Answer 10.14

Table 1. The RxC contingency table.

	Use Oracon		
	Yes	No	Row Total
Case	6	111	117
Control	8	387	395
Column Total	14	498	512

Answer 10.20

The point estimate is 0.0526. The 95% confidence interval is (0.0145, 0.1293).

```
. cii proportions 76 4, exact

                                                         Binomial exact
    Variable |        Obs  Proportion    Std. err.       [95% conf. interval]
-------------+---------------------------------------------------------------
             |         76    .0526316    .0256139        .0145246    .1293092
```

Answer 10.21

Table 2. The RxC contingency table.

	Otorrhea at 2 weeks		
	Yes	No	Row Total
Ear drops	4	72	76
Observation	41	34	75
Column Total	45	106	151

Since there is the smallest expected value = 45 × 75 / 151 = 22 > 5, we will use chi-square test.

Let $p_1 = the\ prevalence\ of\ otorrhea\ for\ the\ ear\ drop\ group, p_2 =$
$the\ prevalence\ of\ otorrhea\ for\ theobservation\ group.$

The hypothesis is given by: $H_0: p_1 = p_2; H_1: p_1 \neq p_2$.

Answer 10.22

We input the frequencies as indicator variables and perform the chi-square test in stata.

```
. tabi 4 72 \ 41 34, chi2

           |          col
       row |         1          2 |     Total
-----------+----------------------+----------
         1 |         4         72 |        76
         2 |        41         34 |        75
-----------+----------------------+----------
     Total |        45        106 |       151

          Pearson chi2(1) =  44.0402   Pr = 0.000
```

Since the p-value = 0.00<0.05, we will reject the null hypothesis and conclude that there is a significate difference in the prevalence of otorrhea between the ear drop group and the intervention group.

Answer 10.38

We first generate an indicator variable for the pancreatic secretions. The variable ps_inc equals to 1 if pancreatic secretion post value > pre value.

```
. generate ps_inc=.
(398 missing values generated)

. replace ps_inc=1 if Pansecpt > Pansecpr
(77 real changes made)

. replace ps_inc=0 if Pansecpt <= Pansecpr
(321 real changes made)
```

Then we generate 2x2 contingency table.

```
. tabulate Hormone ps_inc if Hormone ==1 |  Hormone == 4
```

Hormone	ps_inc 0	1	Total
1	25	5	30
4	31	7	38
Total	56	12	68

The smallest expected value = 12 × 30 / 68 = 5.29 > 5. Therefore, we will use a chi-square test.

Let $p_1 = the\ rate\ of\ increased\ pancreatic\ secretions\ for\ the\ saline\ group, p_2 = the\ rate\ of\ increased\ pancreatic\ secretions\ for\ the\ secretin\ group.$

The hypothesis is given by: $H_0: p_1 = p_2; H_1: p_1 \neq p_2$.

We used the chi-square test in stata.

```
. tabulate Hormone ps_inc if Hormone ==1 |  Hormone == 4, chi2
```

Hormone	ps_inc 0	1	Total
1	25	5	30
4	31	7	38
Total	56	12	68

```
          Pearson chi2(1) =   0.0355   Pr = 0.851
```

Since the p-value = 0.851>0.05, we cannot reject the null hypothesis and conclude that there is no significate difference in the rate of increased pancreatic secretions between the saline and secretin group.

Answer 10.104

Let $p_1 = 10 - year\ incidence\ rate\ in\ the\ control\ group, p_2 = 10 - year\ incidence\ rate\ in\ the\ PMH\ group.$

$$p_1 = 1 - \left(1 - \frac{200}{10^5}\right)^{10} = 0.01982$$

$$p_2 = 1 - \left(1 - \frac{240}{10^5}\right)^{10} = 0.02374$$
$$\alpha = 0.05$$

We calculate the estimated sample size in stata.

```
. power twoproportions 0.01982 0.02374, test(chi2)

Performing iteration ...

Estimated sample sizes for a two-sample proportions test
Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.0039  (difference)
           p1 =    0.0198
           p2 =    0.0237

Estimated sample sizes:

            N =    43,528
  N per group =    21,764
```

The total estimated sample size is 43,528.

Answer 10.105

```
. power twoproportions 0.01982 0.02374, test(chi2) n1(20000) nratio(1)

Estimated power for a two-sample proportions test
Pearson's chi-squared test
H0: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
            N =    40,000
           N1 =    20,000
           N2 =    20,000
        delta =    0.0039  (difference)
           p1 =    0.0198
           p2 =    0.0237

Estimated power:

        power =    0.7660
```

The estimated power for 20,000 samples in each group is 76.6%.

Answer 10.125

We will use the McNemar's exact test, since the number of discordant pairs is 10+5=15<20.

$$n_d = 15, n_a = 5$$

The exact two-sided p-value is 0.301>0.05. Therefore, we cannot reject the null hypothesis and cannot determine whether the new test is better or worse than the standard test.

Answer 10.139

We have a 2xk contingency table constructed, and we are interested in trend over k binomial proportions. Therefore, we can use the Chi-square trend test.

Answer A1

Table 3. The RxC contingency table.

	Adverse events		
Drugs	Yes	No	Row Total
Aspirin Alone	20	537	557
Aspirin + Ticlopidine	3	543	546
Column Total	23	1080	1103

Since the smallest expected value = 23 × 546 / 1103 = 11.39 > 5, we will use the chi-square test.

Let $p_1 = the\ proportion\ of\ adverse\ events\ for\ the\ Aspirin\ group, p_2 =$ $the\ proportion\ of\ adverse\ events\ for\ the\ Aspirin\ and\ Ticlopidine\ group.$

The hypothesis is given by: $H_0: p_1 = p_2; H_1: p_1 \neq p_2$.

We use the chi-square test in stata.

```
. tabi 20 537 \ 3 543, chi2
```

row	col 1	2	Total
1	20	537	557
2	3	543	546
Total	23	1,080	1,103

```
          Pearson chi2(1) =  12.4901   Pr = 0.000
```

Since the p-value = 0.000 < 0.05, we can reject the null hypothesis and conclude that there is an association between drug regimen and the number of adverse events.