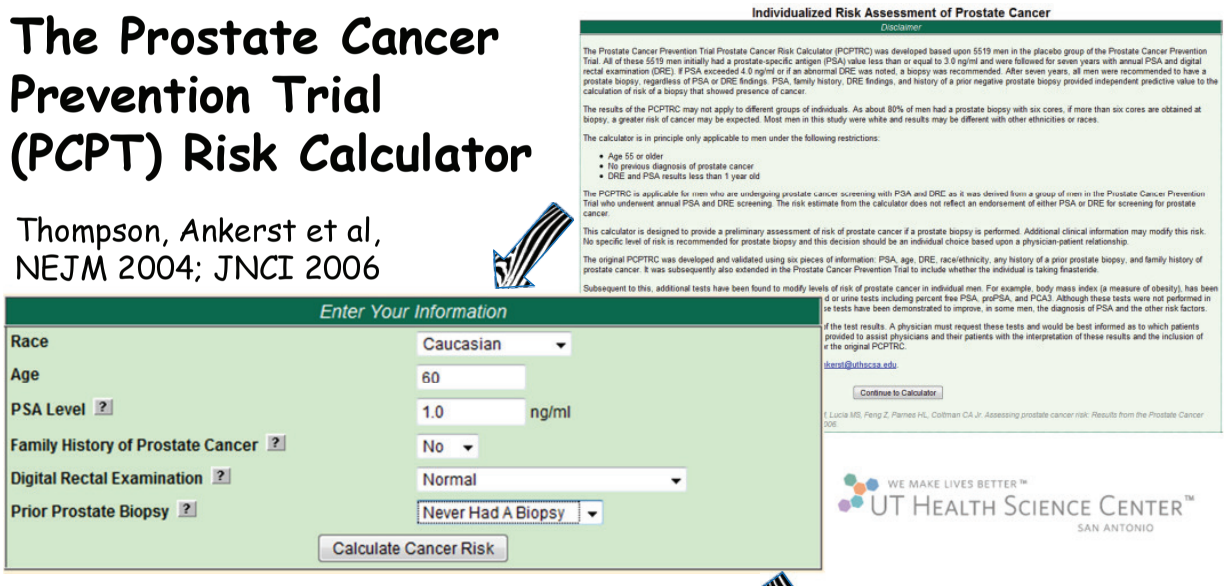
Developments in prostate cancer risk prediction tools in response to changes in clinical landscape

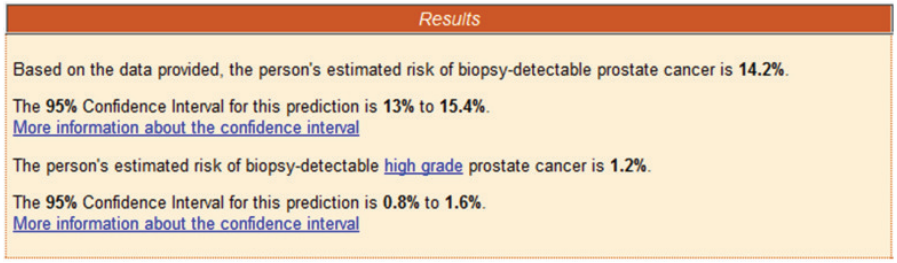
Donna Pauler Ankerst
Mathematics, Technische Universitaet Muenchen,
Urology and Epidemiology/Biostatistics, University of Texas Health Science Center at San Antonio, San Antonio, Texas

The Prostate Cancer Prevention Trial (PCPT) Risk Calculator

Thompson, Ankerst et al,
NEJM 2004; JNCI 2006



www.prostate-cancer-risk-calculator.com



Definitions/formulas

PSA: enter **prostate-specific antigen** in ng/mL
DRE: enter 1 if **digital rectal examination** is abnormal, 0 otherwise
FAMHIST: enter 1 if there is a first-degree family history of prostate cancer, 0 otherwise
PRIORBIOP: enter 1 if there has been one or more prior biopsies performed (all negative for prostate cancer), 0 otherwise
AA: enter 1 for African American, 0 otherwise
AGE: enter age in years

Separate logistic regressions

For Prostate Cancer Risk:
PCA = -1.7968 + 0.8488 × log(PSA) + 0.2693 × FAMHIST + 0.9054 × DRE -0.4483 × PRIORBIOP
Prostate Cancer Risk = 1/[1+exp(-PCA)]
For High-grade Prostate Cancer (Gleason ≥ 7) Risk:
HG= -6.2461 + 1.2927 × log(PSA) + 0.0306 × AGE + 1.0008 × DRE
+ 0.9604 × AA -0.3634 × PRIORBIOP
High-grade Risk = 1/[1+exp(-HG)]

Completion of randomized trials and studies have brought about a change in the clinical landscape since 2006

- Recognition that validation of a risk tool is a property of both the tool and the cohort to which it is applied.
- Discovery and validation of new biomarkers for the early detection of prostate cancer.
- Quantification of a significant degree of **overdetection**: detection and treatment of prostate cancers that would not have caused mortality in the patient's lifetime.

Completion of randomized trials and studies have brought about a change in the clinical landscape since 2006

- Recognition that validation of a risk tool is a property of both the tool and the cohort to which it is applied.
- Discovery and validation of new biomarkers for the early detection of prostate cancer.
- Quantification of a significant degree of **overdetection**: detection and treatment of prostate cancers that would not have caused mortality in the patient's lifetime.

Prostate Biopsy Collaborative Group (PBCG)

Andrew Vickers
Memorial Sloan-Kettering Cancer Center

- 7 European, 3 US biopsy cohorts
- 25,772 biopsies from 23,070 patients
- 8,503 prostate cancers

AIM: Validation is a property of BOTH the prediction tool and the cohort to which it is applied.

Vickers et al., Clinical Cancer Research, 2010

VOLUME 29 · NUMBER 22 · AUGUST 1 2011

JOURNAL OF CLINICAL ONCOLOGY | E D I T O R I A L S

Prediction Models: Revolutionary in Principle, But Do They Do More Good Than Harm?

Andrew J. Vickers, *Memorial Sloan-Kettering Cancer Center, New York, NY*

See accompanying article on page 2959; listen to the podcast by Dr Cooperberg at www.jco.org/podcasts

It can sometimes seem as though we are drowning in prediction models. Every month brings a multitude of newly published risk calculators and nomograms to add to the multitude already in the literature—there are more than 100 prediction models on prostate cancer alone[1]—and Web sites such as www.nomograms.org, www.nomogram.org, and www.cancernomograms.org continue to proliferate. As such, it is easy to become somewhat inured to prediction modeling and thus to forget that it constitutes an important shift in the way that medicine is practiced.

prostate cancer have a risk threshold of 20%; those who would do no more than four have a risk threshold of 25%. Patients with risks higher than those thresholds from a prediction model would accordingly be advised to consider biopsy.

Similarly, predicted risk can be used to individualize decision making. Given a patient averse to biopsy, the results of a prediction model can be used as part of shared decision making. Informing a patient that his risk is 60% versus 26% is far more conducive to decision making than reporting a PSA of 11 ng/mL versus one of

Table 1. Description of study cohorts

Name of cohort	Location	Type of cohort	Biopsy algorithm		Biopsy scheme	Prior screening
			Indication for biopsy	Decision for biopsy a clinical decision?		
ERSPC Göteborg Round 1	Sweden	Screening	PSA ≥3 ng/mL	No	6-core*	No
ERSPC Göteborg Rounds 2-6	Sweden	Screening	PSA ≥3 ng/mL	No	6-core*	Yes
ERSPC Rotterdam Round 1	The Netherlands	Screening	PSA ≥3 ng/mL or ≥4 ng/mL, depending on year	No	6-core*	No
ERSPC Rotterdam Rounds 2-3	The Netherlands	Screening	PSA ≥3 ng/mL or ≥4 ng/mL†	No	6-core*	Yes
ERSPC Tarn Round 1	France	Screening	PSA ≥3 ng/mL	Yes	Primarily 10- to 12-core	Mixture
SABOR	San Antonio, TX	Screening	PSA ≥2.5 ng/mL, abnormal DRE, or family history	Yes	10- to 12-core	Mixture
Cleveland Clinic	Cleveland, OH	Clinical	Elevated PSA, abnormal DRE, rapid rise in PSA	Yes	Primarily ≥8-core	Mixture
ProtecT	United Kingdom	Screening‡	PSA ≥3 ng/mL	No	10-core	No
Tyrol	Austria	Screening‡	PSA ≥1.25 ng/mL, percent free PSA, abnormal DRE	Most men with elevated PSA were biopsied	6-, 10-, or 10- to 15-core†	Mixture
Durham VA	Durham, NC	Clinical	Elevated PSA, abnormal DRE	Yes	6-, 10-, or 12-core †	Mixture
PCPT	U.S.	Screening	PSA ≥4ng/mL or abnormal DRE for "for cause" biopsies; end of study biopsy offered to all men	In the case of "for cause" biopsies	Primarily 6-core	Yes

Externally validate the PCPTRC/HG by 3 criteria

Steyerberg E. Clinical Prediction Models, Springer, 2009

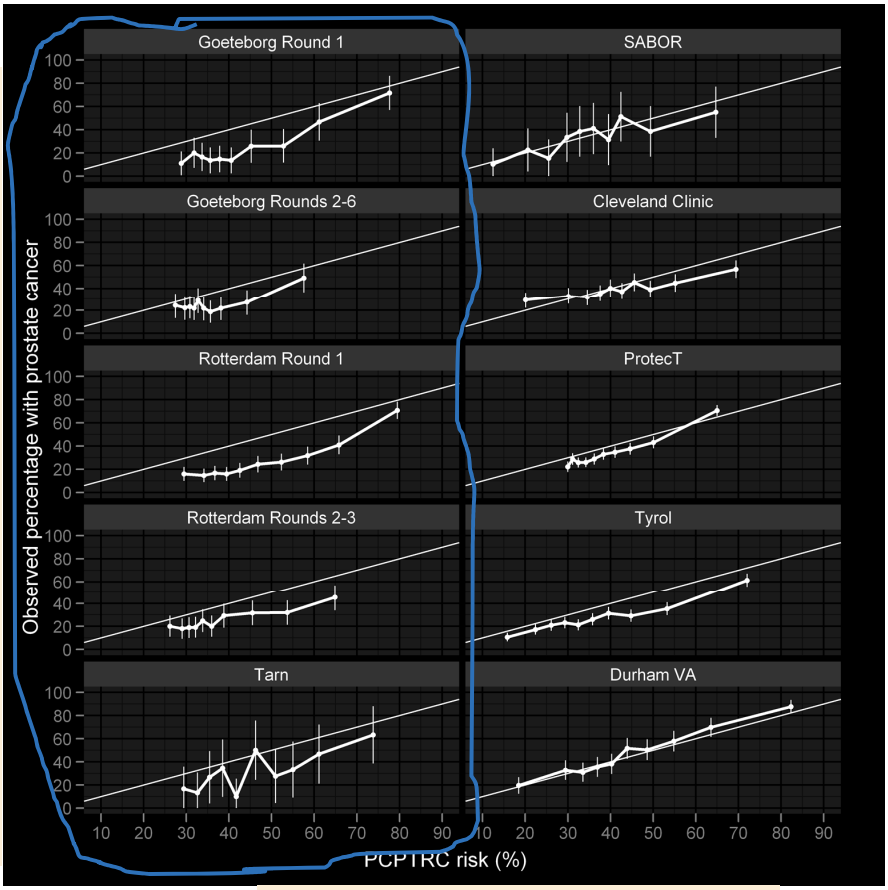
1.) **Calibration**: How close are predicted risks to observed risks? Calibration curves
2.) **Discrimination**: How well does a risk prediction discriminate between those with and without the disease? Area underneath the receiver operating characteristic curves
3.) **Clinical net benefit**: Decision-curve analysis that compares the net benefit of using a risk prediction tool to refer patients to biopsy versus referring all or no patients to biopsy. Net benefit curves

There are many more, some such as the Brier score, combine multiple metrics; these 3 are most seen in Urology.

Calibration of the PCPTRC

Plot of predicted risks on x-axis versus observed risks (grouped by deciles) on y-axis. Diagonal line indicates perfect fit.

The PCPTRC overpredicts for the European cohorts at all levels of low to high risk.



Ankerst et al, World J Urol, 2012

Discrimination of the PCPTRC

Area underneath the receiver-operating-characteristic curve (AUC) gives the probability that for a randomly selected cancer case and control, the cancer case would have a higher PCPTRC risk. It varies from 50% (no better than random guessing) to 100% (perfect).

AUC varies from 56.2% to 72.0%, a bigger range than any new biomarker has ever moved an AUC.

Cohort (*n*)	Discrimination AUC PCPTRC (%) (*P*-value for comparison to the AUC of PSA)
ERSPC Goeteborg Round 1 ($n = 740$)	72.0 (<0.0001)
ERSPC Goeteborg Rounds 2–6 ($n = 1,241$)	56.2 (<0.0001)
ERSPC Rotterdam Round 1 ($n = 2,895$)	70.0 (<0.0001)
ERSPC Rotterdam Rounds 2–3 ($n = 1,494$)	61.0 (0.15)
ERSPC Tarn ($n = 298$)	66.7 (0.07)
SABOR, US ($n = 392$)	65.4 (0.20)
Cleveland Clinic, US ($n = 3,286$)	58.8 (<0.0001)
ProtecT, UK ($n = 7,324$)	63.9 (0.14)
Tyrol, Austria ($n = 5,644$)	66.7 (<0.0001)
Durham VA, US ($n = 2,419$)	71.5 (<0.0001)

Clinical net benefit

value

Get a biopsy → **a** if he has cancer (**p**)
b if he does not (**1-p**)

No biopsy → **c** if he has cancer (**p**)
d if he does not (**1-p**)

The threshold p_t that a patient should use is the one where his expected value of getting a biopsy is equal to the expected value of not getting the biopsy.

Value of a biopsy = Value of no biopsy

$$p_t a + (1 - p_t) b = p_t c + (1 - p_t) d$$

$$p_t (a - c) = -(1 - p_t)(b - d)$$

Clinical net benefit

value

Get a biopsy → **a** if he has cancer (**p**)
b if he does not (**1-p**)

No biopsy → **c** if he has cancer (**p**)
d if he does not (**1-p**)

$$p_t (a - c) = -(1 - p_t)(b - d)$$

Fix $a - c = 1$, then $b - d = -\dfrac{p_t}{1 - p_t}$.

- a – c = "value of a true positive result": if you **have cancer**, relative weight of biopsy versus not.
- b – d = "value of a false positive result": if you **do not have cancer**, relative weight of biopsy versus not.

Net benefit

Expected benefit of using risk $> p_t$ to decide for biopsy =

$$P(\text{true positive}) \times \text{Value of true positive} + P(\text{false positive}) \times \text{Value of false positive}$$

$$\frac{\#\text{True positives}(p_t)}{\text{N}} \times 1 \quad + \quad \frac{\#\text{False positives}(p_t)}{\text{N}} \times \left(- \frac{p_t}{1 - p_t} \right)$$

$$\frac{\#\text{True positives}(p_t)}{\#\text{Cancer cases}} \frac{\#\text{Cancer cases}}{\text{N}} - \frac{p_t}{1 - p_t} \frac{\#\text{False positives}(p_t)}{\#\text{Non - cancers}} \frac{\#\text{Non - cancers}}{\text{N}}$$

$$\text{Sensitivity}(p_t)(\text{Prevalence}) - \frac{p_t}{1 - p_t}[1 - \text{Specificity}(p_t)](1 - \text{Prevalence}).$$

Since prevalence, sensitivity and specificity vary by cohort, net benefit will also vary by cohort.

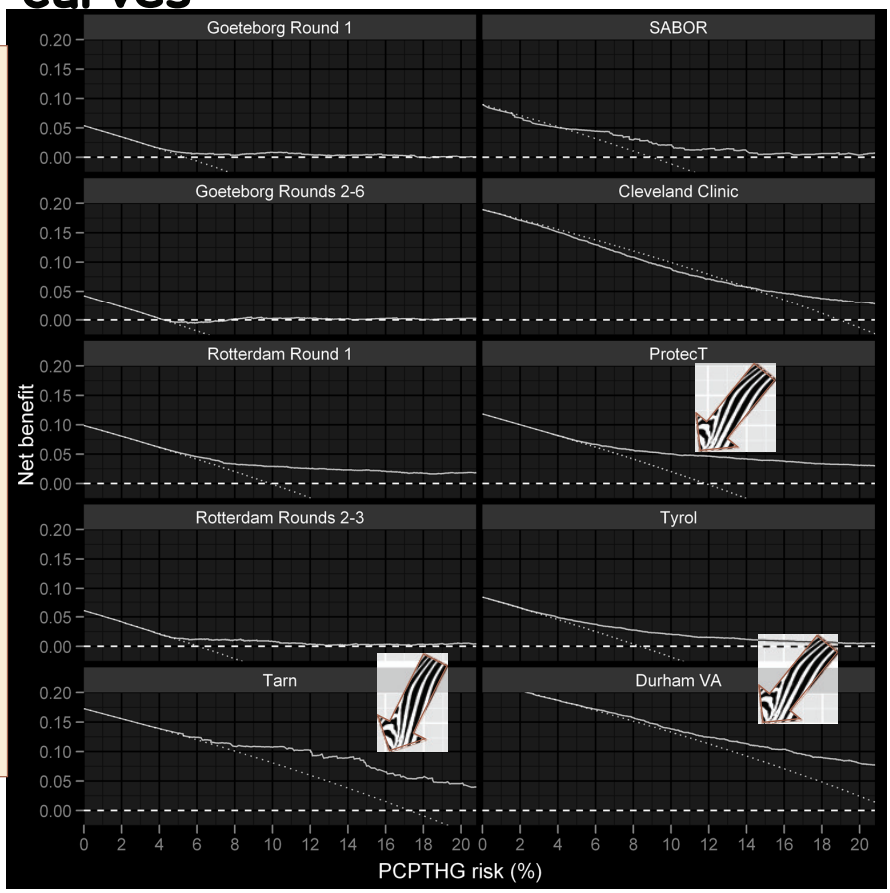
Net benefit curves

Solid line: use PCPTHG for high-grade prostate cancer (when risk exceeds threshold) to decide on biopsy

Dotted line: do biopsy regardless

Dashed line: do not biopsy regardless

In three cohorts, benefit of using the PCPTHG for risk thresholds > 8%.



Completion of randomized trials and studies have brought about a change in the clinical landscape since 2006

- Recognition that validation of a risk tool is a property of both the tool and the cohort to which it is applied.
- Discovery and validation of new biomarkers for the early detection of prostate cancer.
- Quantification of a significant degree of **overdetection**: detection and treatment of prostate cancers that would not have caused mortality in the patient's lifetime.

Updating an existing risk tool

- Prostate cancer biomarker research is dynamic.
- New markers are discovered/tested/validated.
- Cannot measure these markers retrospectively on the 5519 participants of the PCPT.

Problem to be solved

- How to update a risk calculator built on one cohort with a new risk factor measured on a different cohort?

Solution

- Bayes theorem

From prior to posterior risk

X = PCPT Risk factors: PSA, DRE, family history, prior biopsy
Y = New markers

Posterior Odds Cancer(Y,X) = Likelihood Ratio(Y|X) × Prior Odds Cancer (X)

within a given strata of X, how much more likely is the new marker to be observed in cases rather than controls; estimated from a separate study to PCPT

from PCPT risk calculator

$$\frac{P(Cancer \mid X, Y)}{P(No\ Cancer \mid X, Y)} = \frac{P(Y \mid X, Cancer)}{P(Y \mid X, No\ Cancer)} \times \frac{P(Cancer \mid X)}{P(No\ Cancer \mid X)}$$

Confidence,prediction intervals for posterior risk by delta rule.

Appropriate distribution for new markers

PCPT logistic regressions

Single continuous marker

X = PCPT Risk factors: PSA, DRE, family history, prior biopsy
Y = log(PCA3)

$$\frac{P(Cancer \mid X, Y)}{P(No\ Cancer \mid X, Y)} = \frac{P(Y \mid X, Cancer)}{P(Y \mid X, No\ Cancer)} \times \frac{P(Cancer \mid X)}{P(No\ Cancer \mid X)}$$

Linear regressions of Y on X in cancer cases and controls separately.

PCPT logistic regression

$$\frac{\frac{1}{\sqrt{\sigma^2_{cancer}}}\exp\left\{-\frac{1}{2\sigma^2_{cancer}}(Y-\mu_{cancer})^2\right\}}{\frac{1}{\sqrt{\sigma^2_{no\ cancer}}}\exp\left\{-\frac{1}{2\sigma^2_{no\ cancer}}(Y-\mu_{no\ cancer})^2\right\}}$$

$$\exp(\beta' X)$$

$$\beta' X = -1.7968 + 0.8488\log(psa) + 0.2693famhist + 0.9054dre - 0.4483priorbiop$$

$$\mu_{cancer} = 1.1926 - .0836\log(psa) + .0376age + .1055dre + .0658priorbiop$$

$$\mu_{no\ cancer} - .6915 - .1137\log(psa) + .0577age - .3345dre + .1260priorbiop$$

$$\sigma_{cancer} = 1.0366$$

$$\sigma_{no\ cancer} = 1.0191$$

Multiple markers → Multivariate normal distributions
Ankerst et al, J Urol 2009,
Ankerst et al, Biometrical J 2012

Single nucleotide polymorphisms

X = PCPT Risk factors: PSA, DRE, family history, prior biopsy; we believe that mutations are inherited or occur before X and so do not need to condition on X.

Y = SNP with published genotype or allele frequencies (example A,G).

$$\frac{P(\text{Cancer} \mid X, Y)}{P(\text{No Cancer} \mid X, Y)} = \frac{P(Y \mid X, \text{Cancer})}{P(Y \mid X, \text{No Cancer})} \times \frac{P(\text{Cancer} \mid X)}{P(\text{No Cancer} \mid X)}$$

Published GWAS study

PCPT

$$\frac{\left(\pi^{AA}_{cancer}\right)^{I(Y=AA)}\left(\pi^{GA,AG}_{cancer}\right)^{I(Y=GA,AG)}\left(\pi^{GG}_{cancer}\right)^{I(Y=GG)}}{\left(\pi^{AA}_{no\ cancer}\right)^{I(Y=AA)}\left(\pi^{GA,AG}_{no\ cancer}\right)^{I(Y=GA,AG)}\left(\pi^{GG}_{no\ cancer}\right)^{I(Y=GG)}}$$

$$\approx \left(\frac{\hat{\pi}^{AA}_{cancer}}{\hat{\pi}^{AA}_{no\ cancer}}\right)^{I(Y=AA)}\left(\frac{\hat{\pi}^{GA,AG}_{cancer}}{\hat{\pi}^{GA,AG}_{no\ cancer}}\right)^{I(Y=GA,AG)}\left(\frac{\hat{\pi}^{GG}_{cancer}}{\hat{\pi}^{GG}_{no\ cancer}}\right)^{I(Y=GG)}$$

$$\exp(\beta' X)$$

$$\beta' X = -1.7968 + 0.8488\log(\text{psa}) + 0.2693\text{famhist} + 0.9054\text{dre} - 0.4483\text{priorbiop}$$

Multiple SNPs not in LD → multiply LR's

Completion of randomized trials and studies have brought about a change in the clinical landscape since 2006

- Recognition that validation of a risk tool is a property of both the tool and the cohort to which it is applied.
- Discovery and validation of new biomarkers for the early detection of prostate cancer.
- Quantification of a significant degree of **overdetection**: detection and treatment of prostate cancers that would not have caused mortality in the patient's lifetime.

PCPTRC 2.0: A response to overdetection concerns

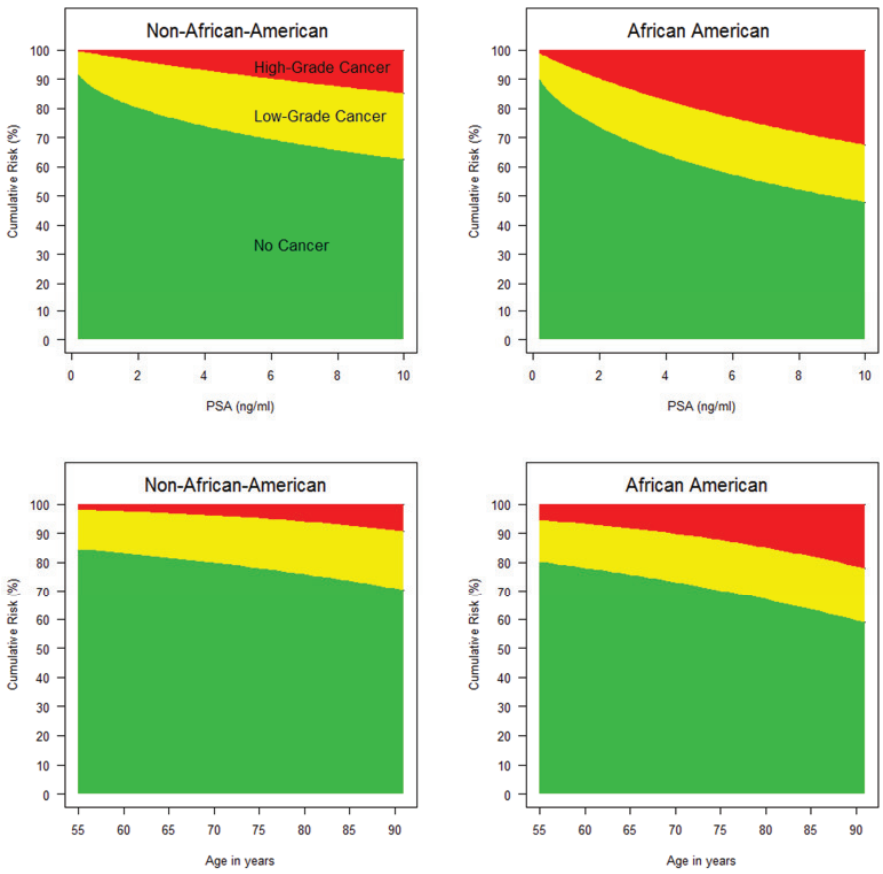
Risk of three outcomes if biopsy were to be performed:

- No cancer
- low-grade prostate cancer (Gleason grade < 7)
- high-grade prostate cancer (Gleason grade ≥ 7)

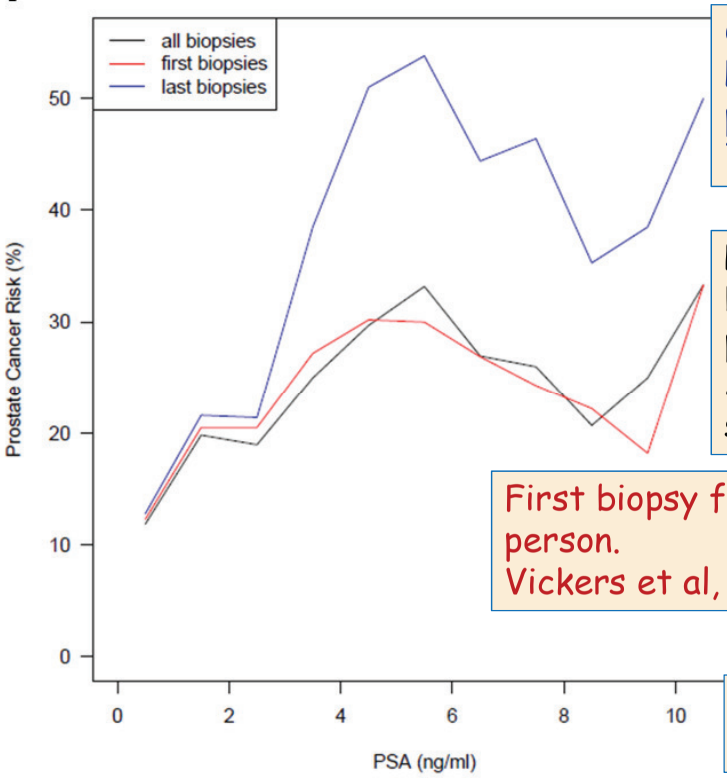
6664 PCPT biopsies	No Cancer	Low-grade cancer	High-grade cancer
N	5468	942	254
% of N	82.1	14.1	3.8

Use nominal logistic regression; fits better than ordinal logistic regression

Final chosen model had same risk factors as the orginal PCPTRC



Empirical smoothed risks according to PCPT biopsies selected for the denominator



Original PCPTRC, Last biopsy for each participant Thompson et al, 2006.

PCPTRC 2.0: all biopsies for each person. Ankerst et al, submitted.

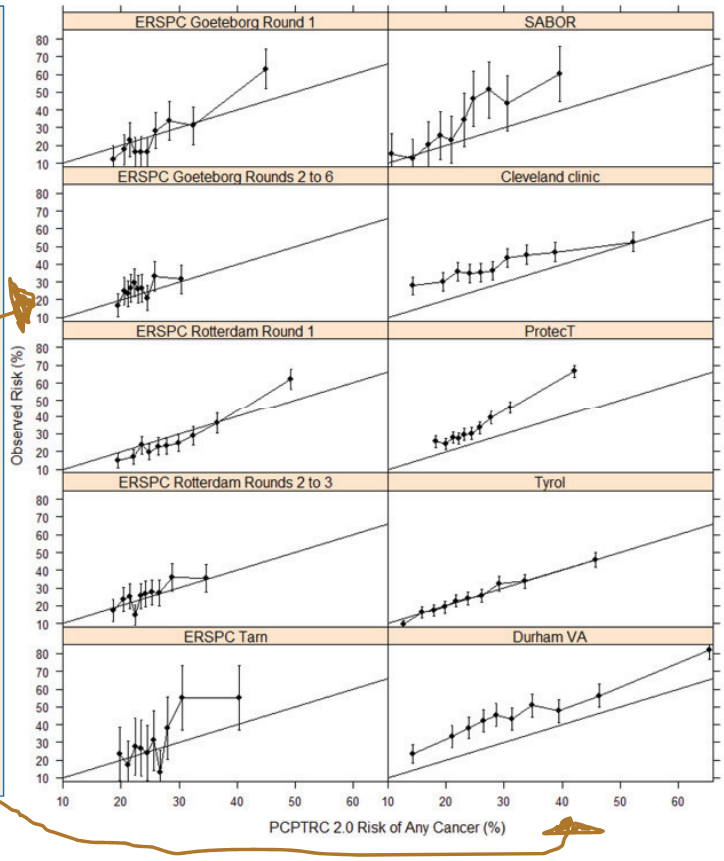
First biopsy for each person. Vickers et al, 2013.

Changing denominator improved calibration.

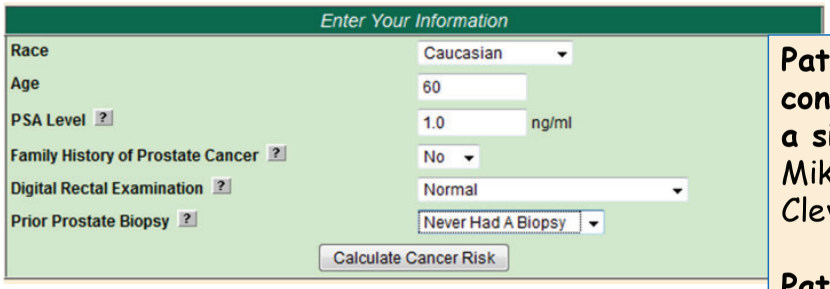
Calibration of PCPTRC 2.0

No longer over-fitting for the European screening cohorts on the left.

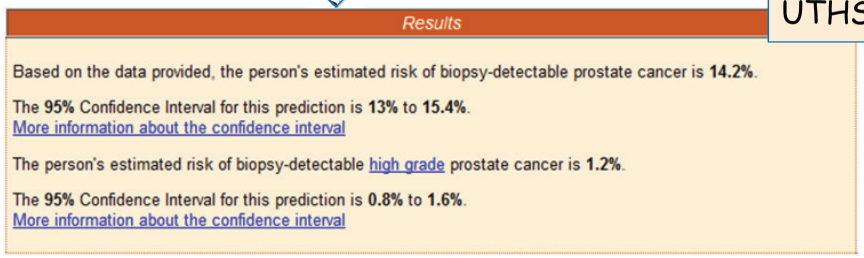
Under-fitting for the clinical cohorts on the right—that is to be expected since these men are referred with symptoms.



A more complicated message (low- versus high-grade disease) requires a more simple explanation. www.prostate-cancer-risk-calculator.com



PCPTRC 1.0

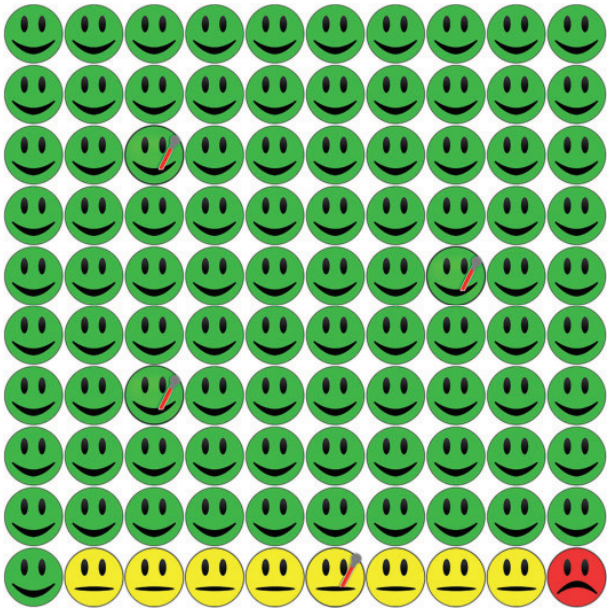


Patients do not understand confidence intervals, want a single number.
Mike Kattan,
Cleveland Clinic

Patients don't even want a number.
Ian M. Thompson,
UTHSCSA

PCPTRC 2.0: A new more patient-friendly display of results.

PCPTRC 2.0



Please consult your physician concerning these results. Click **here** to watch a video overview of these results.

Based on the provided risk factors a prostate biopsy performed would have a:

1% chance of high-grade prostate cancer,

8% chance of low-grade cancer,

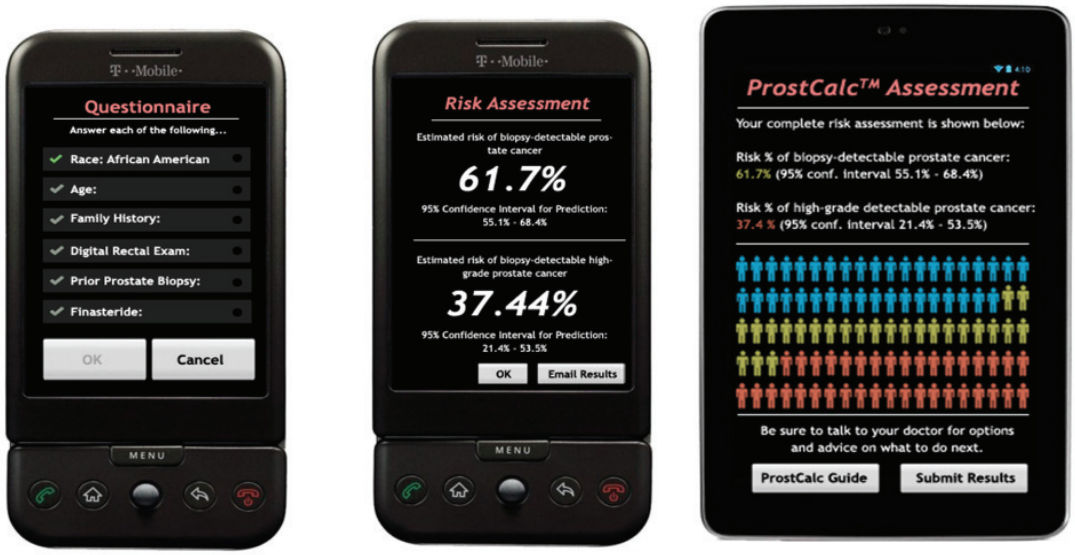
91% chance that the biopsy is negative for cancer.

About 2 to 4% of men undergoing biopsy will have an infection that may require hospitalization.

PCPTRC 2.0: 27 minute video

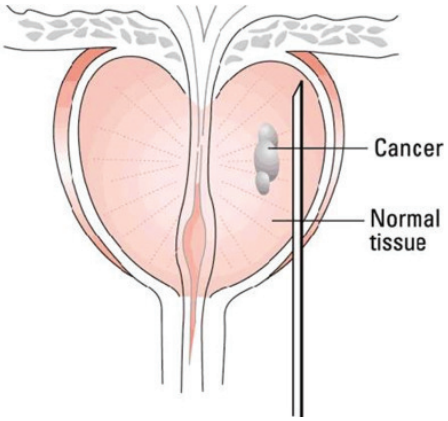
http://youtu.be/LZC1erZV5Gc

Texas CIPRIT Grant in collaboration with Engineers at RICE



A handheld device because doctors use these when talking to patients and are not often at their computer.

Current work: changes in clinical practice, what happens when your cohort becomes outdated?



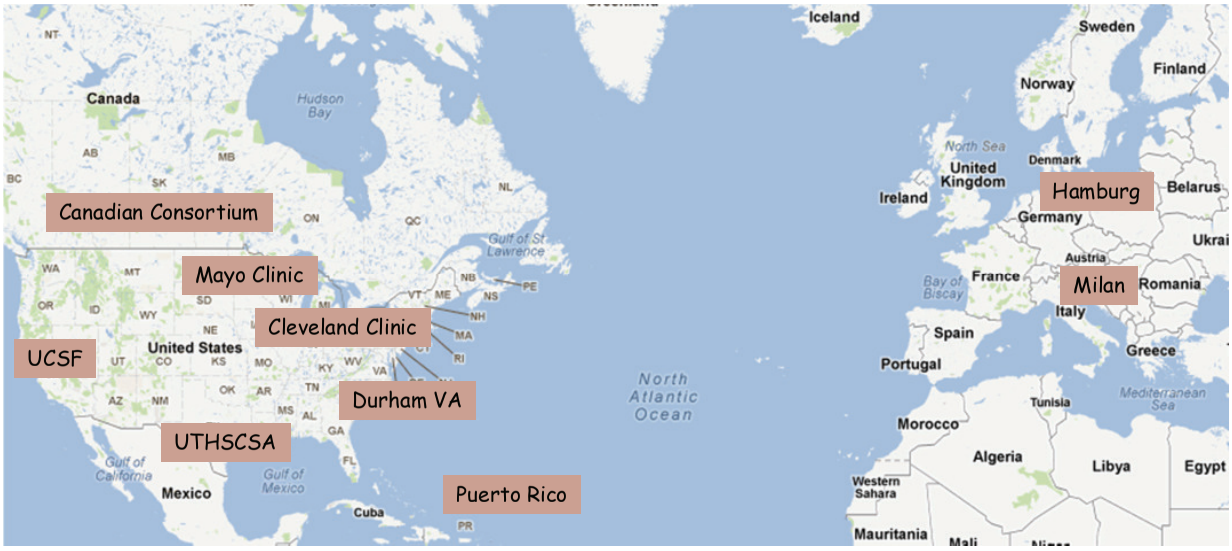
The PCPT cohort was collected from the late 1990's through 2004.

The PCPT protocol for the biopsy procedure was a 6-core sample, but modern practice collects 12- or even more cores.

It has been documented that a higher number of cores increases the likelihood of detecting cancer and high-grade cancer.

Prostate Biopsy Collaborative Group PBCG 2.0

Cheaper to build a new house or keep fixing up the old one?



Data elements: Same as before but now ask if ever had a prior elevated PSA test; insert causal inference to model ascertainment bias (collaboration with Univ. Michigan)

Steyerberg recalibration versus Bayesian methods

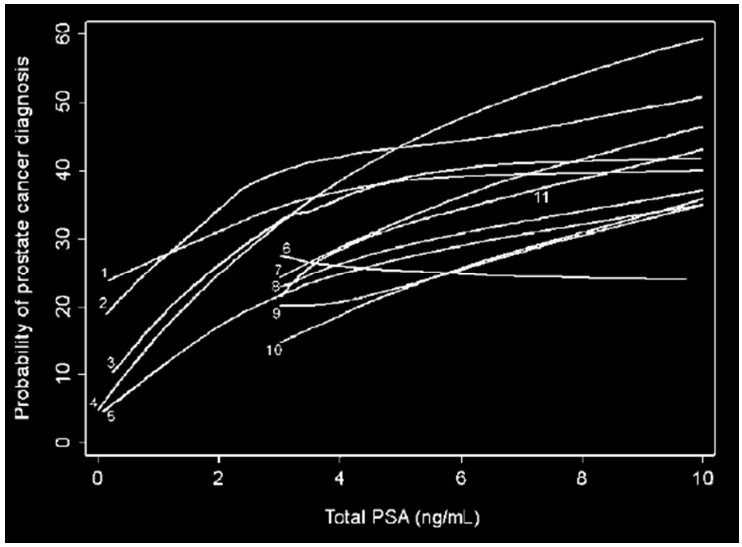
Methods for recalibrating an existing risk prediction tool on a new data set each year will be compared to building from scratch:

Recalibration in the large: Use log PCPTRC 2.0 risk as offset and estimate new intercept in nominal logistic regression (NLR)
Recalibration: NLR to estimate new intercepts and slopes for log PCPTRC 2.0 risk as single covariate
Revision: Same as recalibration but allow individual risk factors to enter separately as covariates
Bayesian: Use prior to posterior updating on parameters
Bayesian likelihood ratio: Use PCPTRC 2.0 as prior odds. 31

Big Data needs Big Cleaning, Smart Statisticians:

Lots of biopsies running through Electronic Medical Records overnight but no one is processing it.

Empirical risk curves according to PSA across 11 cohorts in the PBCG 1.0.
Vickers et al., Clinical Cancer Research, 2010



Need for automatic downloading, quality-checking, updating of a risk tool, correction for ascertainment bias to remove the cohort effect.

Acknowledgements

TUM: Sebastian Bock, Josef Hoefler
UTHSCSA: Russell MacShane, Robin Leach, Ian M. Thompson

And Dornbirn, Austria