**Report elaborazioni statistiche richieste dalla dott.ssa Mesina il 20.05.2015**

**( a cura di Ugo Lancia)**

La richiesta della dott.ssa Mesina nella mail del 20 maggio 2015 è stata quella di valutare la significatività statistica per una serie di dati allegati in un file excel dal titolo Rheumanagement grafici e tabelle.xlsx, allegato alla mail.

**Descrizione del materiale fornito e commento generale**

Nel file sono presenti delle tabelle e dei grafici derivati dalle stesse.

**Punto 1**. La prima tabella fornita è la seguente:

|  |  |
| --- | --- |
| **tab. 1** |  |
| descrizione casistica( 579 paz) | |
| % patologia | |
| AR (n) | 52,8 (306) |
| SpA (n) | 47,2 (273) |
| % genere | |
| M (n) | 37 (214) |
| F (n) | 63 (365) |
| % terapia | |
| bDMARDs (n) | 46,5% (269) |
| sDMARDs (n) | 53,5% (310) |
| % eventi (129 paz) | |
| evento si | 22,3% |
| evento no | 77,7% |

Si tratta di una tabella descrittiva di una casistica di 579 pazienti.

**Punto 2.** Il grafico 1 corrisponde ad una tabella (di seguito) che necessita di una trasformazione in tabella di contingenza 2x2.

|  |  |  |
| --- | --- | --- |
| divisione di genere per patologia | | |
|  | M % | F % |
| AR | 23 | 77 |
| SpA | 45 | 55 |

**Punto 3**. La tabella 2 (di seguito) fornisce una suddivisione dei pazienti con eventi per patologia e di seguito secondo due categorie di trattamento.

|  |  |
| --- | --- |
| **tab. 2** |  |
| descrizione casistica eventi (129 paz) | |
| AR | 56,6 (73) |
| SpA | 43,4 (57) |
| bDMARDs | 11,6% (15) |
| sDMARDs | 88,4% (114) |

Per poter procedere ad un’analisi bisogna suddividere la tabella fornita in due tabelle 2 x 2 integrando i dati della tabella 2 con quelli generali della casistica suddivisi secondo la presenza o meno dell’evento cardiovascolare.

**Punto 4**. Il grafico 3 suddivide i pazienti in terapia con biologici e che hanno avuto l’evento cardiovascolare in due sottogruppi secondo il tipo di farmaco biologico. Si può creare una tabella 2x2 ed effettuare un’analisi chisquare.

|  |  |
| --- | --- |
| **grafico 3** |  |
| eventi in paz a bDMARDs (5,6%) | |
| ETN (8/146) | 5,5% |
| ADA (7/82) | 8,5% |

**Punto 5.** Il grafico 4 fa riferimento alla suddivisione per sesso dei pazienti con diagnosi di SpA in percentuali. Per poter valutare l’applicabilità di un test bisogna prima risalire ai conteggi dei pazienti.

|  |  |  |
| --- | --- | --- |
| **grafico 4** |  |  |
| divisione di genere per eventi | | |
|  | M % | F % |
| SpA | 63 | 37 |
| SpA in bDMARDs | 83,3 | 16,7 |

**Punto 6**. I dati del grafico 5 si riferiscono ad una suddivisione dei pazienti di cui si hanno dati relativi alla presenza o meno di ipertensione. Mancano informazioni relativi alla distribuzione dei dati mancanti e sulla loro interpretazione e i dati relativi alla presenza di ipertensione nel gruppo di pazienti senza evento cardiovascolare. Si può procedere alla ricerca di questi dati in altro materiale fornito in precedenza e successivamente valutare l’applicabilità di procedure statistiche.

|  |  |
| --- | --- |
| **grafico 5** |  |
| ipertensione | |
| totale (249) | 43% |
| AR | 40% |
| SpA | 46% |
| con eventi |  |
| totale (59) | 45,7% |
| AR | 39,7% |
| SpA | 52,6% |
| nei paz ipertesi |  |
| con eventi | 23,7% |
| senza eventi | 76,3% |

**Punto 7**. La tabella 3 descrive aspetti relativi ai pazienti affetti da eventi suddivisi in gruppo AR e gruppo SpA. Dopo l’ispezione descrittiva si può valutare se procedere o meno alla produzione di tabelle utilizzabili per l’analisi statistica.

|  |  |  |
| --- | --- | --- |
| **tab 3** |  |  |
| descrizione casistica eventi per patologia | | |
|  | AR | SpA |
| eventi | 23,8% (73) | 20,9% (57) |
| età media all'evento | 68,5 | 62,1 |
| durata di malattia media all'evento | 11,4 | 9,3 |
| sDMARDs | 86% (64) | 89,5% (51) |
| bDMARDs | 14% (9) | 10,5% (6) |
| tempo a bDMARDs | 3,8 | 2 |

**Punto 8.** Il grafico 8 fa riferimento alla suddivisione dei pazienti con eventi (129) secondo la tipologia di evento (C o V) e il gruppo di appartenenza (AR o SpA). Si può derivare una tabella 2x2 su cui applicare un test del chi quadrato.

|  |  |  |  |
| --- | --- | --- | --- |
| **grafico 6** |  |  |  |
| distribuzione tipologia eventi | | | |
|  | totali | AR | SpA |
| eventi C (83) | 56% | 56,6% | 43,4% |
| eventi V (65) | 44% | 55,4% | 44,6% |

**RISULTATI:**

**Punto 1**. Si tratta di una tabella descrittiva generale non adatta di per sè per l’analisi statistica; la tabella comunque fornisce dati numerici utili per integrare altre tabelle e renderle analizzabili.

**Punto 2**.

|  |  |  |  |
| --- | --- | --- | --- |
| divisione di genere per patologia | | |  |
|  | M | F | Totale |
| AR | 70 | 236 | 306 |
| SpA | 123 | 150 | 273 |
| Totale | 193 | 386 | 579 |

L’analisi statistica chisquared è risultata altamente significativa: sesso e diagnosi risultano fortemente associati. Una rappresentazione visuale è fornita di seguito all’output dell’analisi statistica.

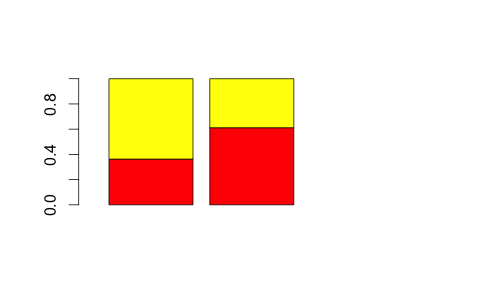
Pearson's Chi-squared test with Yates' continuity

correction

data: mymatr0

X-squared = 30.948, df = 1, p-value = 2.651e-08

Barplot standardizzato con presenza di sesso M (rosso) su totale, nel gruppo AR (prima barra) e nel gruppo SpA (seconda barra).



**Punto 3.**

Si procede alla suddivisione della tabella 2 in due differenti tabelle.

Prima tabella derivata dalla tabella 2.

|  |  |  |  |
| --- | --- | --- | --- |
| patologia | evento | Non evento | Totale |
| AR | 73 | 233 | 306 |
| SpA | 57 | 216 | 273 |
| Totale | 130 | 449 | 579 |

Risulta esserci una difformità nel numero di eventi riportato (129) e la somma degli eventi presenti nei due gruppi. Si procede comunque all’analisi utilizzando la somma di 130 eventi.

Il test di indipendenza chisquared ha evidenziato l’assenza di significatività statistica (vedi output di seguito) che comunque si confermerebbe anche considerando 129 eventi.

Pearson's Chi-squared test with Yates' continuity

correction

data: mymatr

X-squared = 0.57341, df = 1, p-value = 0.4489

Seconda tabella derivata dalla tabella 2.

|  |  |  |  |
| --- | --- | --- | --- |
| trattamento | evento | Non evento | Totale |
| bDMARD’s | 15 | 254 | 269 |
| sDMARD’s | 114 | 196 | 310 |
| Totale | 129 | 450 | 579 |

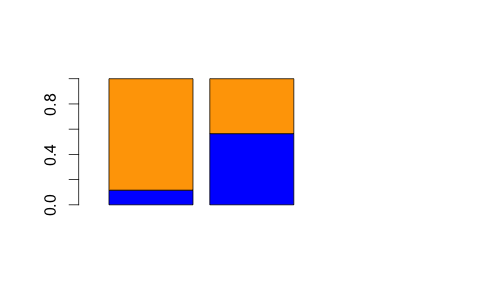
Pearson's Chi-squared test with Yates' continuity

correction

data: mymatr2

X-squared = 79.163, df = 1, p-value < 2.2e-16

Risultato altamente significativo, come si evince dall’output, evidenziato in forma grafica di seguito come barplot standardizzato (presenza di eventi cardiovascolari in blu, nei due gruppi, trattato con farmaci biologici il primo e con farmaci tradizionali il secondo).



**Punto 4.**

Tabella derivata dai dati del grafico 3.

|  |  |  |  |
| --- | --- | --- | --- |
| Tipo di farmaco biologico | evento | Non evento | Totale |
| ETN | 8 | 146 | 154 |
| ADA | 7 | 82 | 89 |
| Totale | 15 | 228 | 243 |

Anche se vi è una differenza percentuale tra gli eventi nei due gruppi non vi è alcuna significatività.

Pearson's Chi-squared test with Yates' continuity

correction

data: mymatr3

X-squared = 0.3099, df = 1, p-value = 0.5777

**Punto 5.**

La prima fase è stata quella di risalire ai conteggi:

pazienti di sesso M e F con SpA: 273\*0.63 e 273\*0.37, 172 e 101 rispettivamente

La suddivisione dei conteggi dei pazienti con SpA con eventi si desume dalla tabella 3 che si tratta di 6 pazienti in trattamento con biologici e affetti da SpA ma non vi è la suddivisione per sesso.

Essendo comunque il numero sicuramente molto basso possiamo affermare con certezza la non applicabilità di qualsivoglia test statistico.

**Punto 6.**

L’analisi della tabella sembra dire che i dati disponibili sulla presenza di ipertensione sono circa il 43 % del campione totale e, limitatamente ai pazienti con eventi circa il 46 %. Un dato che si desume dalla tabella è che la presenza di ipertensione nei pazienti con eventi è del 23.7 % (14 su 59). Non essendo indicata nella tabella il numero di pazienti con ipertensione nei pazienti senza eventi non è possibile applicare una statistica.

Un altro problema significativo è la quantità di dati mancanti (superiore al 50 %) e la mancanza di indicazioni sulla random missingness.

**Punto 7.** La possibilità di applicare test statistici per proporzioni o test chisquare richiederebbe la formazione di 10 diverse tabelle di contingenza, ma la semplice ispezione descrittiva consente di affermare con sicurezza che non vi sarebbero valori significativi.

**Punto 8.** Il grafico 8 fa riferimento alla suddivisione dei pazienti con eventi (129) secondo la tipologia di evento (C o V) e il gruppo di appartenenza (AR o SpA).

**Tabella derivata dal grafico 6.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gruppo AR | Gruppo SpA | Totale |
| Eventi C | 46 | 37 | 83 |
| Eventi V | 29 | 36 | 65 |
| Totale | 75 | 54 | 129 |

L’analisi ha mostrato l’assenza di significatività statistica.

Pearson's Chi-squared test with Yates' continuity

correction

data: mymatr4

X-squared = 1.2981, df = 1, p-value = 0.2546

**CONCLUSIONI**

Di tutte le analisi fattibili con i dati disponibili e con le numerosità relative, due hanno evidenziato una significatività statistica:

* l’associazione tra sesso e diagnosi (AR o SpA): sesso femminile significativamente più rappresentato nei pazienti affetti da AR rispetto ai pazienti affetti da SpA.
* la minore presenza di eventi cardiovascolari nei pazienti in trattamento con farmaci biologici rispetto ai pazienti in trattamento con farmaci tradizionali.

Per completezza si segnala che non è stato possibile effettuare una delle analisi (pazienti affetti da ipertensione), sia per la presenza di dati mancanti in più del 50 % del campione, sia (pure assumendo la random missingness) per la mancanza del dato sul numero di pazienti affetti da ipertensione nel gruppo senza eventi cardiovascolari.