

**Testimony of Peter J. Levine
President, Correlogic Systems, Inc.
before the
Subcommittee on Environment, Technology and Standards
Committee on Science
U.S House of Representatives
Port Jefferson, New York
June 22, 2002**

Introduction

My name is Peter J. Levine. I am President of Correlogic Systems, Inc., a bioinformatics company that has developed a promising new technology for the early detection of cancer and other diseases.

Before I begin to discuss the work Correlogic is doing, I would like to thank the Chairman, Congressman Grucci and the Subcommittee for the important work they are doing on the critical issue of the high incidence of breast cancer in Port Jefferson, New York and elsewhere in Nassau and Suffolk Counties. I grew up nearby in Huntington, New York, and I know from my own personal experiences how cancer can impact the lives of the people of this area.

I would like to thank Congressman Israel, for his leadership in support of cancer research and his ongoing efforts to ensure Medicare and other insurance coverage of effective technologies. In particular, his bill, H. Con. Res. 385, now with 134 cosponsors, is a dramatic demonstration of congressional support for this research and for the availability and accessibility of diagnostic testing.

Major Developments in the Early Detection of Cancer

Today, I want to tell the Subcommittee about new technologies for the early detection of cancer. While these technologies have not been used to explore possible causes for the high breast cancer rates in this region, these technologies will address, *in the very near future*, the critical issue of early detection. This technology can be used to help identify the possible causes.

I am testifying this morning in my capacity as the President of Correlogic Systems, Inc., a bioinformatics company that has developed scientifically validated methodologies for the early detection of various cancers through the use of high throughput bioassays and pattern discovery software. The technologies we have developed have a wide variety of applications for the creation of diagnostic "models," biomarker discovery, disease detection and new drug discovery processes.

In the 1990's new "protein separation" technologies – devices that measure the molecular weight and intensity of the proteins in a patient's blood (sera) -- became available. Researchers had identified individual proteins from sera for many years. The new technologies provided researchers with unprecedented volumes of *data* about the proteins in sera. But, standing alone, the new technologies provided limited *information*. Now, presented with as many as 20,000 individual data points from a single drop of blood, researchers still continued to look for specific, individual proteins that might serve as a biomarker.

A New Approach

Correlogic, in cooperation with the FDA and NCI Clinical Proteomics Program, developed a patent-pending process entitled "A Process for Discriminating between Biological States Based on Hidden Patterns from Biological Data." The central premise of the invention is that subtle *patterns* of proteins, rather than individual biomarkers, can be analyzed to detect disease at the earliest and most treatable stage. For example, rather than looking for the increased presence of a single protein like PSA for prostate cancer, or CA-125 for ovarian cancer, our technology evaluates thousands of proteins simultaneously to identify subtle changes that reflect the health of the patient.

A New Technology

However, this concept – looking at the composition of the haystacks rather than looking for the needle in the haystack – required a higher order of analysis than was available when we initiated our research.

In the spring of 2000, Correlogic's Chief Scientist, Dr. Ben Hitt invented an algorithm, "A Heuristic Method of Classification", also known as the Knowledge Discovery Engine™. With this algorithm, Correlogic built an analytical software system designed to validate the concept that *patterns* of proteins can be used for the early detection of disease.

Working with Dr. Emanuel Petricoin of the FDA and Dr. Lance Liotta of the NCI, and using sera from prostate cancer patients, we concluded that it was possible to make a diagnostic assessment based upon patterns of proteins. We then extended the concept to research on ovarian and breast cancers as well as other disease states.

The Results

In February 2002, the results of our work with the FDA and the NCI on the early detection of ovarian cancer were "fast-track" published in the British medical journal, *The Lancet* (see Appendix A). I have included a copy of this article for the record. The reported results were remarkable. *From a single drop of blood, we were able to detect 100% of the patients with cancer – including all Stage 1 cancers, when treatment is most effective.*

While the results were exciting on a number of levels, the fact that 100% of all Stage 1 cancers were identified is particularly important because women diagnosed in Stage 1 have a 5-year survival rate of 95%. The five-year survival rate for late-stage diagnosis is 25%. Currently, nearly two-thirds of women diagnosed with ovarian cancer over the age of 50 are detected in that late stage. Clearly, early detection and treatment equal survival.

The Process

How exactly did we obtain the results reported in the *Lancet*, and can they be applied to other disease states, including breast cancer? The attached diagram (Appendix B) provides an overview of the process.

We started by taking a small blood sample from each patient from two defined groups: Group 1, in yellow on the slide, are patients with clinically diagnosed ovarian cancer; and, Group 2, in blue on the slide, a mixture of patients unaffected by ovarian cancer, some with other gynecological conditions such as ovarian cysts, and others “normal.”

The blood from each patient is spun to separate the serum from the blood cells. The serum is then placed on a chip that attracts and binds certain categories of proteins. The chip is then placed into a mass spectrometer, and a laser is used to ionize the proteins on the chip. The proteins and/or protein fragments fly down a vacuum tube, and their time of flight is measured, yielding the molecular weight and intensity of the proteins and/or protein fragment. The result of this process is 15,200 data points per patient.

The next step is to find a pattern within each of the 15,200 data points per patient that will effectively and consistently distinguish the cancer patients from the unaffected patients. This was not a simple task. In fact, it is a nearly impossible task.

For ovarian cancer, we found that five features out of the 15,200 formed a pattern, or “model” that would separate the cancer patients from the unaffected. What this means is that we had to explore $15,200^5$ — or approximately 8,113,000,000,000,000,000 — possible combinations to find a model that worked. If all of the conventional computing power on the Earth were applied to this task, it would take millions of years to sort through. Our program, Proteome Quest™ (Appendix B, Step 2), accomplished this task in a far more manageable time frame, and created two diagnostic models: one of ovarian cancer patients and one of healthy patients.

In Appendix B, Step 3, and in Appendix C, we show a three-dimensional representation of the diagnostic “models” as taken from the actual ovarian cancer data. It is important to note that this is simply a representation. The pattern discovery and pattern matching that is at the heart of the process is not a visual matching. Rather, the diagnostic “model” is a computational model that exists in “N” dimensional space. For ovarian cancer, it is a five-dimensional model.

At this point in the process – the creation the diagnostic model – we had “trained” Proteome Quest™. We then repeated the process on 116 blinded samples. That is, samples from patients whose condition was unknown to us. This was the testing, or validation of the diagnostic model. Each patient’s protein patterns were compared to the diagnostic model created during the training process (See Appendix B, Step 4).

The results, as noted earlier, were that we correctly identified *all* of the patients with cancer, 50 of 50, including 18 patients in Stage 1, the organ *confined* stage. We had 3 false positives out of 66, resulting in an overall positive predictive value of 94.3% (See Appendix D). The most widely used biomarker for ovarian cancer CA-125, has a positive predictive value of less than 10% as a single marker and about 20% with the addition of trans-vaginal ultrasound.

We are currently developing standard operating procedures and clinical trial protocols in preparation for clinical trials for the ovarian cancer test. We expect to commence these trials with the FDA and NCI in the coming months.

Beyond Ovarian Cancer

Can this approach to early detection be applied to other cancers? I am pleased to report that the answer is yes.

Our original work was on prostate cancer. We obtained similar results, and these results are pending publication.

We have also begun our exploration of breast cancer. Our preliminary results, which are not yet published, but appear in our patent application, used nipple aspirants. The results are promising, but much work needs to be done.

Under the terms of our Cooperative Research and Development Agreement (CRADA) with the FDA/NCI Clinical Proteomics Program, we will be focusing our attention on the development of early diagnostic tests for ovarian, prostate and breast cancer. We will also examine the application of the technology to pancreatic, colon, lung and other cancers.

The Connection to Long Island

We are currently talking with a number of Long Island institutions about their participation in the trials and other research efforts, and we are encouraged by their interest in participating.

Approximately 300 Long Island women are diagnosed with ovarian cancer each year. There is a significant connection between the incidence of breast cancer and ovarian cancer. According to the National Breast Cancer Coalition, women who have had breast cancer before the age of 50 are twice as likely to develop ovarian cancer. So our work on ovarian cancer may be of particular interest to Long Island women.

As I described, our work on breast cancer, is underway.

Correlogic's technology has many applications. In addition to early detection, it may also be used to guide researchers seeking the causes of increased cancer incidences on Long Island. Studying women with breast cancer on Long Island, the technology may be used to determine whether there are unique patterns in the protein expressions among these individuals. With that information, we may be able to better determine if the higher rate of cancer has a cause specific to Long Island.

Conclusion

When our study was published in *The Lancet* in February, Correlogic and our colleagues at the FDA/NCI Clinical Proteomics Program received a great deal of attention from the press, the scientific and business communities. News about the new technology also struck a cord with the general public. Their heart-felt outpourings for more information on this new technology is a reflection of the hunger that people have for good news in the area of cancer research.

The implications for public health are far reaching – from a single drop of blood, we may be able to determine the entire state of a person's health, reducing the need for unnecessary biopsies and surgery on one hand, and initiating treatment earlier when a disease is present.

I would again like to thank Congressman Israel, for his leadership in support of cancer research and his ongoing efforts to ensure Medicare and other insurance coverage of effective technologies. I also want to thank Congressman Grucci for focusing attention on the critical issue of the cancer "hot spots" here on Long Island. As we continue our research in cancer and other disease states, we will make every effort possible – consistent with good science - to see these tests through the clinical trial process quickly. To that end, we appreciate the Congressional support we've received to date to make these technologies available to the public as soon as possible.

Thank you again for the opportunity to present the results of our research to the Subcommittee.