Research design & study execution workshop series

Session 9

OCTOBER 28, 2015

Quick review of Sessions 1-8

- How to identify a “good” research question
- Common study designs: Pros & cons
- Selecting appropriate study subjects
- Understanding variables types and their measurement
- Nuts and bolts of good data management

Case study: Football-related injuries

Describing data:
Statistical and graphical methods

Sonnad S. Describing data: Statistical and graphical methods. Radiology 2002; 225:622-628.

Primary goal of statistics

To collapse data into easily understandable summaries

Football injury study

6-17 year old male patients who visited the ED at the main hospital for a football-related injury in 2014

(n=338 exams)

Numerical summaries

1. Count how often each value occurs
2. Find the center (mean, median, mode)
3. Measure the spread
4. Examine the distribution

Frequency tables

Count how often each value occurs

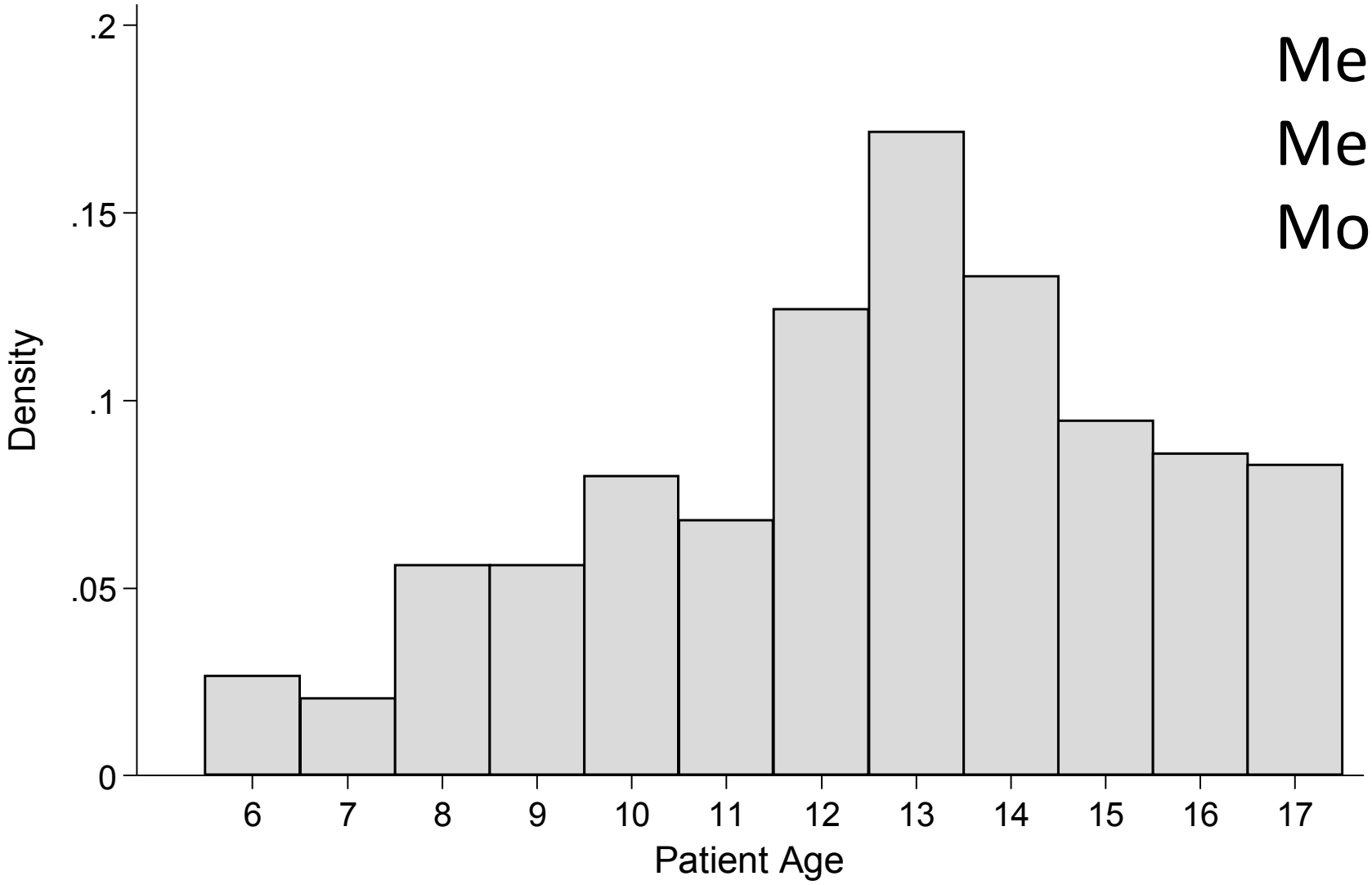
Patient Age	Freq.	Percent	Cum.
6	9	2.66	2.66
7	7	2.07	4.73
8	19	5.62	10.36
9	19	5.62	15.98
10	27	7.99	23.96
11	23	6.80	30.77
12	42	12.43	43.20
13	58	17.16	60.36
14	45	13.31	73.67
15	32	9.47	83.14
16	29	8.58	91.72
17	28	8.28	100.00
Total	338	100.00	

Age category	n	%
6-11 years old	104	31%
12-17 years old	234	69%
Total	338	100%

Find the center

Mean, median & mode

Measure	Definition	Excel formula
Mean	Sum of all observations/total number of observations	=AVERAGE(range of cells)
Median	Number in the middle of the data set	=MEDIAN(range of cells)
Mode	Most common number in the data set	=MODE(range of cells)



Mean = 12.6 years

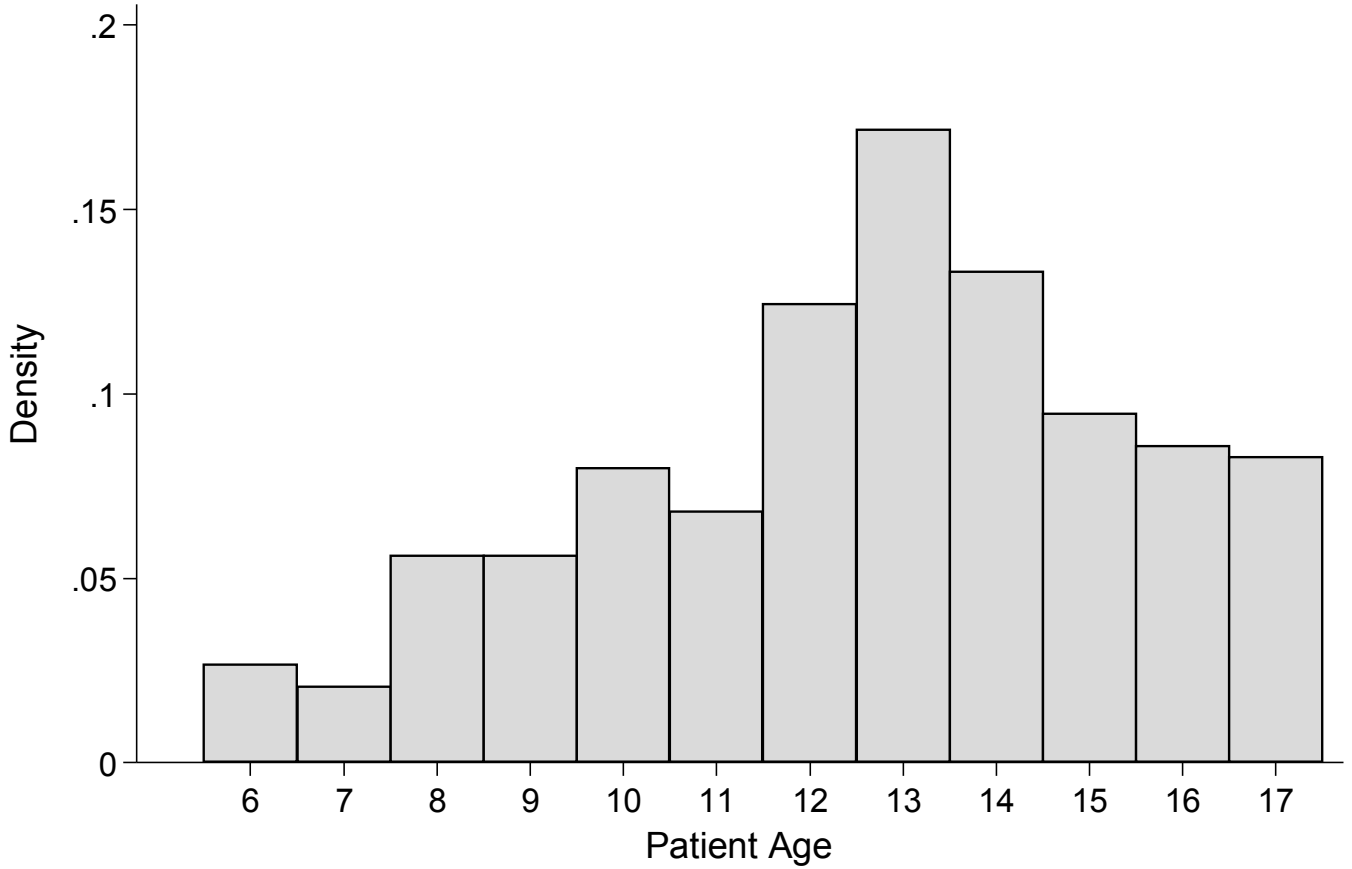
Median = 13 years

Mode = 13 years

Measure the spread

(tells more of the story)

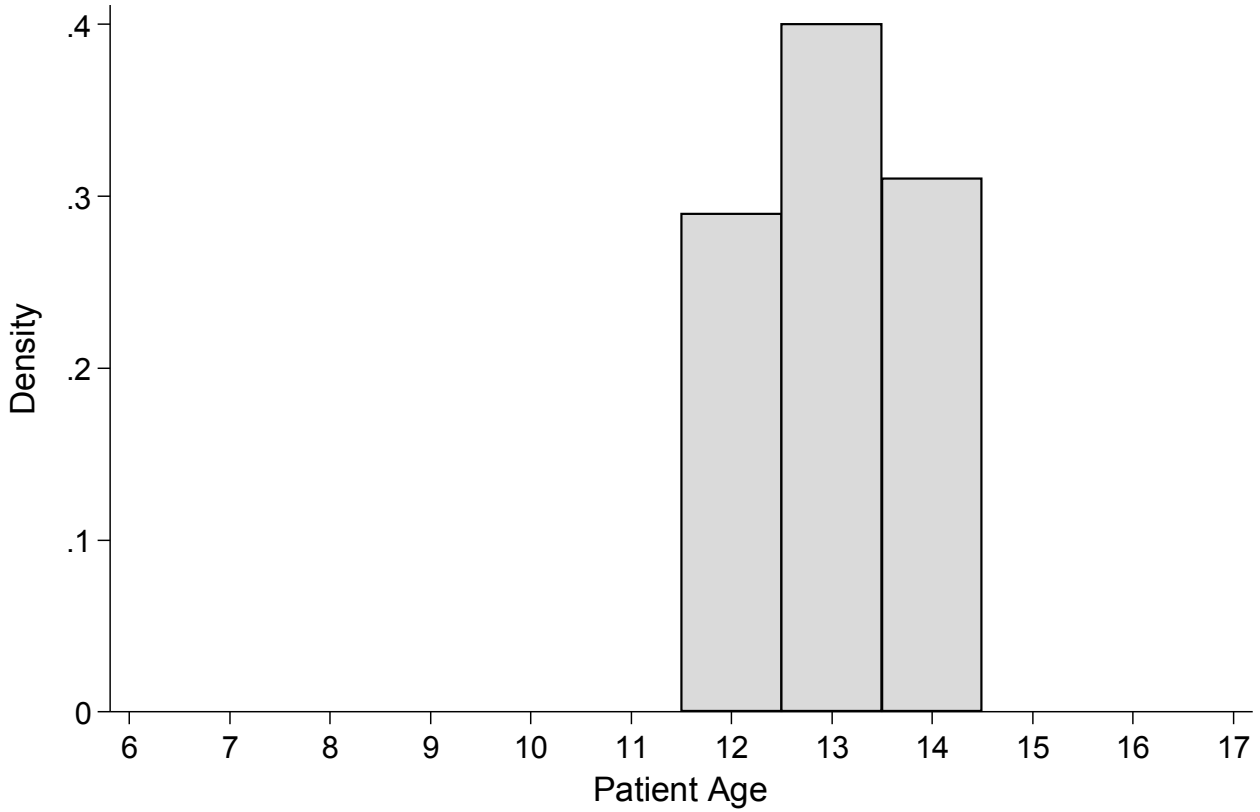
Center	Spread	Reporting format
Mean	Standard deviation (sd)	Mean ± sd
Median	Interquartile range (IQR)	50^{th} percentile (25^{th}, 75^{th} percentile)



Patient age

12.6 ± 2.8 years *
13 (11, 15 years) **

* Mean ± sd
** Median (IQR)



Patient age

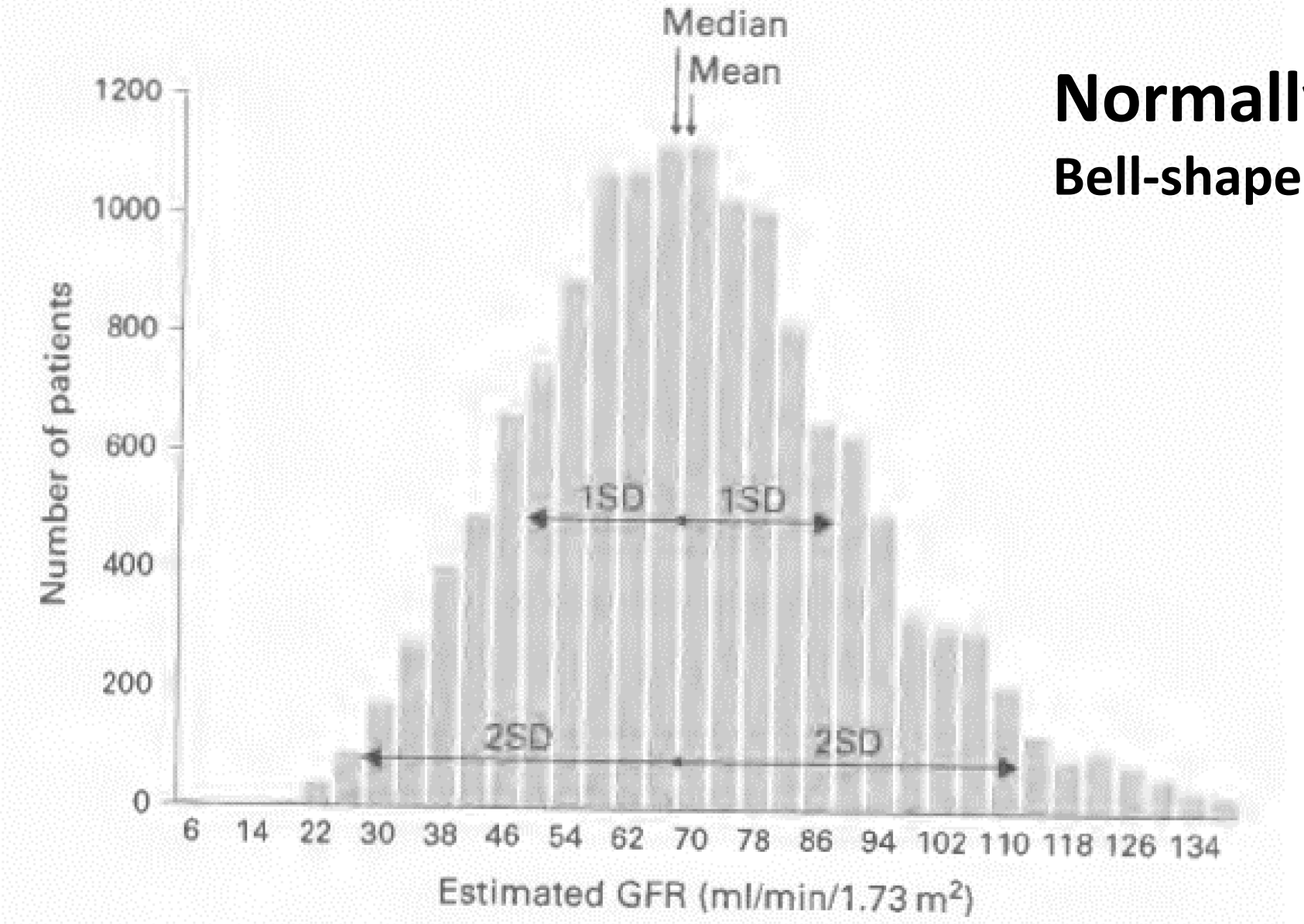
13.0 ± 0.8 years *
13 (12, 14 years) **

* Mean ± sd
** Median (IQR)

Examine the distribution

Is it normal (symmetric)?

Is it skewed (asymmetric)?

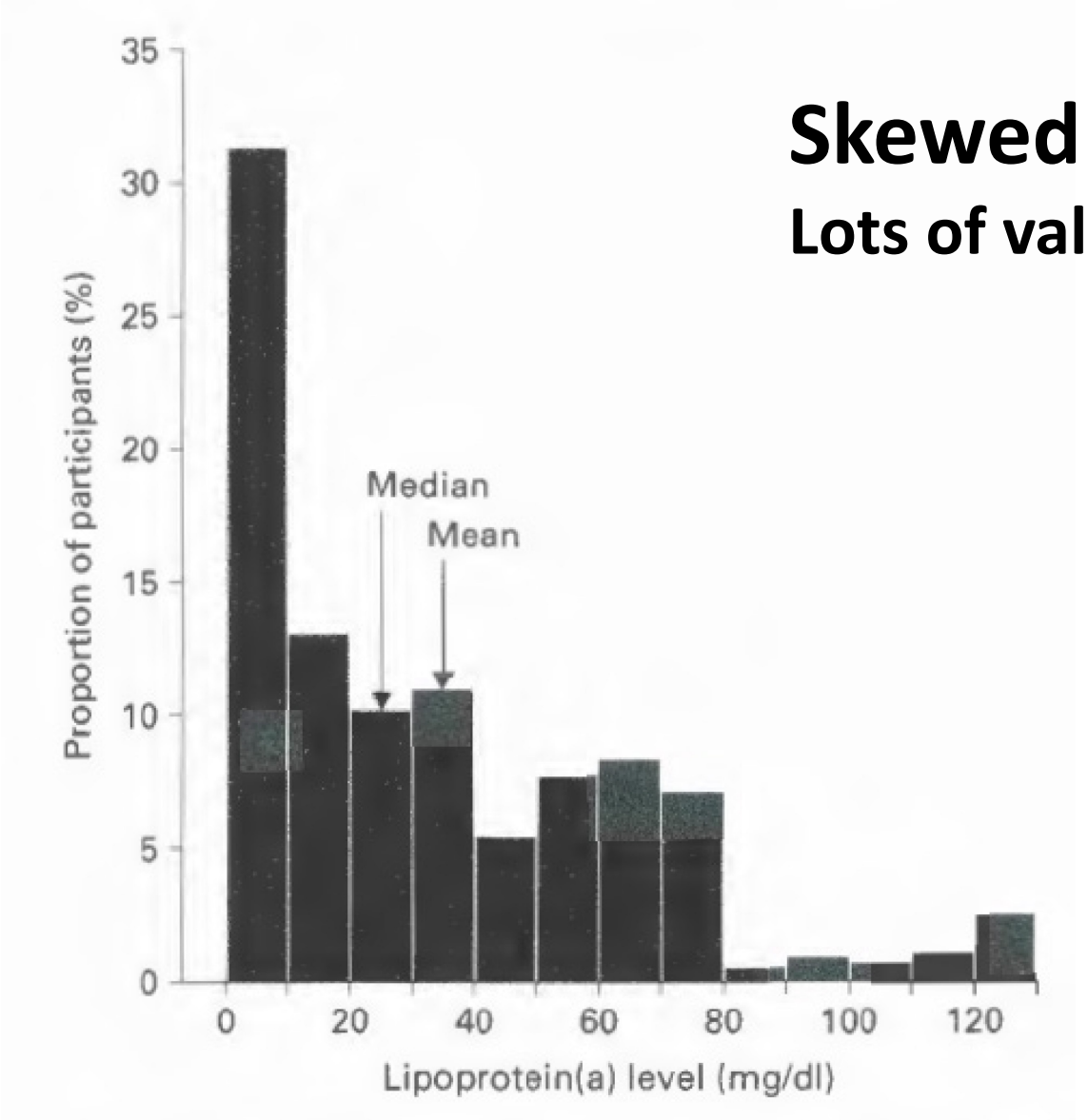


Normally distributed

Bell-shaped curve

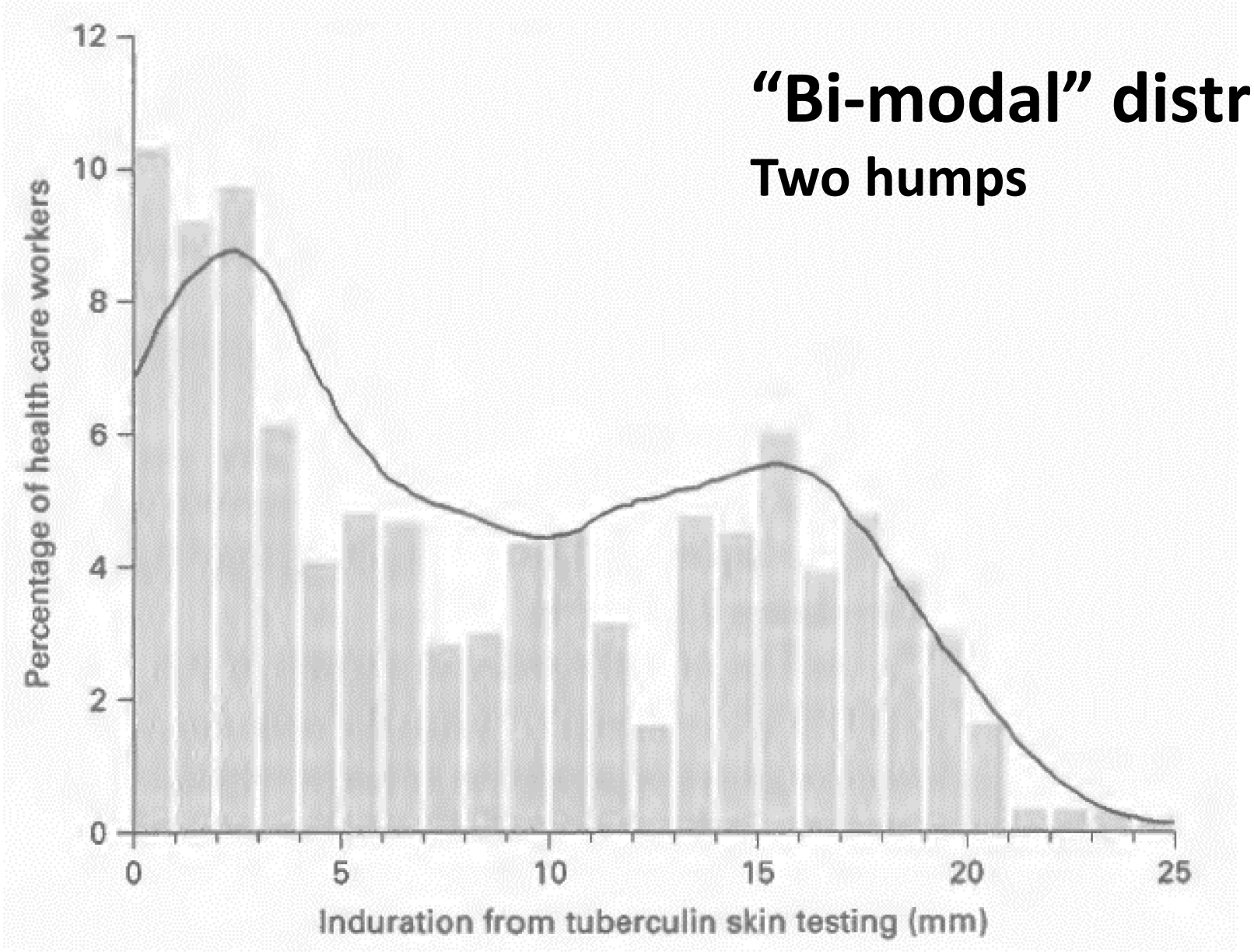
Skewed distribution

Lots of values on one end or the other



“Bi-modal” distribution

Two humps



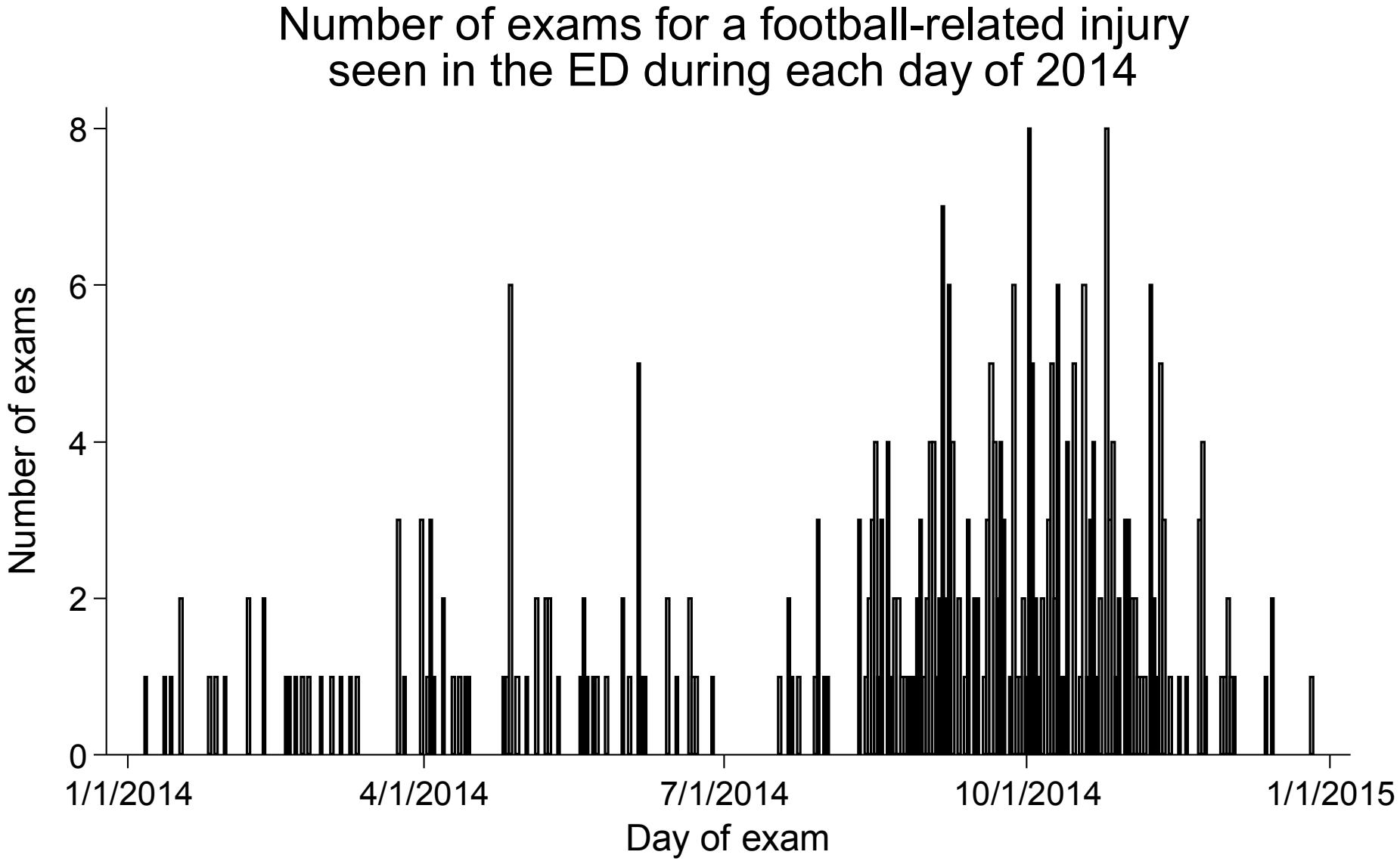
Graphical summaries

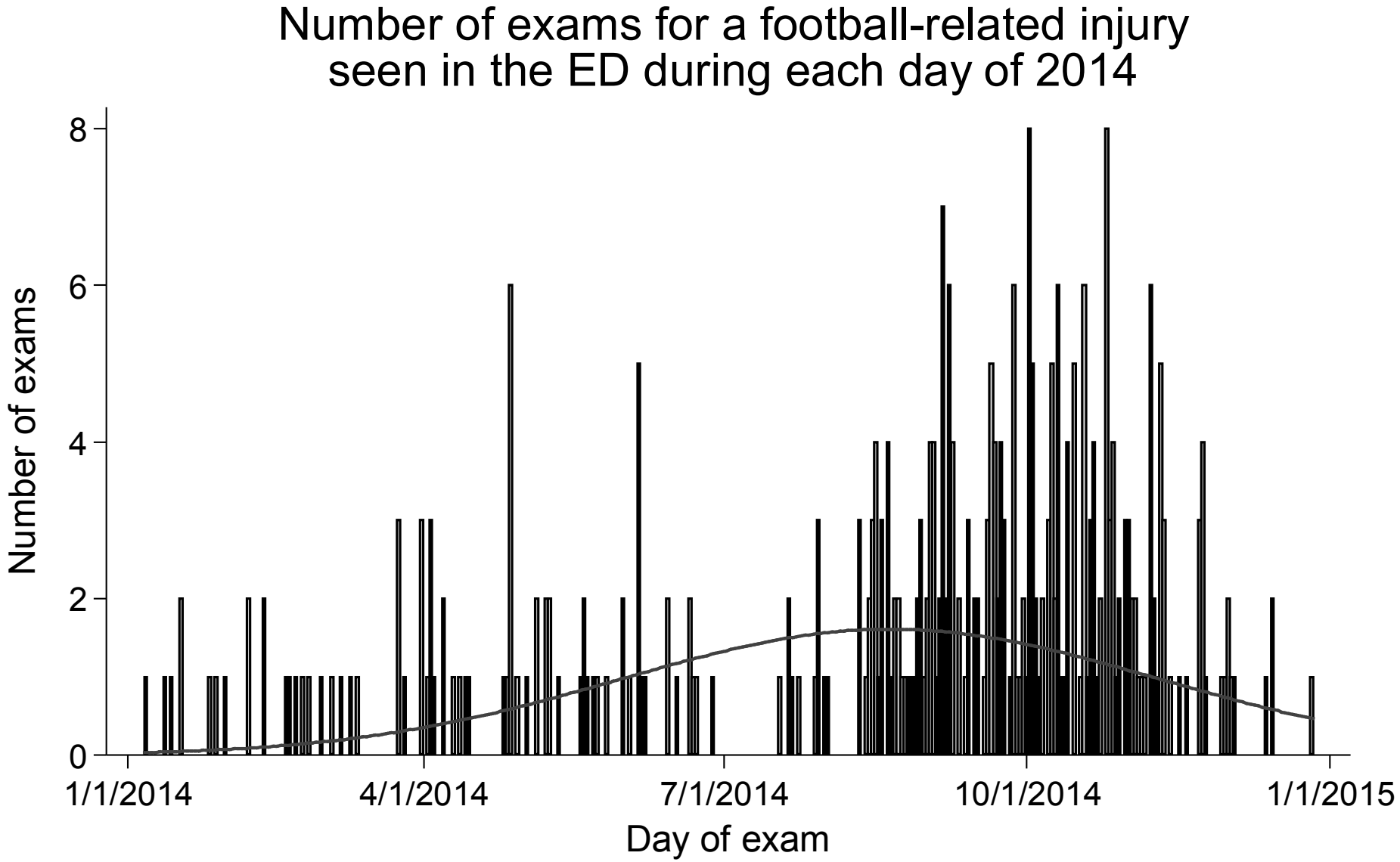
1. Histograms
2. Box plots
3. Bar graphs
4. Line graphs

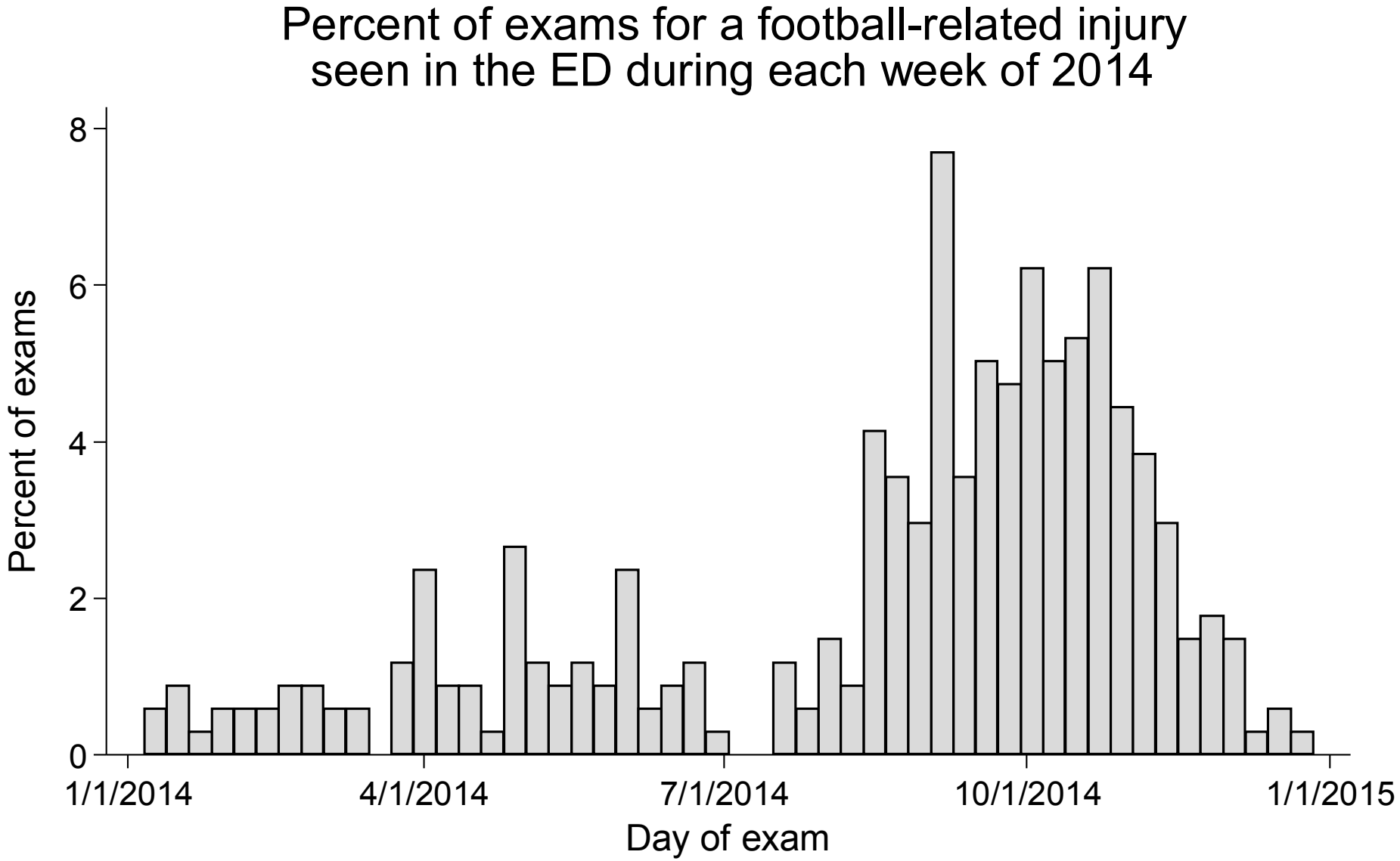
Histograms

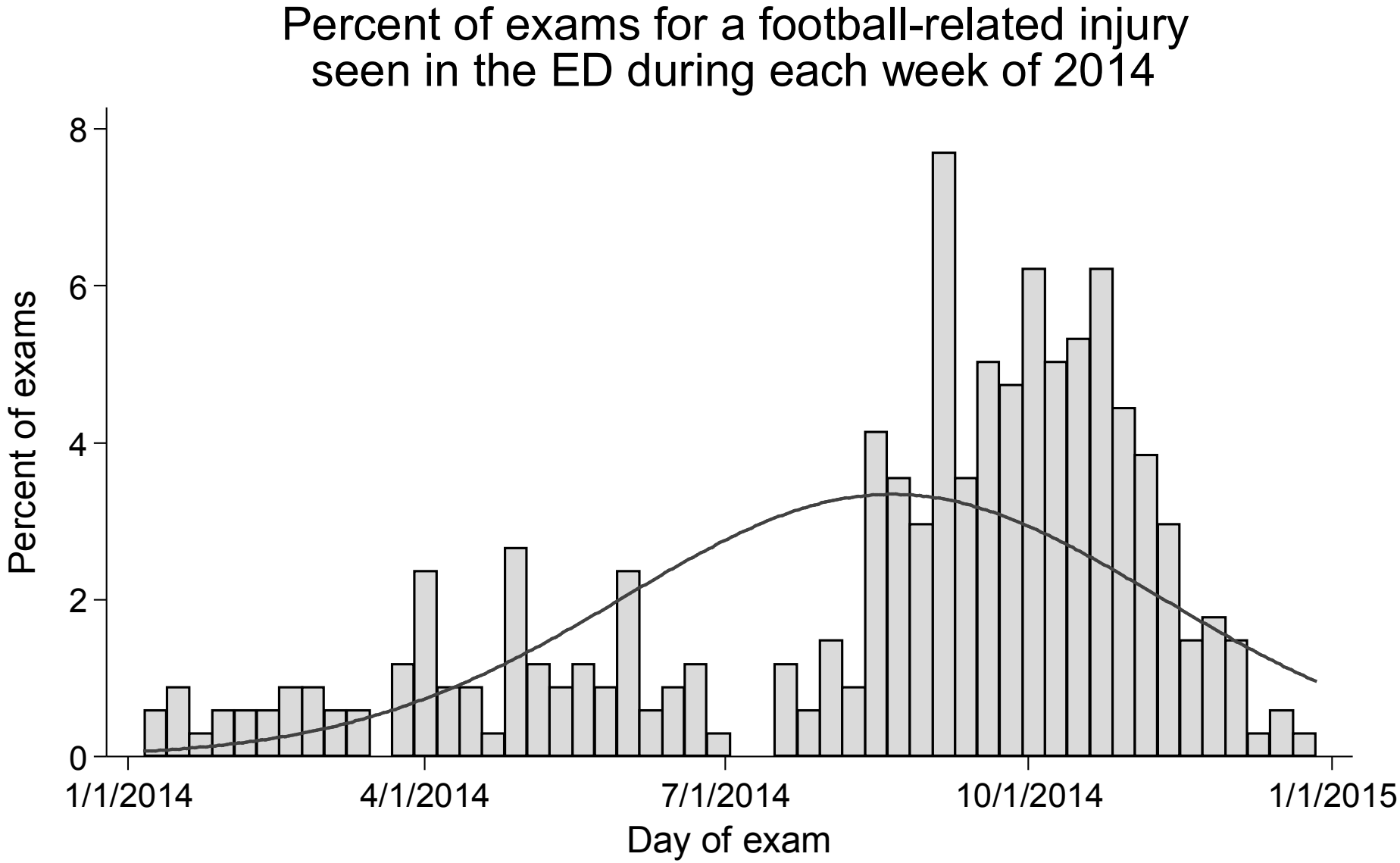
Used to represent a sampling distribution

(x-axis units of equal size)


Number of exams for a football-related injury seen in the ED during each day of 2014
Number of exams
0
2
4
6
8
1/1/2014
4/1/2014
7/1/2014
10/1/2014
1/1/2015
Day of exam


Number of exams for a football-related injury seen in the ED during each day of 2014
Number of exams
0
2
4
6
8
1/1/2014
4/1/2014
7/1/2014
10/1/2014
1/1/2015
Day of exam


Percent of exams for a football-related injury seen in the ED during each week of 2014
Percent of exams
0
2
4
6
8
1/1/2014
4/1/2014
7/1/2014
10/1/2014
1/1/2015
Day of exam

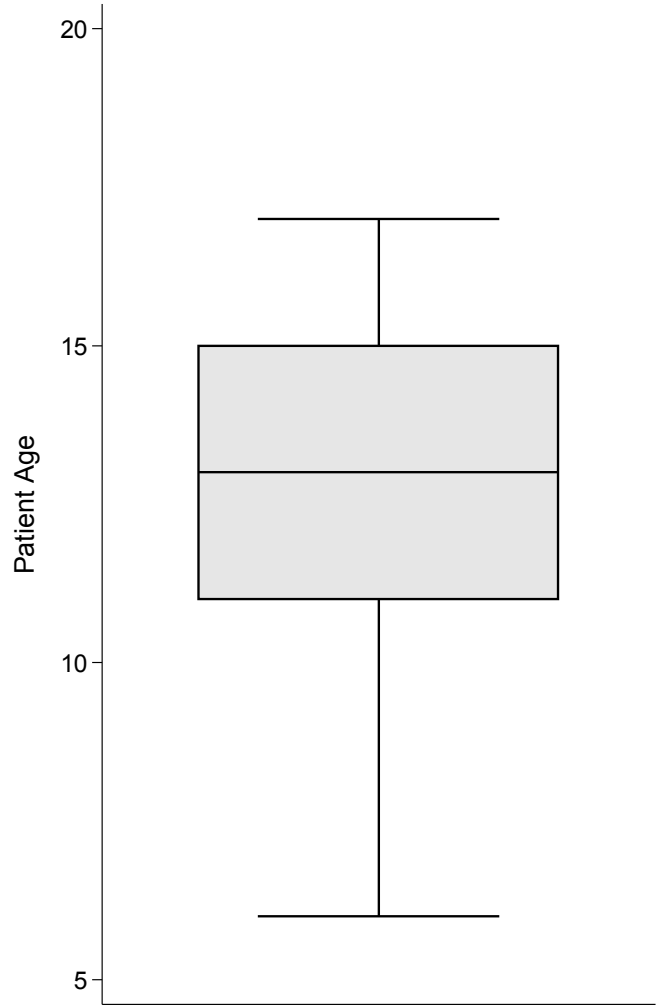

Percent of exams for a football-related injury seen in the ED during each week of 2014
Percent of exams
0
2
4
6
8
1/1/2014
4/1/2014
7/1/2014
10/1/2014
1/1/2015
Day of exam

Box plots

Used to examine spread in the data

(within or across variables)

Anatomy of a box plot



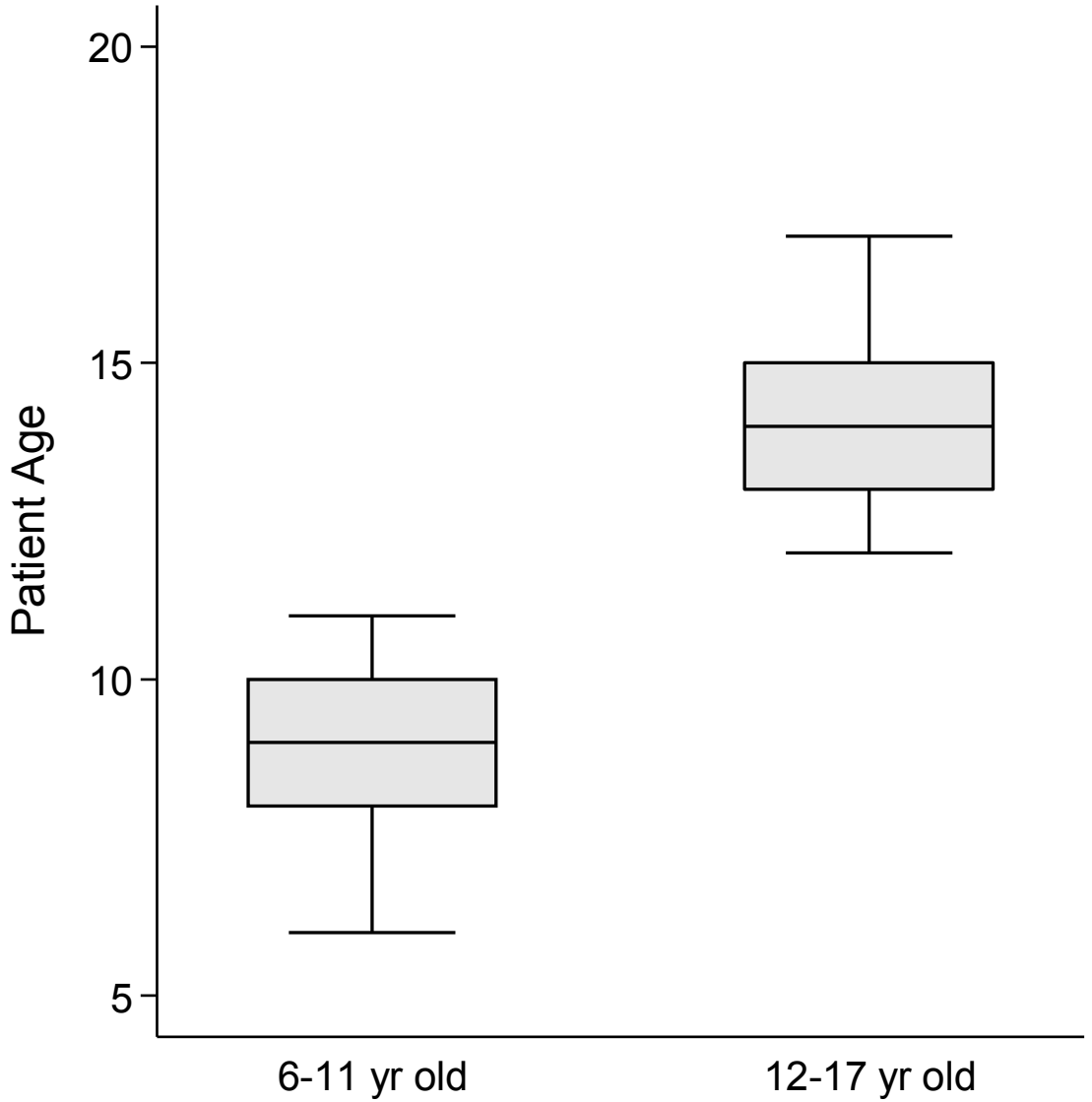
Maximum

75th percentile

50th percentile

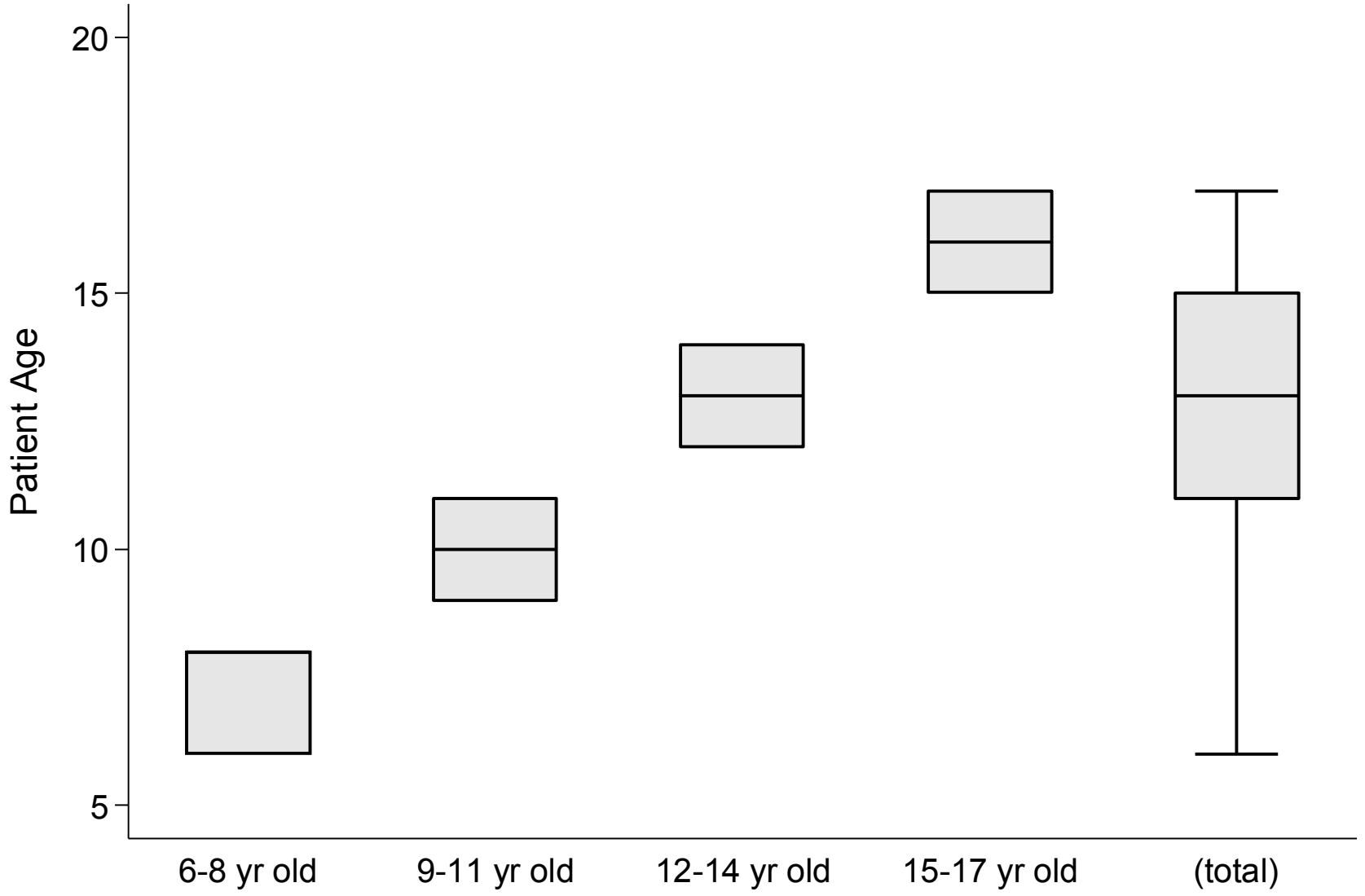
25th percentile

Minimum



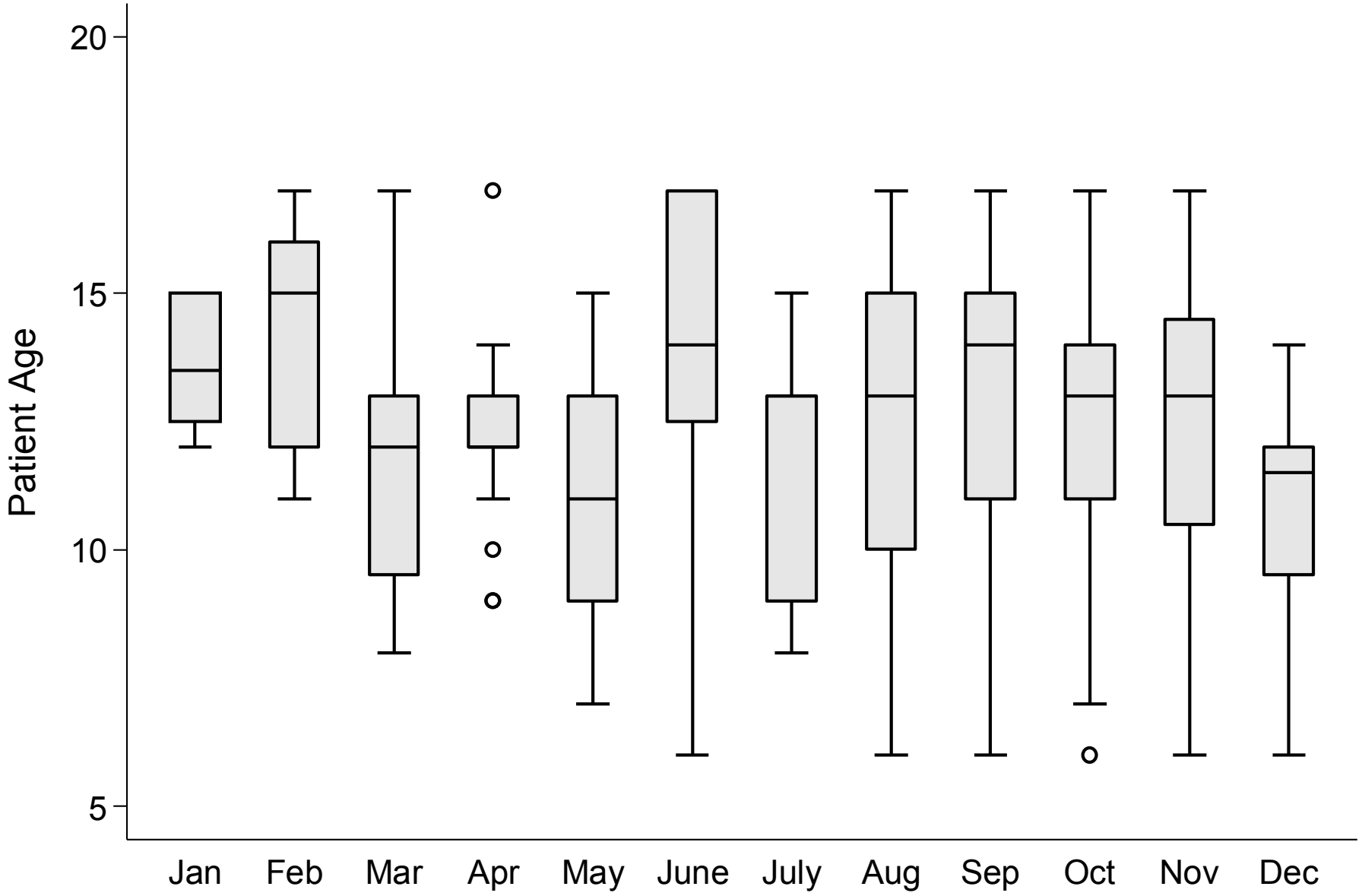
Median (IQR) = 13 (11, 15)
Range: 12-17

Median (IQR) = 9 (8, 10)
Range: 6-11



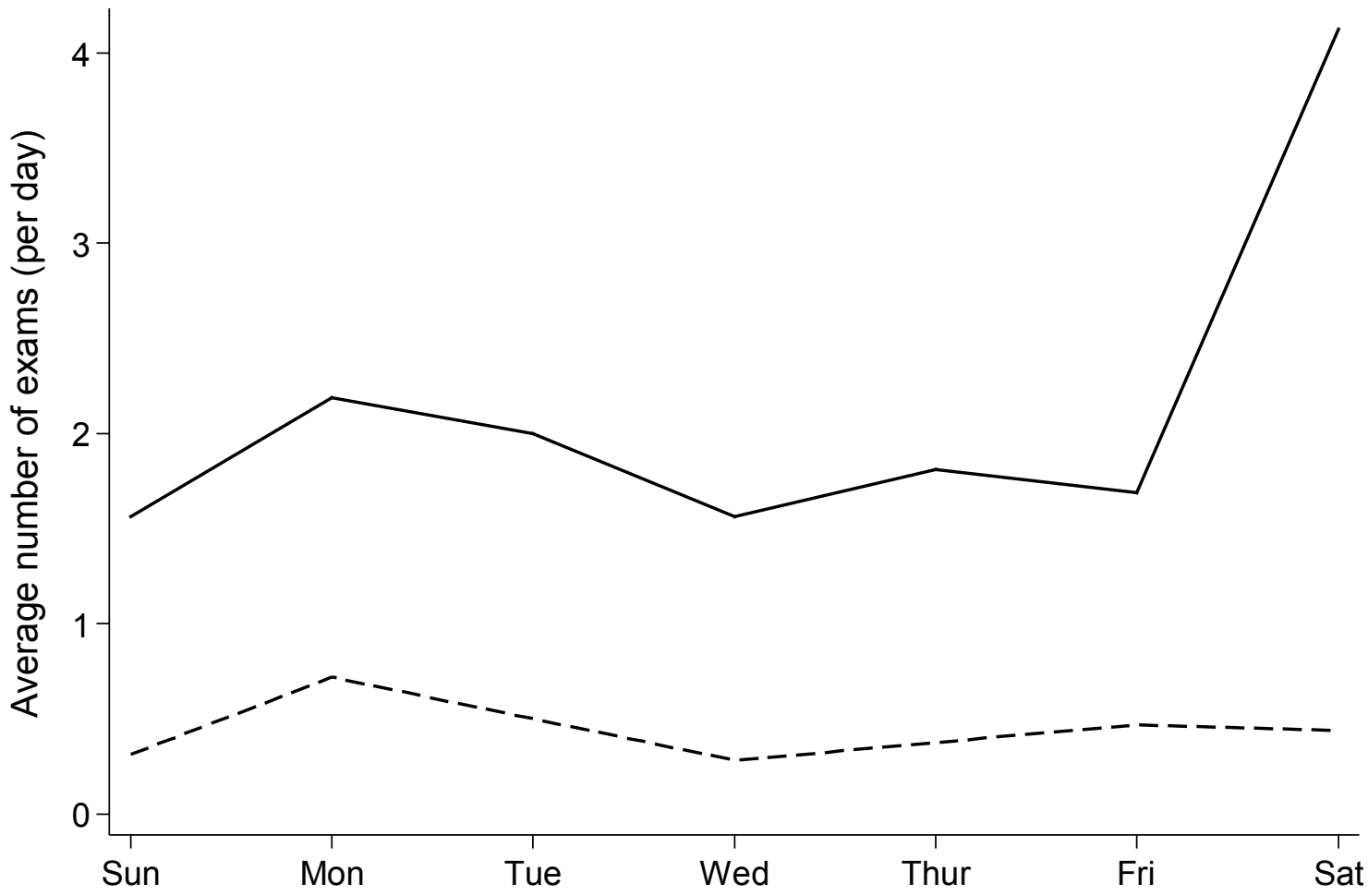
Median (IQR) = 13 (11, 15)

Range: 6-17


Patient Age
20
15
10
5
Jan
Feb
Mar
Apr
May
June
July
Aug
Sep
Oct
Nov
Dec

Line graphs

Used to show change over time

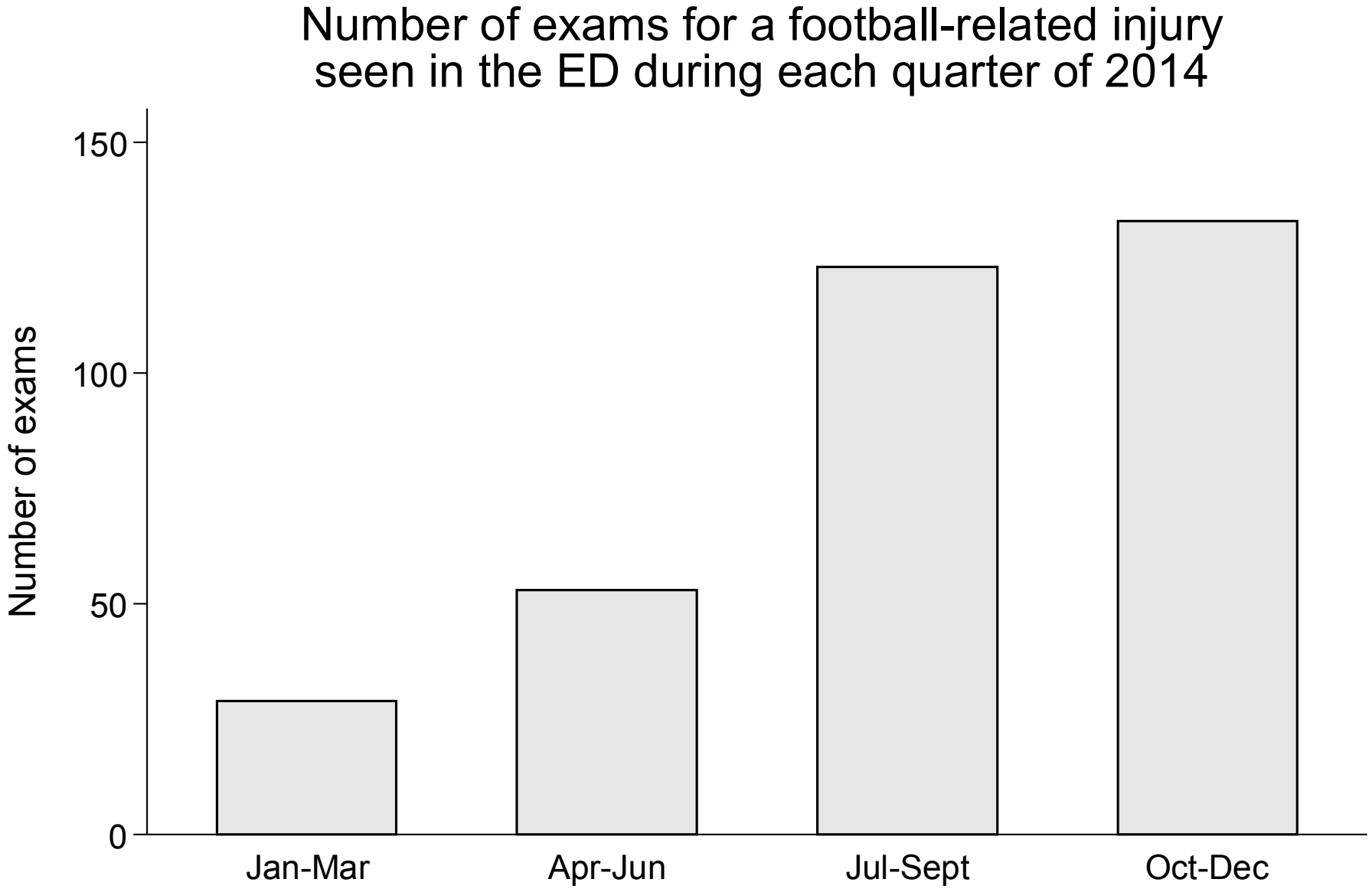


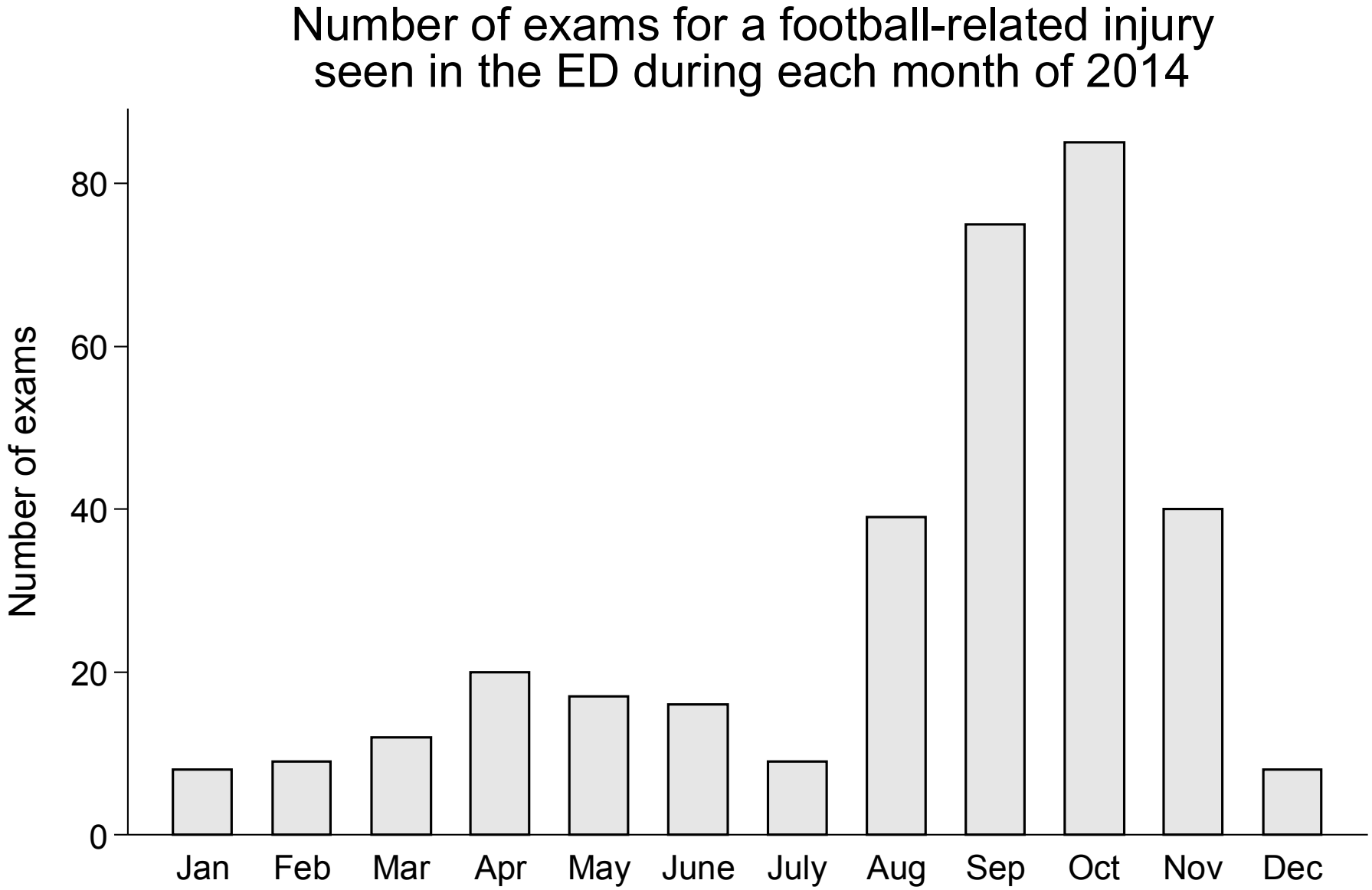
Fall football season

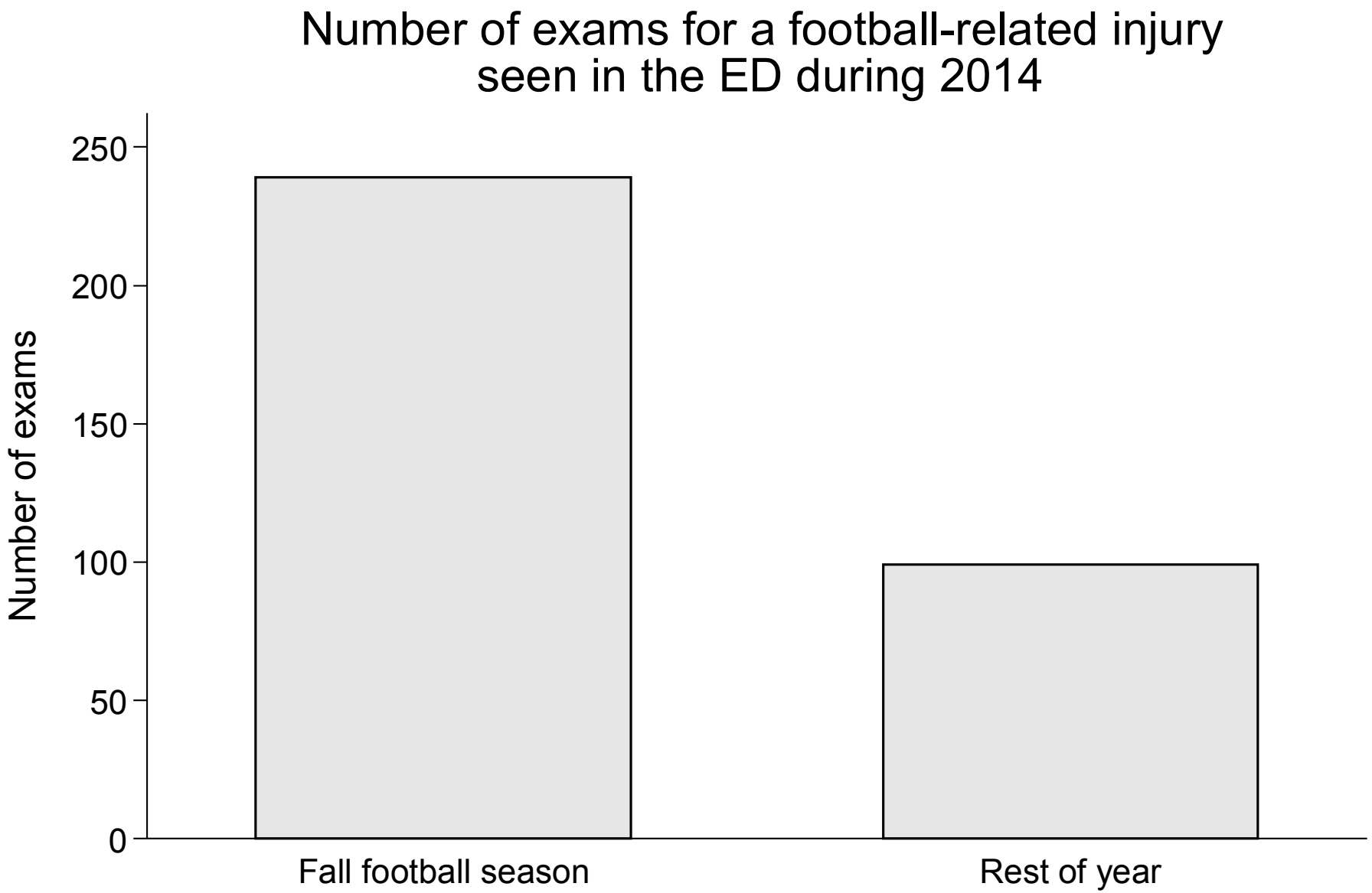
Rest of year

Bar graphs

Used to compare values across groups (within or across variables)


Number of exams for a football-related injury seen in the ED during each quarter of 2014
Number of exams
150
100
50
0
Jan-Mar
Apr-Jun
Jul-Sept
Oct-Dec


Number of exams for a football-related injury seen in the ED during each month of 2014
Number of exams
0
20
40
60
80
Jan
Feb
Mar
Apr
May
June
July
Aug
Sep
Oct
Nov
Dec


Number of exams for a football-related injury seen in the ED during 2014
Number of exams
250
200
150
100
50
0
Fall football season
Rest of year

Best practices

- Consult recent issues of your target journal for examples
- Read the instruction to authors page (requirements/restrictions)
- Decide what key messages you want to highlight

- Learn from the masters and then be as creative as you can (within limits!)