pcptwebappendix. October 27, 2009 **BIOSTATISTICS MS B09-049**

Supplementary material to Transparency and reproducibility in data analysis: The Prostate Cancer Prevention Trial

STUART G. BAKER

[1]Biometry Research Group, Division of Cancer Prevention,National Cancer Institute, EPN 3131, 6130 Executive Blvd, MSC 7354, Bethesda MD 20892-7354, USA

AMY K DARKE

[2]Statistical Center, Fred Hutchinson Cancer Research Center Seattle, WA

PAUL PINSKY

[3]Early Detection Research Group Division of Cancer Prevention, National Cancer Institute, Bethesda MD

HOWARD L. PARNES

[4]Prostate and Urologic Cancer Research Group , Division of Cancer Prevention, National Cancer Institute, Bethesda MD

BARNETT S. KRAMER

[5]Office of Disease Prevention, National Institutes of Health, Bethesda MD

*email: sb16i@nih.gov

1. INTRODUCTION

In this web appendix, we describe the clinical background and document important definitions, classification decisions, derivations, and calculations.

2. CLINICAL BACKGROUND

The primary endpoint of the PCPT was the difference in period prevalence of all prostate cancers between the finasteride and study arms, and it was for this endpoint that the Data and Safety Monitoring Committee recommended stopping the trial and reporting the results. There was a highly significant 25% reduction in overall prostate cancer in the finasteride study group. However, detailed analysis of tumors *detected on biopsy* showed an excess of high-grade prostate tumors (a Gleason score of 7 or higher) in the finasteride arm versus the placebo arm (Thompson and others, 2003), dampening enthusiasm for finasteride. There was also concern that the excess high-grade prostate tumors detected on biopsy in the finasteride arm could be artifactual because finasteride shrinks the prostate making sampling-based detection of high-grade prostate cancer more likely in the finasteride arm even if the true prevalence of high-grade prostate cancer were the same in both arms (Scardino 2003). Fortunately data were also collected on grade of tumors in a subset of persons who subsequently received surgery for prostate cancer. A less biased endpoint is high-grade prostate *determined by surgery*.

A major challenge with analyzing the data using this endpoint of high-grade prostate cancer on surgery is the complexity of handling two types of missing data: missing in biopsy and missing in surgery following biopsy, both of which were likely dependent on whether or not a biopsy was recommended. An additional challenge was that the data were messy, due to complications including hormone treatment, transurethral resection of the prostate, cystoprostatectomy.

3. DEFINITION OF POSITIVE BIOPSY RECOMMENDATION

We defined a positive biopsy recommendation $(a = 1)$ as either an elevated PSA and abnormal DRE, only an abnormal DRE, only an elevated PSA, or only an outside recommendation. (The level of PSA defined as elevated differed in the two randomization groups because the study designers wanted to achieve similar biopsy rates

in the randomization groups and consequently needed to adjust for the effect of finasteride on PSA levels). We classified all others as a negative biopsy recommendation $(a = 0)$, including those without biopsy recommendation because they received transurethral resection of the prostate (41 persons). The variable for biopsy recommendation referred to the last testing before the end of the study or a biopsy.

4. CLASSIFYING PROBLEMATIC SUBJECTS INTO CATEGORIES

There were 142 persons in PCPT who did not fit neatly into our categories. We classified these persons in the manner described below.

(a) The 2 persons who had prostate cancer at the time of randomization were discarded from the analysis because they were ineligible for the trial.

(b) The 28 persons diagnosed on surgery outcome but missing a biopsy result were classified as having a biopsy grade that was the same as the surgery grade.

(c) The 32 persons missing biopsy and surgery outcomes who were diagnosed with prostate cancer in a procedure other than biopsy or surgery (29 via transurethral resection of the prostate, 1 on an autopsy, 2 on cystoprostatectomy) were classified as missing for both biopsy and surgery outcomes.

(d) The 1 person with squamous cancer of the prostate, and the 2 persons with tumors too small to grade on biopsy were classified as low-grade.

(e)The 28 persons not graded on surgery because of prior hormone treatment were classified as missing on surgery (or low or high-grade in a sensitivity analysis).

(f) The 1 person not graded on biopsy because of hormone treatment was classified as missing on biopsy and surgery (or low or high-grade in a sensitivity analysis).

(g) The 3 persons with no cancer found on surgery but a Gleason score on biopsy were classified for surgery outcome according to the Gleason score on biopsy.

(h) The 1 person with no cancer found on surgery and missing a biopsy outcome was classified as missing on biopsy and surgery.

(i) The 46 persons whose cancer was diagnosed at the screening sites but not confirmed by the central laboratory were classified as missing biopsy and surgery

(j) Information on cancer endpoint after the closing date of June 23, 2003 was not included in the analysis.

With these classifications, we analyzed data from 9457 persons in the placebo group and 9423 persons in the finasteride group. See Tables 1 and 2 below for results of sensitivity analysis with classification of persons receiving hormone treatment.

5. FORMULATION OF THE LIKELIHOOD

Although we did not need to formulate the likelihood to derive maximum likelihood estimates, the formulation may be of interest to some readers. Under this model, the joint probability of prostate cancer outcome on surgery d, prostate cancer outcome on biopsy y, biopsy recommendation a, receiving biopsy, and receiving surgery conditional on randomization group x is

$$pr(D=d, Y=y, A=a, M_Y=1, M_D=1|x) = \beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ (1-\pi_{xa})(1-\gamma_{xay}).$$

The joint probability of prostate cancer outcome on biopsy y, biopsy recommendation a, receiving biopsy, and missing surgery conditional on randomization group x is

$$pr(Y=y, A=a,\ M_Y=1, M_D=0\,|x) = \sum_{d=0}^{2} \beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ (1-\pi_{xa})\ \gamma_{xay}.$$

The joint probability of biopsy recommendation a, missing biopsy, and missing surgery conditional on randomization group x is

$$pr(A = a,\ M_Y = 0, M_D = 0|\ x)\ \ =\ \ \sum_{y=0}^{2}\ \sum_{d=0}^{2}\beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ \pi_{xa.}$$

Based on these joint probabilities, the likelihood kernel is

$$L = \prod_{x=0}^{1}\ \prod_{a=0}^{1}\ \prod_{y=1}^{2}\ \prod_{d=1}^{2}\ \{\beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ (1-\pi_{xa})(1-\gamma_{xay})\}^{k_{xayd}}$$
$$\prod_{x=0}^{1}\ \prod_{a=0}^{1}\ \prod_{y=0}^{2}\ \{\ \sum_{d=0}^{2}\beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ (1-\pi_{xa})\ \gamma_{xay}\}^{n_{xay}}$$
$$\prod_{x=0}^{1}\ \prod_{a=0}^{1}\{\sum_{y=0}^{2}\ \sum_{d=0}^{2}\ \beta_{d|x}\ \lambda_{y|xd}\ \phi_{a|xyd}\ \pi_{xa}\}^{m_{xa}}.$$

6. DERIVATION OF MAXIMUM LIKELIHOOD ESTIMATES

We present details for the derivation of the maximum likelihood estimates based on perfect fit estimates. The first step was to check that the model is saturated, meaning it has the same number of independent parameters as independent cell counts. An independent parameter is a parameter that cannot be computed from the other parameters. An independent cell count is a cell count that cannot be determined from the other cell counts and the fixed totals. Because all persons with surgery had prostate cancer on biopsy and surgery, the following parameters equal zero: $\lambda_{0|x1}, \lambda_{0|x2}, \phi_{1|x01},\ \phi_{1|x02},$ $\phi_{1|x10},$ and $\phi_{1|x20}$. Therefore for each value of $x,$ there are 15 independent parameters. Because $\beta_{0|x} + \beta_{1|x} +\ \ \beta_{2|x} = 1,\ \lambda_{1|xy} +\ \lambda_{2|xy} = 1,$ and $\phi_{1|xyd}\ + \phi_{2|xyd} = 0$, for $\{y, d\} \in\ \{\{0,0\},\ \{1,1\},\ \{1,2\},\ \{2,1\},\ \{2,2\}\}$, one possible set of independent parameters is: $\beta_{1|x},\ \beta_{2|x},\ \lambda_{1|x1},\ \lambda_{1|x2}, \phi_{1|x00}, \phi_{1|x11},\ \phi_{1|x12},\ \phi_{1|x21}, \phi_{1|x22},\ \pi_{x0},\ \pi_{x1},$ $\gamma_{x01}, \gamma_{x02},\ \gamma_{x11},$ and γ_{x12}. For each value of x there are also 15 independent counts: 2 for $\{m_{xa}\}$, 6 counts for $\{n_{xay}\}$, and 8 for $\{k_{xayd}\}$ minus 1 because the total for x is fixed. Therefore the model is saturated. Because the model is saturated, maximum likelihood estimates can be easily obtained by setting observed counts equal to expected values,

$$\beta_{d|x}\ \lambda_{y|xd}\,\phi_{a|xyd}\,(1-\pi_{xa})(1-\gamma_{xay})\,N_x\ =\ k_{xayd}, \tag{A.1}$$

$$\sum_d \beta_{d|x}\,\lambda_{y|xd}\,\phi_{a|xyd}\,(1-\pi_{xa})\,\gamma_{xay}\,N_x\ = n_{xay}, \tag{A.2}$$

$$\sum_y \sum_d \beta_{d|x}\,\lambda_{y|xd}\,\phi_{a|xyd}\,\pi_{xa} N_x = m_{xa.} \tag{A.3}$$

and solving. Adding (A.1) and (A.2) and summing over d gives

$$\sum_d \beta_{d|x}\ \lambda_{y|xd}\,\phi_{a|xyd}\,(1-\pi_{xa})\,N_x\ =\ k_{xay+}\ +\ n_{xay.} \tag{A.4}$$

Adding (A.1), (A.2), and (A.3) and summing over y and d gives

$$\sum_y \sum_d \beta_{d|x}\ h_{y|xd}\,\phi_{a|xyd}\,N_x\ =\ k_{xa++}\ +\ n_{xa+}\ + m_{xa}. \tag{A.5}$$

Summing (A.5) over a gives

$$\sum_y \sum_d \beta_{d|x}\ \lambda_{y|xd}\,N_x\ =\ k_{x+++}\ +\ n_{x++}\ + m_{x+}. \tag{A.6}$$

From (A.3) and (A.6) $\widehat{\pi}_{xa} = \frac{m_{xa}}{k_{xa++}+n_{xa+}+m_{xa}}$. (A.7)

From (A.4) and (A.6) $\widehat{r}_{xay} = \frac{n_{xay}}{k_{xay+}+n_{xay}}$. (A.8)

From (A.1), $\beta_{d|x}\,\lambda_{y|xd}\,\phi_{a|xyd} = \frac{k_{xayd}}{(1-\pi_{xa})(1-\gamma_{xay})\,N_x}$. (A.9)

Summing both sides of (A.9) over a and y and substituting (A.7) and (A.8) gives

$$\widehat{\beta}_{d|x}\ =\ \sum_a \sum_y \frac{k_{xayd}}{(1-\widehat{\pi}_{xa})\,(1-\widehat{\gamma}_{xay})\,N_x}. \tag{A.10}$$

7. COMPUTATION OF CONFIDENCE INTERVALS

We used the following procedure to compute confidence intervals for the estimated relative risk. Applying the Multinomial-Poisson transformation (Baker, 1994) gives

$$\begin{aligned}\widehat{var}(\widehat{\beta}_{2|x}) \ =\ &\sum_{x=0}^{1}\ \sum_{\mathrm{a=0}}^{1}\ \sum_{\mathrm{y=1}}^{2}\ \sum_{\mathrm{d=1}}^{2}\left(\frac{\partial\widehat{\beta}_{2|x}}{\partial k_{xayd}}\right)^2 k_{xayd}\ +\ \sum_{x=0}^{1}\ \sum_{\mathrm{a=0}}^{1}\ \sum_{\mathrm{y=0}}^{2}\ \left(\frac{\partial\widehat{\beta}_{2|x}}{\partial n_{xay}}\right)^2 n_{xay}\\ &+\ \sum_{x=0}^{1}\ \sum_{\mathrm{a=0}}^{1}\ \left(\frac{\partial\widehat{\beta}_{2|x}}{\partial m_{xa}}\right)^2 m_{xa},\end{aligned} \tag{B.1}$$

where the derivatives are readily computed using software for symbolic algebra. Using the delta method, the 95% confidence interval for the unadjusted relative risk is

$(exp[\, log(RR) \, - 1.96 \, se], \; exp[\, log(RR) \, + 1.96 \, se),$

where $se \, = \sqrt{var\{log(RR)\}}$ and $var\{log(RR)\} = \sum_{x=0}^{1} \frac{\hat{var}(\hat{\beta}_{2|x})}{\hat{\beta}_{2|x}^2}$.

8. ADJUSTMENT FOR BASELINE VARIABLES

We used a propensity-to-be-missing score to adjust likelihood-based analyses for many baseline variables in the presence of missing outcomes (Baker et al 2006). To create a propensity-to-be-missing score for these data, we first fit a logistic regression for missing surgery as a function of the baseline variables of family history (first degree relative with prostate cancer or not), race (white or not), and age (four categories) along with interactions of age with family history and age with race. Next, for each person we substituted the parameter estimates into the logistic regression with the variables for that person. This yielded a propensity-to-be-missing score for each person.

Next, we grouped the propensity-to-be-missing scores into strata. To avoid zero counts, we used only three strata when high-grade prostate cancer was defined as a Gleason score greater than or equal to 7, and only two strata when high-grade prostata cancer was defined as Gleason score greater than or equal to 8. For each randomization group x in stratum $i \, = \, 1, \, 2, \, ... \, , I,$ we computed the fraction of persons in each stratum of the propensity-to-be-missing score, which we denote $w_{i|x}$, where $\Sigma_i \, w_{i|x} = 1$. For each stratum i, we computed via equation (4) in the text, $\hat{\beta}_{2|xi}$, the estimated probability of high-grade prostate cancer in propensity-to-be-missing stratum i for randomization group x. Then we computed the adjusted relative risk,

$$RR^* = \frac{\Sigma_i \, \hat{\beta}_{2|1i} \, w_{i|1}}{\Sigma_i \, \hat{\beta}_{2|0i} \, w_{i|0}} \; .$$

Based on the delta method, the estimated variance of the logarithm of RR^* is

$$\widehat{var}\{log(RR^*)\} = \sum_{x=0}^{1} \frac{var(\widehat{\beta}_{2|x})\, w_{i|x}^2}{(\Sigma_i\, \widehat{\beta}_{2|xi}\, w_{i|x})^2} + \frac{\widehat{\boldsymbol{\beta}}_{2|x}\ \widehat{var}(w_x)\ \widehat{\boldsymbol{\beta}}_{2|x}^T}{(\Sigma_i\, \widehat{\beta}_{2|xi}\, w_{i|x})^2}$$

where

$var(\widehat{\beta}_{2|x})$ is computed from (B.1),

$w_x = (w_{1|x}, w_{2|x}, ..., w_{(I-1)|x})$,

$\widehat{var}(w_x)$ is the multinomial variance-covariance matrix for w_x,

and $\widehat{\boldsymbol{\beta}}_{2|x} = (\widehat{\beta}_{2|x1} - \widehat{\beta}_{2|xI},\ \widehat{\beta}_{2|x2} - \widehat{\beta}_{2|xI}, ..., \widehat{\beta}_{2|x(I-1)} - \widehat{\beta}_{2|xI})$.

The 95% confidence interval the adjusted relative risk is

$(exp[\ log(RR^*)\ - 1.96\ se^*],\ exp[\ log(RR^*)\ + 1.96\ se^*)$,

where $se^* = \sqrt{\widehat{var}\{log(RR^*)\}}$,

9. REPRODUCING CALCULATIONS

For readers interested in reproducing the basic calculations, we present the following details.

High-grade is Gleason score of 7 or higher

We first consider the calculations when high-grade prostate cancer was defined as a Gleason score of 7 or higher (Table 1). The estimated probabilities of missing a biopsy among persons in the placebo group were

$$\begin{aligned} \widehat{p}_{00} &= \tfrac{3955}{8248}=0.48, \quad \text{for persons without a biopsy recommendation,} \\ \widehat{p}_{01} &= \tfrac{215}{1209}=0.18, \quad \text{for persons with a biopsy recommendation,} \end{aligned} \qquad \text{(C.1)}$$

indicating the anticipated strong dependence on whether or not the person was recommended for biopsy. The estimated probabilities of missing surgery among persons in the placebo group were

$$\widehat{r}_{001} = \frac{417}{519}=0.80, \text{ for persons without a biopsy recommendation who were low-grade,}$$
$$\widehat{r}_{002} = \frac{78}{99} =0.79, \text{ for persons without a biopsy recommendation who were high-grade,}$$
$$\widehat{r}_{011} = \frac{249}{365}=0.68, \text{ for persons with a biopsy recommendation who were low-grade,}$$
$$\widehat{r}_{012} = \frac{104}{159}=0.65, \text{ for persons with a biopsy recommendation who were high-grade.} \quad \text{(C.2)}$$

Thus, in the placebo group, the estimated probabilities of missing surgery among persons biopsied depended more strongly on whether not the person had received a biopsy recommendation than whether the person was low-grade or high-grade at biopsy. Based on equation (4) in the text, (C.1), (C.2), the estimated probability of high-grade prostate cancer among persons in the placebo group is

$$\widehat{\beta}_{2|0} = \frac{185.74 + 117.75 +126.30 +3161.75}{9457} = 0.063. \quad \text{(C.3)}$$

Now consider the finasteride group. The estimated probabilities of missing a biopsy among persons in the finasteride group were

$$\widehat{p}_{10} = \frac{4169}{8342}=0.50, \text{ for persons without a biopsy recommendation,}$$
$$\widehat{p}_{11} = \frac{214}{1081}=0.20, \text{ for persons with a biopsy recommendation,} \quad \text{(C.4)}$$

again indicating strong dependence on whether or not the person was recommended for biopsy. The estimated probabilities of missing surgery among persons in the finasteride group who received a biopsy were

$\hat{r}_{101} = \frac{230}{286}$=0.80, for persons without a biopsy recommendation who were low-grade,

$\hat{r}_{102} = \frac{75}{96}$=0.78, for persons without a biopsy recommendation who were high-grade,

$\hat{r}_{111} = \frac{145}{212}$=0.68, for persons with a biopsy recommendation who were low-grade,

$\hat{r}_{112} = \frac{132}{197}$=0.67, for persons with a biopsy recommendation who were high-grade (C.5)

Thus in the finasteride group, as with the placebo group, the estimated probabilities of missing surgery among persons biopsied depended more strongly on whether not the person had received a biopsy recommendation than on whether the person was low or high grade at biopsy. Based on equation (4) in the text, (C.4), and (C.5), the estimated probability of high-grade prostate cancer among persons in the finasteride group is

$$\hat{\beta}_1 = \frac{61.26 + 127.94 + 90.74 + 207.84}{9423} = 0.052. \quad \text{(C.6)}$$

From (C.3) and (C.6) and the variance formula in Appendix B, the estimated relative risk for high-grade prostate cancer is

$$RR = \frac{0.052}{0.063} = 0.83 \text{ with 95\% confidence interval of (0.65, 1.05).} \quad \text{(C.7)}$$

High-grade is Gleason score of 8 or higher

We now consider the calculations when high-grade prostate cancer was defined as a Gleason score of 8 or higher (Table 2). The estimated probabilities of missing a biopsy were the same as in the primary analysis. The estimated probabilities of missing surgery among persons in the placebo group were

$\hat{r}_{001} = \frac{488}{608}$=0.80, for persons without a biopsy recommendation who were low-grade,

$\hat{r}_{002} = \frac{7}{10}$=0.70, for persons without a biopsy recommendation who were high-grade,
$\hat{r}_{011} = \frac{316}{472}$=0.67,for persons with a biopsy recommendation who were low-grade,
$\hat{r}_{012} = \frac{37}{52}$=0.71, for persons with a biopsy recommendation who were high-grade. (C.8)

Based on equation (4) in the text and (C.8), the estimated probability of high-grade prostate cancer among persons in the placebo group was

$$\hat{\beta}_0 = \frac{0 + 0 + 14.72 + 42.16}{9457} = 0.0061. \quad \text{(C.9)}$$

The estimated probabilities of missing surgery among persons in the finasteride group were

$\hat{r}_{101} = \frac{291}{362}$=0.80, for persons without a biopsy recommendation who were low-grade,
$\hat{r}_{102} = \frac{14}{20}$=0.70, for persons without a biopsy recommendation who were high-grade,
$\hat{r}_{111} = \frac{231}{338}$=0.68, for persons with a biopsy recommendation who were low-grade,
$\hat{r}_{112} = \frac{46}{71}$=0.65, for persons with a biopsy recommendation who were high-grade. (C.10)

Based on equation (4) in the text, (C.4), and (C.10), the estimated probability of high-grade prostate cancer among persons in the finasteride group was

$$\hat{\beta}_1 = \frac{10.19 + 6.66 + 23.63 + 46.03}{9423} = 0.0092. \quad \text{(C.11)}$$

Based on equation (4) in the text, (C.9), and (C.11), the unadjusted estimated relative risk is

$$RR = \frac{0.0092}{0.0061} = 1.53 \text{ with 95\% confidence interval (0.85, 2.75).} \quad \text{(C.12)}$$

ADDITIONAL REFERENCE

SCARDINO, P. T. (2003). The prevention of prostate cancer--the dilemma continues. *New England Journal of Medicine* **349**, 297-299.

SENSITIVITY ANALYSIS: HORMONE TREATMENT

Table 1. *Estimated relative risks for high-grade prostate cancer defined as Gleason score 7 or above*

	Estimated relative risk (95% confidence interval)		
our method	Persons with hormone treatment classified as	Unadjusted estimate based on Table 1	Estimate adjusted for family history, race, and age
	missing	0.83 (0.65, 1.05)	0.82 (0.64, 1.06)
	low-grade	0.79 (0.62, 1.01)	0.77 (0.59, 1.01)
	high-grade	0.82 (0.65, 1.03)	0.82 (0.64,1.04)
Redman and others (2008)	0.73 (0.56,0.96)		
Pinsky and others (2008)	0.84 (0.58, 1.06)		

Table 2. *Estimated relative risks for high-grade prostate cancer defined as Gleason score 8 or above*

	Estimated relative risk (95% confidence interval)		
our method	Persons with hormone treatment classified as	Unadjusted estimate based on Table 2	Estimate adjusted for family history, race, and age
	missing	1.53 (0.85, 2.75)	1.40 (0.71, 2.76)
	low-grade	1.41 (0.76, 2.61)	1.36 (0.44, 4.15)
	high-grade	1.23 (0.75, 2.01)	1.24 (0.78, 1.98)
Redman and others (2008)	1.25 (no confidence interval reported; said to be imprecise)		
Pinsky and others (2008)	1.39 (0.79, 2.50)		