

Measuring and Improving Radiologists' Interpretative Performance on Screening Mammography

Karla Kerlikowske, MD

Diana Buist, PhD

Patricia Carney, PhD

Berta Geller, EdD

Diana Miglioretti, PhD

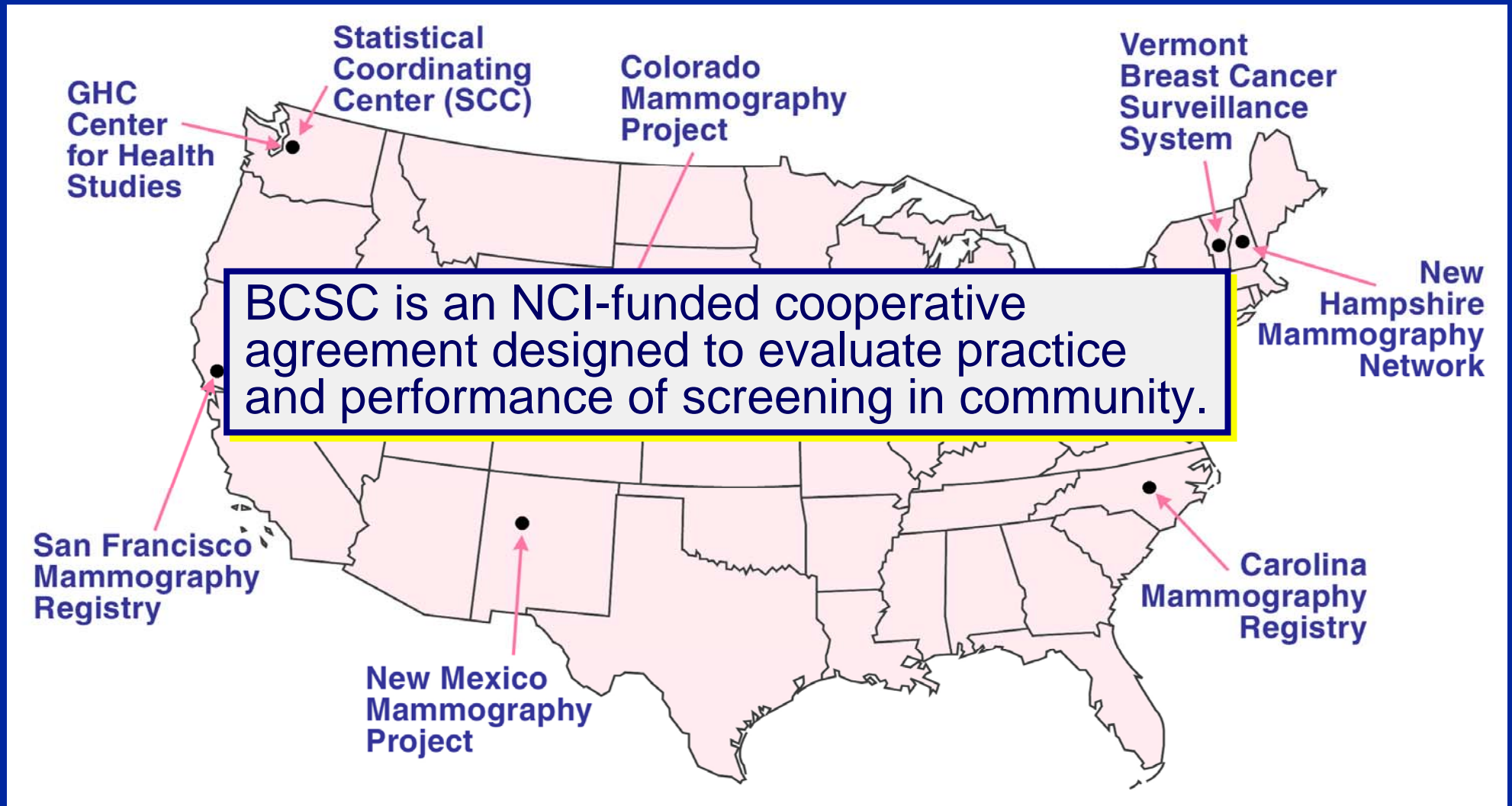
Robert Rosenberg, MD

Bonnie Yankaskas, PhD

Institute of Medicine Report

- ***Improving Breast Imaging Quality Standards***
 - Technical quality of mammography in the U.S. has improved since implementation of the Mammography Quality Standards Act, mammography interpretation remains quite variable.
- **We plan to study how best to reduce variability and improve interpretive performance among US radiologists.**

Breast Cancer Surveillance Consortium



Size of the Pooled BCSC Data Resource

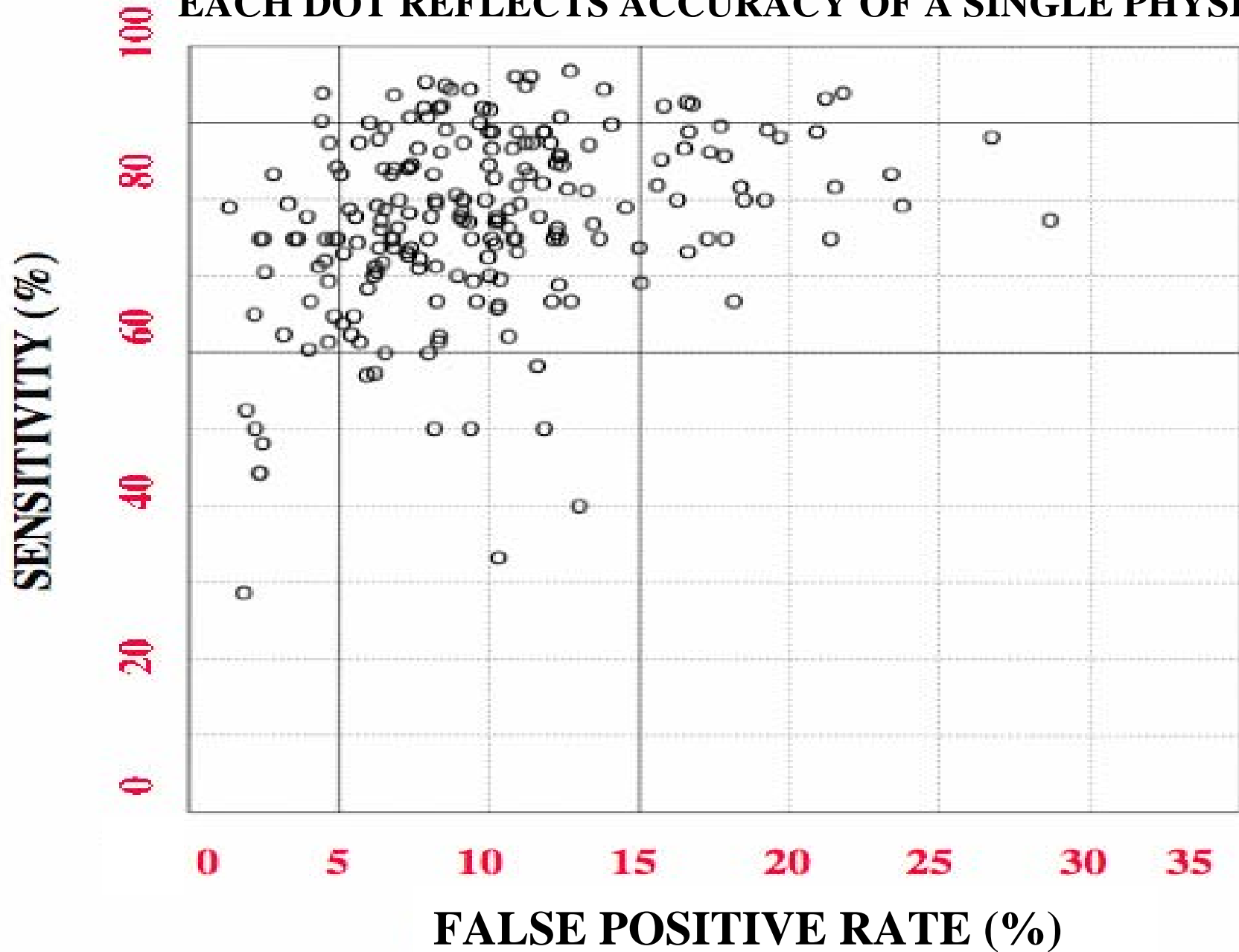
- **Based on 1996-2004 examinations**
 - Total mammograms = 5.4 million
 - Total women = 1.9 million
- **Cancer data**
 - Invasive cancers = 59,303
 - In situ cancers = 10,858
- **Number of radiologists**
 - 972

Mammography performance is highly variable across radiologists in BCSC

Variability in Screening Mammography Performance

Measure	No. Radiologists	Average (Range) %	Reference
Recall rate	344	10 (1 – 25)	Rosenberg, In press
PPV2	330	25 (4 – 52)	Rosenberg, In press
Sensitivity	208	77 (29 – 97)	Smith-Bindman, 2005
Specificity	209	90 (71 – 99)	Smith-Bindman, 2005
	124	90 (74 – 98)	Barlow, 2004

EACH DOT REFLECTS ACCURACY OF A SINGLE PHYSICIAN



Recall rate 2-fold higher in U.S. versus U.K.

Age	Recall per 100			Cancer per 1000		
	<u>UK</u>	<u>BCSC</u>	<u>CDC</u>	<u>UK</u>	<u>BCSC</u>	<u>CDC</u>
50-54	3.9	8.7	8.0	3.8	2.6	2.8
55-59	3.6	8.3	7.0	4.9	3.6	3.5
60-64	3.4	7.9	6.7	5.9	3.9	3.7

Physician Characteristics Associated with Clinical Screening Performance

Characteristic	Association with Performance	Reference
Years of Experience	<p>↓ FP, no Δ TP</p> <p>↓ FP, ↓ TP</p> <p>↓ FP</p>	<p>Smith-Bindman, 2005</p> <p>Barlow, 2004</p> <p>Elmore, 2002</p>
Volume	<p>↓ FP, no Δ TP</p> <p>↑ FP, ↑ TP</p> <p>↓ FP, no Δ CDR</p> <p>↓ FP, no Δ or ↑ CDR</p> <p>↑ PPV, no Δ CDR</p>	<p>Smith-Bindman, 2005</p> <p>Barlow, 2004</p> <p>Théberge (Quebec), 2005</p> <p>Kan (BC), 2000</p> <p>Coldman (Canada), 2006</p>
Screening Focus	<p>↓ FP, ↓ TP</p> <p>no Δ FP or TP</p>	<p>Smith-Bindman, 2005</p> <p>Barlow, 2004</p>
Specialists	<p>↓ FP, ↑ TP</p>	<p>Sickles, 2002</p>

Goals of ACS-NCI Project

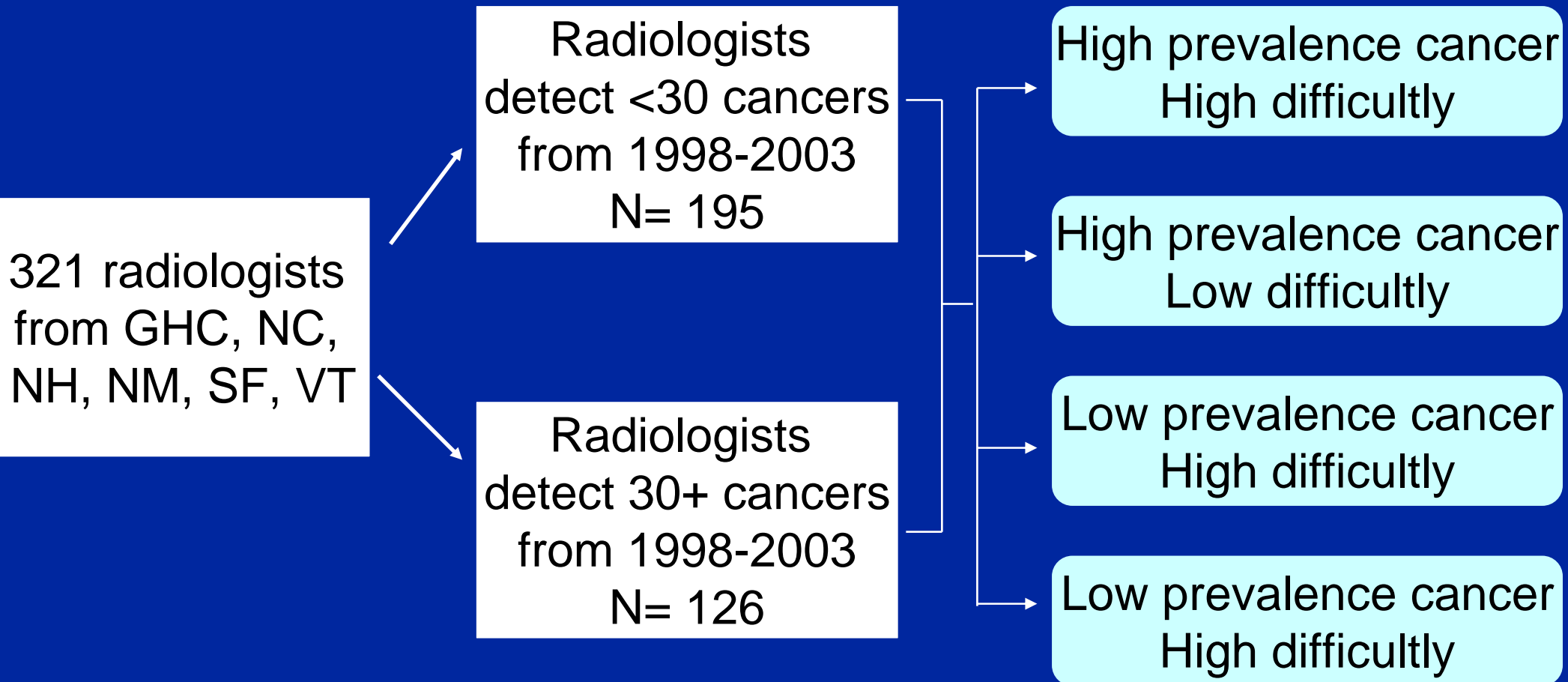
- Determine the effects of radiologists' interpretive volume on clinical performance measures.
- Create and evaluate assessment test sets that consist of representative screening mammograms from community practice.
- Develop and pilot test innovative educational programs designed to improve radiologists' mammography interpretation skills.

Radiologists in BCSC by Average Annual Interpretive Volume--1998-2003

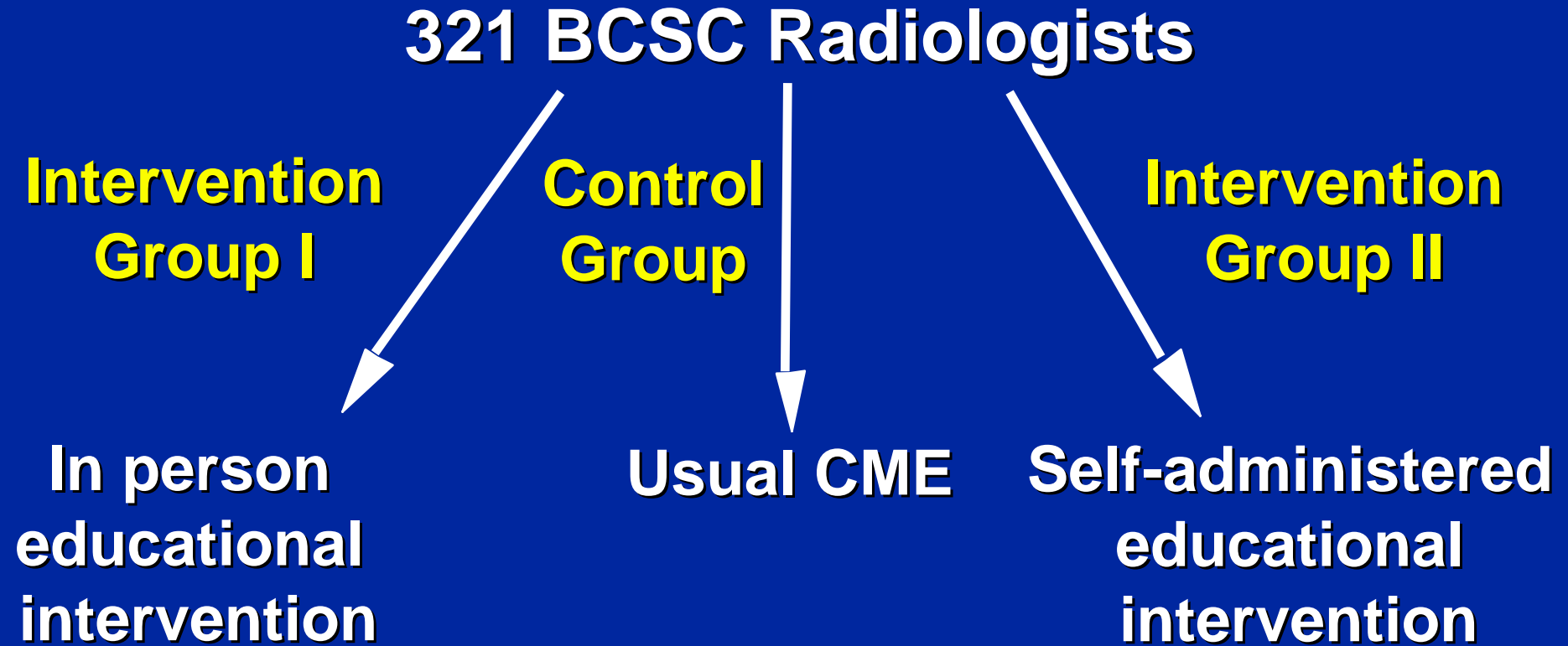
Average annual volume	Total
<480	76
480-999	66
1000-1999	114
2000-2999	33
3000+	32
<i>Total</i>	<i>321</i>

- **Survey Radiologists**
 - **FAVOR survey**
 - **Verify volume at non-BCSC facilities**
 - **Verify prevalence of double reads**
 - **Collect physician and facility characteristics**

Examine if Performance on Test Set Reflects Performance in Clinical Practice



Randomized Controlled Trial



- 1. Improvement in performance on test set**
- 2. Improvement in actual performance**

Summary

- **Variability exists in performance measures among radiologists in U.S.**
- **Understanding factors that explain variability in performance measures is important**
- **Need a means to assess interpretative skills that is associated with performance in clinical practice**
- **Need a means to improve interpretative skills**

Acknowledgments

- **National Cancer Institute**
- **American Cancer Society**
- **Women and Radiologists that have contributed information to BCSC**

Thank you