

Homework # 2 Solutions

Biostatistics 210

Your colleague performed a logistic regression model using prior1995 as a predictor.

- (1) *Can you verify the association?* Yes, prior to 1995 about 3% of kids died and after 1995 about 2% of kids died. This gives the following output out of a logistic regression

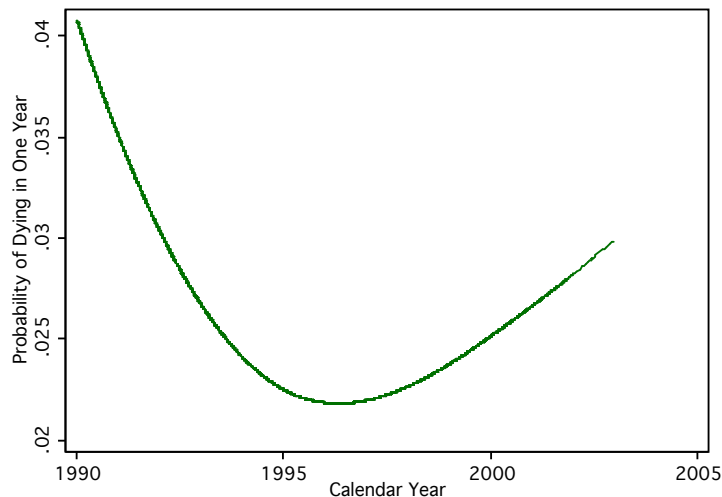
```
. logistic died p

Logistic regression                               Number of obs   =       7949
                                                LR chi2(1)      =       4.92
                                                Prob > chi2     =       0.0265
Log likelihood = -956.9725                       Pseudo R2      =       0.0026

-----+-----
      died | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
prior1995 |   1.371822   .1970383    2.20  0.028     1.035232     1.81785
-----+-----
```

which is statistically significant.

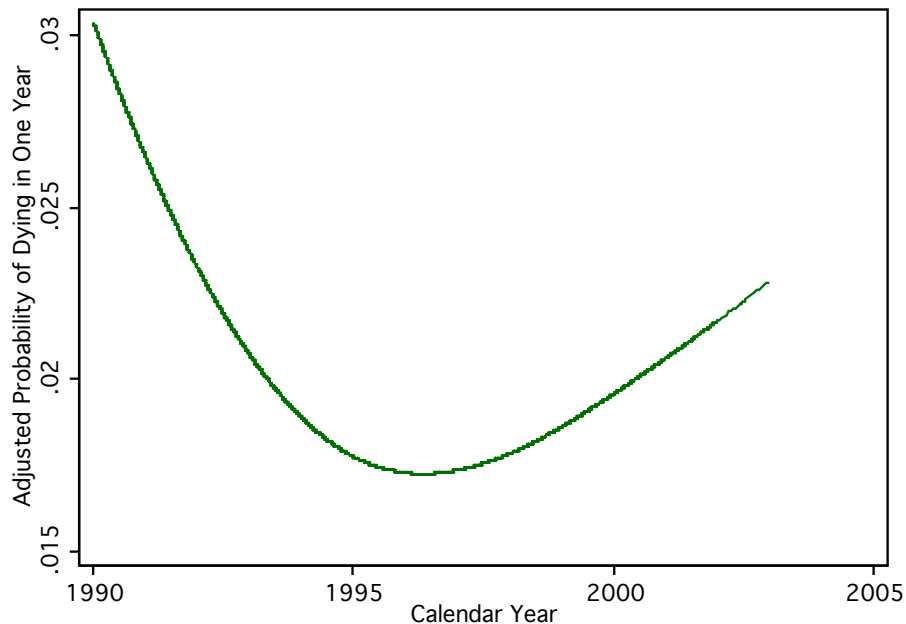
- (2) *Given your colleague's hypothesis, how would you expect the risk of death to change over the years?* It's implicit but not stated that we would expect to see a stable risk of death over the years prior to 1995 that it might decrease as the intervention is rolled out and then perhaps be stable after that. Any other pattern would not fit with the results above.
- (3) *Explore this using some techniques which discussed. You might also want to make sure you have taken into account factors which have changed over the year (recipient age and living v. cadaveric tx).* I used a cubic spline fit to year with knots at 1990 1995 and 2000 and got the following fit.



We find that mortality was falling prior to 1996 and that it has increased slightly since then. I was also able to get adjusted survival using the following set of commands

```
mkspline cube_cts = year_cts, cubic knots(1990 1995 2000) displayknots
logistic died cube_cts* age ttype
gen out22 = _b[_cons] + _b[ttype]*0.47 + _b[age]*11.7 + _b[cube_cts2]*cu-
be_cts2 + _b[cube_cts1]*cube_cts1
gen pr = exp(out22)/(1+exp(out22))
twoway (line pr year_cts, sort), ytitle(Adjusted Probability of Dying in One
Year) xtitle(Calendar Year) scheme(s1color)
```

The adjusted graph shows the same pattern



- (4) *Which analysis appears to be the most informative?* The spline fit, preferably the adjusted one is vastly more informative than the binary fit of years. It gives us some sense of how mortality changes flexibly with time.
- (5) *Do you conclude that the guideline implementation has improved one year survival?* It's difficult to see the implicit hypothesis of improved survival after guideline implementation contradicted **more**. We see that mortality fell before the guidelines and may have increased after. Hence, any improvement after 1996 seems to reflect changes that started much early. It is an important lesson in how binary fits have the potential to greatly mislead if taken entirely at face value.