National Cancer Institute

Division of Cancer Prevention

The Early Detection Research Network

Conducting Research to Identify, Test, and Validate Cancer Biomarkers

FIFTH REPORT • DECEMBER 2011

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

Acknowledgments

This report was composed by the Cancer Biomarkers Research Group of the National Cancer Institute Division of Cancer Prevention Program Staff:

Sudhir Srivastava, Ph.D., M.P.H., M.S., Chief of the NCI Cancer Biomarkers Research Group, EDRN Program Coordinator, Director of Administrative Activities

Annalisa Gnoleba, M.S.A., Program Specialist, Cancer Biomarkers Research Group

Jacob Kagan, Ph.D., M.Sc., Program Director, Prostate and Other Urological Cancers

Karl Krueger, Ph.D., Program Director, Lung and Upper Aerodigestive Cancers

Kathy Marshall, M.A., Project Manager

Richard Mazurchuk, Ph.D., Program Director, Prostate and Other Urological Cancers

Christos Patriotis, Ph.D., Program Director, Breast and Gynecological Cancers

Jo Ann Rinaudo, Ph.D., Program Director, Colorectal and Other Gastrointestinal Cancers

Lynn Sorbara, Ph.D., Program Director, Lung and Upper Aerodigestive Cancers

Paul Wagner, Ph.D., Program Director, Colorectal and Other Gastrointestinal Cancers

Wendy Wang, Ph.D., M.Sc., Program Director, Breast and Gynecological Cancers

Dr. Srivastava appreciates the review of the manuscript and valuable suggestions by Karen Anderson, M.D., Ph.D., Arizona State University; Dan Chan, Ph.D., The Johns Hopkins University; Martin Sanda, M.D., Harvard School of Medicine; Ziding Feng, Ph.D., Fred Hutchinson Cancer Research Center; Joshua LaBaer, M.D., Ph.D., Arizona State University; Larry Norton, M.D., Memorial Sloan-Kettering Cancer Center and Weill Medical College of Cornell University; Steve Skates, Ph.D., Massachusetts General Hospital; and Ian Thompson, M.D., University of Texas Health Science Center San Antonio.

Additional editing and production coordination were conducted by Gwen Moulton, GM Authoring Services, and Kara Smigel Croker, Communications Manager, NCI Division of Cancer Prevention.

National Cancer Institute

Division of Cancer Prevention

The Early Detection Research Network

Conducting Research to Identify, Test, and Validate Cancer Biomarkers

FIFTH REPORT • DECEMBER 2011

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Foreword

The essential purpose of early detection of cancer is to reduce mortality and morbidity while minimizing the risks of screening and associated treatment. As molecularly-informed research moves us closer to progressively more specific interventions with less toxicity, the Early Detection Research Network (EDRN) is focused on finding new and improved methods to noninvasively and accurately detect potentially life-threatening cancers at their earliest stages.

As genomic knowledge accumulates, general guidelines for cancer screening may be supplanted by more targeted testing methods using validated cancer biomarkers to screen individuals at differing levels of risk. However, the deluge of molecular data increasingly challenges the ability of researchers and medical practitioners to find reliable ways to stratify the many kinds of cancer and degrees of cancer risk.

Accordingly, the EDRN laboratories are bridging the critical steps in the process of moving a biomarker from discovery to clinical use by validating both its analytical and clinical performance. Analytical performance measures the reproducibility, precision, and accuracy of the assays. Clinical validation of a biomarker or panel of biomarkers determines its capacity to accurately distinguish patients with cancer from those without, or to detect preclinical cancer. The full range of expertise for the development of early detection biomarkers requires a multidisciplinary effort, and the EDRN was created to integrate the efforts of the necessary spectrum of disciplines.

The EDRN established an innovative five-phase approach for cancer biomarker development and delineated a coherent and comprehensive set of guidelines for study design in the discovery, evaluation, and initial clinical development of biomarkers. In order to be of clinical value biomarkers must be reliable and reproducible in testing; highly sensitive and specific; quantitative; readily obtained by non-invasive methods; part of the causal pathway for disease; and have high predictive values for clinical disease. Ideally, identified biomarkers may also serve as targets for chemopreventive interventions.

An important goal is to identify molecular fingerprints of tumors and preneoplastic tissue abnormalities that can be used to predict clinical behavior and distinguish harmful progressive tumors detected by cancer screening tests from indolent tumors that do not require therapy. This would maximize the benefits of cancer screening while limiting the harms. All screening tests can introduce risks, and biomarkers will undoubtedly as well. But with the rigorous framework of the EDRN laboratories and clinical centers, the known harms of screening, such as overdiagnosis and false positive and false negative results, can be minimized. The multidisciplinary structure and rigorous process of the EDRN provide an innovative, integrated approach to biomarker discovery and validation of biomarkers for early cancer detection. The fruits of this process over the relatively short history of the EDRN come through in this progress report.

Barnett Kramer, M.D., M.P.H.
Director, Division of Cancer Prevention
National Cancer Institute

Contents

Introduction

Early detection of cancer can dramatically improve outcomes. Finding breast and colon cancers when they remain localized results in a significant decrease in mortality. The Early Detection Research Network (EDRN) is helping make this an achievable goal for more cancers.

In 2000, NCI's Division of Cancer Prevention created EDRN, an investigator-driven network designed to conduct translational research that identified markers both for the early detection of cancer and for cancer risk. EDRN focuses on the goal of creating validated biomarkers ready for large-scale clinical testing and eventual application. Without a doubt, real progress has been made—and is being made, by this consortium of more than 300 investigators and 40 private sector and academic institutions. These scientists represent diverse disciplines, including genomics, proteomics, metabolomics, bioinformatics and public health.

EDRN is at the forefront of technology-driven research on the discovery and use of biomarkers for the early detection of cancer. By identifying and validating biomarkers, such as novel proteins or changes in gene expression, it is possible to measure an individual's disease risk, the development and progression of disease, and the response to therapy. Ultimately, EDRN research will aid in prevention and in early therapeutic intervention, based on early detection of disease.

Today, EDRN is a nationwide, interdisciplinary group of established partnerships among scores of institutions and hundreds of individuals working to advance the science for public health benefit.

Research collaborations take place within an environment of teamwork across different disciplines and laboratories focused on achieving common goals, such as:

- Developing and testing promising biomarkers and technologies to obtain preliminary information to guide further testing;
- Evaluating promising, analytically proven biomarkers and technologies, such as measures of accuracy, sensitivity, specificity and, when possible, potential predictors of outcomes or surrogate endpoints for clinical trials;
- Analyzing biomarkers and their expression patterns to serve as background for large, definitive validation studies;
- Collaborating with academic and industrial leaders to develop high-throughput, sensitive assay methods;
- Conducting early phases of clinical and epidemiological biomarker studies; and
- Encouraging collaboration and dissemination of information to ensure progress and avoid fragmentation of effort.

EDRN is a leader in defining and using criteria for the validation of biomarkers—an essential condition for scientific progress. While myriad proteins and genes have been linked to a variety of cancers, acceptable biomarkers must be: reliable and reproducible in testing; highly sensitive and specific; quantitative; readily obtained by non-invasive methods; part of the causal pathway for disease; capable of being modulated by a chemopreventive agent; and a high predictive value for clinical disease.

Executive Summary

The National Cancer Institute's Early Detection Research Network (EDRN) is the leading national initiative for finding cancer early using molecularly informed research. Since it was created in 2000, the EDRN has been setting standards for how to discover, validate appropriate samples, and translate biomarkers into clinical use for risk and early cancer detection, diagnosis and prognosis.

Improved methods for early detection are vital to reducing morbidity and mortality due to cancer. A primary cause of poor survivability is that many cancers are detected late, after they have metastasized to distant sites. Once a cancer has spread, it is more difficult to eradicate; therapies for late-stage disease are still not successful for nearly all cancer types. The mortality rates from cancers where screening tools are available are lower than from cancers for which no viable screening tools exist.

Even for cancers where screening tools currently exist, there is room for improvement, either in the accuracy of the tests or in making them more acceptable to patients. For instance, despite widely available screening methods that can detect early stage colon cancer, only about 40% of newly diagnosed colon cancers are localized. Consequently, EDRN investigators are committed to finding new and improved methods to noninvasively and accurately detect cancers at their earliest stages.

Thousands of research articles published in the last decade have fueled expectations that effective biomarker-based diagnostics would rapidly take form. However, much of the literature includes studies, which were conducted without appropriate study designs and population statistics. In reality, only a handful of biomarkers were approved by the Food and Drug Administration (FDA) during that time, several of which had completed the EDRN validation process. It is therefore important to have an infrastructure such as EDRN, which systematically assesses reported biomarkers and selects the truly promising ones for transition through rigorous validation for clearly defined clinical utilities. Today, at least 300 candidate biomarkers are positioned to move forward through the EDRN validation process. The state of biomarker research has changed dramatically since the inception of EDRN (see Box).

State of Biomarker Research Prior to and After EDRN

Prior to EDRN

- Fragmented studies, with discoveries using convenience samples
- Results of studies not generalizable
- Lack of Standard Operating Procedures for sample collection and study designs
- Studies compromised by chance, bias and confounders
- Lack of evidence for the claimed clinical use

After EDRN

- Clinically annotated samples for discovery
- Roadmap for biomarker discovery and validation using EDRN five-phase guidelines and PRoBE design
- Well-designed multi-center, multi-disciplinary validation studies to minimize chance, bias, confounders
- Well-designed Standard Reference Sample Sets to quickly evaluate biomarkers for intended clinical uses
- Adoption of EDRN-developed guidelines and concept of validation throughout the biomarker research community
- Adoption of EDRN-developed study-design evaluation criteria by the biomarker community and the NIH study sections

Beyond the issue of safety, the field had been comparatively unregulated until recently, and numerous potential markers emerge in this environment. The EDRN has adopted a series of validation benchmarks to effectively compare one candidate/technology against another. This helps the field to avoid numerous competing claims of being "the biomarker of choice," the notion of which arises simply from marketplace competition or differences between laboratories. The EDRN approach facilitates well-designed clinical studies that have an increasing hierarchy of complexity.

EDRN developed the first roadmap for guiding the scientific process of discovery, development and validation of cancer biomarkers. These guidelines are needed because molecules specific for early stage cancer that are "easy" to detect are very hard to find and verify. The EDRN adopted a five-phase schema in 2001[1] and its comprehensive infrastructure gave researchers with promising biomarkers a place to accurately assess them from the point of discovery to clinical validation. Strategies and methods for finding biomarkers with next generation technologies that query at the systems level with more sensitive detection capabilities moved forward.

Work by EDRN investigators to validate putative biomarkers has shown that many of the assays used in previous years were either not reproducible, or not sensitive and specific enough to warrant further consideration. EDRN has developed and standardized a number of assays commonly used in research and clinical application (Table 1). Biomarkers with both high sensitivity and specificity are desired for accurate diagnosis of disease. The sensitivity of a medical test refers to the fraction of individuals with a disease who test positive for the disease. The specificity of a medical test is the fraction of individuals without the disease who test negative for the disease. No medical tests are 100% sensitive or specific, and some people with or without the disease will not be identified correctly.

Some of the EDRN's major accomplishments to date include:

- More than 1,000 biomarkers identified. For example, for Triple-Negative Breast Cancer alone, more than 100 new biomarkers have been identified, including circulating proteins, autoantibodies and miRNAs.
- More than 300 prioritized biomarkers ready for verification and validation studies. For example, 125 to 150 candidate biomarkers for early ovarian cancer detection, not previously tested, are being subjected to verification and validation (see Part II, Chapter 1).
- Eight validation trials completed: MSA for bladder cancer; DCP and AFP-alpha 3 for liver cancer; ProPSA, PCA3, and MS proteomic profile assay for prostate cancer; methylation marker panel for esophageal cancer; Annexin 1 and 2 for lung cancer; and circulating protein markers for ovarian cancer (see Part III, Chapter 2). A number of EDRN-developed biomarkers have been adopted by the diagnostic community for clinical use (Table 2).
- Eight large validation studies underway (see Part III, Chapter 2).

[1] Pepe M S, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001;93:1054-61.

Table 1. Biomarker Assays Developed and Standardized

Assays	Application
Validation of bleomycin-induced chromosomal breakage in lymphocytes	Biomarker of Lung Cancer susceptibility
Validation of 3.4KB mitochondrial DNA deletion	Biomarker for Prostate Cancer Risk
Development of high-density breast and prostate tissue microarrays	Testing of ISH and other molecular probes
Validation of SOPs for Microsatellite Instability (MSA) and DNA methylation assays	Biomarkers for Bladder Cancer
Validation of saliva-based mRNA assay	Biomarkers of Oral Cancer
Validation of proteomic prostate-specific biomarkers, including percent proPSA and other PSA isoforms	Biomarkers for Improving PSA Screening of Prostate Cancer
Urine PCA3 assay	Detection of Prostate Cancer
Urine/TMA assay for T2S:ERG fusion	Detection of Prostate Cancer
FISH assay for T2S:ERG fusion	Detection of Prostate Cancer
Aptamer-based assay	Detection of Lung Cancer
Proteomic panel for Lung Cancer	Detection of Lung Cancer
OVA1™ test for Ovarian Cancer	Differential Diagnosis of Benign Pelvic Mass from Ovarian Cancer
ROMA algorithm for Ovarian Cancer	Differential Diagnosis of Benign Pelvic Mass from Ovarian Cancer
Vimentin methylation in stool	Detection of Colon Cancer
SOPs for blood, sera, plasma, urine, stool	Standard Reference Sample Sets
8-oxyguanine DNA glycosylase (8OGG); Alkyl-adenine DNA Gycosylase; APE1 Endonuclease	Measuring DNA repair capacity for Lung Cancer risk

- Two validation studies planned (see Part III, Chapter 2).
- Ten valuable clinical specimen reference sets available for rapid and inexpensive testing of biomarkers (see Part IV, Chapter 3).
- More than 28 patents and licenses developed, indicative of the robustness of studies in which diagnostic companies are willing to invest further.
- More than 15 collaborations with biotechnology and diagnostic companies.
- More than 1,450 collaborative papers generated.
- Seven major workshops conducted, attracting on average 400 participants each (see Significant Activities table in Appendix).

Future short-term goals of the research network are to:

- Augment the ability of commonly used screening tests to detect major epithelial cancers (colon, breast, cervical, lung and prostate), and facilitate co-development of diagnostic tests for prevention or therapeutic interventions (theranostics).
- Evaluate biomarker discovery, development and validation, and collaborate with the NCI Cancer Intervention and Surveillance Modeling Network (CISNET) on integrating cost-benefit models in discovery and development.
- Create well-defined consensus standards and guidelines for biomarker development, validation and qualification using the Translational Research Working Group-developed Device Pathway schema to reduce uncertainty in discovery and development.

Testing Biomarkers to Make Sure They Work: EDRN as a "Brake" and an "Accelerator"

The OvaSure screening test for ovarian cancer which made its way to market use in June 2008 had to be withdrawn from the market for the lack of scientific rigor during testing.

After a week on the market, the Society of Gynecologic Oncology called for additional research on the effectiveness of the test, which was manufactured by LabCorp. Scientists at the Canary Foundation said that the reported LabCorp findings were overly optimistic, in part because they failed to take into account effects of a key protein in the panel, prolactin. An EDRN investigator also could not replicate the findings in a well-designed study using an appropriate cohort of samples.

The primary mistake made by the OvaSure producers was to calculate the PPV using the prevalence of cancer only from the study population, when it should have been determined based on the accepted prevalence of ovarian cancer for all post-menopausal women. By August 2008, the Food and Drug Administration declared that the test could harm public health because of the lack of adequate validation. OvaSure was removed from the market in October 2008.

Buchen L. Cancer: missing the mark. Nature. 2011 Mar 24;471(7339):428-32.

Visintin I, Feng Z, Longton G, et al. Diagnostic markers for early detection of ovarian cancer. Clin Cancer Res. 2008 Feb 15;14(4): 1065-72.

McIntosh M, Anderson G, Drescher C, et al. Ovarian cancer early detection claims are biased. Clin Cancer Res. 2008 Nov 15;14(22):7574; author reply 7577-9. Epub 2008 Oct 23.

Greene MH, Feng Z, Gail MH. The importance of test positive predictive value in ovarian cancer screening. Clin Cancer Res. 2008 Nov 15;14(22):7574; author reply 7577-9. Epub 2008 Oct 23.

Table 2. Adaption of EDRN-Supported Assays

Detection/ Biomarker Assay	Discovery	Refine/ Adapt for Clinical Use	Clinical Validation	Clinical Translation
Blood proPSA		✓	✓	FDA IVD pending review
Urine PCA3		✓	✓	FDA IVD pending review
Urine/TMA assay for T2S:ERG fusion for Prostate Cancer	✓	✓	✓	CLIA in process
FISH for T2S:ERG fusion for Prostate Cancer	✓	✓	✓	In CLIA Lab
Aptamer-based markers for Lung Cancer		✓	✓	In CLIA Lab
Proteomic panel for Lung Cancer		✓	✓	In CLIA Lab
OVA1™ for Ovarian Cancer		✓	✓	FDA Approved
SOPs for Blood (Serum, Plasma), Urine, Stool,		✓		Frequently used by biomarker research community
Vimentin methylation marker for Colon Cancer		✓	✓	In CLIA Lab
ROMA algorithm for CA125 and HE4 Tests for Pelvic Mass Malignancies		✓	✓	FDA Approved
Blood/DCP and AFP-L3 for Hepatocellular Carcinoma		✓	✓	FDA Approved
Blood GP73 for Hepatocellular Carcinoma	✓	✓	✓	Together with AFP-L3 used in China for monitoring/ risk assessment of cirrhotic patients for HCC

In the longer term, the EDRN aims to:

- Integrate the genetic, cell signaling and biochemical pathways with biomarker discovery efforts to have a broader applicability across different tumor types.
- Determine the potential of novel network- and pathway-based markers to detect and diagnose cancer, using a systems approach to diagnosis, prevention and therapeutic strategies.
- Develop new serum- and tissue-related methods for early detection and diagnosis to identify clinically significant diseases and predictions of clinical outcomes, with or without conventional tissue examination, by utilizing currently available biomarker tests.
- Expand collaborative efforts and shared resources to improve the capacity to conduct biomarker development and validation trials.
- Examine genome-wide chromosomal instability (i.e., chromosome copy gain or loss and loss of heterozygosity) and genome-wide association studies to predict progression from benign to malignant cancers via characterization of these regions of the genome.
- Integrate The Cancer Genome Atlas (TCGA) findings into biomarker discovery, verification and validation.
- Address the issue of over-diagnosis.

Improving Tumor Modeling with Biomarkers: Flexible Sigmoidoscopy

A model used to estimate the cost-effectiveness of flexible sigmoidoscopy screening for colorectal cancer in the U.S. will likely become more common as new biomarkers are introduced. Such mathematical models are built to simulate tumor growth over time, either at the cellular level or as representative of tumor size.

The models are calibrated to observed data, typically from longitudinal population studies of tumor growth and clinical disease. Real or hypothetical screening or diagnostic tests are then introduced, with sensitivity and specificity based on their ability to detect either markers that increase in the blood in proportion to tumor burden, resolution for detecting tumors of a particular size, or ability to detect metabolic changes.

Based on known or hypothesized distributions representing ranges in the rates of tumor onset and growth, the models are then run for a population either with or without disease to generate presumed life histories both in the absence and presence of testing. The outcomes (such as stage at diagnosis, tumor response rate, cancer-specific or overall survival, and medical care costs) are then compared under the test and no-test scenarios.

Knudsen A B, Lansdorp-Vogelaar I, Rutter C M, et al. Cost-effectiveness of computed tomographic colonography screening for colorectal cancer in the Medicare population. J Natl Cancer Inst. 2010;102,1238-52

In summary, EDRN has made significant progress in biomarker discovery, development and validation, and established a number of study standards that have been adopted by the greater biomarker research community. Standard Reference Sets have become a valuable tool for cheaply and effectively triaging promising biomarkers from those that are not clinically useful. EDRN has established a number of productive collaborations with extramural scientists, overseas investigators, and research foundations. In collaboration with stakeholders, EDRN is constantly evaluating the program focus to include research on companion imaging with molecular markers, as well as to address challenges related to over-diagnosis and the detection of interval cancers. Finally, EDRN is poised to bring many biomarkers to clinical fruition by working with industrial partners and the Food and Drug Administration.

NCI's Early Detection Research Network Wins NASA Award

The National Aeronautics and Space Administration (NASA) has awarded NCI's Early Detection Research Network Informatics Team the NASA Honors Award for Group Achievement for the innovative and pioneering use of NASA data system technologies to promote data and specimen sharing among cancer biomarker researchers.

The award was presented to EDRN teams from NASA's Jet Propulsion Laboratory, NCI, Dartmouth, and the EDRN Data Management and Coordinating Center at Fred Hutchinson Cancer Research Center.

Under the leadership of Dr. Sudhir Srivastava, chief of NCI's Cancer Biomarkers Research Group in the Division of Cancer Prevention, the EDRN Informatics Team pioneered and deployed a highly distributed national data system to implement cancer biomarker data analysis and research. This project is a flagship example of successful technology infusion and transfer between agencies.

The informatics platform has enabled EDRN science teams to share and integrate diverse datasets and software platforms for complex biomarker research, including discovery and validation of cancer biomarkers.

Incorporating a Systems Approach to Biomarker Research: Discovery of Gene Fusions in Solid Tumors

As data rapidly accumulate from genomic and proteomic analyses, algorithms to combine information from multiple biomarkers are in development.

The systems approach was successfully used to discover biomarkers from high-volume expression data along with a metabolite whose expression was also regulated by gene fusion products. In this case, it pertained to chromosomal translocations, which have been long known to occur in various types of hematopoietic and soft tissue tumors, but only a few have been documented in solid cancers.

A chromosomal rearrangement implicated as a potential cause of some prostate cancers was discovered through analyzing a database called Oncomine, a collection of data from cancer studies across the globe that screened the activity of thousands of genes.

By surveying Oncomine for over-expressed genes in prostate tumors, investigators identified two candidate genes that encode transcription factors, ERG and ETV1. The Oncomine database integrates 132 gene expression data sets representing 10,486 microarray experiments.

While the specific roles of these gene fusions are unknown, they may be important in prostate cancer development or progression. This belief is based on data obtained from cultured primary prostate cancer cell lines that showed this gene fusion in 23 of the 29 cell lines examined in the study.

These findings led EDRN investigators to study the downstream effect of these gene fusions on phenotypic expression and identified a metabolite, sarcosine, likely involved in the etiology, progression, and aggressive behavior of prostate cancer.

Additionally, components of the sarcosine synthesis pathway were found to have potential, not only as diagnostic/prognostic biomarkers of prostate cancer, but also as targets for developing new therapeutic modalities.

Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. 2005;310,644-8.

Hampton T. Tool helps cancer scientists mine genes. JAMA. 2004;292,2073.

Part One

Evidence-Based Research on Biomarkers