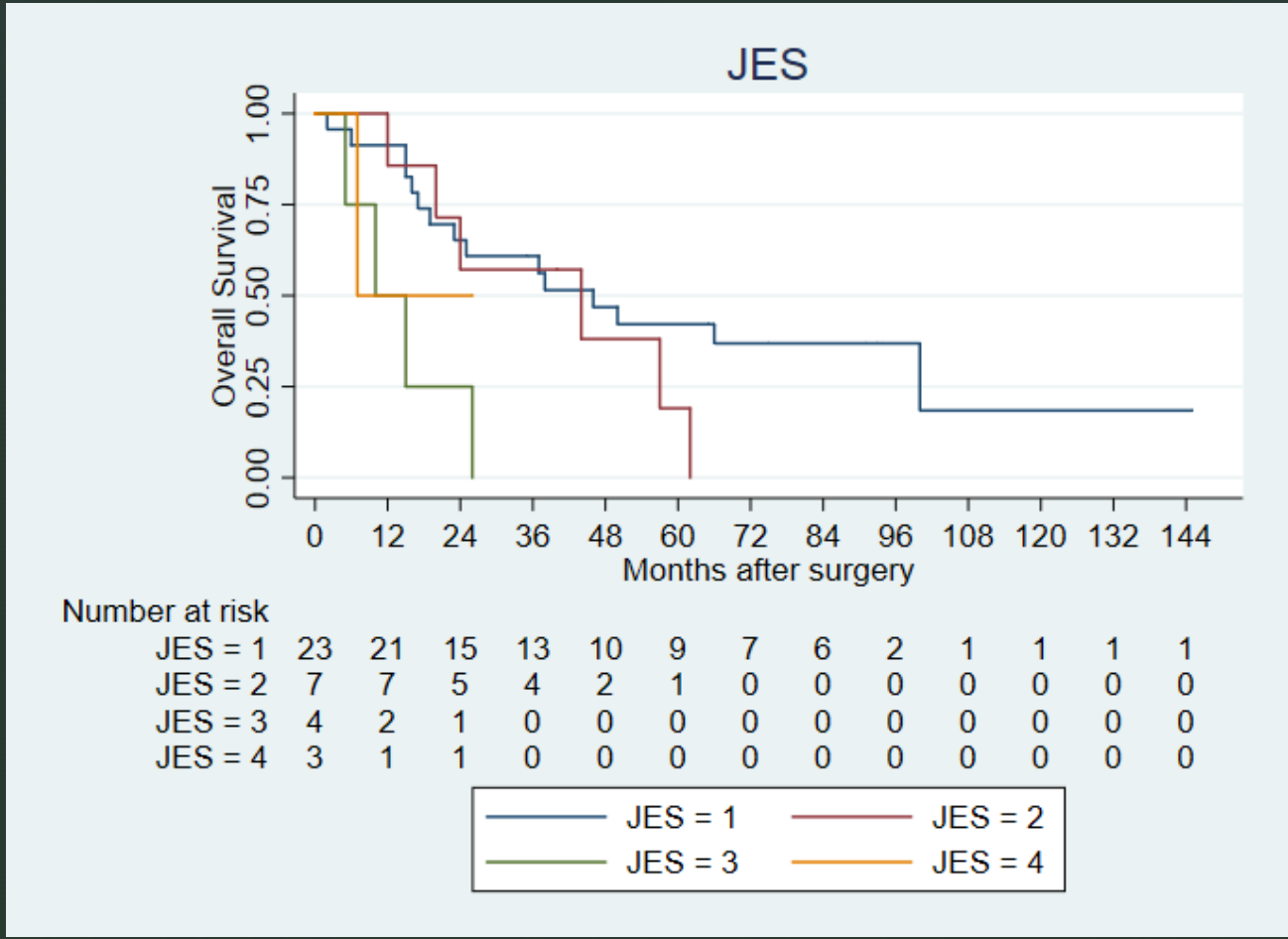
Figure 2. Crude and adjusted odds ratios of being currently depressed (Reference group: Permanent Ostomy)



Logistic Regression Models: Current depression status (depression coded present when reporting ≥ 4 pts.)

- Adjusted for age, sex, and race, odds of being currently depressed were *24%* and 76% lower in the *anastomosis* and temporary ostomy groups than the permanent ostomy group.

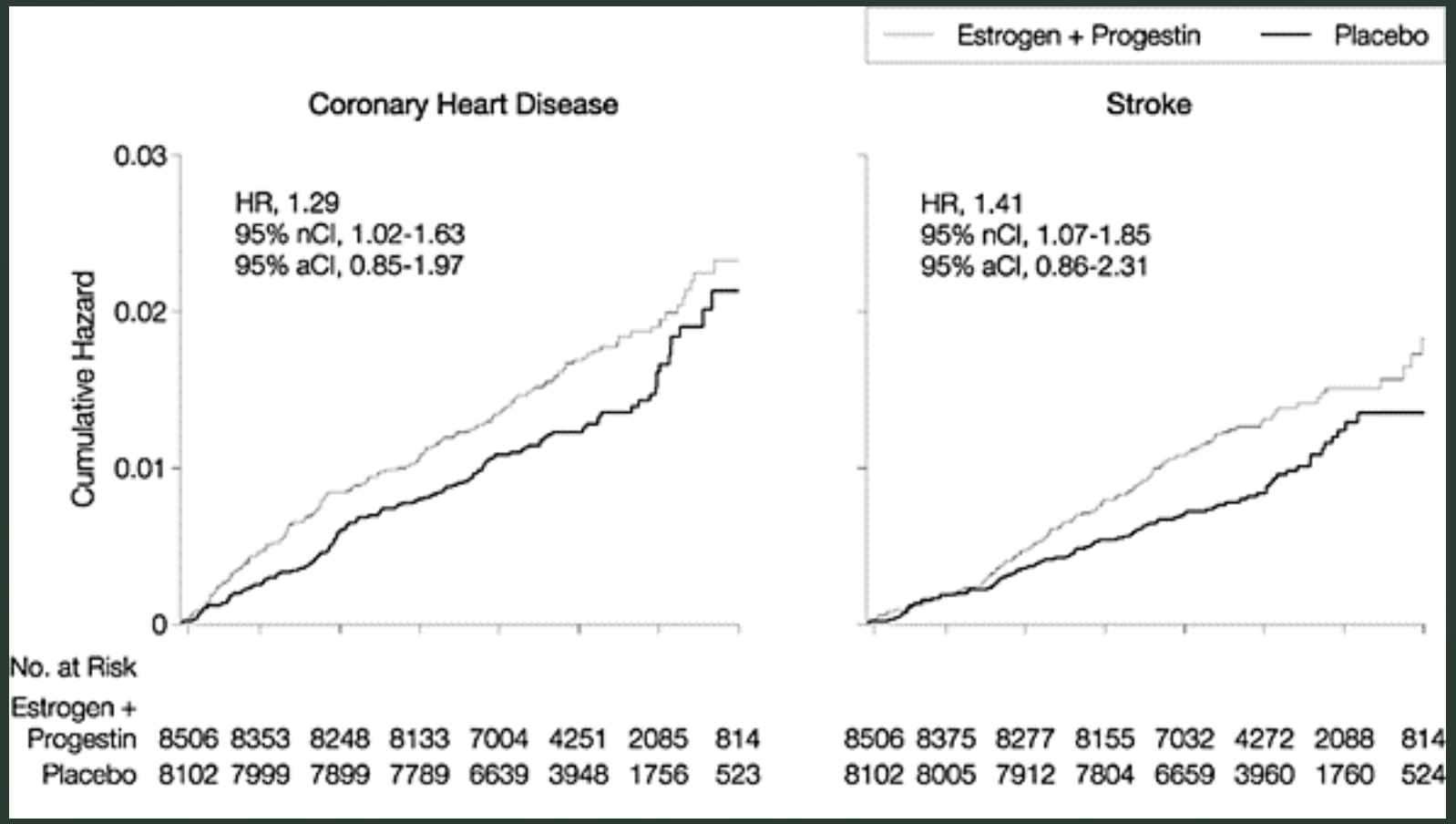
Figure: Kaplan Meier Curve



Showing survival function

Figure: Kaplan Meier Curve (Cont.)

Cumulative hazard for clinical outcomes



Adapted from “Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women” by Women’s Health Initiative Investigators, retrieved from doi:10.1001/jama.288.3.321

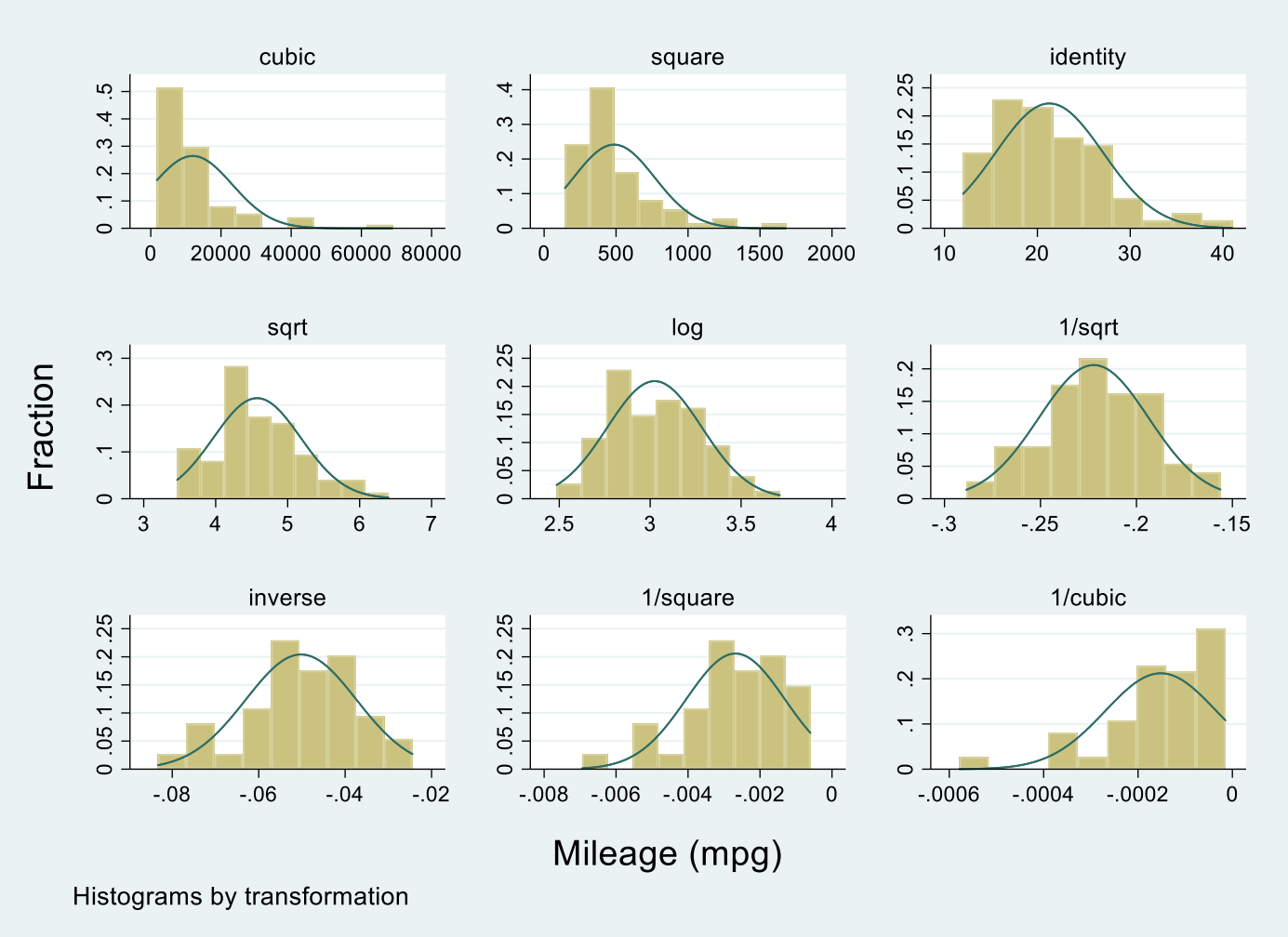
Create an Analysis Set

- After planning the statistical analysis
- Very important step before conducting statistical analysis
- R-program: complete.cases() syntax
- STATA: create new variable indicating the observations should be included in the analysis
  - Using if syntax
  - Using regression syntax
- SPSS: select case (https://www.youtube.com/watch?v=TiMk-4yFC24)

Q & A
Sharing Experiences

- MUSTS: SAVE a copy of original data set and BACK UP frequently
- Q & A
- Experiences:
  - Computer crashed with no back up, only summary statistics left
  - Data codebook not clear for a new collaborator
  - Typos in data entry: rx names
  - Inconsistency in terminology: brand name vs. generic rx, diseases

Data Transformation



Validity and Reliability Study

Table 1 – describing participants' characteristics

Table 1 Characteristics of 306 participants with drug hypersensitivity experience

Characteristics	Mean (SD)
Age (years)	46.3 (15.9)
Time to complete questionnaire (min)	7.44 (5.90)
Self-reported comprehensibility of questionnaire	7.57 (2.47)
	N (%)
Sex	
Females	225 (76.0)
Males	71 (24.0)
Pharmaceutical classes[a]	
Non-steroidal anti-inflammatory drugs and other pain reliefs	149 (34.9)
Beta-lactam antibiotics	129 (30.2)
Sulfonamide antibiotics	13 (3.0)
Other antibiotics[b]	71 (16.6)
Anti-epileptic drugs	11 (2.6)
Other drugs	54 (12.6)

[a]Each participant can be hypersensitive to more than one drug classes.
[b]Including antituberculosis drugs.

Adapted from "Reliability and validity of the Thai Drug Hypersensitivity Quality of Life Questionnaire: a multi-center study" by Chongpison et al. (2018)