

薬機能オントロジー (Drug Interaction Ontology) :分子機能と生体現象の関連化

吉川澄美 松村和美 小長谷明彦

理化学研究所ゲノム総合科学研究センター(GSC) ゲノム情報科学グループ
〒230-0045 横浜市鶴見区末広町 1-7-22 E209
{sumi, konagaya@gsc.riken.jp}

要旨

薬物-生体物質相互作用 (DBI:Drug Biomolecule Interaction) は、薬物間相互作用のみならず、薬効の作用機作においても重要な概念である。そこで、薬物と生体物質の相互作用は薬理知識の使い回し可能な要素単位とみなせる立場のもと、これらを形式化し薬機能オントロジー (DIO:Drug Interaction Ontology) として関係データベースへ実装した。このオントロジーの利用例として薬物間相互作用の予測を示す。さらに分子レベルの相互作用と生物機能・現象を繋ぐ3項関係の基礎概念について論じる。

1 はじめに

分子生物学やゲノム科学の進展に伴い、個々の薬物が生体内でどのように作用するか詳細に解明できるようになった。新薬が販売されるまでに、ターゲット分子や主要な代謝経路などの基本的な作用はある程度調べられているが、市場に出はじめて見出される反応もある。その原因は、動物実験での限界、管理下で行われる臨床試験からの知見の限界などに由来し、その背景には遺伝子多型や環境要因を含めた発現状態の多様性で説明できる場合もある。また古くから使用されている医薬品においても、実際に生体内で作用する物質の種類や特異性が多様であることがわかってきた。このように、薬物と生体物質との相互作用は薬効や副作用の機構、さらに反応の個人差を説明する上で基本的な知識であり、薬物間相互作用 (薬の飲み合わせによって引き起こされる様々な反応) においても同様である。そこで、薬物と生体物質の相互作用を使い回し可能な情報単位として扱うことにより、専門家が行う推論を機械処理で支援できるのではないかとという想定にもとづいて、薬物-生体物質相互作用の知識表現としてオントロジーを設計した[1]-[3]。ここで扱う「相互作用」とは生体で起こる物質間の作用関係を広く捉え、物理化学的な直接作用 (酵素反応、受容体への結合、分子輸送、結合による分子機能阻害など) に加えて間接的作用 (薬物による遺伝子発現など) も含む。

個々の薬物の生体内での挙動の情報が文献上、または専門家のドメイン知識として蓄積されているとはいえ、解明されているのはまだ一部分である。しかしながらそれらの量的・質的に制限のある情報を駆使して、臨床現場での薬物選択の判断や基礎研究における仮説設計の段階で、専門家はさまざまな推論を行っている。推論においては同レベルの情報を関連化 (たとえば、2つの物質を分子相互作用という同レベルで関連化) させるのみならず、異なるレベルの情報を関連化する操作が行われる。ここでのレベルとは、生物学的階層¹におけるレベルを主に想定している。

例えば、薬物が生体内に取り込まれることによって引き起こされる現象と分子レベルの相互作用との関連化、疾病状態の現象と“正常”状態の現象の関連化などが挙げられる。ここで言う「現象」は巨視的に見た生物機能であり、「いずれは複数の分子レベルの相互作用の連鎖として説明できる対象」という立場をとる。Karpの立場に立った integrated function (相互作用のネットワーク) と local function (分子レベルの個々の相互作用) の関係に相当する[4]。

人が行う思考や推論においては、「現象」の概念を示す専門用語を操作することによって、まだ解明されていない分子レベルの相互作用情報を補完すると同時に、分子間相互作用の複雑な連鎖から特定の関心範囲を切り出し、それに意味を持たせている。このように人の知識としては分子レベルの相互作用と

¹ ここでの「生物学的階層」は、生体内に存在する分子、細胞構成器官、細胞、組織、器官、またはそれらを組み合わせた系、個体、個体集団 という階層をさす。

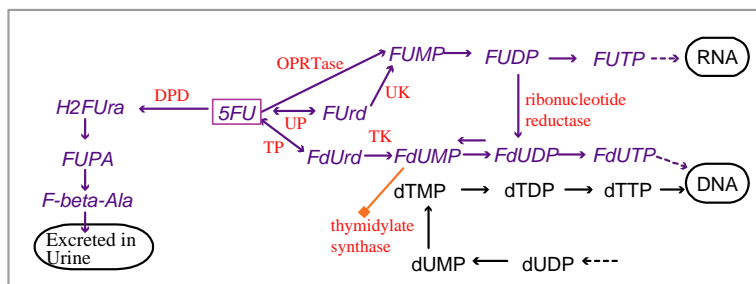
現象レベルの概念は密接に関わりあってはいるが、Gene Ontology [5][6]では両者は molecular function と biological process の別々の領域として完全に切り離されている。本オントロジーの設計においては分子レベルと現象レベルの関連化を考慮し、分子レベルの相互作用のみならず現象の情報も同時に扱えるような基本概念に注目し、概念間の関係を形式化した。

本報告では、まず薬機能オントロジーの基本要素を紹介し、次に薬物間相互作用の推論への応用例を示す。さらに、分子レベルの相互作用と生物機能または現象概念との関連化についてUMLS²[7]におけるセマンティックネットワークとの比較も交えて議論する。

2 薬機能オントロジー

2.1 情報単位

薬理学や分子生物学において一般的に行われている、図による分子レベルの反応経路（相互作用の連鎖）の表現例を示す。図表 1 は抗癌剤、5FU(5 fluorouracil)の作用機構/代謝経路を示している。このよ

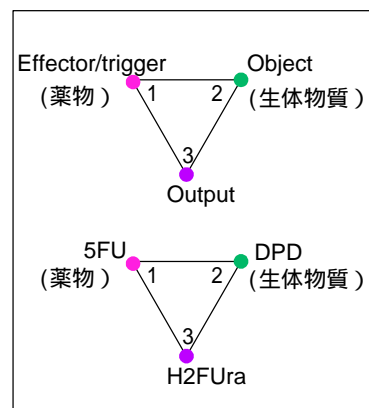


図表 1: 5FU の代謝経路と薬効メカニズム

うな表現方法においては特に厳密な表記規則が設けられているわけではなく、書き手の文脈依存の意味の割当てが行われている。この例では、文字列は物質名またはその略称であり、紫色で書かれたものは外来物質（薬物）由来であり、黒で書かれたものは内在性物質である。矢印は物質の変遷を表し、赤い文字はその反応に介在する酵素を示す。5FU の位置から右側の反応経路は 5FU の主薬効となる Thymidylate synthase の阻害や、

塩基のアナログとして DNA や RNA 合成に使用され、結果的に阻害作用に繋がる機構を示す。それに対して左側は最終的に尿中へ排出される、いわゆる解毒化の経路が示されている。

このような模式図においては一連の反応が示されているが、情報の基本単位はノード（物質）を繋ぐ個々の反応（相互作用）である。タンパク質相互作用などの情報モデルにおいては、物質 2 つから成る 2 項関係として扱われるものが多いが、我々のモデルにおいては 3 項関係を基本単位とする（図表 2 上）。すなわち、相互作用に関わる名付け可能実体³（典型的には物質名）へ 1) effector (trigger, または刺激) 2) object 3) output のいずれかを割り当てる。1)と2)の違いは相対的なものであり、2)の方が1)よりも生体（または注目している生体内部位/細胞種）において定常的に存在し、1)の存在は反応を起こすきっかけとなると解釈されるものである。従って、薬物-生体物質相互作用において薬物または代謝物などの薬物由来物質は（本来生体には存在しないので）常に 1)に割り当てる。



図表 2: 3 項関係とその例

5FU が酵素 DPD (dihydro pyrimidine dehydrogenase) によって H2Ura に変換される反応を 3 項関係概念の例として、図表 2 の下部に示した。これは薬剤が基質となる酵素反応であるが、この反応様式のみならず、酵素阻害反応、薬剤と受容体の関係、薬剤とトランスポーターの関係などいずれにおいてもこの 3 項関係を基本概念とする。

特徴的なのは、各反応様式における役割と 3 項の役割を一致させていない点である。例えば、酵素反応における、基質、酵素、生成物という「反応様式における役割」を「3 項関係の役割」へ対応させてい

² NLM (National Library of Medicine) が提供する UMLS (Unified Medical Language System) は医学関係の概念と語彙を扱い、そこから派生した MeSH のシソーラスが MEDLINE のインデックスに利用されていることでも有名である。

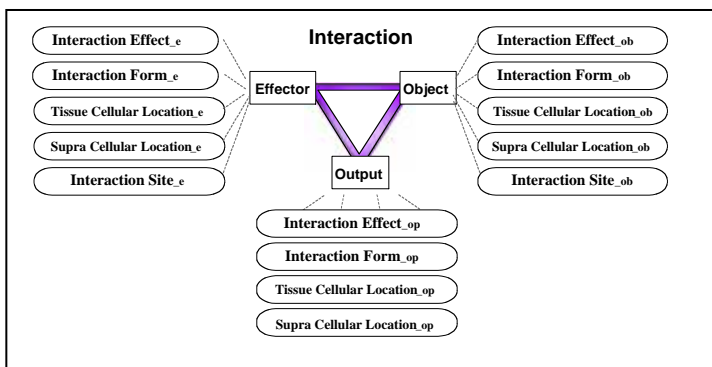
³ Nameable entity または named entity。名前(この場合、多くは専門用語)を付けて、示すことができるモノやコト。この文脈では主に物質名だが、4 章で述べるように生物機能・現象概念も含む。

る訳ではない（酵素が薬剤で、反応する生体物質が基質となる場合もある）。生体反応（相互作用）の基本概念として3項関係を設けることの利点は、異なる相互作用様式を同じ基盤で扱うことが可能となることが第一に挙げられる。「代謝経路」「神経伝達経路」「情報伝達経路」「免疫反応経路」あるいは「中枢神経系」「循環器系」などとして人が括ることができる分類も、実際の生体反応においては相互に複雑に絡み合っている。各経路や系を別々に扱うのは実際の生体反応を反映しそこなうことになる。3項関係とその属性で相互作用を記述することにより、個々の反応の物理化学的な詳細記述を可能にしなが、語彙やタグの複雑な操作を回避して反応経路をたどることができる。機械処理における単純化だけでなく、人が直感的にたどる反応経路のイメージとの対応が取りやすいことも特徴である。

3項関係の概念的背景は、生体物質の機能は他分子との関わりで記述するという立場にもある⁴。相互作用に関わる機能は一義的に行わず、相互作用の相手（または刺激要因）とその反応場所によって多様に存在しそれぞれを3項関係で記述する。生体反応の情報が蓄積するに従って、同じ生体分子の働きは同じではなく、さまざまな細胞で異なる反応に関わり、刺激や生体環境の相違によって異なる挙動をしめすことがわかってきている。3項関係によりこのような複雑な分子機能の記述へ柔軟に対応できる。

2.2 相互作用の記述

一般的に「相互作用」の概念は反応前の状態、反応励起状態（または遷移状態）、反応後の状態という時間の流れを伴い、焦点を当てる時間的フレームは研究分野や目的によっても異なる。本モデルでは遷移状態レベルの記述も可能だが、反応前と反応後の分子の相対的な存在量と機能を主に扱う。各3項へ割り当てられた物質は、反応前後の相対的存在(量)情報と機能情報、存在場所などの各

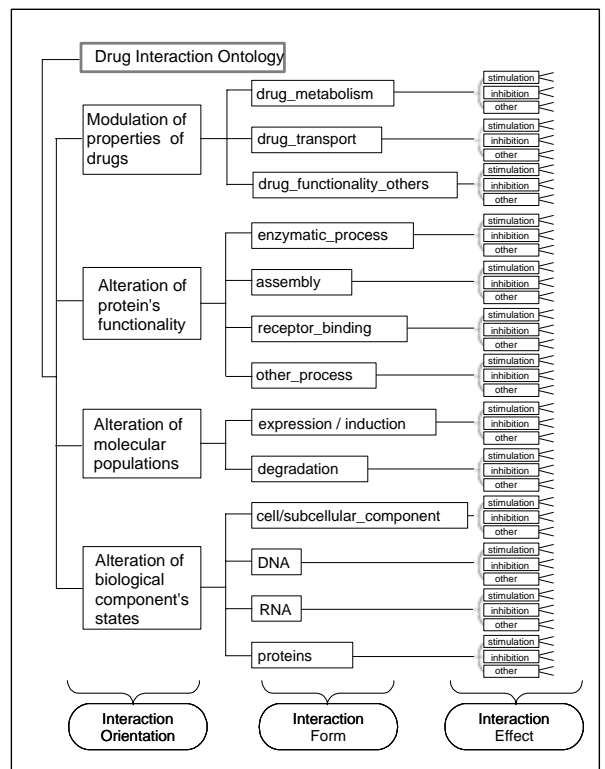


図表 3: 3項関係とその属性

属性を持つ（図表 3）。

3項関係の3番目の位置においては、「相互作用によって生成されるモノ」⁵（そしてモノに附随する機能の発生）として規定できるが、1番目の位置と2番目の位置においては、インスタンスごとに記述する必要がある。分子の存在情報に加えて、機能面での相互作用概念は1) 方向性 (interaction orientation)、2) 様式 (interaction form) 3) 効果 (interaction effect) の3つから成り立つと想定し、図表 4 に構造化した。

5-FU における例を図表 5 に示した。左側は解毒化に進む反応で、右側は細胞毒性の方向へ進む反応である。2者は同じ酵素反応様式で、似たような分子存在様式をもたらすが、機能面における効果（セマンティックな効果）の記述は異なる。それは各反応を完全に単独で扱っているの



図表 4: 薬物-生体物質相互作用の概念構造

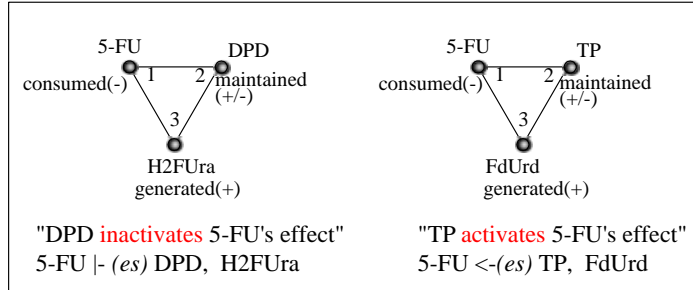
⁴ アミノ酸配列のような単一の分子としての構造的な情報は、物質の属性として named entity のテーブルに格納している。

⁵ 生物機能・現象概念の場合は「引起されるコト」

なく、薬効や解毒のような現象に繋がる反応の流れの部分としての情報を付与しているためである。

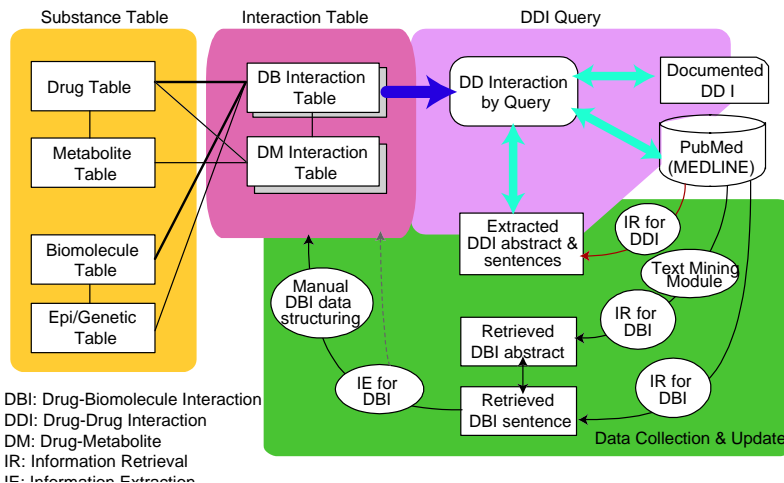
酵素反応のみならず、物理化学的な反応の記述とセマンティックな機能に関する記述を区別している点もこの情報モデルの特徴である。例えば「燐酸化」の多くは、「活性化」と対応するが、中には燐酸化状態を不活性とみなされる生体分子もある。

相互作用における自然言語の表現は多様であり、同じ表現でも反応様式依存の解釈があるので、機械処理のため、3種類の概念要素付与した記号で形式化した。図表5の例では、< (または>) や|が反応の効果を表し、との相対的位置で反応の方向性を示す。また反応様式は添字で示す(es; 酵素反応)。このような記号化した相互作用を一階述語論理でクエリーに組み込み、二つの相互作用が組み合わさった場合の集約された効果を導くことを可能にした。次節で薬物間相互作用における適用例を示す。



図表 5: 分子の存在パターンと意味的機能の相違

3 薬機能オントロジーを用いた薬物間相互作用の推論



図表 6: Drug Interaction Knowledge Base

関係データベースに薬機能オントロジーを実装し、知識ベースを構築した (図表 6: Substance table, Interaction table 部分)

薬物-生体物質相互作用と薬物とその代謝物のテーブルからクエリーを使用して薬物間相互作用の可能性を導く。相互作用のテーブル内の値一致によって関連づけられた2つの相互作用、この例では同じ酵素を相互作用の相手とする2種類の薬物の関係が導かれる。すなわち、この2種類の薬物を同時に、または連続して投与した際にどのような影響を生じる可能性があるかを推論する。クエリーは各相互作用様式と効果の組み合わせに依存したルールを採用する。図表8ではこの例で使用した基本ルールを模式的に示した。

図表7はブラウザ上での薬物間相互作用の検索結果画面を示す。まず、特定の薬物名を入力すると(sorivudine : SRV)その物質自身、またはその代謝物と相互作用する生体物質のリストが提示される(画面省略)。さらに、特定の生体物質(DPD)を選択すると、その物質を介して最初の薬物と相互作用する可能性のある他の薬物が検索される。

図表 7: 薬物相互作用の検索結果例

図表8は相互作用の連結ルールを示す。左側の表は連続的結合 (Sequential conjunction) のルールを、右側の表は多目標結合 (Multi-target conjunction) のルールを示している。各ルールには結合パターン、符号 (Sign)、およびコメント (Comment) が記載されている。

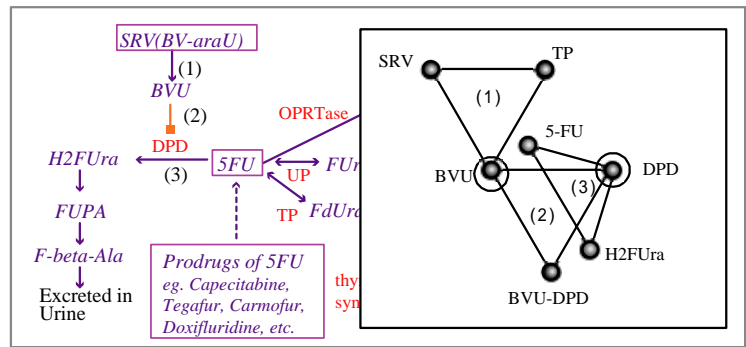
Sequential conjunction	Converged Effect		Confluent conj.	Converged Effect	
$D_1 - P - D_2$	Sign	Comment	$D_1 > P$	Sign	Comment
$\rightarrow \rightarrow$	$D_1 \rightarrow D_2$	D_1 may facilitate D_2		cX	May cause synergistic effect. See note.
$\rightarrow \rightarrow$	$D_1 \rightarrow D_2$	D_1 may facilitate D_2		cX	See note.
$\rightarrow \rightarrow$	$D_1 \rightarrow D_2$	D_1 may inhibit D_2		c~	See note.
$\rightarrow \rightarrow$	$D_1 \rightarrow D_2$	D_1 may inhibit D_2			
	Note: Scheme for each combination pattern of interaction form should be referred (omitted)		Multi-target conj.	Converged Effect	
			$P < D_1, D_2$	Sign	Comment
				mX	May interfere each other. See note.
				mX	See note.
				m~	See note.

図表 8: 相互作用の連結ルール

図表 7 の結果で示した薬物間相互作用は、図表 9 の 1), 2), 3) 部分の連結に由来する。各相互作用を三角形で表し(名前の一致で) 3 つが連なっている様子が示されるが、実際の検索結果は 5FU のプロドラック(代謝されると 5FU となる薬物) へも対応しているので、4 つまたはそれ以上の三角形(個別の相互作用) が連なって結果的に離れた位置にある薬物同士が関連づけられ、その影響について推論できたことになる。

Sorivudine の販売時 1994 年には

sorivudine と 5-FU との薬物間相互作用の危険性を直接的扱った文献はなかったが、それ以前に独立に報告されていた薬物-生体物質の文献情報 [8]-[11]のみでもそれらを 3 項関係で表現し、関連づけることによって薬物間相互作用が導かれることが示せた。



図表 9: 検索結果と対応する反応経路・3 項関係の連鎖

このように機械処理で薬物間相互作用の可能性を網羅的に調べることは、臨床情報がない、または少ない新規医薬品において特に有用だと考えられる。またすでに臨床現場で使用されている医薬品においても、原因不明の相互作用や副作用の機構の仮説の生成に役立つだろう。薬物間相互作用のみならず、今後解明が進むと想定される遺伝子多型や環境要因による特定生体物質(多くは遺伝子産物) の欠損や発現異常が個人に見出された際、その影響を受けやすい薬剤と影響を受けにくい薬剤の選別にも役立つことが期待できる。

薬物間相互作用については、機械的に導かれた薬物の組み合わせがすでに文献上報告されているかどうか、またその重要度を確認することも重要である。そのために、2 種類の薬物名をプローブとして MEDLINE やすでに集めた薬物間相互作用の文献セットを相手に検索するモジュールの組み合わせを作成中である(図表 6 右部分)。

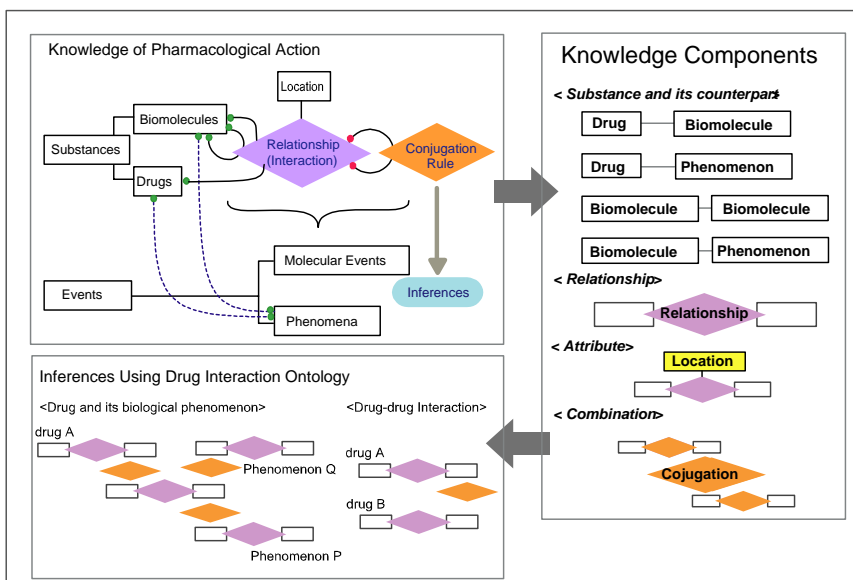
4 使い回し可能なドメイン知識要素と生物機能・現象概念

4.1 薬理学分野の使い回し可能な知識要素

前述の例は物理化学的な分子同士の直接的な作用(酵素反応・酵素阻害) を記述したものである。しかし、薬物による生体反応全体をながめると、直接的な分子の作用で解明されている部分は限られている。

分子レベル情報の欠落は随所にみられるがそれらを補いながら人は推論し、情報をつなぎ合わせている(図表 10)。例えば、薬物による酵素の発現誘導のように、発現機構は部分的にしか分らなくても、発現物質への作用として扱い、次の分子レベルの連鎖へ繋ぐことができる。したがって、この情報モデルにおいても遺伝子発現の直接的分子レベルの中間反応連鎖を省略し、抽象化した(間接的) 反応様式としての「遺伝子発現」の 3 項関係で記述することにより、その後の反応へと繋ぐことを意図している。

薬理学や医学・生物学の領域で



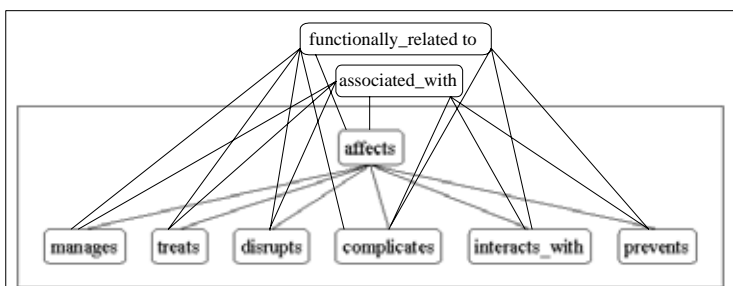
図表 10: ドメイン知識要素を使った推論

このような間接的な作用概念による関連化は、分子レベル同士を関連化するだけでなく、むしろ多くは現象や生物機能に割りあてた専門用語を用いて行われる。

例えば、いくつかの高脂血症治療薬の副作用として知られている“紋筋溶解症”について考察してみる。その発症機構についてはいくつかの説があり、また各プロセスに関わるすべての反応を分子レベルで記載できるわけではないが、その過程を省略して薬物が引き起こし得る現象(この場合は横紋筋溶解症)と薬物名との関連化が行われている。また、筋肉組織の損傷・溶解は、血中のミオグロビン分子や、クレアチンキナーゼの増加という分子レベルの作用を引き起こし、さらに血中ミオグロビン分子の増加は腎機能障害をもたらす。このように、細胞レベルやそれ以上の生物学的階層に関連する生物機能や現象間、またはそれらと分子レベルの概念との関連化はドメイン知識として使い回しの対象、すなわち三項関係で記述できると考えられる。

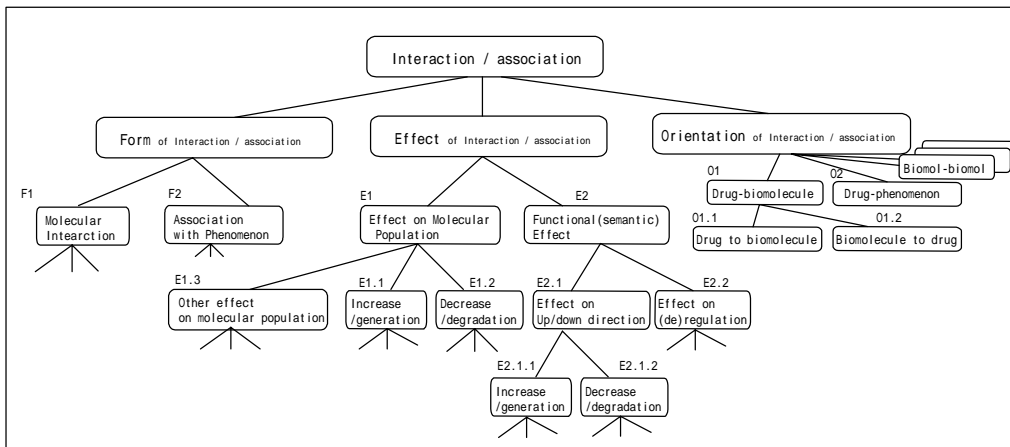
4.2 分子レベルと現象を貫く関係を記述する基本概念

前述のとおり、Gene Ontology では molecular function と biological function の二つの世界は切り



図表 11: UMLS semantic network に使用される関係語の例

離されているが、UMLS では大項目間で概念間のつながりを semantic network で表現している。そこでは 50 余りの関係概念が提供されているが、その中から、生物機能(現象)や物質の関係にも使用される affect 周辺の階層を図表 11 に示す。affect 以下のものについても affect のみならず、その上位にある associated_with や functionally_related to を親として持ち、具体的に関連化されている概念と対応させても若干煩雑な印象を受ける。



図表 12: 分子レベルと現象を貫く関係概念の構造化

また、物質間の関係はほとんどすべてが interact_with へ対応づけられてそれ以上の詳細化・形式化はされてない。

前述の薬物-生体物質相互作用における関係の形式化(図表 4)は言うなれば、interact_with に含まれる分子間相互作用を個別のインスタンスの具体的な記述のために詳細化したものと位置づけることができよう。

UMLS における、物質と現象(生体機能、生理学、病理学的概念)との関係は、affects, associated_with, treat, disrupts, complicates, prevents, cause, diagnoses etc. などと多様である。われわれは、薬物-生体物質相互作用で用いた相互作用概念は分子間の関係のみならず、分子と現象、または現象間の関係にも適用できるのではないかという立場から相互作用概念の基本構造を整理し、図表 12 にその根幹部を示した。図表 4 はこの基本構造の中から薬物-生体物質相互作用における部分を切り出したことになる。

この形式化の特徴は、UMLS においては関係を示す用語に内包させていた方向性の概念を独立させた点、分子相互作用で採用した作用(効果)の基本概念を導入した点にある。分子機能の相対的な up と down があるのと同様に、現象においても専門用語の操作によって up と down の概念が(すべての事例ではないが)導入できるという立場である。言い換えると、up や down などの作用を示す関係概念

の整理は現象に関する専門用語（概念）の整理と相互関連化（例えば、physiological function と pathological function）につながると考えられる。

自然言語における作用を示す関係概念を担うのは、多くの場合動詞またはそれらを名詞化、形容詞化したものである。われわれはそれらを「作用語」と呼び、薬物-生体物質相互作用の3項関係に対応した情報抽出に向けて、MEDLINEの抄録文における使用パターンを分析した[12]。さらに薬物や生体物質と生物機能や現象を示す専門用語との関係における作用語の役割を調べることによって分子レベルと現象概念の関係の知識抽出の可能性について検討している。

5 まとめ

薬物-生体物質相互作用を3項関係で表現する薬機能オントロジーを設計した。3項関係を使い回し可能な情報単位として扱い、相互作用の連結の基本ルールを定義した。これにより異なる反応様式のものと同じ基盤で連結し、反応連鎖を機械処理で生成し、さらに分子レベルの情報と現象情報とを繋ぐことを可能にした。

6 引用文献

- [1] Yoshikawa S, Konagaya A. DIO: Drug Interaction Ontology-Application to Inferences in Possible Drug-drug Interactions. 2003 In: Proceedings of The 2003 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS '03), Las Vegas, Nevada, USA. pp231-236, 2003.
- [2] 吉川澄美, 小長谷明彦. 薬物と生体物質の相互作用オントロジーに基づく薬機能知識ベースの設計. 臨床評価, Vol.29, No.2・3, pp.275-286, 2002.
- [3] Yoshikawa S. Formal Representation of Pharmaceutical Function Based on Drug Interaction Ontology. Doctorial Dissertation submitted to Japan Advanced Institute of Science and Technology September 2003.
- [4] Karp PD. An ontology for biological function based on molecular interactions. *Bioinformatics* 2000; 16: 269-85.
- [5] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology, The Gene Ontology Consortium. *Nature Genet* 25: 25-29.
- [6] <http://www.geneontology.org/>
- [7] <http://www.nlm.nih.gov/research/umls/umlsmain.html>
- [8] Desgranges C, Razaka G, Clercq ED, Herdewijn P, Balzarini J, Drouillet F, Bricaud H. Effect of (E)-5-(2-bromovinyl)uracil on the catabolism and antitumor activity of 5-fluorouracil in rats and leukemic mice. *Cancer Res* 1986; 46: 1094-101.
- [9] Nishimoto T, Yokoyama N, Sakamoto H, Ohata K, Kohno S, Murai K and Tatsumi H. Studies on the metabolic fate of Brovavir (YN-72) II: Metabolism in rats. *Iyakuhin Kenkyu*.1990; 21: 378-89.
- [10] Ogiwara T, Miue H, Nakamaru M, Ohtsuka T and Kumahara Y. Phase I clinical study of YN-72 (BV-araU, Brovavir). *Jpn Pharmacol Ther* 1990; 18: 507-23.
- [11] Diasio, R. B., Beavers, T. L., and Carpenter, J. T. Familial deficiency of dihydropyrimidine dehydrogenase. Biochemical basis for familial pyrimidinemia and severe 5-fluorouracil-induced toxicity. *J.Clin.Investig.*1988; 81: 47-51, 1988.
- [12] 吉川澄美, 松村和美, 長嶋剛史, 保坂順子, 小長谷明彦. 薬物相互作用の固有表現分析. In.言語処理学会第10回年次大会併設ワークショップ「固有表現と専門用語」発表論文集 pp. 17-20, 2004.