Published in final edited form as:

Med Decis Making. 2019 February ; 39(2): 130–136. doi:10.1177/0272989X18819792.

Cancer screening markers: A simple strategy to substantially reduce the sample size for validation

Stuart G. Baker

National Cancer Institute

Abstract

Background.—Studies to validate a cancer prediction model based on cancer screening markers collected in stored specimens from asymptomatic persons typically require large specimen-collection sample sizes. A standard sample size calculation targets a true positive rate (TPR) of 0.8 with a 2.5% lower bound of 0.7 at a false positive rate (FPR) of 0.01 with a 5% upper bound of 0.03. If the probability of developing cancer during the study is $P=0.01$, the specimen-collection sample size based on the standard calculation is 7600.

Methods.—The strategy to reduce the specimen-collection sample size is to decrease both the lower bound of TPR and the upper bound of FPR while keeping a positive lower bound on the anticipated clinical utility.

Results.—The new sample size calculation targets a TPR of 0.4 with a 2.5% lower bound of 0.10 and an FPR of 0.0 with a 5% upper bound of 0.005. With $P=0.01$, the specimen-collection sample size based on the new calculation is 1800 instead of 7600.

Limitations.—The new sample-size calculation requires a minimum benefit-cost ratio (number of false positives traded for a true positive). With $P=0.01$, the minimum cost-benefit ratio is 5, which is plausible in many studies.

Conclusion.—In validation studies for cancer screening markers, the strategy can substantially reduce the specimen-collection sample size, substantially reducing costs and making some otherwise infeasible studies now feasible.

Introduction

The first step in the pipeline to develop new cancer screening modalities is the discovery and validation of cancer screening markers to predict the development of cancer in asymptomatic persons.^{1,2} Cancer screening markers can take a variety of forms. For example, investigators are considering the following pancreatic cancer screening markers: genomic markers in the blood, cyst fluid, and pancreatic juice, proteomic markers including CA-19, glycoproteins, and exosomes, and imaging markers involving radiography and radiomics.³

Correspondence to: Stuart G. Baker, Division of Cancer Prevention, National Cancer Institute, 9609 Medical Center Dr, Room 5E606, MSC 9789, Bethesda, MD 20892-9789 USA; sb16i@nih.gov.

Notes

The statements contained herein are solely those of the author and do not represent or imply concurrence or endorsement by the NIH. The study sponsor had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication

The discovery phase involves creating a cancer prediction model using markers tested in specimens collected from cases and controls. Usually in the discovery phase, investigators select as cases persons with symptomatic cancer and select as controls persons without symptomatic cancer. This choice of cases and controls facilitates rapidly obtaining results and avoids the concern (with molecular assays) of leaving an insufficient amount of specimen for validation. However, as discussed later, there are drawbacks related to the lack of relevance for the early detection of cancer. The cancer prediction model should include risk factors, such as age and family history, in addition to the cancer screening markers.

The validation phase involves evaluating the performance of the cancer prediction model derived from the discovery phase in an independent validation sample of asymptomatic persons from whom specimens are collected and stored. The collection of specimens can occur in either (i) a large existing cohort or a clinical trial of asymptomatic people that is typically not designed to investigate the specific cancer or (ii) a custom-built smallest practical cohort of asymptomatic persons with target follow-up for a specific cancer. The focus here is on the latter case, such as the Alliance of Pancreatic Cancer Consortia for Early Detection.³ After specimen collection and follow-up to ascertain cancer incidence, investigators test for the markers in stored specimens from all cases (persons who developed cancer) and random sample of controls (persons who did not develop cancer). For each person with stored specimens tested for the markers, investigators apply the cancer prediction model derived from the discovery phase to the markers and predictors associated with the validation sample. The result is a marker score for each person in the validation sample.

A standard summary of the performance of the cancer prediction model is the Receiver-Operating Characteristic (ROC) curve, which plots true positive rate (TPR) versus false positive rate (FPR) for various cutpoints of the marker score. A marker score larger than a specified cutpoint is called positive; otherwise it is called negative. TPR, which is also called sensitivity, is the fraction of cases that are positive for the corresponding cutpoint. FPR, which is also called one minus specificity, is the fraction of controls that are negative for the corresponding cutpoint.

In this context validation means finding a combination of FPR and TPR on the ROC curve in the validation sample which indicates a sufficiently promising marker. A major difficulty with validation in a custom-built cohort of asymptomatic persons is the large sample size needed for specimen collection because the endpoint of clinical cancer is a rare event in asymptomatic persons. This article develops a strategy to substantially reduce the specimen-collection sample size (the total number of stored specimens) for validation.

Specimen-collection and marker-testing sample sizes

The marker-testing sample size is the number of stored specimens tested for the marker. The specimen-collection sample size is related to the marker-testing sample size as follows.

Let x_{case} denote the marker-testing sample size for cases. The expected number of specimens needed to obtain x_{case} marker-testing samples is $E_{case} = x_{case}/P$, where P is the

probability of developing the cancer of interest during the study. The problem with using E_{case} for the specimen collection sample size is that it does account for random variability. A more relevant quantity than E_{case} is N_{case} , the specimen-collection sample sized needed to obtain with 95% probability at least x_{case} case specimens for marker-testing. N_{case} solves the following equation for the cumulative distribution function (CDF) for a binomial random variable evaluated at x_{case} ,

$$\text{CDF}[\text{Binomial}\{N_{case}, P\}, x_{case}] = 0.95, \quad (1)$$

where $\text{CDF}[\text{Binomial}\{n, p\}, x] = \sum_{i=0}^x \{n!/(i!(n-i)!)\} p^i (1-p)^{n-i}$ is the probability that a Binomial random variable with probability p of a single outcome yields x or fewer outcomes among n events. Computation of N_{case} in equation (1) involves the numerical tabulation of various values for N_{case} .

Let $x_{control}$ denote the marker-testing sample size for controls. The expected number of specimens needed to obtain $x_{control}$ marker-testing samples is $E_{control} = x_{control} / (1-P)$. Let $N_{control}$ denote the specimen-collection sample sized needed to obtain with 95% probability at least $x_{control}$ control specimens for marker-testing. $N_{control}$ solves the following equation for the CDF for a binomial distribution evaluated at $x_{control}$,

$$\text{CDF}[\text{Binomial}\{N_{control}, (1-P)\}, x_{control}] = 0.95. \quad (2)$$

The specimen-collection sample size, N , is the larger of $N_{control}$ and N_{case} . The formulas for E_{case} and $E_{control}$ indicate that for a small P , as with cancer, E_{case} is usually much larger than $E_{control}$. Correspondingly, N_{case} is usually much larger than $N_{control}$. Consequently, the specimen-collection sample size depends on x_{case} and not on $x_{control}$, which is important for the sample size reduction strategy.

Standard specimen-collection sample size

For the standard specimen-collection sample size, consider the method of Baker and Kramer⁴, which computes the marker-testing sample size by specifying target values

$$FPR_{target} = 0.01 \text{ with } FPR_{upper} = 0.03,$$

$$TPR_{target} = 0.80 \text{ with } TPR_{lower} = 0.70,$$

where FPR_{upper} corresponds to the 2.5% upper bound for the estimated FPR, and TPR_{lower} corresponds to the 2.5% lower bound for the estimated TPR. They computed the marker-testing sample size in controls as 110 because $\text{CDF}[\text{Binomial}\{110, 0.01\}, 0.03 \times 110] = 0.975$. The standard error of TPR evaluated at $TPR_{target} = 0.80$ is $\{0.8(1-0.8)/n\}^{1/2}$. Setting this standard error equal to 0.05, they computed a marker-testing sample size for cases of $n =$

64, which they rounded up to 70. A slightly different calculation based on half the width of the confidence interval, $1.96 \{0.8(1-0.8)/n\}^{1/2} = .80-.70$, gives a marker-testing sample size for cases of 62. Applying equation (1) to the marker-testing sample size for cases of 62, with $P=0.01$, yields the standard specimen-collection sample size of $N_{STD} = 7600$.

Reduced specimen-collection sample size

The reduced specimen-collection sample size is based on the following decision-analytic formulation. Let C denote the cost of a false positive (the anticipated harm from side effects of an incorrect prediction of cancer). Let B denote the benefit of a true positive (the anticipated reduction in cancer mortality from a correct prediction of cancer). Following fundamental work by Peirce,⁵ the anticipated clinical utility of the marker score (evaluated at a cutpoint) is the expected benefit of a true positive minus the expected cost of a false positive,

$$U = TRP \times P \times B - FPR \times (1 - P) \times C. \quad (3)$$

Pepe et al.⁶ recommended choosing target false and true positive values, FPR_{target} and TPR_{target} , to ensure a positive anticipated clinical utility.

A useful desideratum that incorporates the variability of estimates is a positive lower bound on the anticipated clinical utility. Let TPR_{lower} denote a lower bound for the true positive rate and FPR_{upper} denote an upper bound for the false positive rate. The lower bound on the anticipated clinical utility is defined as

$$U_{lower} = TPR_{lower} \times P \times B - FPR_{upper} \times (1 - P) \times C. \quad (4)$$

A positive lower bound for the anticipated clinical utility ($U_{lower} > 0$) requires that

$$TPR_{lower}/FPR_{upper} > \{(1 - P)/P\} \times (C/B). \quad (5)$$

Because equation (5) implies $B/C > \{(1 - P)/P\} / (TPR_{lower}/FPR_{upper})$, the target values for FPR_{upper} and TPR_{lower} apply only when the benefit-cost ratio is larger than

$$(B/C)_{min} = \{(1 - P)/P\} / (TPR_{lower}/FPR_{upper}). \quad (6)$$

$(B/C)_{min}$ is the minimum number of false positives traded for a true positive to ensure a positive lower bound on the anticipated clinical utility.

Investigators can choose any pair of FPR_{upper} and TPR_{lower} that satisfies equation (5) to yield a positive lower bound on the anticipated clinical utility. The key insight is that decreasing both FPR_{upper} and TPR_{lower} while satisfying equation (5) can substantially decrease the sample size. Decreasing FPR_{upper} involves estimating FPR more precisely,

which corresponds to increasing $x_{control}$. Decreasing TPR_{lower} involves estimating TPR less precisely, which corresponds to decreasing x_{case} . Because the specimen-collection sample size with small P depends on x_{case} and not $x_{control}$, decreasing x_{case} decreases the specimen-collection sample size N .

With this motivation, the new specimen-collection sample size calculation specifies target values of

$$FPR_{target} = 0.0 \text{ with } FPR_{upper} = 0.002 \text{ or } 0.005,$$

$$TPR_{target} = 0.40 \text{ with } TPR_{lower} = 0.10, 0.15, \text{ or } 0.20.$$

Here FPR_{upper} corresponds to the 5% upper bound for the estimated FPR because it is one-sided. As with the standard sample size calculation, TPR_{lower} corresponds to the 2.5% lower bound for the estimated TPR. The exact values for these targets are arbitrary, but values for TPR_{target} and TPR_{lower} that are too small may not be convincing.

Although $TPR_{target}=0.40$ may seem small, $TPR_{target}=0.40$ at $FPR_{target}=0$ yields a positive anticipated clinical utility, $U=0.40$ P B, which is 40% of the expected benefit of a correct prediction. Another way to appreciate the desirability of $TPR_{target}=0.40$ at $FPR_{target}=0$ as a target for classification performance is to compute the corresponding positive and negative predictive values. In this context, the positive predictive value (PPV) is the probability of developing clinical cancer given a positive marker score,

$$PPV = TPR \cdot P / \{ TPR \cdot P + FPR(1 - P) \}. \quad (7)$$

In this context, the negative predictive value (NPV) is the probability of not developing clinical cancer given a negative marker score,

$$NPV = (1 - FPR)(1 - P) / \{ (1 - TPR)P + (1 - FPR)(1 - P) \}. \quad (8)$$

With $TPR_{target}=0.40$ and $FPR_{target}=0.0$, equations (7) and (8) yield $PVP_{target}=1$ for all P and $NVP_{target}=0.98$ at P=0.03. These positive and negative predictive values indicate outstanding prediction.

While not an ideal target, the lower bound of $TPR_{lower}=0.10$ is still worthwhile. With $FPR_{target}=0$, it yields a positive anticipated clinical utility, $U=0.10$ P B, with PPV=1 for all P and NPV=0.97 for P=0.03. With $FPR_{upper}=0.005$, it yields a positive anticipated clinical utility by construction, with PPV=0.38 at P=0.03 and NPV=0.97 at P=0.03.

The marker-testing sample size for controls is the value of $x_{control}$ that solves the following equation corresponding to the 5% upper bound for the estimated FPR,

$$(1 - FPR_{upper})^{x_{control}} = 0.05. \quad (9)$$

For $FPR_{upper} = 0.002$, $x_{control} = 1500$. For $FPR_{upper} = 0.005$, $x_{control} = 600$.

The marker-testing sample size is the value for x_{case} that solves the following equation for one-half the width of a 95% confidence interval,

$$1.96 \times \left[\left\{ \left(TPR_{target} (1 - TPR_{target}) \right) / x_{case} \right\}^{1/2} \right] = TPR_{target} - TPR_{lower}. \quad (10)$$

For $TPR_{lower} = 0.10, 0.15, 0.20$, $x_{case} = 11, 15$, and 24 , respectively. Tables 1 and 2 show $(B/C)_{min}$, the specimen-collection sample, the marker-testing sample sizes, the standard specimen collection sample, and predictive values for various values P and TPR_{lower} , with $FPR_{upper} = 0.005$ for Table 1 and $FPR_{upper} = 0.002$ for Table 2. As shown in Table 1, for $TPR_{lower} = 0.10$ with $P = 0.01$, $(B/C)_{min} = 5$ and the specimen-collection sample size is $N = 1800$, a substantial reduction from the standard specimen-collection sample size of 7600.

For the smallest sample sizes, the first step is to consult Table 1 with $TPR_{lower} = 0.10$ to determine $(B/C)_{min}$. If the ratio of anticipated benefits to anticipated costs is than larger than $(B/C)_{min}$, the corresponding specimen-collection sample size in Table 1 applies. Otherwise, it is necessary to consider either a higher risk group (larger P), $TPR_{lower} = 0.15$ or $TPR_{lower} = 0.20$, or $FPR_{upper} = 0.002$.

Example 1: Sample size with pancreatic cancer screening markers

Consider using markers and risk factors to predict, for the purpose of screening, pancreatic cancer among high-risk asymptomatic persons age 50 or older with diabetes. For this high-risk group, the probability of developing pancreatic cancer in the next 3 years is $P = 0.01$.⁷ For $TPR_{lower} = 0.10$ and $P = 0.01$, Table 1 shows $(B/C)_{min} = 5$. To determine if the ratio of anticipated benefits to anticipated costs exceeds this minimum, it is necessary to consider the consequences of a positive marker score. A likely scenario is that a positive marker score would be followed by endoscopic ultrasound imaging. A study of endoscopic imaging found the following probabilities of adverse events among possibly overlapping categories: 44% with pancreatitis, 41% with infections, 8% with bleeding, and 1% with perforations.⁸ The benefit of early intervention following a positive marker score for pancreatic cancer is unknown, but even a small improvement in survival would be important. Based on these considerations, it is likely that the ratio of anticipated benefits to anticipated costs would be larger than $(B/C)_{min} = 5$, implying that one would exchange at least 5 false positives for 1 true positive. Based on Table 1, with $P = 0.01$ and $(B/C)_{min} = 5$, the new specimen-collection sample size is $N = 1800$, which is substantially less than the standard specimen-collection sample size of $N_{STD} = 7600$.

Example 2: Sample size with ovarian cancer screening markers

Consider using various markers and risk factors to predict, for the purpose of screening, invasive epithelial ovarian or tubal cancer (iEOC) among asymptomatic persons. In a four-year study involving asymptomatic women ages 50 or over, the probability of developing iEOC was $P=133/46237=0.003$.⁹ For $TPR_{lower}=0.10$ and $P=0.003$, Table 1 shows $(B/C)_{min}=17$. To determine if the ratio of anticipated benefits to anticipated costs would likely exceed this minimum, it is necessary to consider the consequences of a positive marker score. A positive marker score for ovarian cancer would be followed by surgery, usually laparoscopic bilateral salpingo-oophorectomy.⁹ For every detection of (iEOC), there are approximately 4 surgeries with no iEOC detection and a complication rate of 4.5%.⁹ Recent studies found no reduction in ovarian cancer mortality resulting from current ovarian cancer screening.¹⁰ These considerations are consistent with a previous benefit-cost ratio for markers for ovarian cancer screening⁶ of $B/C=9$, which says that 9 false positives would be traded for 1 true positive. Because B/C is smaller than $(B/C)_{min}$, it is not worthwhile to implement a trial with $TPR_{lower}=0.10$ and $P=0.003$ and $FPR_{upper}=0.005$. A reasonable strategy is to consider $FPR_{upper}=0.002$ in Table 2, yielding a specimen-collection sample size is $N=6000$, which is substantially less than the standard specimen-collection sample size of $N_{STD}=25500$.

Discussion

The methodology does not require costs and benefits to be measured on the same scale because the focus is on trading false positives (with their associated costs) for true positives (with their associated benefits), as opposed to constructing a combined metric of benefits and costs. By analogy, one can trade apples for oranges without assigning a monetary price to each fruit.

To apply this methodology, it is generally *not* necessary to specify the ratio of anticipated benefits to anticipated costs. Instead it is only necessary to perform a simpler task, namely to determine if the ratio of anticipated benefits to anticipated costs is larger than $(B/C)_{min}$, the *minimum* ratio of anticipated benefits to anticipated costs associated P and TPR_{lower} . For perspective, it is worth noting that standard sample size calculation implicitly requires a positive ratio of anticipated benefits to anticipated costs in specifying acceptable target values for FPR and TPR.

The target values of $FPR_{target}=0.0$ and $TPR_{target}=0.40$ suggests the following simple data analysis. First, identify the cutpoint at which $FPR=0$, which is the largest cutpoint among the controls. Second, compute the lower bound of TPR at this cutpoint and deem the cancer prediction model promising if this TPR lower bound is larger than TPR_{lower} .

The analysis corresponds to selecting a FPR and TPR values from the two-dimensional space of acceptable FPR and TPR values specified by an ROC curve. In fact, a more sophisticated analysis plots the ROC curve and reports the optimal point as the point where the slope of the ROC curve equals $(1-P)/P \times (C/B)$. Basic results in decision analysis show that FPR and TPR at this point maximizes U for a concave ROC curve.^{11,12} If the upper bound of FPR at the optimal point is less than or equal to FPR_{upper} and the lower bound of

TPR at this point is greater than or equal to TPR_{lower} , deem the cancer prediction model as promising. Because FPR is small, the conclusions from the two types of analyses should be similar because the slope of the ROC curve at the optimal point is likely close to the ratio of TPR to FPR at the optimal cutpoint.

As mentioned previously, the discovery phase typically uses specimens collected from the persons with symptomatic cancer versus no cancer. The limitation of this usual discovery phase is that it assumes that markers that are good at discriminating between symptomatic cancers and no cancer are will also be good at predicting cancer in asymptomatic persons. To avoid this limitation, if it all possible, investigators should implement the ideal discovery phase involving stored specimens from asymptomatic persons who develop cancer versus asymptomatic persons who do not develop cancer. Consequently, the strategy to the reduce specimen-collection sample size in the validation phase could also substantially reduce the specimen-collection sample size for the ideal discovery phase.

The most important determinant of specimen collection sample is P. If investigators can identify a relevant high-risk group (so P is relatively large) they can obtain smaller sample sizes than otherwise the case. The large reduction in sample size with this new method could allow for prospective testing of markers instead of testing markers in stored specimens, thereby avoiding problems from the degradation of markers in stored samples.

Acknowledgment

The author thanks Mathew Young for helpful comments.

Funding

This research was supported by the National Institutes of Health.

References

1. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y. Phases of marker development for early detection of cancer. *J Natl Cancer Inst* 2001;93(14):1054–1061. [PubMed: 11459866]
2. Baker SG. Improving the marker pipeline to develop and evaluate cancer screening tests. *J Natl Cancer Inst* 2009; 101:1116–1119. [PubMed: 19574417]
3. Young MR, Wagner PD, Ghosh S, Rinaudo JA, Baker SG, Zaret KS, Goggins M, Srivastava S. Validation of markers for early detection of pancreatic cancer summary of The Alliance of Pancreatic Cancer Consortia for Markers for Early Detection Workshop. *Pancreas* 2018; 47(2): 135–141. [PubMed: 29346214]
4. Baker SG, Kramer BS, Srivastava S Markers for early detection of cancer: Statistical issues for nested case-control studies. *BMC Med Res Method* 2002;2:4.
5. Peirce CS. The numerical measure of the success of predictions. *Science*, 1884, 4, 453–454.
6. Pepe MS, Janes H, Li CI, Bossuyt PM, Feng Z, Hilden J. Early-phase studies of markers: What target sensitivity and specificity values might confer anticipated clinical utility? *Clin Chem* 2016 62(5):737–42. [PubMed: 27001493]
7. Chari ST, Leibson CL, Rabe KG, Ransom J, de Andrade M, Petersen GM. Probability of pancreatic cancer following diabetes: a population-based study. *Gastroenterology* 2005 129,:504–11. [PubMed: 16083707]

8. Lee KH, Kim EY, Cho J, Kang D, Bang S, Kim HK, et al. Risk factors associated with adverse events during endoscopic ultrasound-guided tissue sampling. PLoS One 2017;12(12):e0189347. [PubMed: 29236743]
9. Menon U, Ryan A, Kalsi J, Gentry-Maharaj A, Dawnay A, Habib M, et al. Risk algorithm using serial marker measurements doubles the number of screen-detected cancers compared with a single-threshold rule in the United Kingdom collaborative trial of ovarian cancer screening. J Clin Oncol 2015;33:2062–71. [PubMed: 25964255]
10. Pinsky PF, Yu K, Kramer BS, Black A, Buys SS, Partridge E, et al. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15years follow-up. Gynecol Oncol 2016;143:270–275. [PubMed: 27615399]
11. Baker SG and Kramer BS. Peirce, Youden, and Receiver Operating Characteristic curves. Am. Stat 2007;61:343–346.
12. Metz CE. Basic Principles of ROC Analysis, Semin Nucl Med, 1978;VIII:283–298.

Table 1.Specimen-collection sample sizes for $FPR_{upper} = 0.005$

Specified quantities			Derived quantities related to sample size					Predictive values	
TPR_{lower}	P	x_{case}	$(B/C)_{min}$	N_{case}	$N_{control}$	N	N_{STD}	PPV	NPV
0.10	0.03	11	2	600	630	630	2500	0.38	0.973
	0.02		3	900	620	900	3800	0.29	0.982
	0.01		5	1800	620	1800	7600	0.17	0.991
	0.005		10	3600	610	3600	15300	0.09	0.995
	0.003		17	6000	610	6000	25500	0.06	0.997
	0.002		25	9100	610	9100	38200	0.04	0.998
	0.001		50	18200	610	18200	72000	0.02	0.999
0.15	0.03	15	2	700	630	700	2500	0.48	0.974
	0.02		2	1100	620	1100	3800	0.38	0.983
	0.01		4	2300	620	2300	7600	0.23	0.991
	0.005		7	4600	610	4600	15300	0.13	0.996
	0.003		12	7600	610	7600	25500	0.08	0.997
	0.002		17	11500	610	11500	38200	0.06	0.998
	0.001		34	23000	610	23000	72000	0.03	0.999
0.20	0.03	24	1	1100	630	1100	2500	0.55	0.976
	0.02		2	1600	620	1600	3800	0.45	0.984
	0.01		3	3300	620	3300	7600	0.29	0.992
	0.005		5	6700	610	6700	15300	0.17	0.996
	0.003		9	11200	610	11200	25500	0.11	0.998
	0.002		13	16800	610	16800	38200	0.07	0.998
	0.001		25	33700	610	33700	72000	0.04	0.999

P is the probability of developing cancer during the study.

$(B/C)_{min}$ is minimum benefit cost ratio (number of false positives traded for a true positive) to ensure a positive lower bound for the anticipated clinical utility.

x_{case} is the marker-testing sample size for cases.

N_{case} is the specimen-collection sample size for cases (rounded to the nearest 100).

$N_{control}$ is the specimen-collection sample size for controls (rounded to the nearest 10).

N is the final specimen-collection sample size = $\max(N_{case}, N_{control})$ for the new calculation.

N_{STD} is the final specimen-collection sample size for the standard calculation that depends only on P in this table.

PPV is the positive predictive value for P with TPR_{low} and $FPR_{upper}=0.005$.

NPV is the negative predictive value for P with TPR_{low} and $FPR_{upper}=0.005$.

Table 2.Specimen-collection sample sizes for $FPR_{upper}=0.002$

Specified quantities			Derived quantities related to sample size					Predictive values	
TPR_{lower}	P	x_{case}	$(B/C)_{min}$	N_{case}	$N_{control}$	N	N_{STD}	PPV	NPV
0.10	0.03	11	1	600	1560	1560	2500	0.61	0.973
	0.02		1	900	1550	1550	3800	0.51	0.982
	0.01		2	1800	1530	1800	7600	0.34	0.991
	0.005		4	3600	1520	3600	15300	0.2	0.995
	0.003		7	6000	1510	6000	25500	0.13	0.997
	0.002		10	9100	1510	9100	38200	0.09	0.998
	0.001		20	18200	1510	18200	72000	0.05	0.999
0.15	0.03	15	1	700	1560	1560	2500	0.7	0.974
	0.02		1	1100	1550	1550	3800	0.6	0.983
	0.01		2	2300	1530	2300	7600	0.43	0.991
	0.005		3	4600	1520	4600	15300	0.27	0.996
	0.003		5	7600	1510	7600	25500	0.18	0.997
	0.002		7	11500	1510	11500	38200	0.13	0.998
	0.001		14	23000	1510	23000	72000	0.07	0.999
0.20	0.03	24	1	1100	1560	1560	2500	0.76	0.976
	0.02		1	1600	1550	1600	3800	0.67	0.984
	0.01		1	3300	1530	3300	7600	0.5	0.992
	0.005		2	6700	1520	6700	15300	0.33	0.996
	0.003		4	11200	1510	11200	25500	0.23	0.998
	0.002		5	16800	1510	16800	38200	0.17	0.998
	0.001		10	33700	1510	33700	72000	0.09	0.999

P is the probability of developing cancer during the study.

$(B/C)_{min}$ is minimum benefit cost ratio (number of false positives traded for a true positive) to ensure a positive lower bound for the anticipated clinical utility.

x_{case} is the marker-testing sample size for cases.

N_{case} is the specimen-collection sample size for cases (rounded to the nearest 100).

$N_{control}$ is the specimen-collection sample size for controls (rounded to the nearest 10).

N is the final specimen-collection sample size = $\max(N_{case}, N_{control})$ for the new calculation.

N_{STD} is the final specimen-collection sample size for the standard calculation that depends only on P in this table.

PPV is the positive predictive value for P with TPR_{low} and $FPR_{upper}=0.002$.

NPV is the negative predictive value for P with TPR_{low} and $FPR_{upper}=0.002$.