Federal Big Data Working Group Meetup

MEDLINE Demo

Fredrik Salvesen
January 7 2014
Graphs and Traditional Technologies

- **Square peg, round hole:**
  - Current technology does not support efficient representation, storage, and interaction with complex graph structures
  - Traditional relational models only add to an already complex structure
  - Traditional hardware approaches do not support efficient access to highly interconnected graphs

- **You don’t know what you don’t know:**
  - Efficient relational schemas require prior knowledge of the relationships between database fields
  - Updating and modifying schemas frequently introduces delays and errors

- **You can’t partition the problem:**
  - Distributed computing solutions are good…If your problem can be easily partitioned
  - Graphs are not predictable; accessing graph nodes across large clusters can be unwieldy at best and does not work at scale
The YarcData Approach

Business Challenge:

Real-time, Interactive Analytics on Large Graph Problems

- Large Shared Memory Architecture Up to 512 TB
- XMT2 Massively Multi-Threaded Processors 128 Threads
- Scalable IO Up to 350TB per Hour
Current Big Data approaches focus on SEARCH

What is the RIGHT answer to a question?
But the high value Analytics is in DISCOVERY

What is the RIGHT question to ask?
“In the amount of time it takes to validate one hypothesis, we can now validate a 1000 hypotheses – increasing our success rate.” – Ilya Shmulevich, Institute of Systems Biology
By offloading Graph Analytics

Data Warehouse

Hadoop

Other Big Data Appliances

Existing Analytic Environment(s)

Export Analytic/Relationship Results

Import Graph Datasets

User/App Visualization

Programmatic Interfaces

Import

Graph

Datasets

User/App

Visualization

Export Analytic/

Relationship

Results

Existing Analytic Environment(s)

Programmatic Interfaces
Discovery/Graph Analytics is everywhere…

**Government/Security**
- Patterns of Activity Analytics
- Cyber Threat Discovery
- Tax Fraud Discovery
- Crime Prediction

**Healthcare**
- Personalized Treatment
- Fraud Detection
- Efficacy of Care
- Adverse Event Clustering
- Disease Prediction

**Energy/Resources**
- Location Discovery
- Field Production Analysis
- Contingency Analysis
- Climate Modeling

**Telecom/Media**
- Influencer Discovery
- Churn Analytics
- Behavior Analytics

**Life Sciences**
- Drug Discovery
- Drug Repurposing
- Clinical Trial Mining

**Financial Services**
- Market Sensing
- News/Trading Analytics
- Counterparty/Risk
- Insider Threat
- AML/Compliance
Effectively Exploiting Big Data with Semantics: A Pilot Project

Thomas C. Rindflesch, Ph.D., FACMI

National Library of Medicine
The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.
Pilot project: Components

- Data and metadata in RDF (2.2 billion triples)
  - MEDLINE citations (22,496,051)
- Effective computing infrastructure
  - YarcData Urika graph appliance
- Semantics
  - Metadata: 65,465,539 semantic predications in RDF
    - Produced with SemRep
  - Application to manipulate predications
    - Semantic MEDLINE
Semantic processing: SemRep

- Manipulate extracted information
  - Not just documents
- Bridge the gap between
  - Language (text)
  - Meaning
- Computable representation of meaning
  - Semantic predications
trastuzumab was added to adjuvant chemotherapy in patients with HER2-overexpressing breast cancer. Notwithstanding the significance of this molecular target, the discovery of the estrogen receptor (ER) may be of even greater importance. Although tamoxifen has long been considered the hormonal therapy of choice for patients with estrogen-responsive breast cancer, accumulating clinical data suggest the new generation of aromatase inhibitors (AIs) is more effective and less toxic. With the availability of new information, guidelines have been updated and reformulated regarding the use of AIs as first-line hormonal therapy in postmenopausal women with ER-positive breast cancer. This paper, a product of the ongoing advances in oncology, incorporates two distinct, yet important, features of oncology: first, clinical concepts related to hormone-dependent breast cancer and second, pharmacoeconomic evaluation of the anitestrogen tamoxifen and the new generation of antiaromatase agents. TI - An alpha-fetoprotein-derived peptide reduces the uterine hyperplasia and increases the antitumour effect of tamoxifen. AB - Tamoxifen (Tam) is effective for the treatment and prevention of breast cancer. However, it has toxic drawbacks and has limited duration utility because, over time, human tumours become refractory to Tam. Recently, a new nontoxic peptide, alpha-fetoprotein-derived peptide (AFPep) has been proposed for the treatment and prevention of breast cancer. The purpose of this paper is to determine whether combining AFPep with Tam would increase efficacy and reduce toxicity in experimental models of breast cancer. Low doses of AFPep and Tam were more effective in combination than either agent alone against breast cancer growth in cell culture, in tumour-xenografted mice, and in carcinogen-exposed rats. Alpha-Fetoprotein-derived peptide interfered with Tam-induced uterine hyperplasia in immature mice, and showed no toxic effects. Unlike Tam, AFPep did not inhibit binding of oestradiol (E2) to oestrogen receptor (ER). Thus, these two agents utilise different mechanisms to interfere with ER functionality, yet work cooperatively to reduce breast cancer growth and alleviate Tam's troubling toxicity of uterine hyperplasia and appear to be a rational combination for the treatment of ER-positive breast cancer.

Tamoxifen (Tam) has been the mainstay of hormone therapy in breast cancer for over two decades. However, it has toxic drawbacks and has limited-duration utility because, over time, human tumours become refractory to Tam. Recently, a new nontoxic peptide, alpha-fetoprotein-derived peptide (AFPep) has been proposed for the treatment and prevention of breast cancer. Low doses of AFPep and Tam were more effective in combination than either agent alone against breast cancer growth in cell culture, in tumour-xenografted mice, and in carcinogen-exposed rats. Alpha-Fetoprotein-derived peptide interfered with Tam-induced uterine hyperplasia in immature mice, and showed no toxic effects. Unlike Tam, AFPep did not inhibit binding of oestradiol (E2) to oestrogen receptor (ER). Thus, these two agents utilise different mechanisms to interfere with ER functionality, yet work cooperatively to reduce breast cancer growth and alleviate Tam's troubling toxicity of uterine hyperplasia and appear to be a rational combination for the treatment of ER-positive breast cancer.

The first-generation selective estrogen receptor modulator (SERM) tamoxifen has been the mainstream hormone therapy in breast cancer. Tamoxifen benefits all stages of the disease, but its use increases the risk of uterine cancer and thromboembolic events and it can only be administered for 5 years. Aromatase inhibitors are superior to tamoxifen at advanced stages of disease and as adjuvants; however, because they increase fractures, aromatase inhibitors are unlikely to be used to prevent disease.Raloxifene, a second-generation SERM, leads, like tamoxifen, to approximately 50% fewer cases of invasive breast cancer in high risk women, with a lower incidence of thromboembolic events. Several other SERMs are in development to improve tissue specificity, efficacy and tolerance. Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone--a tissue-targeted precursor of sex steroid formation--offers hope of a physiological tissue-targeted hormone replacement that, combined with a SERM, would simultaneously prevent breast and uterine cancer. 

The first-generation selective estrogen receptor modulator (SERM) tamoxifen has been the mainstream hormone therapy in breast cancer. Tamoxifen benefits all stages of the disease, but its use increases the risk of uterine cancer and thromboembolic events and it can only be administered for 5 years. Aromatase inhibitors are superior to tamoxifen at advanced stages of disease and as adjuvants; however, because they increase fractures, aromatase inhibitors are unlikely to be used to prevent disease. Raloxifene, a second-generation SERM, leads, like tamoxifen, to approximately 50% fewer cases of invasive breast cancer in high risk women, with a lower incidence of thromboembolic events. Several other SERMs are in development to improve tissue specificity, efficacy and tolerance. Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone--a tissue-targeted precursor of sex steroid formation--offers hope of a physiological tissue-targeted hormone replacement that, combined with a SERM, would simultaneously prevent breast and uterine cancer. 

The first-generation selective estrogen receptor modulator (SERM) tamoxifen has been the mainstream hormone therapy in breast cancer. Tamoxifen benefits all stages of the disease, but its use increases the risk of uterine cancer and thromboembolic events and it can only be administered for 5 years. Aromatase inhibitors are superior to tamoxifen at advanced stages of disease and as adjuvants; however, because they increase fractures, aromatase inhibitors are unlikely to be used to prevent disease. Raloxifene, a second-generation SERM, leads, like tamoxifen, to approximately 50% fewer cases of invasive breast cancer in high risk women, with a lower incidence of thromboembolic events. Several other SERMs are in development to improve tissue specificity, efficacy and tolerance. Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone--a tissue-targeted precursor of sex steroid formation--offers hope of a physiological tissue-targeted hormone replacement that, combined with a SERM, would simultaneously prevent breast and uterine cancer.

The first-generation selective estrogen receptor modulator (SERM) tamoxifen has been the mainstream hormone therapy in breast cancer. Tamoxifen benefits all stages of the disease, but its use increases the risk of uterine cancer and thromboembolic events and it can only be administered for 5 years. Aromatase inhibitors are superior to tamoxifen at advanced stages of disease and as adjuvants; however, because they increase fractures, aromatase inhibitors are unlikely to be used to prevent disease. Raloxifene, a second-generation SERM, leads, like tamoxifen, to approximately 50% fewer cases of invasive breast cancer in high risk women, with a lower incidence of thromboembolic events. Several other SERMs are in development to improve tissue specificity, efficacy and tolerance. Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone--a tissue-targeted precursor of sex steroid formation--offers hope of a physiological tissue-targeted hormone replacement that, combined with a SERM, would simultaneously prevent breast and uterine cancer.
Exemestane after non-steroidal aromatase inhibitor for post-menopausal women with advanced breast cancer
Breast carcinoma

Tamoxifen

Associated with

CDKN1A gene

Aromatase Inhibitors

Breast carcinoma

PKA-induced resistance to tamoxifen is associated with an altered orientation of ERα towards co-activator SRC-1. 

BARD1 gene

Breast carcinoma

CDKN1A gene

Tamoxifen

Patients

Prognostic and Predictive Value of CDKN1A Expression in Postmenopausal Women With Early Breast Cancer: Results From the BIG 1-98 Collaborative Groups. 

Individual

Exemestane after non-steroidal aromatase inhibitors for post-menopausal women with advanced breast cancer. 

TI - Exemestane after non-steroidal aromatase inhibitors for post-menopausal women with advanced breast cancer. AB - A retrospective analysis was performed on 31 consecutive locally advanced or metastatic breast cancer patients who commenced exemestane 25mg orally following previous treatment with Tamoxifen and a non-steroidal third-generation aromatase inhibitor (AI). Patients were seen 3 monthly until clinical or radiological disease progression. Median age was 64 years (range 34-90yrs). The average number of recurrences between AI failure and exemestane treatment was three (range 1-5). There were two complete responses (CR), four partial responses (PR), two stable disease (SD) and twelve patients who progressed. 

9%. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value.

Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone—a tissue-targeted precursor of oestradiol (E2)—may be of even greater importance. Although tamoxifen has locoregional efficacy in breast cancer, it is not fully effective and has toxic effects related to uterine hyperplasia that result in the withdrawal of the second generation SERM tamoxifen. A new generation of aromatase inhibitors (Ais) is more effective and less toxic than tamoxifen. Unlike tamoxifen, the tissue-specificity of some Ais is being developed. 

Low doses of AFPep and Tam were more effective when combined in vivo than either agent alone. However the effect of tamoxifen on ER function may be of even greater importance. Although tamoxifen has long been considered the hormonal therapy of choice for patients with estrogen-responsive breast cancer, accumulating clinical data suggest the new generation of aromatase inhibitors (Ais) is more effective and less toxic than tamoxifen. Unlike tamoxifen, the tissue-specificity of some Ais is being developed. 

Intracranial ependymoma is a highly lethal neoplasm, occurring in infancy and childhood. 

Invasive breast cancer in high risk women, with a lower incidence of thromboembolic events. Several other SERMs are in development to improve tissue specificity, efficacy and tolerance. Raloxifene shows protection against vertebral fractures similar to bisphosphonates; however, no significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone—a tissue-targeted precursor of oestradiol (E2)—may be of even greater importance. Although tamoxifen has long been considered the hormonal therapy of choice for patients with estrogen-responsive breast cancer, accumulating clinical data suggest the new generation of aromatase inhibitors (Ais) is more effective and less toxic than tamoxifen. Unlike tamoxifen, the tissue-specificity of some Ais is being developed. 

Central Pathology Office received material for 6,549 patients (82%), of which 79% were assessable (6,291 patients). Prognostic and predictive value of the receptors were evaluated among 3,950 assessable patients. 

The likelihood of breast cancer metastasis was 1.7-fold higher (95% CI 1.16 - 2.65) in patients who received trastuzumab. CONCLUSION: The results from this meta-analysis are sufficiently compelling to consider 1 year of adjuvant trastuzumab treatment for women with HER-2-positive EBC based on the risk; benefit ratio demonstrated in these studies. Adequate assessment of HER2/neu status is critical, and careful cardiac monitoring is warranted because of cardiac toxicity. Clinical trials should be designed to answer these questions. 

Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value.

TRAIL-AUGMENTED RESPONSE TO TREATMENT OF INTRACRANIAL EPENDYMOMA. AB - A randomized phase II study comparing intracranial doses of TRAIL (25 and 50 mg/m^2) with or without intravenous methylprednisolone in 16 patients with newly diagnosed, biopsy-proven ependymoma was conducted. 

No significant effect has been observed on nonvertebral fractures. Many SERMs are in development for prevention and treatment of osteoporosis. As breast cancer metastasizes early and advanced disease cannot be cured, prevention is essential. To avoid the concerns about the use of traditional hormone replacement therapy, dehydroepiandrosterone—a tissue-targeted precursor of oestradiol (E2)—may be of even greater importance. Although tamoxifen has long been considered the hormonal therapy of choice for patients with estrogen-responsive breast cancer, accumulating clinical data suggest the new generation of aromatase inhibitors (Ais) is more effective and less toxic than tamoxifen. Unlike tamoxifen, the tissue-specificity of some Ais is being developed. 

The MEDLINE, EMBASE, CANCERLIT and Cochrane Library databases, and abstracts published in the annual proceedings were systematically searched for evidence. Relevant reports were reviewed by two reviewers independently and the references from these reports were searched for additional trials, using guidelines set by QUOROM statement criteria. RESULTS: Pooled results from that five randomized trials of adjuvant Trastuzumab showed a significant reduction of mortality (p<0.00001), recurrence (p=0.00001), metastases (p=0.00001) and second tumors other than breast (p=0.007) as compared to no adjuvant Trastuzumab patients. There were more grade 3 or 4 cardiac events in patients treated with Trastuzumab (25/4555 = 0.56%) versus no trastuzumab (12/4562 = 0.26%). The likelihood of cardiac toxicity was 2.45-fold higher (95% CI 1.89 - 3.16) in patients treated with trastuzumab as compared to no adjuvant trastuzumab patients (p=0.00001). The likelihood of bone metastases was 1.9-fold higher (95% CI 1.16 - 2.65) in patients who received trastuzumab. CONCLUSION: The results from this meta-analysis are sufficiently compelling to consider 1 year of adjuvant trastuzumab treatment for women with HER-2-positive EBC based on the risk; benefit ratio demonstrated in these studies. Adequate assessment of HER2/neu status is critical, and careful cardiac monitoring is warranted because of cardiac toxicity. Clinical trials should be designed to answer these questions. 

Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value. 

Among patients with centrally assessed ER >/=10% and PgR >10%, 73 found to have more than 10% positive cells, and eight had 1% to 9%. Of 105 tumors locally ER-negative, 73 were found to have no staining, and 54 had only 1% to 9% positivity. Discordance was more marked for PgR than ER. Patients with tumors reclassified centrally as ER-negative, or as hormone receptor-negative, had poor DFS. Centrally assessed ER and PgR showed prognostic value.
**Summarized predications**

Aromatase Inhibitors $\xrightarrow{TREATS}$ Breast carcinoma

Tamoxifen $\xrightarrow{TREATS}$ Breast carcinoma

CDKN1A gene $\xrightarrow{ASSOCIATED\_WITH}$ Breast carcinoma

CDKN1A gene $\xrightarrow{STIMULATES}$ BARD1 gene

Tamoxifen $\xrightarrow{TREATS}$ Patients

Breast carcinoma $\xrightarrow{PROCESS\_OF}$ Individual
Tamoxifen TREATS Breast carcinoma

Breast carcinoma ASSOCIATED_WITH CDKN1A gene

CDKN1A gene STIMULATES BARD1 gene

BARD1 gene TREATS Aromatase Inhibitors

Aromatase Inhibitors TREATS Breast carcinoma
Tamoxifen TREATS Breast carcinoma
SemMed: **Link to text**

- **Tamoxifen** Treats **Breast carcinoma**
An alpha-fetoprotein-derived peptide reduces the uterine hyperplasia and increases the antitumour effect of tamoxifen.

Andersen TT, Georgekutty J, Defreest LA, Amaratunga G, Narendran A, Lemanski N, Jacobson HI, Bennett JA.

Center for Cardiovascular Sciences, Albany Medical College, Albany, NY 12208, USA. anderst@mail.amc.edu

Tamoxifen (Tam) is effective for the treatment and prevention of breast cancer. However, it has toxic drawbacks and has limited-duration utility because, over time, human tumours become refractory to Tam. Recently, a new nontoxic peptide, alpha-fetoprotein-derived peptide (AFPep) has been proposed for the treatment and prevention of breast cancer. The purpose of this paper is to determine whether combining AFPep with Tam would increase efficacy and reduce toxicity in experimental models of breast cancer. Low doses of AFPep and Tam were more effective in combination than either agent alone against breast cancer growth in cell culture, in tumour-xenografted mice, and in carcinogen-exposed rats. alpha-Fetoprotein-derived peptide interfered with Tam-induced uterine hyperplasia in immature mice, and showed no
Exploiting Semantic MEDLINE

- Guide research
  - Hypothesis generation
- Discern trends
  - Where is research headed?
  - Where should it be headed?
- Methodology: Discovery browsing
  - Cooperative reciprocity between system and human
    - Inspect graph for “interesting” concept
    - Use selected concept to seed another query
    - Iterate until satisfied
Use case: Inflammation and cancer

- With some exceptions, therapy is not effective
  - Has not progressed significantly in 60 years
- Scientific basis
  - Traditionally – cancer cells
  - More recently – non-cancer cells (immune system)
- Immune system and cancer
  - Connection noted in 1863 (Virchow)
  - But not exploited until recently
- Goal: look for trends in cancer immunotherapy
Acknowledgments

- Michael J. Cairelli, D.O.
- Marcelo Fiszman, M.D., Ph.D.
- Halil Kilicoglu, Ph.D.
- Graciela Rosemblat, Ph.D.
- Dongwook Shin, Ph.D.
- YarcData team
  - Aaron Bossert
  - Fredrik Salvesen
  - Tim White
NIH/NLM
Tom Rindflesch
tcr@nlm.nih.gov

YarcData
Tim White
Tim@YarcData.com
Fredrik Salvesen
Fredrik@Salvesen.me