* 流行病學原理：資料分析 bias (systemic error)

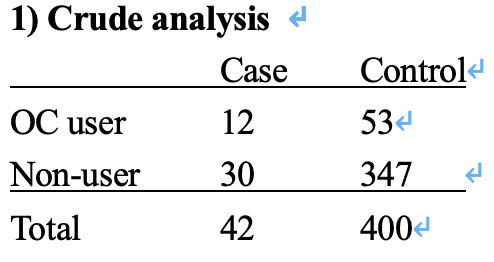
流行病學資料分析 – SAS

Bias (Systemic error)

學號：b07401048

系級：醫學五

姓名：賴柏瑞

* 1. Crude analysis :
     1. Creating data
        1. Data to create
        2. Code：

*/\* creating data \*/*

**data** crude;

input oc$ group$ count;

datalines;

user case 12

user control 53

non\_user case 30

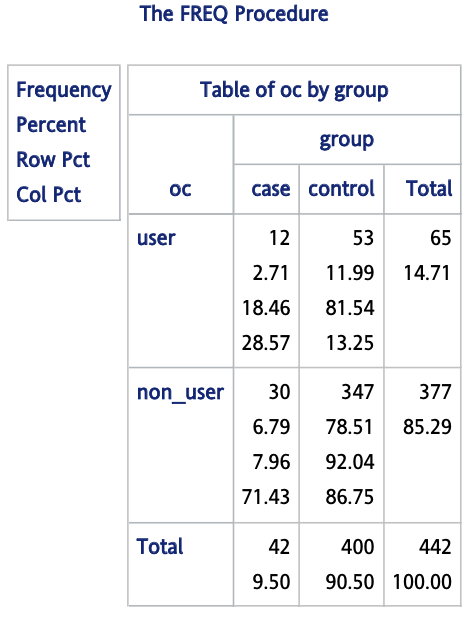
non\_user control 347

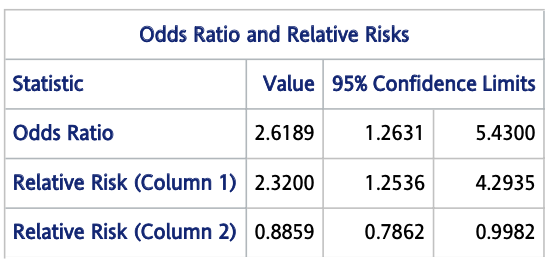
;

**run**;

**proc** **print** data = crude;

**run**;

* + - 1. Data created：
    1. Crude odd ratio
       1. Result：
          1. Odds ratio為2.6189
          2. 95% CI ：1.2631-5.4300，達統計上顯著。
       2. Figure：



* + - 1. Description：
         1. 分為case組（發生venous thrombosis）與control組（沒有發生venous thrombosis）。比較兩組曾服用Oral-contraceptive（exposure）的比例。
         2. 將case組與control組服用OC的比例相除為odd ratio。Odd ratio為2.6189，95% CI：1.2631-5.4300，達統計上顯著。
         3. 也就是說，發生venous thrombosis的病人暴露於OC的比例顯著高於沒有發生venous thrombosis的人，且為2.6189倍。
         4. 在sampling rate independent of exposure的前提下，exposure odd ratio等於risk ratio。也就是說，在抽樣與暴露相互獨立的前提下，服用oral-contraceptive的人與沒有服用oral-contraceptive的人發生venous thrombosis的risk ratio是2.6189。
         5. 此odd ratio沒有經過分層處理，故稱為crude analysis。
      2. Code：

*/\* q1 : crude analysis \*/*

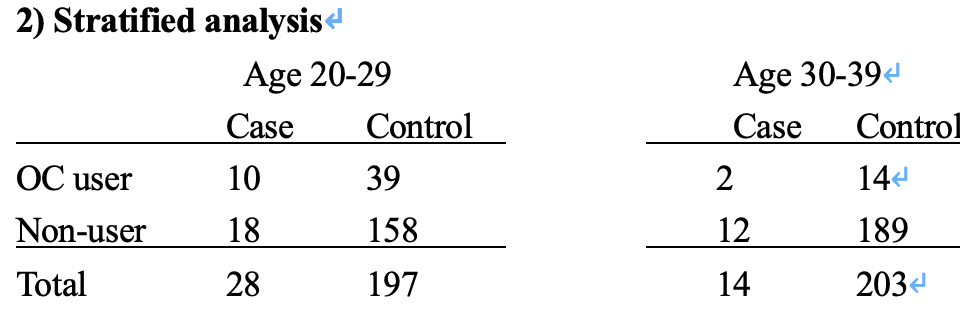
title "q1 crude analysis";

**proc** **freq** data = crude order = data;

tables oc \* group / chisq or;

weight count;

**run**;

* 1. Stratified analysis：
     1. Creating data
        1. Data to create
        2. Code：

*/\* q2 : stratified analysis \*/*

title "q2 stratified analysis";

*/\* creating data \*/*

**data** stratified;

input age$ oc$ group$ count;

datalines;

20\_29 user case 10

20\_29 user control 39

20\_29 non\_user case 18

20\_29 non\_user control 158

30\_39 user case 2

30\_39 user control 14

30\_39 non\_user case 12

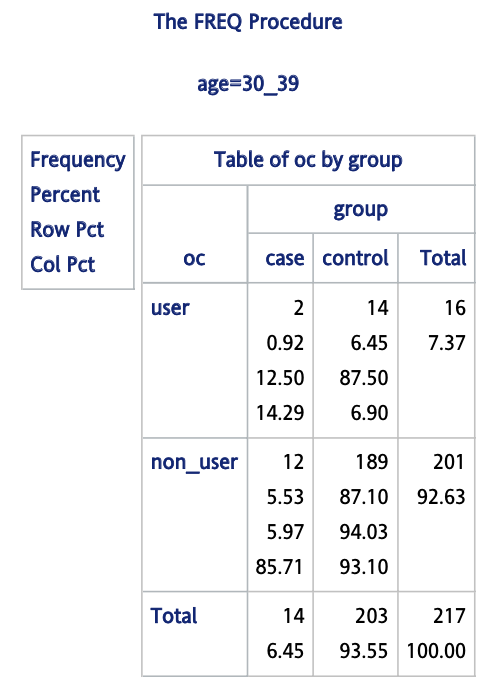
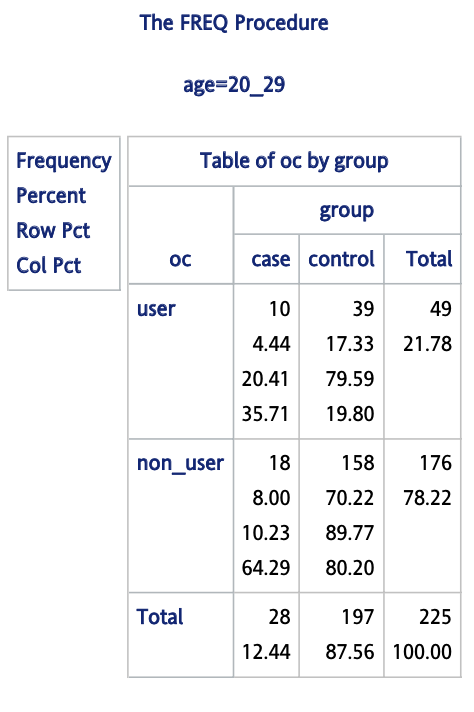
30\_39 non\_user control 189

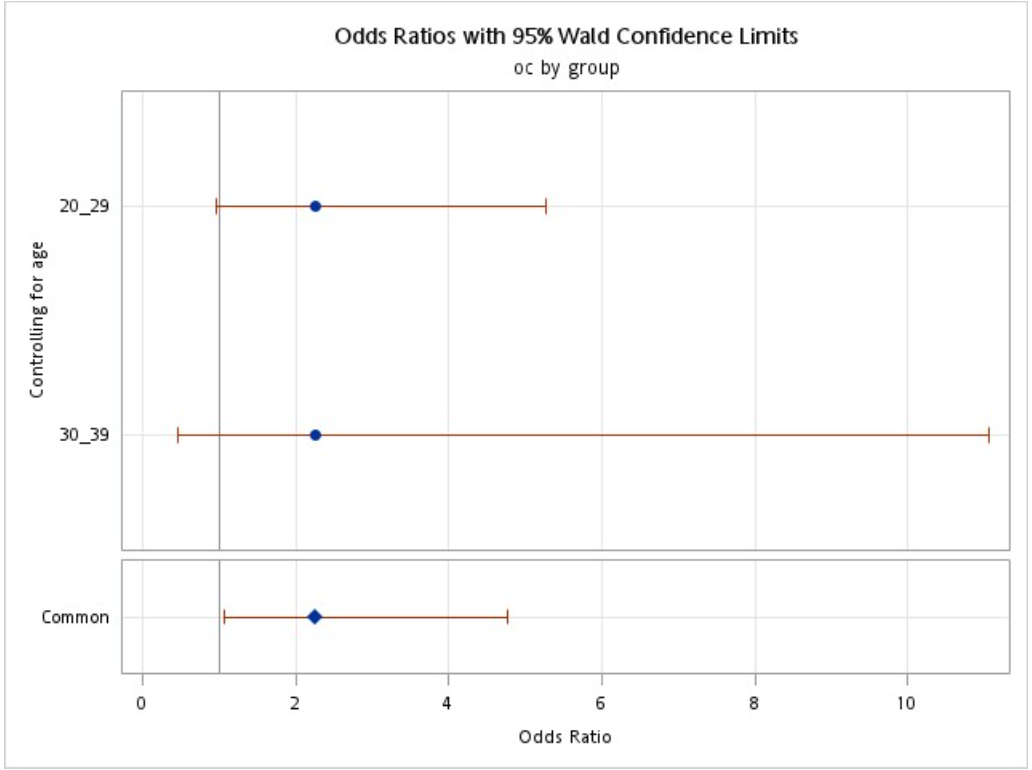
;

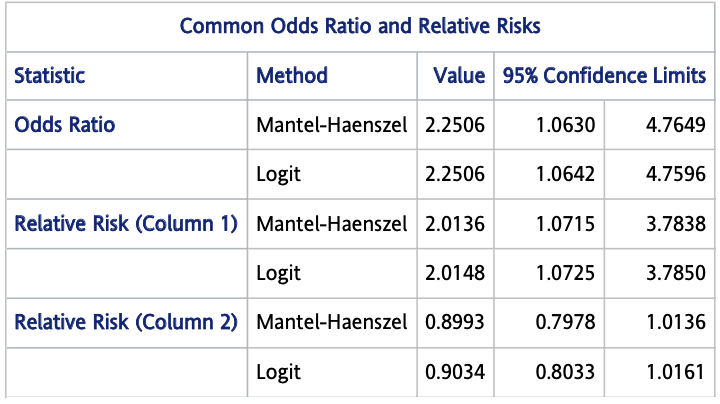
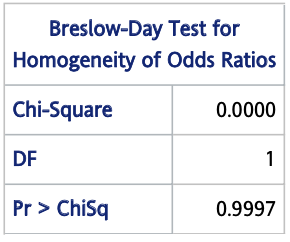
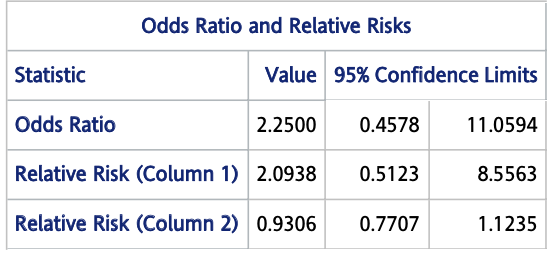
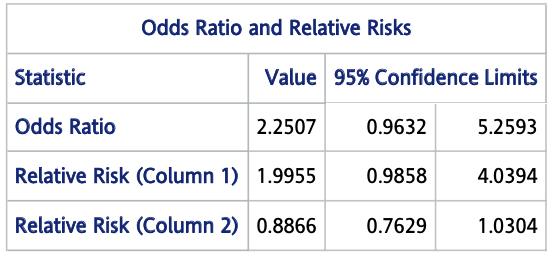
**run**;

**proc** **print** data = stratified;

**run**;

* + - 1. Data created：
    1. Stratum-specific-ORs
       1. Result：
          1. 20-29歲組的odd ratio為2.2507，95% CI：0.9632 – 5.2592，未達統計上顯著。
          2. 30-39歲組的odd ratio為2.2500，95% CI：0.4578 – 11.0584，未達統計上顯著。
          3. 以Breslow-Day test檢定兩stratum-specific odd ratio是否相等。檢定結果p-value為0.9997，兩stratum-specific odd ratio無顯著差異。
          4. 因此以Mantel-Haenszel method計算pooled odd ratio。Common odd ratio為2.2506，95% CI：1.0630 – 4.7649，達統計上顯著。
       2. Figures：





* + - 1. Code：

*/\* stratum-specific odd ratio \*/*

title "q2-a";

**proc** **freq** data = stratified order = data;

tables oc \* group / chisq or cmh;

weight count;

by age;

**run**;

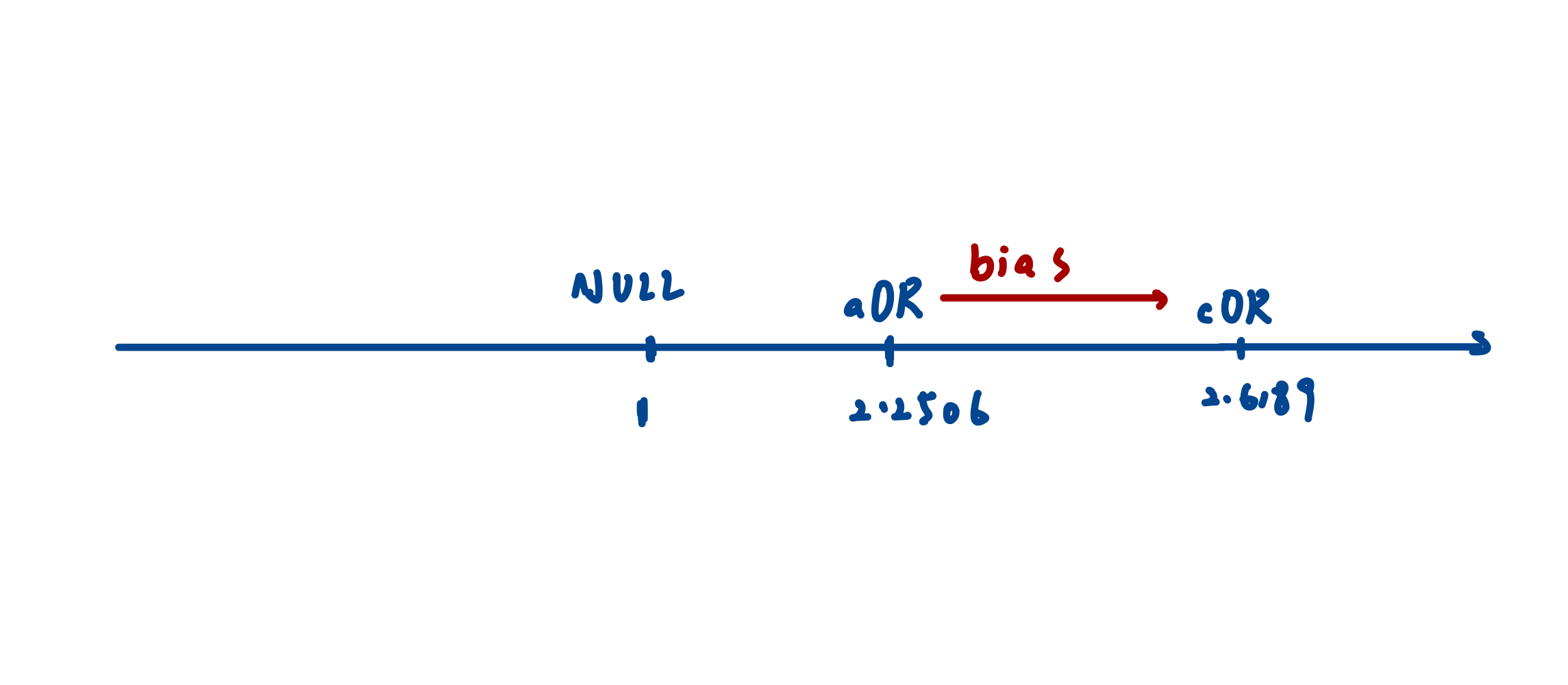
**proc** **freq** data = stratified order = data;

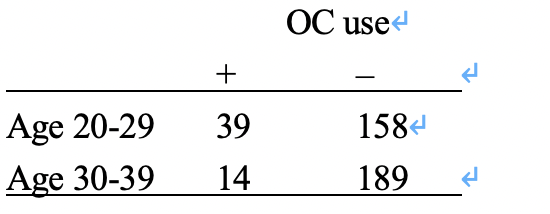
tables age \* oc \* group / chisq or cmh ;

weight count;

**run**;

* + 1. Comparing stratum-specific ORs and crude ORs
       1. 如第一題之結論，crude odd ratio為 2.6189
       2. 如上題之結論，adjusted odd ratio為2.2506。
       3. 方向：away from NULL，導致overestimate the effect of exposure。



* + 1. Whether age is a confounder for the association between OC use and venous thrombosis？
       1. Confounding effect：
       2. 以10% rule判斷，則age符合confounding的定義。
  1. OC use in source population among controls ：
     1. Creating data
        1. Data to create
        2. Code：

*/\* creating data \*/*

**data** cf\_e;

input age$ oc$ count;

datalines;

20\_29 user 39

20\_29 non\_user 158

30\_39 user 14

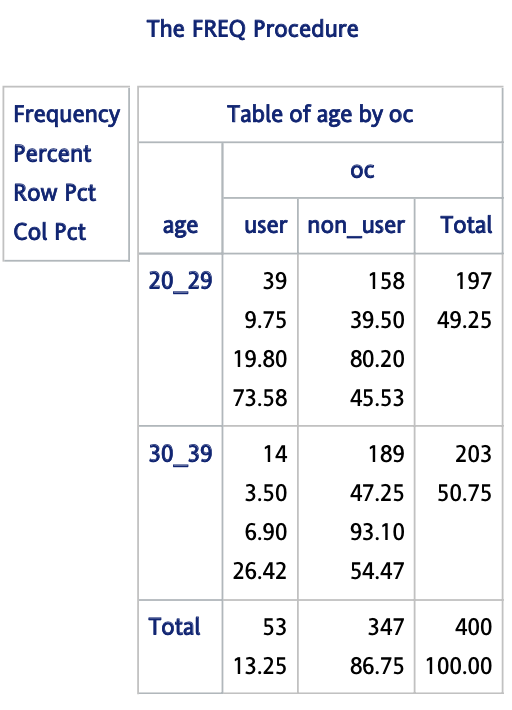
30\_39 non\_user 189

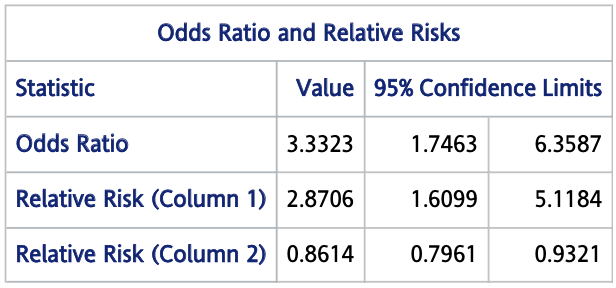
;

**run**;

**proc** **print** data = cf\_e;

**run**;

* + - 1. Data created：
    1. Odd ratio：
       1. Result：
          1. 以是否曾使用OC分組，分為user組與non-user組，比較兩組年齡於20-29歲及30-39歲的比例。
          2. Odd ratio為3.3323，95% CI：1.7463 – 6.3587，達統計上顯著。
          3. 也就是說，在抽樣與暴露無關的前提下，20-29歲的人使用OC的risk是30-39歲的3.3323倍，且具顯著差異。
       2. Figures：



* + - 1. Code ：

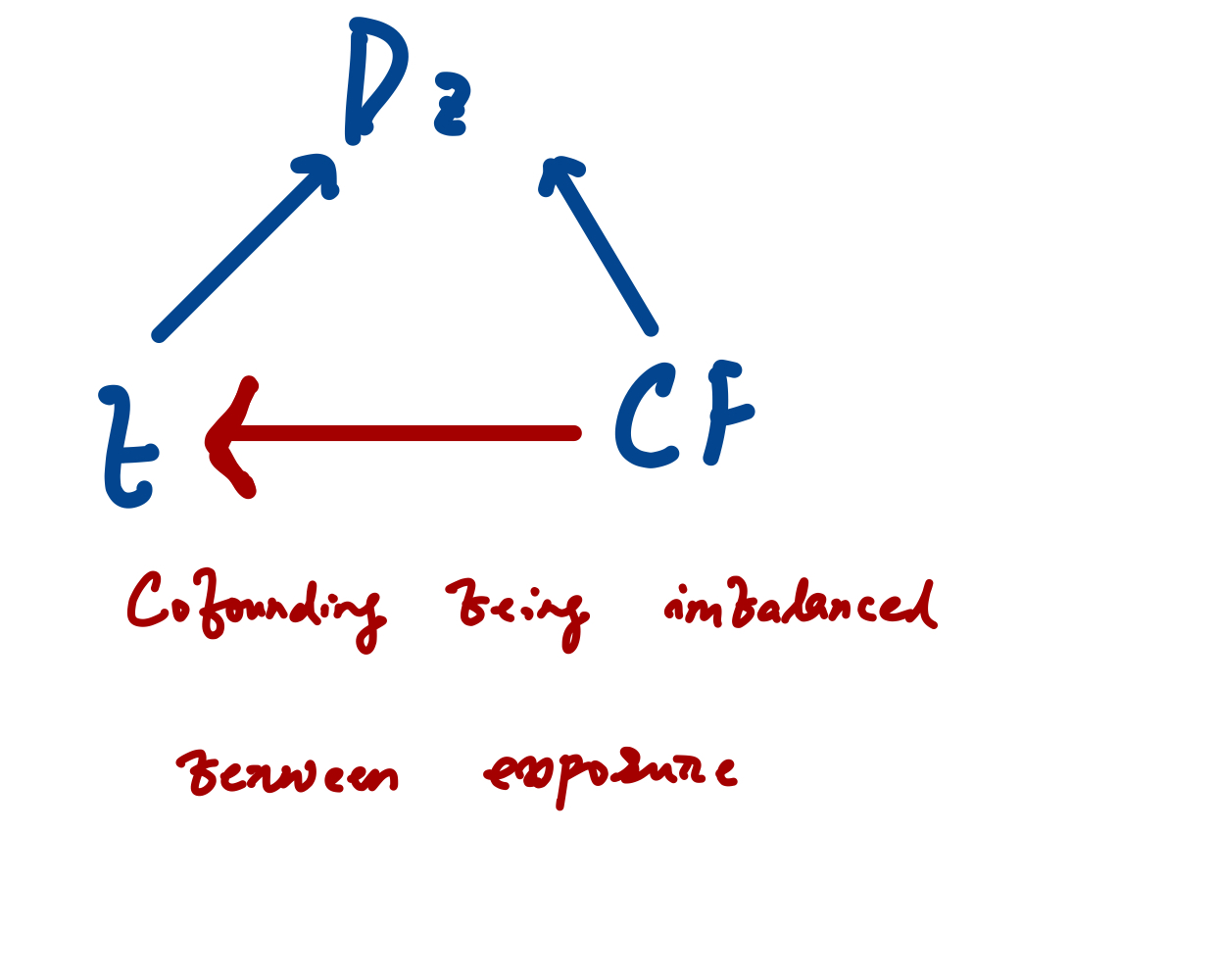
*/\*q3 (a) odd ratio \*/*

**proc** **freq** data = cf\_e order = data;

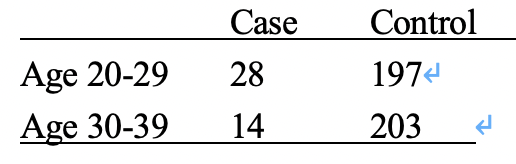
tables age \* oc / chisq or;

weight count;

**run**;

* + 1. Draw the triangle to show what does this OR indicate
    2. Which property of CF does this OR indicate

Ans：confounding must be imbalanced between exposure。

* 1. Age as an independent risk factor
     1. Creating data
        1. Data to create
        2. Code：

*/\* q4 : cf\_dz\*/*

title "q4 cf\_dz ";

*/\* creating data \*/*

**data** cf\_dz;

input age$ group$ count;

datalines;

20\_29 case 28

20\_29 control 197

30\_39 case 14

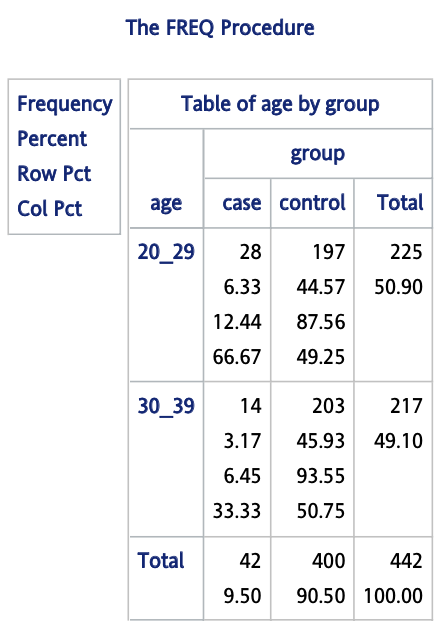
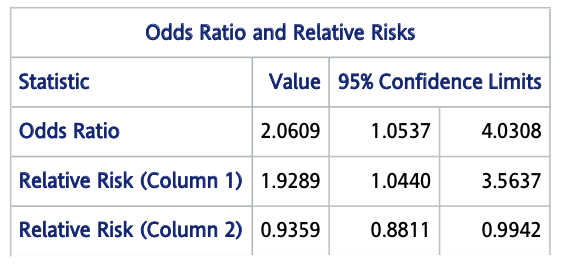
30\_39 control 203

;

**run**;

**proc** **print** data = cf\_dz;

**run**;

* + - 1. Data created：
    1. Odd ratio：
       1. Result：
          1. 兩組odd ratio為2.0609，95% CI：1.0537 – 4.0308，達統計上顯著。
          2. 若符合抽樣與年齡相互獨立的前提，則20-29歲發生venous thrombosis的rate是30-39歲發生venous thrombosis的rate的2.0609倍，且達統計上顯著。
          3. 代表age與venous thrombosis有association，可能是cause或是a proxy for a cause。
       2. Figures：
       3. Code：

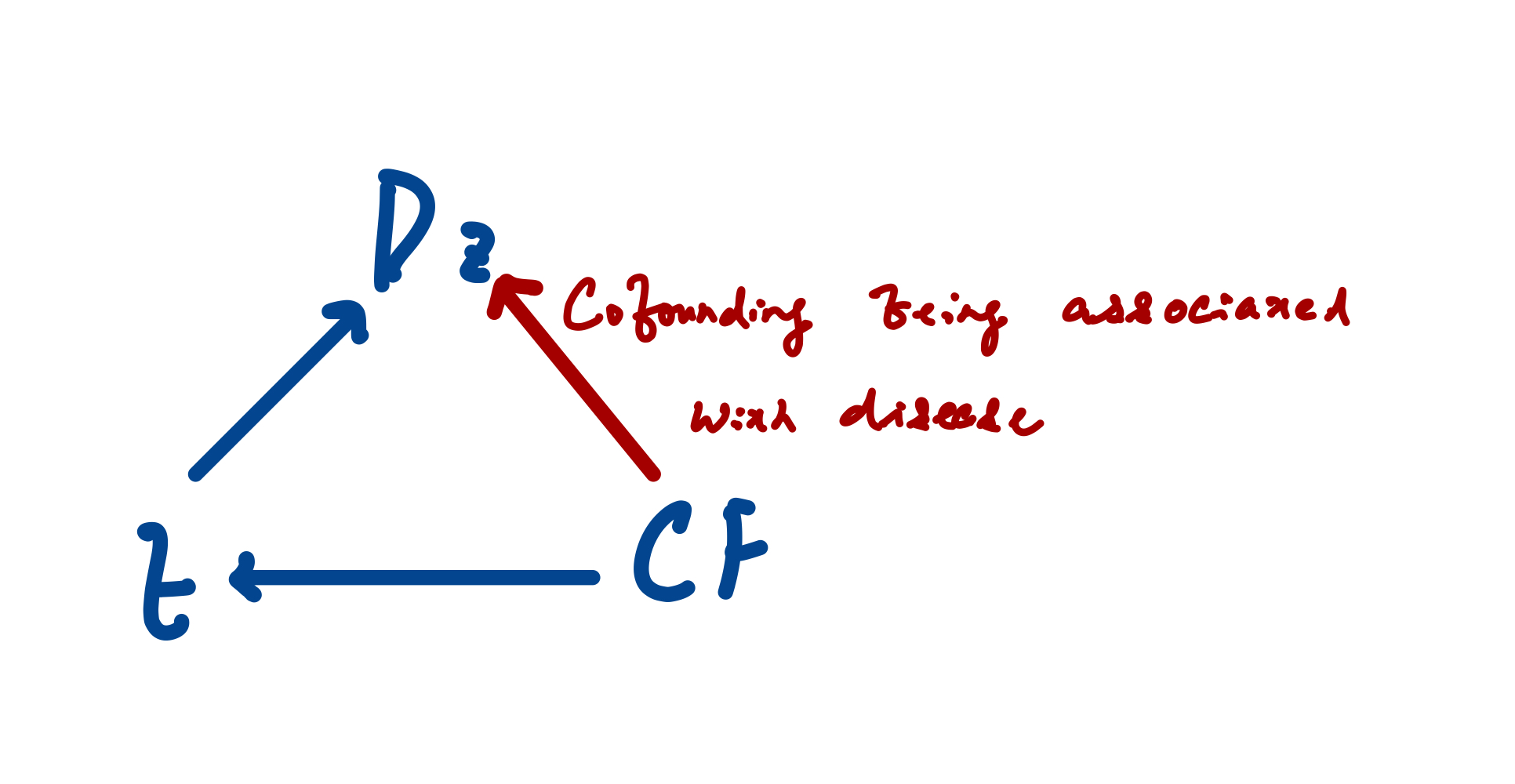
*/\* odd ratio \*/*

**proc** **freq** data = cf\_dz order = data;

tables age \* group / chisq or;

weight count;

**run**;

* + 1. What does this OR indicate in the triangle？
    2. Which property of CF does this OR indicate？

Ans：confounding must be associated with disease。