



Stata in Health Research: from everyday's clinical questions to major studies

Sofia Baptista

Faculdade de Medicina da Universidade do Porto

Unidade de Saúde Familiar Serpa Pinto



Stata and Me

UAB

**Universitat Autònoma
de Barcelona**





Statistical Software in Health Research articles

Table 2 Statistical Software Mentioned in HSR Articles, 2007-2009, by Authorship

| | <u>U.S. Authorship</u> | | <u>Non-U.S. Authorship</u> | | <u>Total</u> | |
|--|------------------------|---------|----------------------------|---------|--------------|---------|
| | Number | Percent | Number | Percent | Number | Percent |
| Total included articles | 791 | 100.0 | 86 | 100.0 | 877 | 100.0 |
| Included articles mentioning software (% of articles) | 481 | 60.8 | 54 | 62.8 | 535 | 100.0 |
| Included articles mentioning software (% distribution) | 481 | 89.9 | 54 | 10.1 | 535 | 100.0 |
| Total number of software mentions (% distribution) | 574 | 90.1 | 63 | 9.9 | 637 | 100.0 |
| Average number of software mentions per article | 1.2 | – | 1.2 | – | 1.2 | – |
| Articles in which Stata was used (% of articles)* | 238 | 49.5 | 8 | 14.8 | 246 | 46.0 |
| Articles in which SAS was used (% of articles)* | 202 | 42.0 | 26 | 48.1 | 228 | 42.6 |
| Articles in which SUDAAN was used (% of articles)* | 33 | 6.9 | 0 | 0.0 | 33 | 6.2 |
| Articles in which SPSS was used (% of articles)* | 17 | 3.5 | 14 | 25.9 | 31 | 5.8 |
| Articles in which other software was used (% of articles)* | 84 | 17.5 | 15 | 27.8 | 99 | 18.5 |

* Note: percentages add up to more than 100% because some articles mentioned the use of more than one statistical software application.



Stata – Advantages for Health Research

- Better oriented for health research
- Robust, versatile and easy software
- User – written commands
- Competitive price



Stata and everyday's clinical questions





Stata and everyday's clinical questions

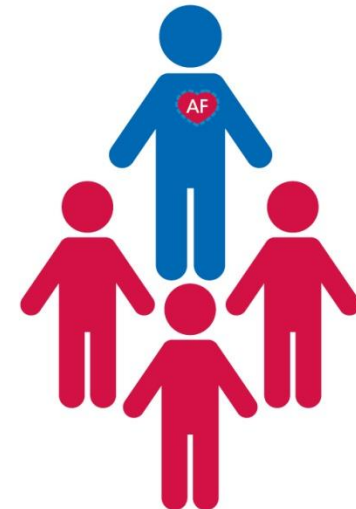
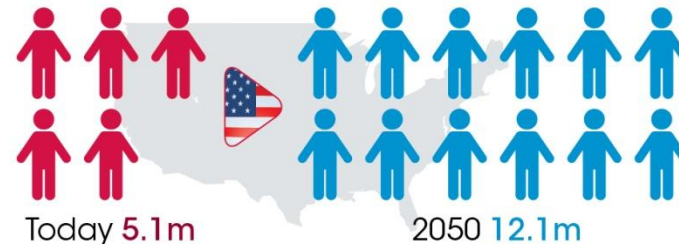
- Digital medical records / Easy to reach databases
- Assessing clinical practice status
- Attaining clinical goals / guidelines oriented practice
- Implementing changes



Stata and everyday's clinical questions: an example



In the **US** the prevalence is projected to be more than **DOUBLED** by **2050**



ONE IN **FOUR**
ADULTS AGED OVER **40** DEVELOPS
AF IN THEIR LIFETIME

~2% GENERAL POPULATION AFFECTED BY AF
~140,000,000 WORLDWIDE

References:

Lloyd-Jones DM, Wang TJ, Leip EP *et al.* *Circulation* 2004; 110:1042-6. Stewart S, Murphy N, Walker A, *et al.* *Heart* 2004; 90:286-92. Miyasaka Y, *et al.* *Circulation* 2006; 114:1119-125. Fuster V, Rydn LE, Cannom DS, *et al.* *Circulation* 2006; 114:e257-e354. Marini C, *et al.* *Stroke* 2005;36:1115-1119. Camm AJ, *et al.* *European Heart Journal.* 2012;33, 2719-2747. United Nations (2011) Available at: <http://www.un.org/apps/news/story.asp?NewsID=40257#.Ul17BrJITue> (Last accessed Oct 2012)



Stata and everyday's clinical questions: an example

- Atrial fibrillation is a cardiac arrhythmia which increases risk of stroke, death and serious impairment.
- The decision to treat with anticoagulant drugs should be taken according to a score (CHA₂DS₂VASC).



Stata and everyday's clinical questions: an example

- 1) How many of my patients suffer from atrial fibrillation?
- 2) What is the distribution of the score CHA2DS2VASC?
- 3) Am I prescribing treatment according to their score?



Stata and everyday's clinical questions: an example

- **1) How many of my patients have this diagnosis?**
- Is my sample roughly representative of general population?
- Command **cii wald**
- Prevalence estimation = 1,23%

| Variable | Obs | Mean | Std. Err. | — Binomial Exact — [95% Conf. Interval] | |
|----------|-------|----------|-----------|--|----------|
| | 22350 | .0123043 | .0007374 | .0109003 | .0138371 |



Stata and everyday's clinical questions: an example

- **Summarize**

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|----------|-----|----------|-----------|-----|-----|
| age | 275 | 75.69818 | 11.65 | 34 | 101 |

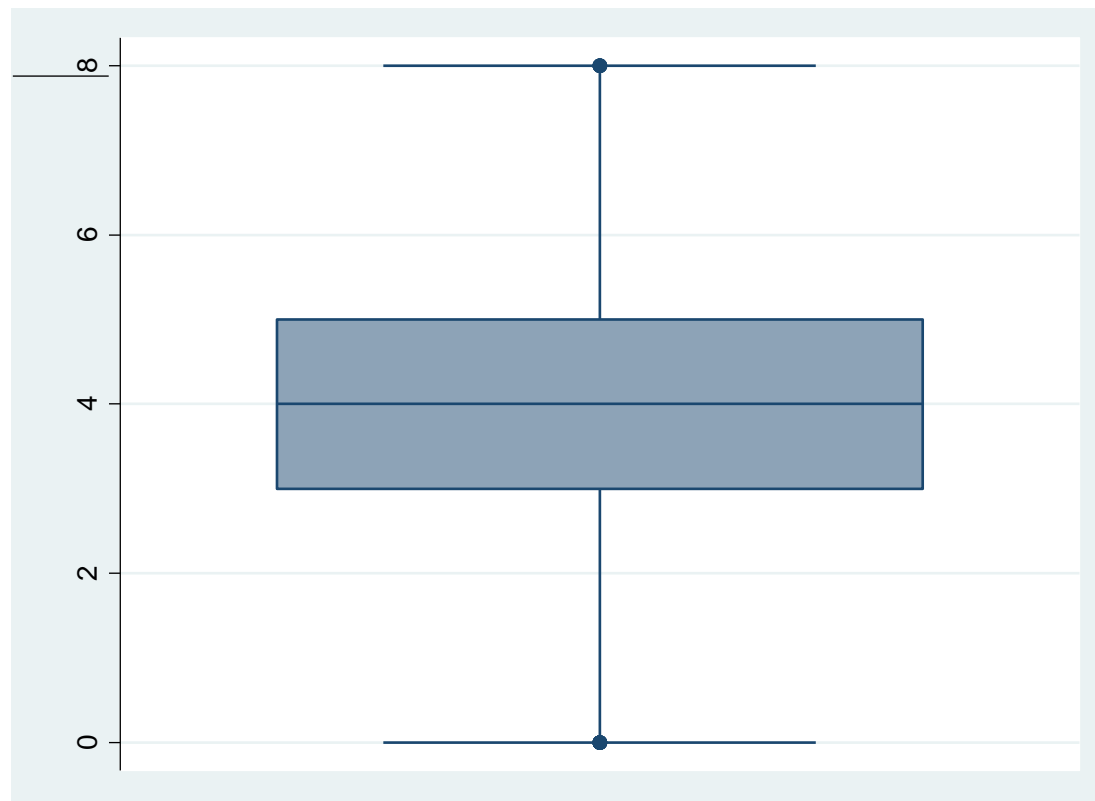
- **Tabulate**

| sex | Freq. | Percent | Cum. |
|-------|-------|---------|--------|
| 0 | 104 | 37.82 | 37.82 |
| 1 | 171 | 62.18 | 100.00 |
| Total | 275 | 100.00 | |



Stata and everyday's clinical questions: an example

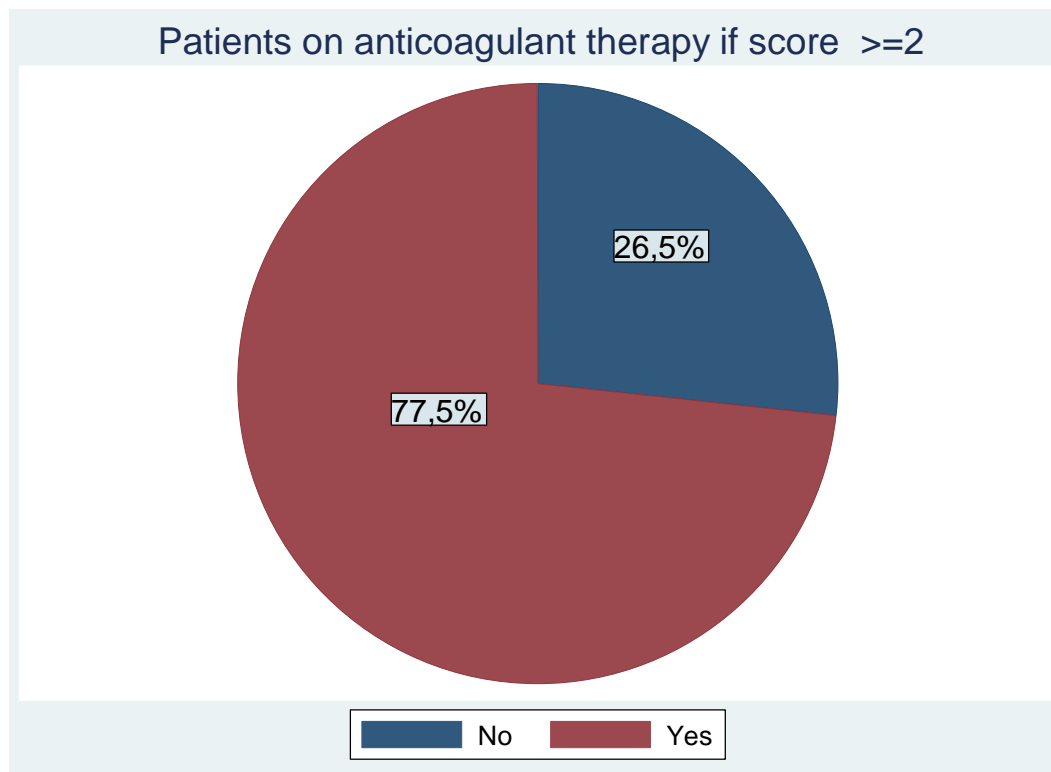
- 2) What is the distribution of the CHA2DS2VASC score?
- graph box





Stata and everyday's clinical questions: an example

- **3) Am I prescribing treatment according to patients' score?**
- **graph pie if CHA2DS2Vasc \geq 2, over(anticoagulation)**





Stata and everyday's clinical questions: an example

67% AF with anticoagulant if (GARFIELD, 2013)

77,5% AF with anticoagulant if CHA2DS-Vasc ≥ 2 (my practice)

- 3 clinical questions / few commands
- Fast intuitive process
- Syntax applicable to another dataset



Other Stata commands

mvmeta, networkplot

