

がんのサーベイランスシステムは、近い将来、がんの発生 (incidence) とがん対策の成果 (outcome) に関連するデータと発がんのリスクがある個人情報とをリンクできる可能性を有している。がんのリスクのある人の行動 (喫煙や運動不足など) を追跡して、がんの発生とがん対策の成果について実証させることが理想的である。しかし、リスクファクターとがん対策の成果を関連付けることは困難で、喫煙のような行動の変化と煙草に関連したがんの発症には長いタイムラグがあったり、人口の移動や長期間の行動変化を追跡することは難しい。

情報通信技術は社会基盤としてその位置を確実なものとし、がんのサーベイランスシステムは益々拡大しているが、現時点では、米国でさえ個人のレベルでの有用性が科学的に証明されておらず、価値ある社会システムをいつ実現できるかを確約していないとしている。しかし、情報通信技術は、がんのサーベイランスを確実に改善し、生活習慣とがん予防対策の成果との間の関連性の検証だけでなく、病気の原因に関する研究や介入研究に向けた仮説構築のための豊富な資源となることは疑いのないことである。

第六項 がん情報通信基盤の可能性

がんの予防やコントロールに対して情報通信技術を応用する場面は数多く存在している。しかし、インターネット技術の拡大に情報流通の促進はありながらも、未だ実質的なインパクトを与えているとは言い難い。がんの予防やコントロールを実現するための情報通信技術の応用には、既述したようにサーベイランスシステムのような社会医学的研究の支援から基礎医学研究者の研究支援、一般向けの生活習慣への情報の提供などの情報介入研究支援など多種多様な適応が考えられている。このため、ある一定の統計モデルに当てはめ検証を行う臨床試験よりも当然複雑なシステムとなる可能性が高い。

このようながん予防に関する研究は、非常に複雑で込み入ったデータ項目と解析を要し、かつ、大量のデータから構成されているので他の分野の研究や対策と比較して、情報通信技術の応用によく適合していると考えられている。つまり、がん情報基盤を整備し統合することは、がんについての知識を増加させ、がんの予防とコントロールに関する科学を飛躍的に進歩させる可能性を有しているともいえる。

Powerful resources available on the NCICB site:

 <p>Infrastructure Build research information systems with open-source software development kit</p>	 <p>Tools Review tools and services built to facilitate your clinical and lab research activities</p>	 <p>Projects Discover cancer research communities created using our tools and infrastructure components</p>	
 <p>About NCICB Learn about our mission, team, initiatives and events</p>	 <p>Downloads Access all available NCICB software components and tools</p>	 <p>Training Explore training opportunities for NCICB applications</p>	 <p>Support Benefit more from NCICB tools and resources with online application support</p>

図 I-9. NCICB のがん対策への情報通信技術の応用プロセス。

第五節 結語

- 米国のがん予防やコントロールに関する研究には、仮説検証という科学的研究方法論が基礎研究以外にも導入され、がんの基礎研究、介入研究、サーベイランス研究、その応用と対策の普及に関する知識の統合化とそれに対する情報通信技術の応用が進められている。
- がん予防とコントロールに関する専門の国立の部門として NCI の中に **Division of Cancer Prevention** と **Division of Cancer Control and Population Sciences** が 1997 年から設置されてがんの化学予防、バイオマーカー、早期発見技術、行動科学、健康サービス、サーベイランスなどの研究を推進している。
- がん予防研究の情報基盤として **Cancer Genetics Network (CGN)** と **Early Detection Research network (EDRN)** があり、前者では家族性腫瘍の症例登録データベースも整備している。その特徴は、基礎研究よりも臨床的にインパクトの大きい研究の支援に重点が置かれていることである。
- がん予防についての介入研究は、既存の薬剤の臨床試験よりも大規模であり、多くは健康人であることから既存の臨床試験システムよりも大規模でデータ収集が既存の手法では困難であることから情報通信技術の応用による効率化と精度管理が強く求められている。
- がん予防の介入方法についても情報通信技術によるリマインダーシステムや Web 技術を用いたがん予防情報の提供などによる行動変容やがんに関する新しいリスクコミュニケーションツールとしての期待が高まっている。
- 米国では、サーベイランスシステムとして全人口の約15%をカバーする **SEER** プログラムというがん登録制があり、**CDC** のデータベースとリンクして全国のがんの疫学情報を提供し、がん対策政策の立案やがん検診のデータベースと統合解析をすることで検診制度の妥当性の検証にも用いている。
- がん対策の情報基盤は、①オープンソースのソフトウェア開発キットを用いた研究情報システムの構築、②臨床あるいは基礎研究の活動を活性化するために作られるレビューツールとサービスの提供、③NCICBで作成したツールや基盤となるコンポーネントで作られたがん研究コミュニティの発見事業の3つの課程の中の最初の基盤の重要な要素の一つとされている。

参考文献:

- [1] Breslow L, Agran L, Breslow DM, Morganstern M, Ellwein L. 1977. Cancer Control: Implications from its History. *Journal of the National Cancer Institute* 59: 671-686.
- [2] Greenwald P, Cullen JW. 1985. The New Emphasis in Cancer Control. *Journal of the National Cancer Institute* 74: 543-551.
- [3] New York City Cancer Committee. 1994. History of the American Society for the Control of Cancer, 1913-1943. New York City Cancer Committee.
- [4] Abrams DB. 1997. A New Agenda for Cancer Control Research: Report of the Cancer

- Control Review Group. Bethesda, MD: National Cancer Institute.
- [5] Shea S, DuMouchel W, Bahamonde L. 1996. A Meta-Analysis of 16 Randomized Controlled Care in the Evaluate Computer-Based Clinical Reminder Systems for Preventive Care in the Ambulatory Setting. *Journal of American medical Informatics Association* 6: 399-409.
 - [6] Bresnick E. 1977. Report of the National Cancer Institute: Cancer Prevention Program Review Group. Bethesda, MD: National Cancer Institute.
 - [7] Glanz K. 1977. Behavioral Research Contributions and Needs in Cancer Prevention and Control: Dietary Change. *Preventive Medicine* 26: S43-S55.
 - [8] Hiatte RA. 1977. Behavioral Research Contributions and Needs in Cancer Prevention and Control: Adherence to Cancer Screening Advice. *Preventive Medicine* 26: S11-S18.
 - [9] Kaluzny AD. 1997. Prevention and Control Research Within a Changing Health Care System. *Preventive Medicine* 26: S31-S35.
 - [10] Lewis FM. 1997. Behavioral Research to Enhance Adjustment and Quality of Life Among Adults with Cancer. *Preventive Medicine* 26: S19-S29.
 - [11] Lichtenstein E. 1997. Behavioral Research Contributions and Needs in Cancer Prevention and Control: Tobacco Use, Prevention and Cession. *Preventive Medicine* 26: S57-S63.
 - [12] Hiatte RA, Rimer BK. 1999. A New Strategy for Cancer Control Research. *Cancer Epidemiology, Biomarkers, and Prevention* 8: 957-964.
 - [13] Mandelblatt JS, Yabroff KR. 1999. Effectiveness of Interventions Designed to Increase Mammography Use: A Meta-Analysis of Provider-Targeted Strategies. *Cancer Epidemiology, Biomarkers, and Prevention* 9: 759-767.
 - [14] Kreuter MW, Stecher VJ, Glassman B. 1999. One Size Dose Not Fit All: The Case for Tailoring Print Materials. *Annals of Behavioral Medicine* 21: 276-283.
 - [15] Rimer BK, Glassman B. 1999. Is There a Use for Tailored Print Communications in Cancer Risk Communications? *Journal of the National Cancer Institute Monographs* 25: 140-148.
 - [16] Long Range Planning Committee. 2000. Final Report of the Long Range Planning Committee: Translating Cancer Research into Cancer Care. Bethesda, MD: National Cancer Institute.

第二章 がん予防に必要な情報基盤技術

第一項 背景と目的

前章で述べたように、米国のがん予防に関する研究推進体制の強化に伴う情報基盤整備は眼を見張るものがある。特に NCI が主導し、がん予防に関する薬物開発と同等の精度で、薬物開発よりも大規模な臨床試験体制を整備し支援していることは注目に値する。特に科学的な研究方法論をがん予防対策にも反映させるという強い意図を背景によみとることができる。これは、単に細胞や動物実験で示唆されたがん予防効果を単に科学的発見として留めるだけでなく、それを社会に還元するために科学的な評価を臨床試験を行って証明するという制度設計のもとに情報基盤も構築されていることが、国内のがん予防に関する情報化とは大きく異なる。

そのキーワードは、前章でも述べたようにがん予防に関する対策を作るための科学的方法論を用いた仮説の作成と臨床試験による科学的な検証にある。また、それを支援するための情報基盤としての「Network」の整備とそれらを管理する国の専属組織の設置にある。

本研究では、NCI が 2002 年に報告したがん予防に関する情報化の具体像とその予算構成を検討することで、がん予防に有効な情報基盤整備の像を明らかにすることを目標とした。

第二項 方法

現状の米国のがん予防に関する制度について文献 NCI' Cancer Informatics Infrastructure [<http://2002.cancer.gov/infbuilding.htm#1>] を中心に調査研究し、国内のがん予防に関する制度とを比較し、両者の情報基盤整備に関する違いを明らかにする。これをもとに、今後のがん予防に有効な情報基盤の概要についてその整備状況に関する調査分析を試みた。

第三項 結果・考察

がん予防に関するデータの標準化は、米国では NCI が主導する NCI' Cancer Informatics Infrastructure の中で一定の基準に基づく知識管理フレームワークである cancer informatics infrastructure と知識を収集、解析、応用、再利用するためのツール群として設計している。これは、がんに関する基礎研究や橋渡し研究、臨床や集団に基づいた研究を行う集団間のインタフェースやがん診療を提供する医療者とのインタフェースやがんに関する情報を必要とする消費者とのインタフェースを開発することを目指している。特に①がんに関する各種情報へ簡単にアクセスを可能とすること。②時間と労力を抜本的に少なくするために情報を生成し、利用すること。③臨床と公衆衛生への介入研究結果の迅速な橋渡しを促進させることに重点を置き、研究とがん医療を行っているコミュニティの融合をはかろうとしている。

その第一ステップとしてあげられているのが、情報基盤を用いた臨床試験の実施プロセスの加速化とされている。臨床試験の実施される期間内で集められるデータを報告するための判断基準や標準の見直しを行い、全てのタイプの臨床試験に関わる標準形式の用語や報告の必要条件を発展させている。この均一化は、結果報告の速度や効率、正確度を向上させることが期待されている。また、医学や科学用語に関する特別なデータベースは、研究者や臨床医を支援し、NCI 情報システムを利用する他の利用者は、探したい情報を見出し理解することになる。

次に情報を統合した新しい研究支援が行われている。Cancer Genome Anatomy Projectはその代表的なプロジェクトであり、異なる科学の規範の中で情報の伝達を阻害していた伝統的な障壁に橋を架ける必須となる機構を提供することが計画されている。

NCIの情報基盤は国民に対して、①コンピュータシステムの互換性の問題の解決すること。②基礎的な発見に関する情報の中で急速に成長しているものを管理すること。③発見とその応用の間に存在する現状のボトルネックを緩和すること。④迅速な発見につながること。⑤コストを削減すること。⑥がんのコミュニティの中の全ての参加者にアクセス可能な情報を作るための方法として電子的にデータを獲得することを目標としている。

1. Clinical Trials Information System:

NCIは、臨床試験における研究者のペーパーワークを減少させ、臨床腫瘍学におけるプロトコルの開発や進捗の確認を研究者により効率的な実施を支援する。このシステムは、全ての臨床試験のライフサイクルと関連し、プロトコルの作成や患者の募集の支援やスクリーニング、臨床試験の実施から臨床試験結果の報告までを支援するために情報を分配する。

- 文書管理ツール:
- リアルタイムの薬剤在庫管理: 臨床試験で使用されている薬剤の在庫管理。
- NCIがサポートしている臨床試験の臨床データ報告機能: 臨床データを薬剤開発や臨床研究を促進させるためのNCIへの報告を規格化することで簡素化すること。
- common data elements や clinical trial enterprise models を用いたシステムとして Clinicaltrials.gov [<http://www.clinicaltrials.gov/>] という Web ベースでの臨床試験を支援するシステムが利用されている。

2. Enhancing Electric Cancer Information Services:

NCIは、XML技術を基本とした情報基盤整備を包括的ながん情報に関する成果物やサービスを PDQ や CancerNet も含めて支援している。

3. Common Scientific Outline:

NCIは、広範囲の科学的で疾患と関連する研究について分類できるツールである。予防関連の項目は下記の通りである。

Prevention

3.1 Interventions to Prevent Cancer: Personal Behaviors that Affect Cancer Risk

Examples of science that would fit:

- Research on determinants of personal behaviors that affect cancer risk, such as diet, physical activity, sun exposure, and tobacco use
- Interventions to change personal behaviors that affect cancer risk

3.2 Nutritional Science in Cancer Prevention

Examples of science that would fit:

- Quantification of nutrients and micronutrients
- Studies on the effect(s) of nutrients or nutritional status on cancer incidence

- Dietary assessment efforts including dietary questionnaires and surveys
- Development, characterization, and validation of dietary/nutritional assessment instruments

3.3 Chemoprevention

Examples of science that would fit:

- Chemopreventive agents and their discovery, mechanism of action, development, testing in model systems, and clinical testing

3.4 Vaccines

Examples of science that would fit:

- Vaccines for prevention, their discovery, mechanism of action, development, testing in model systems, and clinical testing

3.5 Complementary and Alternative Prevention Approaches

Examples of science that would fit:

- Discovery, development, and testing of complementary/alternative prevention approaches such as diet, herbs, supplements, or other interventions that are not widely used in conventional medicine or are being applied in different ways as compared to conventional medical uses
- Hypnotherapy, relaxation, transcendental meditation, imagery, spiritual healing, massage, biofeedback, etc., used as preventive measures

3.6 Resources and Infrastructure Related to Prevention

Examples of science that would fit:

- Informatics and informatics networks; e.g., patient databanks
- Specimen resources (serum, tissue, etc.)
- Epidemiological studies pertaining to prevention
- Clinical trials infrastructure
- Statistical methodology or biostatistical methods
- Centers, consortia, and/or networks
- Education and training of investigators at all levels (including clinicians)

上記のように米国では、がんの予防に関するプロジェクトだけでも 1285 件あることが Cancer Research Portfolio のデータベースから検索できる。その分類項目としては、1. 発がんリスクと関連する個人の生活習慣、2. がんよぼうにおける栄養科学、3. がんの化学予防、4. がんに対するワクチン、5. 相補・代替医療によるアプローチ、6. 予防に関連する資源や基盤に関する研究の6つに分けられている。

4. NCI Research Resources: (<http://resresources.nci.nih.gov/>)

研究者に対して NCI の研究資源を、交渉や知的所有権問題なしで低コストでアクセスし、情報を提供できるようになっている。カテゴリーは100以上あり、下記のように分類されている(図 II-1)。

- ① Animal Resources:
 - ・ Registries of tumors in animals.
 - ・ Animal production labs.
- ② Cancer communication resources:
 - ・ Services, information for patients and providers.
 - ・ Visuals, multimedia.
- ③ Clinical Trials:
 - ・ Clinical trials & cooperative groups.

- Cancer Therapy Evaluation Program.
- ④ Drugs, chemicals and biologicals:
 - Databases of carcinogens, proteins, anti-cancer compounds.
 - Repositories of biological and chemical reagents.
- ⑤ Epidemiological resources and statistics:
 - Database management, exposure, diet, outcome.
 - Surveillance, SEER, other resources.
- ⑥ Family registries/cancer genetics resources:
 - Cancer family registries.
 - Cancer Genetics Network.
- ⑦ Genomic resources:
 - Cancer Genome Anatomy Project.
 - Mammalian Gene Collection.
 - Mouse resources.
- ⑧ NCI intramural research:
 - Research programs on-line.
 - Consultation services.
- ⑨ Research & manufacturing services:
 - Screening and production of compounds.
 - Pre-clinical development of drugs and biologics.
- ⑩ Scientific computing resources:
 - Advanced Biomedical Computing Center.
 - Software for data analysis.
- ⑪ Specimen resources:
 - Human and animal, normal and cancerous specimens.
 - Specimen information.

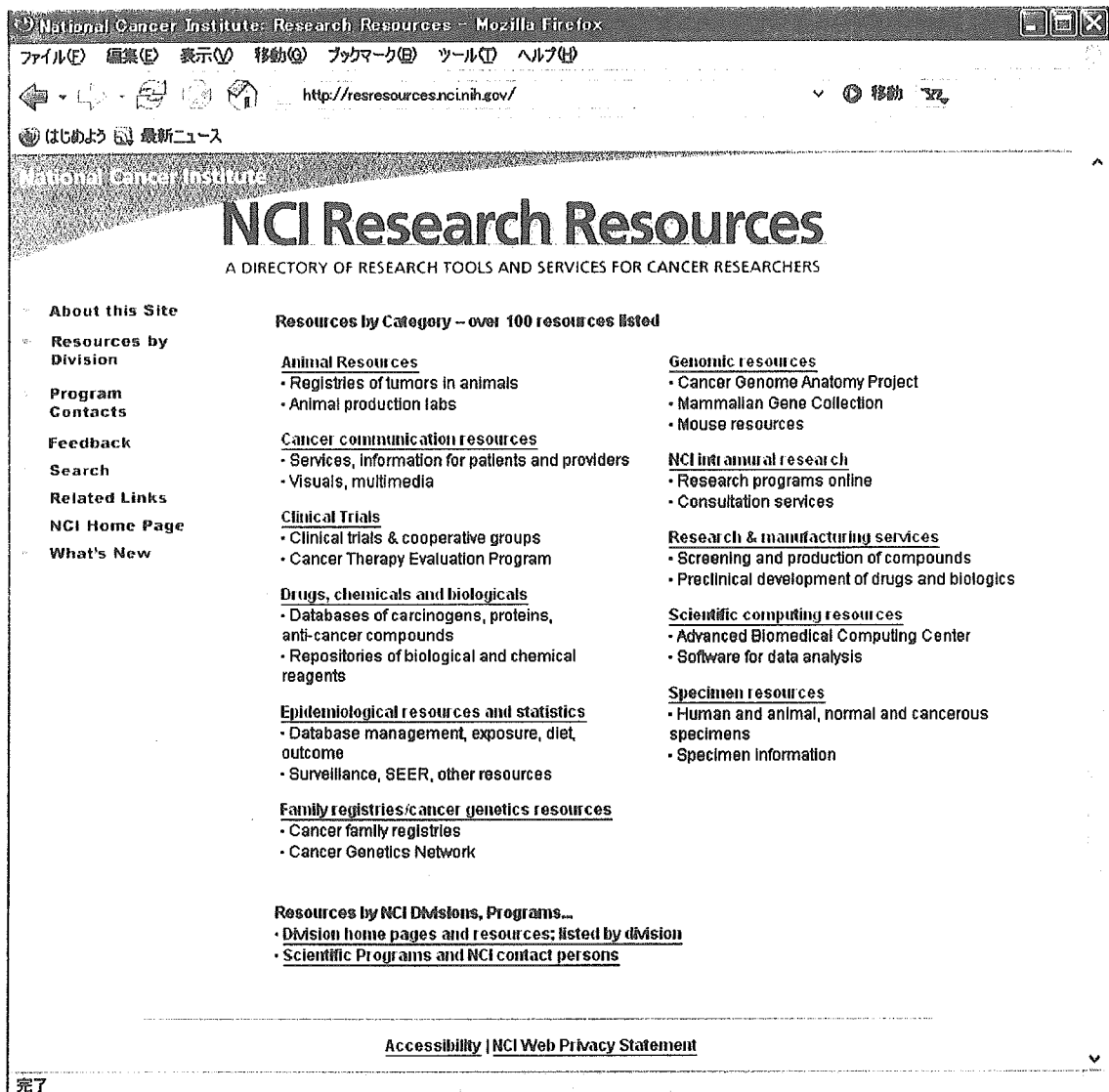


図 II-1. NCI Research Resources のホームページ。

5. Web Access to Cancer Genome Anatomy Project Data:

Cancer Genome Anatomy Project の様々なコンポーネントから発生するデータから豊富でさまざまな採集物の利用により下記が可能となっている(図 II-2)。

- 癌の関連のゲノムデータの統合モデルの開発。
- 研究者が情報のための異なった組織タイプか遺伝子異形の人口における遺伝子発現パターンなどの CGAP データを探ることを許可するポータルサイトの開発。

CGAP - Mozilla Firefox
 ファイル(F) 編集(E) 表示(V) 移動(Q) ブックマーク(B) ツール(T) ヘルプ(H)
 http://cgap.nc.nih.gov/ 移動 Y2
 はじめよう 最新ニュース

CGAP The Cancer Genome Anatomy Project

CGAP HOW TO Data Chromosomes Tissues SAGE Genie RNAi Pathways Tools

CGAP Info

- ◆ Educational Resources
- ◆ Slide Tour
- ◆ Team Members
- ◆ References

CGAP Data

- ◆ Download

Quick Links:








- ◆ ICG
- ◆ NCI Home
- ◆ NCICB Home
- ◆ NCBI Home
- ◆ OCG

The CANCER GENOME ANATOMY PROJECT

The goal of the NCI's Cancer Genome Anatomy Project is to determine the gene expression profiles of normal, precancer, and cancer cells, leading eventually to improved detection, diagnosis, and treatment for the patient. By collaborating with scientists worldwide, CGAP seeks to increase its scientific expertise and expand its databases for the benefit of all cancer researchers.

The CGAP Web Site

Interconnected modules provide access to all CGAP data, bioinformatic analysis tools, and biological resources allowing the user to find "in silico" answers to biological questions in a fraction of the time it once took in the laboratory.

	Genes Gene information, clone resources, SNP500Cancer, GAI, and transcriptome analysis <small>NCBI</small>		Chromosomes FISH-mapped BAC clones, SNP500Cancer, and the Mitelman database of chromosome aberrations
	Tissues cDNA library information, methods, and EST-based gene expression analysis		SAGE Genie Analysis of gene expression using <small>NCBI</small> and <small>NCBI</small> SAGE tag data for both <small>NCBI</small> and <small>NCBI</small>
	Pathways Diagrams of biological pathways and protein complexes, with links to genetic resources for each known protein		Tools Direct access to all analytic and data mining tools developed for the project
	RNAi <small>NCBI</small> RNA-interference constructs, targeted specifically against cancer relevant genes.		

New Initiatives

- ◆ The NCI supports the [Initiative for Chemical Genetics \(ICG\)](#), which is developing a systematic approach for harnessing synthetic chemistry to discover molecular mechanisms in basic cell biology. This approach aims to emulate the success of classical genetics as a discovery platform using chemicals that alter the function of specific proteins in place of mutations. These results will be incorporated into the Gene Info pages of the relevant genes.
- ◆ The OCG/NCI and NHGRI recently convened a workshop to assess the value of a project to catalogue all of the DNA sequence changes that occur in tumorigenesis. A summary of the workshop and its recommendations are available at [Exploring Cancer through Genomic Sequence Comparisons](#). Follow-up information will be posted on the CGAP web site as it becomes available.
- ◆ In July 2005, the NCI and NHGRI convened a workshop in Washington, DC entitled "Toward a Comprehensive Genomic Analysis of Cancer". This meeting brought together physicians, basic scientists, and other members of the U.S. and international cancer communities to assist in outlining the most effective strategies for the development of a successful pilot phase of the human cancer genome characterization project. Additional information about this meeting is reported in the [Executive Summary](#).

完了

図 II-2. Cancer Genome Anatomy Project のホームページ。

6. Web-Based Tools for Information Dissemination:

NCI は、NCI が主導する専用の Web site の迅速な開発と設置、統合、管理を支援する一つのコンピュータプログラムを開発し、実装させてきた。このプログラムは、Cancer Genome Anatomy Project (CGAP) のためのプロトタイプである Web site 設置に使用された。

また、NCIは、オープンソースの情報アーキテクチャのためのプロトタイプを開発した。これにより、複数で依存性のないデータの蓄積や統合を推進することができたとされている。

これらの情報基盤整備にかかった予算を下記に示す。総額で約77ミリオン\$であり、日本円に換算すると85億3千万円/年となる。1993年に国内でも77億円という予算が国のがん対策用スーパーコンピュータの導入経費となり、事業化されたが、ほぼ同額の予算規模となっている。

表 II-1. NCI の情報基盤整備に関する予算項目

SUMMARY		
	Informatics and Information Flow Objectives	2002 Budget Request
1	Expand NCI's informatics infrastructure to enable integration and interface	\$35.3 M (約 38.8 億円)
2	Make information more readily accessible to the entire cancer community.	\$21.0M (約 23.1 億円)
3	Make informatics tools easily accessible and use them to integrate and disseminate information	\$18.0M (約 20.0 億円)
	Management	\$3.0M (約 3.3 億円)
	Total	\$77.3M (約 85 億円)

しかし内容的にみると10年の時代の流れは、大きく計算機・通信科学技術の進歩により内容は異なり、1993年国内で予算化されたがん対策への情報基盤整備に関するもののほとんどがハードウェア経費であったのに対して、NCIでは以下のようになっている。

表 2-2. NCI の情報基盤整備に関する予算概要

2002 Milestones	2002 Budget Request
NCI 支援する研究を主導する施設内や施設間の情報の交換を促進させたり、NCI の Extraordinary Opportunities を通して発生するデータやツールの統合を支援するバイオインフォマティクスセンターの設立。	\$ 20M (約 22 億円)
がんに関する種々の研究活動を支援するための操作原理や基準となるセットに基づいたオープンソースの情報アプリケーションやサービスのツールボックスの開発。	\$ 5M (約 5.5 億円)
コミュニティの間のがんに関する情報やデータの全てのタイプのさらなる交換に、common data element や common vocabulary の使用を達成したり設立させる研究ツールの開発。	\$ 10M (約 11 億円)
国の標準開発プロセスにおける NCI の役割を正式なものとし、情報学の研究ツールの開発に標準仕様を組み込むための NCI 情報標準規格化のための常設のレビューパネルの構築。	\$ 0.25M (約 2750 万円)
NCI がん情報システムの整備:画像や音声、ビデオのようなマルチメディアを含めたリッチフォーマットや modular text document も含めたがんに関する情報資源の広い範囲での作成、統合、アクセスのための強力な促進。この技術的な基盤は、適切にアクセスすることや、がん	\$ 3M (約 3.3 億円)

2002 Milestones	2002 Budget Request
情報への容易に理解を改善することを必須とする。	
約300のがん情報のサマリーと現在進行している1800件の臨床試験の記載内容を含めてNCIの情報成果物を再構成しインデックスを再番する。NCIの再構成したがん情報システムの最大の利点はあらゆるタイプの利用者に対して関心事項に関連した情報を簡単に特定し解釈することができることである。また、異なった情報層の間をシームレスにし、必要に応じ最も妥当なレベルを見出すことを可能とする。	\$ 7.5M (約 8.3 億円)
NCIの包括的な電子的がん情報システムを拡張させた Web ベースの検索基盤とユーザインタフェースの開発。研究成果や臨床試験に関する情報に対して一般から簡単にアクセスすることができるシステムの開発。がんに関する用語の拡張は、単純な検索や専門的な検索のオプションの両方を作成することを支援し、利用者の特別に要求する情報に対してインタラクティブなインタフェースを含む。	\$ 5.0M (約 5.5 億円)
Concept based searching や自然言語処理の使用に対する妥当性の評価の継続。NCIがこのような技術を取り入れることに対して利益があるかどうかを決めるために現存する NCIのがん情報システムに対してパイロット研究を行う。	\$ 1.5M (約 1.65 億円)
NCIが後援していない臨床試験(薬品メーカーが行っているものやヨーロッパの臨床試験組織が行っているもの)情報を自発的に電子登録活動に関する計画を立てる。	\$ 1.0M (約 1.1 億円)
臨床試験の記載やレビューのプロセスを3ヶ月以内に完了することができるように変更すること。	\$ 3.0M (約 3.3 億円)
NCIによって開発されたツールや技術を管理し普及させること機能を拡張させること。	\$ 0.5M (約 5500 万円)
統一され簡単にアクセスでき、臨床試験の募集や全ての参加者に科学的な知識の交換を促進させるための総合的な電子的な臨床試験システムの開発。	\$ 8.0M (約 8.8 億円)
研究コミュニティに科学的で臨床的な情報を提供するための広域の情報アクセスポイントとしての Cancer Genome Anatomy Project (CGAP)のような情報ポータルを開発すること。	\$ 5.0M (約 5.5 億円)
グラントや契約に関する審査を促進するシステムやツールの拡張と NCI cancer research portfolio の管理を簡素化するシステムの拡張。	\$ 3.0M (約 3.3 億円)
NCIの総合的ながん情報システムから情報を直接検索・保存するための情報サービスやNCIの管轄でない Web サイトのための方法を開発提供すること。	\$ 1.5M (約 1.65 億円)

特徴は、① Extraordinary Opportunities と称される既存の研究遂行プロセスのリエンジニアリング 特にバイオインフォマティクスセンターの設置に代表される生命科学における情報をコンピュータ科学を用いた解析とそれによる新しい発見の促進、②既存の臨床試験で使用されている情報処理を電子化することによるペーパーワークの削減と制度管理、③そしてなによりも国民むけのがん情報提供サービスや研究者も含めたコミュニティの支援のためのインターネットと Web 技術を用いた情報基盤整備にある。

特に、そのなかでも注目されるのが、がんに関する情報を電子的に処理できるような「用語の標準化」や「機械可読性を有する統制用語(オントロジー)の作成」にある。つまり、人間のがんに関する情報処理を計算機を用いることにより高速かつ大量にデータ処理を行うことを可能とし、検索や情報伝達にかかる処理時間を大幅に削減させたり、研究者や医療者だけでなく一般の国民特に

がん患者や患者の会に対して、科学的信憑性の高い情報を Web 技術を用いて提供したり、臨床試験情報を提供することで国民の臨床試験に対する参加を呼びかけるシステムを構築している。

国内におけるがんに関する情報提供は平成5年頃から先に紹介した国立がんセンターがん情報サービスによってホームページ上で行われたり、大阪府立成人病センターや各地方がんセンターでも徐々にがんに関する情報コンテンツは充実されているかのように見える。しかし、米国と大きく異なるのは、①がんに関する情報をコンピュータ処理するための計算機可読性を有する日本語における用語の整備が行われていないこと。②がん情報の中でがん検診や治療方針に関する情報提供は盛んに行われているが、がん予防に関する情報提供は Google を検索しても国立がんセンターやNHKの番組のホームページがヒットするが、科学的に信憑性の高い情報を提供しているサイトは殆ど見当たらなかった。

NCI では、上記の予算を用いて <http://www.cancer.gov/prevention/> を充実させ、がん毎の予防法を医療者だけでなく一般向けにもわかりやすく紹介している prevention and Detecting Specific Cancers やがん予防に有益な生活習慣情報の紹介(LIFESTYLE AND PREVENTING CANCER) やがん予防に関する臨床試験情報の紹介(PREVENTION AND DETECTION CLINICAL TRIALS)、がん予防に関するイベント等を紹介する EVENT AND MEETINGS、がん予防や検診に関するプログラムの内容の紹介や高精度のがん予防や発見に関する臨床試験やフェロワーシッププログラムなどの情報を提供する KEY Initiatives in Prevention and Detection, Information for Researchers、がん予防に関する研究を行っている研究施設を基礎研究(Basic Prevention Science, Biometry, Cancer Biomarkers, Chemopreventive Agent Development, Community Oncology and Prevention Trials, Early Detection, Nutritional Science)や乳がんや消化器がん、肺がんなどの臓器がんの専門の研究施設の情報として Research Groups in the Division of Cancer Prevention (DCP), About the Division of Cancer Prevention (DCP)に関する情報が提供され更新されている。

残念ながら国内でこのようがん予防専門のサイトは見出しえなかった。一部がん予防に関するサプリメントと食材についての情報提供は、国立健康・栄養研究所 (<http://www0.nih.go.jp/eiken/>) の「健康食品」の安全性・有効性情報に関する素材情報データベースが有益であった。

国内でのがん予防に有効な情報処理基盤の整備を考えると、NCIのような膨大な予算を用いた予算化と組織化は、本研究では困難である。しかし、現実問題として社会には科学的信憑性の乏しいがん予防に関する情報が流通し、科学的信憑性の高いがん予防に関する知識流通が不足していることは容易に推測できる。

この状況を改善する手段として本研究では、①科学的信憑性の高いがん予防情報の普及を促進させるため、②情報基盤として現在使用されている Web 技術を用いた情報基盤の設計、③今後社会情報基盤として注目されている SemanticWeb と Grid 技術を用いた情報基盤の応用可能性について勘案し、がん予防に有効な情報基盤研究の計画について下記の3つの段階を設定した。

第一段階(がん予防情報処理基盤の設計) : NCI の中で具体化されているがん全体に関する日本語用語の統制化については困難であると考え、本研究では、がん予防をドメインとした用語について概念の構造化をテーマとした。また、バイオインフォマティクス情報のがん予防に関する応用を推進するために、バイオ関連のデータベースとがん予防に関する臨床情報とのデータマイニング等の統合解析を推進するための臨床ゲノム統合データベースの論理設計を目指した。がん予防に有効と思われるデータの殆どは、現段階では大規模なコホート研究から出されている。本研究ではコホート研究に関する情報処理の効率化に関する検討は、他の研究班で研究されていることもあり行わなかった。

第二段階(がん予防のための機械学習技術の応用) : がん予防に有効な情報処理への機械学習

の応用を行うための基盤システムとして、既述の臨床ゲノム情報データベースの開発を目指した。また、がん予防に関する生活習慣改善情報提示の個別最適化に関する応用研究を行うこととした。

第三段階(がん予防に有効な実時間統計処理・推論・予期技術の開発):がん予防に有効な実時間における統計処理や推論、予期技術の開発を具体化するための基盤システムの提案を目指した(図 II-2)。

そこで本研究では、主に第一段階の内容を主に研究を進めていくこととした。

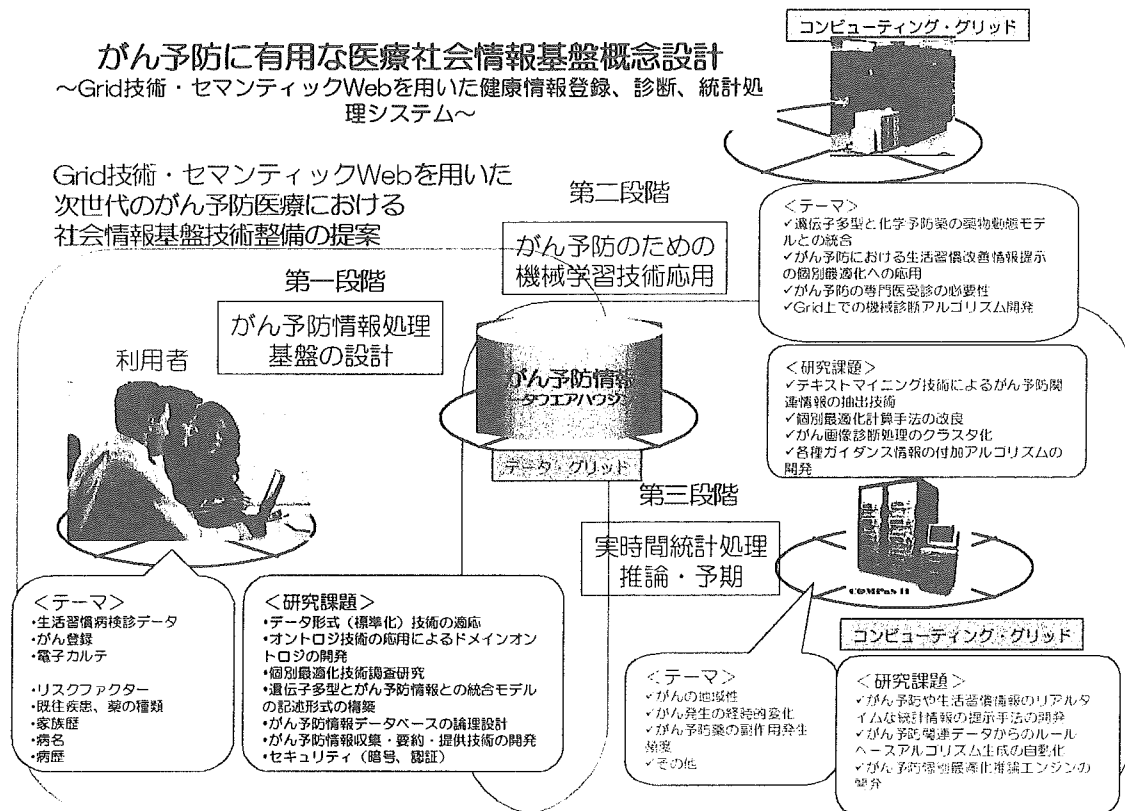


図 II-3. がん予防に有用な医療社会情報基盤に関する概念設計
 (患者の方への科学的信憑性の高いがん予防に関する情報提供を中心とした情報基盤)

第五項 結語

- NCI のがんに関する情報基盤整備のみで約 85 億円予算があることがわかった。
- 研究遂行プロセスのリエンジニアリング、特にバイオインフォマティクスセンターの設置に代表される生命科学における情報をコンピュータ科学を用いた解析とそれによる新しい発見の促進、②既存の臨床試験で使用されている情報処理を電子化することによるペーパーワークの削減と制度管理、③そしてなによりも国民むけのがん情報提供サービスや研究者も含めたコミュニティの支援のためのインターネットと Web 技術を用いた情報基盤整備に重点が置かれていた。
- 上記の情報化を実現するために情報基盤整備としてデータの標準化の基本となる用語の標準化と計算機可読性を有するオントロジーの開発の重要性が極めて高いこと

がわかった。

参考資料:

1. NCI' Cancer Informatics Infrastructure: <http://2002.cancer.gov/infbuilding.htm#1>
2. Clinicaltrials.gov: <http://www.clinicaltrials.gov/>
3. NCI Research Resources: <http://resresources.nci.nih.gov/>
4. 国立がんセンター: <http://www.ncc.go.jp/jp/ncc-cis/pub/index.html>.
5. 大阪府立成人病センター: <http://www.mc.pref.osaka.jp/>
6. NIH: <http://www.cancer.gov/prevention/>.
7. 国立健康・栄養研究所. <http://www0.nih.go.jp/eiken/>.

第三章 がん予防に有効な情報処理基盤に関する検討

III-1. 計算機可読性を有するがん予防に関する用語に関する検討

III-2. がん予防情報データベースの設計

III-3. 遺伝子多型を加味した薬物動態モデル

III-4. 個別最適化がん予防情報提供のための情報システム

III-5. 在宅におけるがん予防情報提供システム

III-6. Grid 技術を用いたがん予防情報検索システム

以下項目毎に記載する。

III-1. 電算機処理可能ながん予防に関する用語に関する検討

A. 目的と背景

NCIのがん予防対策の特徴は、情報通信技術の利活用による新しいがん対策の具現化にあることは既に述べた。特に、その対象は、がんに関する基礎研究者のためのみならず、医療関係者、特にがん治療を受けている方への科学的信憑性の高い情報を提供と情報交換を促進し、新しいコミュニティを創出することを主目的としているようにも捉えることができる。これは、がん予防においても同様で、特にライフスタイルとがんのリスクとの関連性が高く、個人のみならずコミュニティや社会全体の制度との関連が強いがん予防分野における情報通信技術の活用は欧米でも莫大な投資が行われ、新しい社会保障基盤として進化している。

社会情報基盤とは何か。一般に情報通信分野における情報基盤とはネットワークやコンピュータ機器、セキュリティや認証に関する技術をもって考える傾向が強い。現在の国内の情報基盤はここ数年急速に整備されてきたが、どんなに速いネットワークを敷いてもその中を走るデータがなければ意味がない。一方、米国のNLM(年間予算約300億円、スタッフ数約700人)では、1984年から英語圏の医学用語の標準化と計算機可読性を有するオントロジーの開発を行っている。つまりコンピュータやネットワーク網の整備のみならず、その中を走る情報の標準化を行うことで情報処理の効率を格段に向上させることで、人間のがんに対する研究や医療や教育等における意思決定支援を目指している。

本研究では、がん予防の情報処理基盤の重要な要素であるがん予防に関する知識の体系化を図り計算機処理できるため必要となる主にごん予防に関する文献データをもとにしたがん予防に関するドメインオントロジーを開発することを目標とした。

B. 方法

1. オントロジーの開発仕様の作成:

本研究の目的は、生活習慣病の化学予防薬となるターゲットの網羅的探索を可能とするためのドメインオントロジーの開発である。このために本研究では、がんの予防薬のターゲットのオントロジーを用いたシナリオ作成支援を目指した。

オントロジー構築の方法論には、各種あるが本研究では Uschold と King らの方法をとることとした。その理由は、本研究の目的であるがんの化学予防薬ターゲットの網羅的探索がゲノム創薬の中心的过程であり、彼らが提案している Enterprise Ontology の構築方法との類似性が高いことによる。上記も含めて既存のオントロジーについての調査研究対象として NCI の Cancer Ontology に関する調査分析をおこない、がんの化学予防と関連が深い「化学薬品・薬剤」概念の日本語化を行った。

がん予防情報の基盤の中でもがん予防に利用されている日本語の用語体系化は、情報検索やデータベースの構築において重要である。その理由の一つは、がんの一次予防法としては喫煙や飲酒などの生活習慣情報や食事療法、サプリメント情報など既存の医学用の範囲とはことなる用語が使用されていることが予想される。この問題に関する検討は今後の課題であるが、本年度は大阪大学溝口研で開発されたオントロジエディタ「法造」ver4.0 を用いて NCI のオントロジの中でもがん予防と関連の深い「化

学薬品・薬剤」の日本語化を行い、がん予防に関するオントロジ開発の仕様調査のマイルストーンとした。

2. オントロジの構築:

a. オントロジの概略獲得(自然言語を用いたインフォーマルなオントロジの定義を行う。)

i. オントロジの範囲の決定

1)ブレインストーミングによる重要概念の収集。広い範囲にわたる概念に対応する用語が全く構造化されずに収集する。

2)化学予防に関する文献を収集し、アノテーションを作成する。言語構造の解析と修正を行う。これらのアノテーションは XML で記述され、XML データベースで管理する。

ii. 得られた概念の分類と対象となる範囲(ドメイン)の特定。

1)用語には重要度を割り当てて、同義語などを削除しながら重要概念を抽出する。

iii. 対象概念のセットの抽出。

1)抽出の基準は、化学予防薬開発における役割の重要性が意味定義に合意できるものとする。

iv. 概念の名前の決定

1)概念の名前としては一意で不自然でないものを選択する。適当な名前がなければ作成する。

v. 単語名の決定。

1)辞書の意味定義である実際に使われ方の説明を決定する。

b. オントロジの定義。

i. メタオントロジとよばれる実体、関係、状態、ロールという4つの概念を導入して定義する。自然言語による曖昧さをなくすことを目的としてオントロジエディタを用いて概念の意味を定義する。

ii. ID に対応したオントロジのエントリーを作成する。XML ではなくネットワーク構造を用いた推論ができるように RDF (Resource Description Framework)を用いる。

C. 結果

(1) がん予防に関する概念構造の作成

がん予防に関する出版物^{1,2,3,4,5)}に記載されている内容をもとにして先ずがんの一次予防に関連する概念化の前段階として FreeMind を用いてがん一次予防ドメインの中の概念の体系化を試みた(図 III-1.1)。がんの一次予防の場合、がん一次予防に関連する概念として「行為」とその行為の基となる「根拠(エビデンス)」、がんの一次予防に関連する「制度」の3つとした。発がんリスクをあげる環境の改善、行動の改善、体質や体型の改善、発がんに関連する既存の医療行為に分けた。

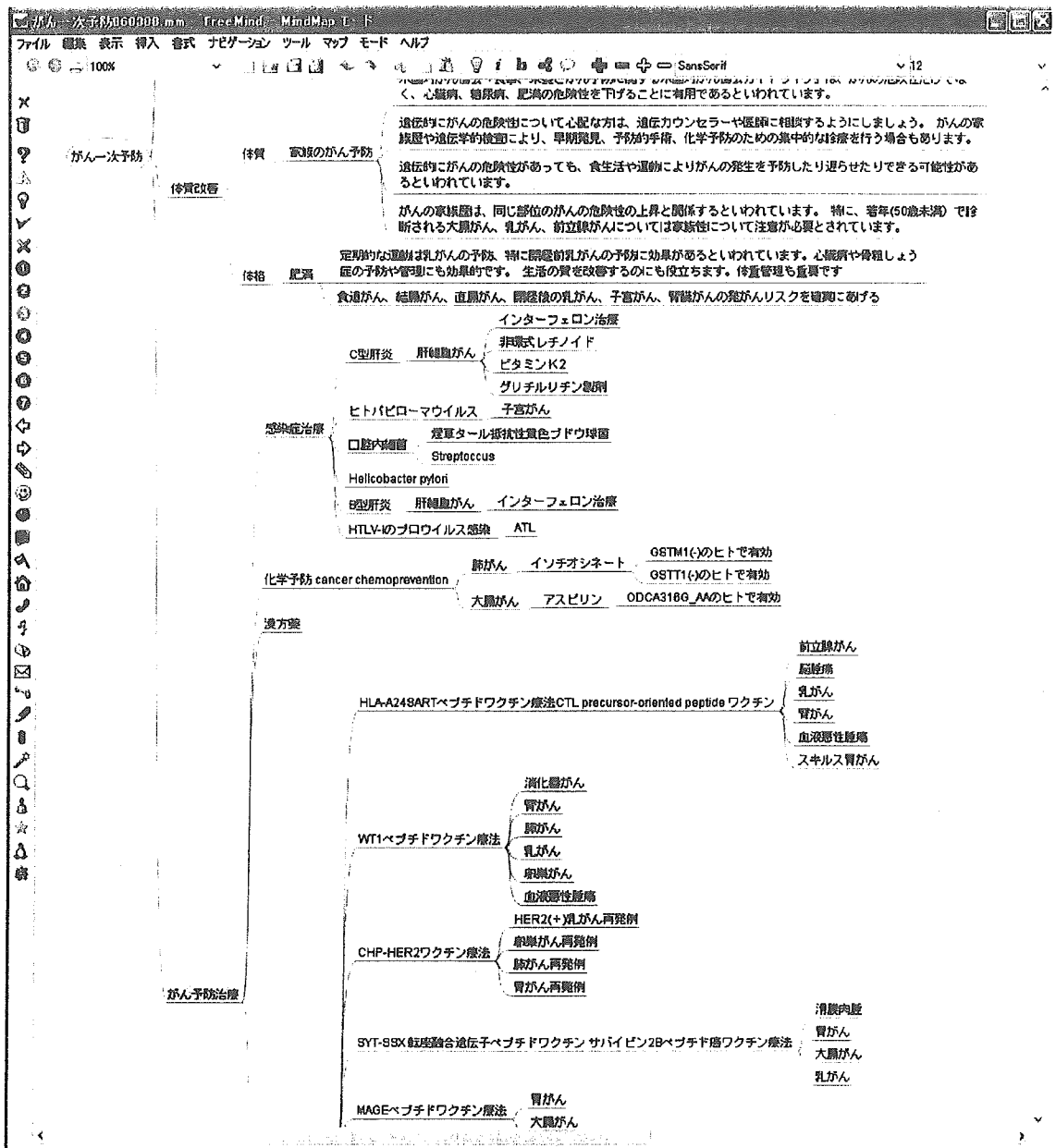


図 III-1.1 FreeMind を用いて作成したがんの一次予防に関する事象の構成例。

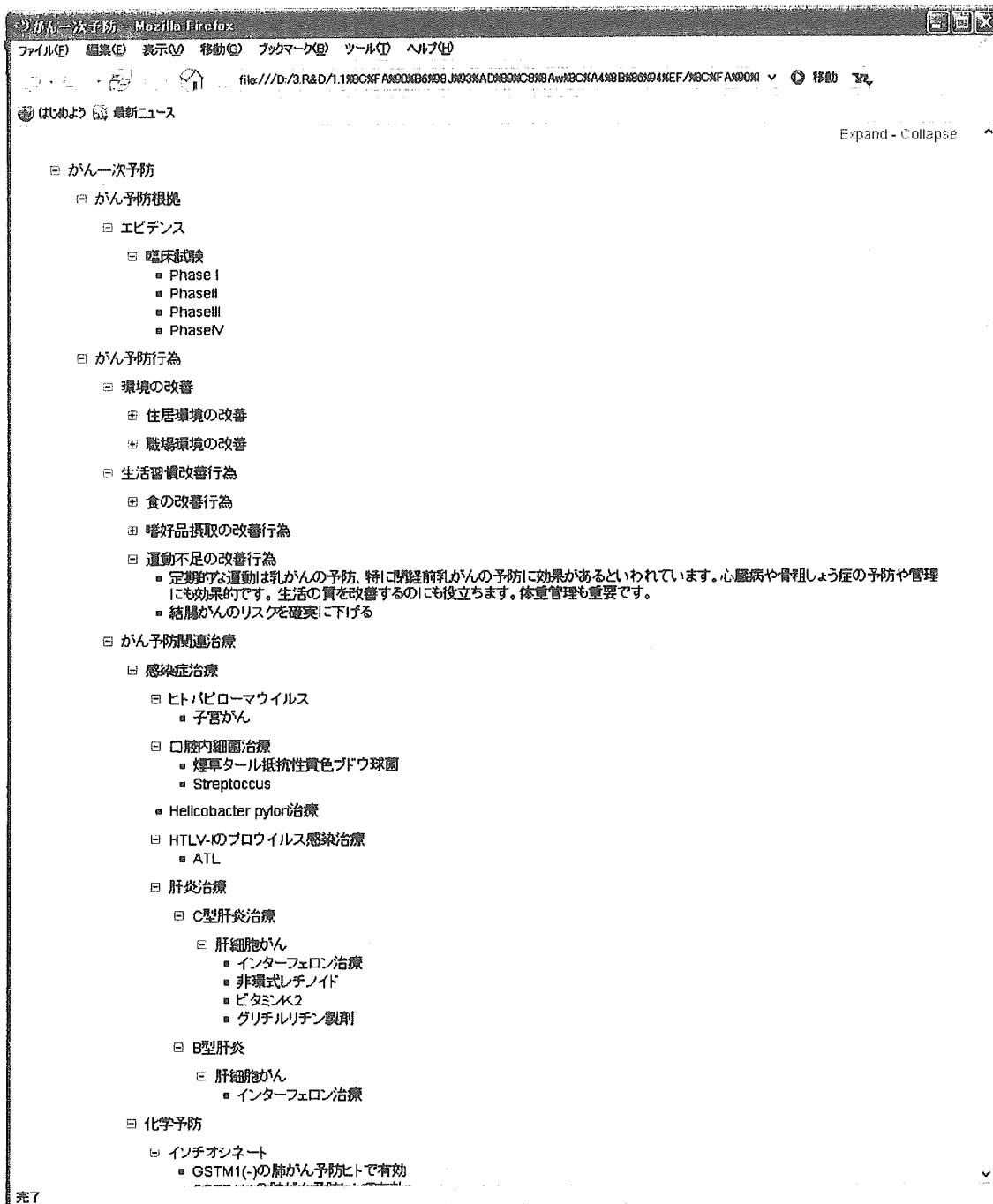


図 III-1.2. FreeMind で XHTML (JavaScript 版) で出力したがんの一次予防に関する事象の構造化表現.

図 III-1.2 JavaScript 版 XHTML を用いて作成したがんの一次予防に関する事象の構成例.

(2) Cancer Ontology の日本語化:

大項目としては、「化学薬品・薬剤の機能」と「化学薬品・薬剤の機能構造」、「使用しなくなった化