

医学における長期仮説の形成と評価

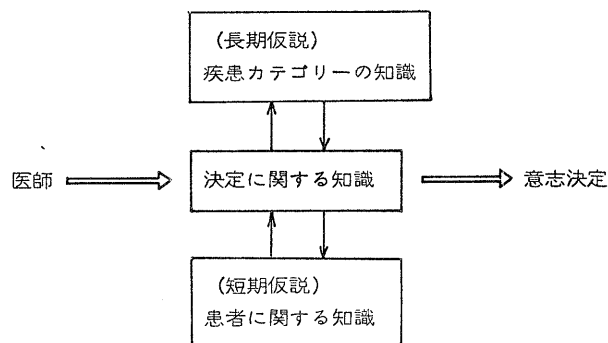
—血液化学データの情報量を事例として

高杉成一 (東大・医学部・医用電子研究施設)

D. A. B. Lindberg (ミズーリ大学・医学部)

I はじめに

臨床医学における意志決定の際に医師がとり扱う知識や情報を分類すると、①疾患カテゴリーや病因、予後など疾病の概念や生体に関する情報、②個々の患者に関する情報、および③これらの情報のとり扱いに関する知識の3つに大別できよう(図ノ)。このうちやや性格の異なる情報のとり扱いに関する知識は、臨床情報を処理して意志決定に至る decision strategy である。この知識は、一般には医師が臨床経験を通じて学習するもので、客観化は容易でない。しかし、最近の認知科学や AIM (Artificial Intelligence in Medicine) 技術の発達により、他の情報や知識とともに解析が進むことが期待されている。



図ノ 診断における知識の相互関係

Black box である decision strategy の大まかな機能としては、次のものをあげることができる。

- ①患者情報を評価し、診断や予後などの仮説を形成、また新しく獲得すべき患者情報を出力する。
- ②疾病や生体に関する知識を評価し、修正する。
- ③ decision strategy 自身を評価し、修正する。

上記①の患者情報に基づく診断思考は、比較的短時間のサイクルでの仮説の逐次的検証にあたるので「短期仮説」の形成過程とみなしうる。新しく獲得すべき情報を求めるには、データのもつ情報量を評価する必要があるが、このような評価システムの構築は臨床経験による学習目標の一つと考えられる。しかし、臨床上の評価関数は単一でなく、評価関数によつて情報量が異なるなど問題は単純でない。データのもつ情報量は、短期仮説によつて二次的に定まるという性質をもっている。一方③の疾病や生体に関する知識は診断のための標本ともみなされるが、比較的長い時間単位では常に改良が加えられ変化を続けているので、「長期仮説」とみなしうる。

いずれにせよ、情報や知識を計量的に評価することは、臨床における意志決定

の機構を解析するための重要なステップと考えられる。本論は Shannon による情報理論が知識評価のモデルとして応用可能か否かを検討しようとしたもので、事例として長期仮説に属する体液調節の熱力学的知識をとり上げた。また、これに先立ち、仮説から知識に移行した医学知識の事例を歴史的な立場から検討した。

II 医学における意志決定機構の客観化の背景

医学における中心課題の遷移を歴史的に眺めると、現代は医療の技術革新の時代ということになる。なかでもコンピュータをはじめとする情報技術の医学への導入はめざましい。意志決定プロセスのシミュレーションや補助はコンピュータ応用の一環として進歩してきたものであるが、この分野の今後の発展方向を探るヒントを得るため、過去の医学の発展経過を展望してみたい。

16世紀の科学ルネッサンス以前の医学は、黒胆汁、黄胆汁などの失調により疾病が発生すると考える体液説にみられるごとく、観念論的な色彩が濃かった。16世紀に入つて自然科学の目で生体を見ようとする動きが活発化した。まず肉眼による解剖学から始まり、顕微鏡を用いたミクロの形態学、さらには機能に着目した生理学へと発展が続いている(図2)。

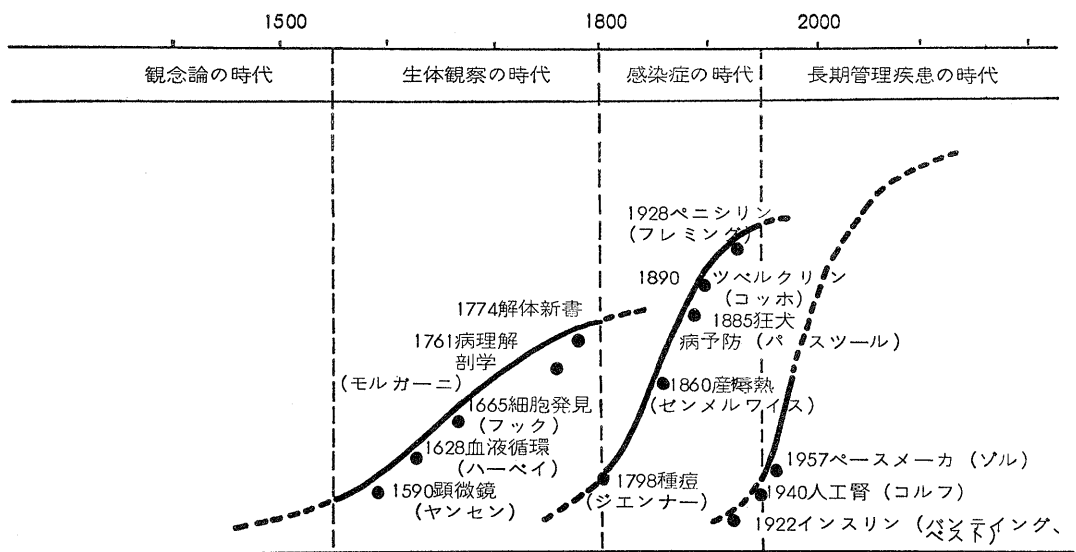


図2 医学における中心テーマの推移

この頃の2-3世紀は、医学が疾病治療に本格的に取り組む以前の生体観察の時代であつたと考えられる。19世紀に入ると伝染病その他の感染症対策が医学の中心課題となつた。1928年にペニシリンが発見されて以来、抗生物質やワクチンが著しく進歩、感染症にかわつて成人病や老年病、免疫疾患などの長期管理疾病対策が主要課題になつてきた。現在は丁度その移行期にあたると考えられる。

長期管理疾患が中心になると医学のイメージはこれまでのものと大幅に変つてしまう。たとえば、元来病気を治すはずの治療は、長期管理疾患においては逆説的に患者数を増加させてしまう。腎不全に対する人工透析がそのよい例で、有効な治療法ほど患者数増加の傾向が強い。治療の意味が疾病の管理に変わっているか

らである。予防も同様に病氣発生の抑制という意味であつたものが、早期からの疾病の管理という意味に変化、治療と同様医療の対象を飛躍的に増加させる傾向もつている。近年における医療の技術革新の努力は、このような逆説的状況を克服しようとする試行錯誤の一環といえる。

意志決定プロセス客観化の意義は、再現性、計量性、普遍性などの追求によつて診断や治療の精度を向上し、アクセシビリティを高めるなどが指摘されているが、とくに個々の患者に応じて疾病を最適に長期管理する知識開発を援助することが重要と考えられる。

III 医学における知識形成の事例

疾病管理の知識開発のための情報源には、①患者の臨床観察データ、②専門医の知識、③基礎医学知識、④制御論など周辺科学知識などがあるが、これまでは、認知科学の問題であるという認識なしに多方面からの研究が行われていた。AIMや認知科学からのアプローチによつて新たなとり扱いが開かれると思われる。AIMはこれまで短期仮説の形成に重点をおいて応用開発が行われてきたが、しだいに長期仮説の形成や生体モデルに基く推論形成にも重点が置かれる傾向にある。医師の短期仮説形成は客観化が容易でないが、長期仮説形成は客観的評価を前提としているという違いがあり、また仮説形成の基本的性格は共通していると思われるので、ここでは確立された医学知識のうち、最初は仮説として形成され検証を受けたいくつかの事例をあげて、仮説形成ならびに評価機構を解析する参考にしたい。

なお、仮説とは高度の信頼性が検証される以前の知識と理解されるが、確立された知識といつても100%の信頼性をもつとは限らないことを考えると、仮説と知識は明確には区別できないことになる。また医学知識には、コントロール・スタディによる検証のステップを経ず直接知識として採用される「観察事実」が多い。往々にしてこの知識に合致しない例が後に見つかるが、その事例は「異常」と判定されることになる。

(1) 血液循環説

心臓を出た血液が体内を循環して再び心臓に戻ってくるという知識は、医学的知識というだけでなくむしろ一般常識である。この説は1628年 William Harvey により提唱されたが当時の医学界から強い反対にあい、採用されるまでに数十年を要している。ハーベイはこの仮説の証明のため計数的なとり扱いをし、心臓の血液拍出量などを求めている。

(2) コッホの三原則

医学知識のなかで、因果に属する事象は検証が難しい。感染症もその一例であるが、Robert Kochは一つの細菌が一つの病気の病原体であることの三条件を提唱した。すなわち1.その病気ではいつもその細菌が見つかること、2.体の外でその細菌が培養できること、3.健康体にその細菌が入るとその病気がおこることである。コッホの三原則は、因果に関する仮説を検証するための一つのモデルと考えられる。

(3) DNAの構造の発見

イギリスのワトソンとクリックによる遺伝物質DNAの分子構造の発見が医学生物学に及ぼした影響はきわめて大きい。ワトソンとクリックは構造発見のためにシミュレーション技法を用いている。1950年代の当時遺伝物質について判っていた知識は、DNAの化学組成にみられる一定の法則性やX線構造解析によ

ラセン構造などの組成情報、およびメンデル以来の遺伝現象の観察による機能情報の2つで、これらをつなぐべき分子構造の知識が欠けていた。そこで彼らはボールとスティック型の分子模型をさまざまに組み立て、二重ラセン模型という「仮説」を提唱した。あまりにもすぐれた仮説であつたため、提唱の始めから「構造の発見」という高い評価を受けている。これはモデルによる推論が成功をおさめた例と考えられる。

この他医学領域における仮説形成には種々のパターンがある。

IV 知識の情報量 — 血液化学データに関する体液調節系の例

いうまでもなく臨床医学において観察データから仮説を導く場合、二重盲検法などコントロール・スタディを用いる統計的アプローチが基本的評価手法となっている。しかし、経験から得る医師の知識は、コントロール情報が明確でないという特徴がある。医学的常識からはずれた「異常」例に遭遇したとき、医師が何らかの知識を獲得すると考えれば、医学的常識がコントロール値として使われていることになる。これは短期仮説が長期仮説をコントロールとして形成されることに対応している。

意志決定機構の客観化に際して、仮説を定量的に評価することが必要と考えられるが、情報理論はこのための有力な手法を提供する可能性がある。そこで体液調節に関する知識を情報論的に評価する試みを行つた。

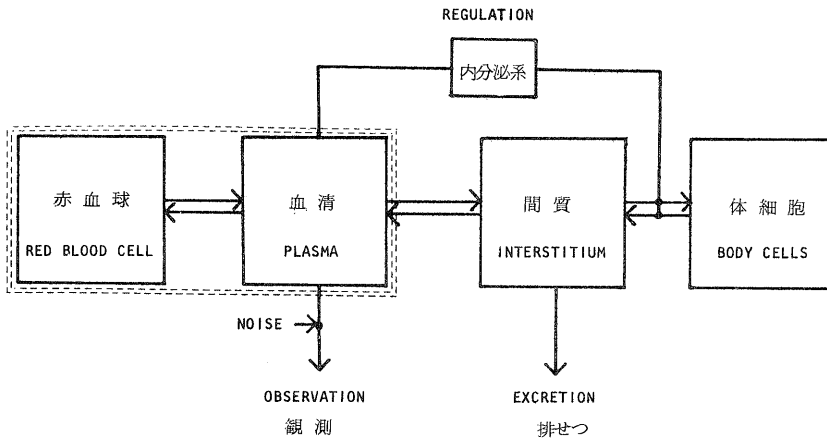


図3 体液調節系の模式図

生体は図3のような多コンパートメント系であり、患者の血清データを得ても他のコンパートメントの情報を完全に把握したことにはならない。ところが、体液調節に関する医学知識を導入すれば、観測情報がより有効に解釈可能となるはずである。ここで情報量の評価を試みた知識は熱力学の法則で、次の要素から成る。

- ① 各溶質の電気化学的ポテンシャルは、赤血球内および血清内で等しい。
- ② 総電荷も等しい。
- ③ 溶質間には特定の化学反応がある。
- ④ 全システムにおいて質量保存則が成り立つ。
- ⑤ 細胞膜にはNa およびKの能動転送がある。

以上のほか各溶質の特性値を組み込んでモデルを構成した。表ノはモデルに採用

した化学成分で、図4はモデルの基本構造である。患者データにあたる入力変数に図ではNa、K、HCO₃をあげているが、モデルに含まれている限り何でもよい。

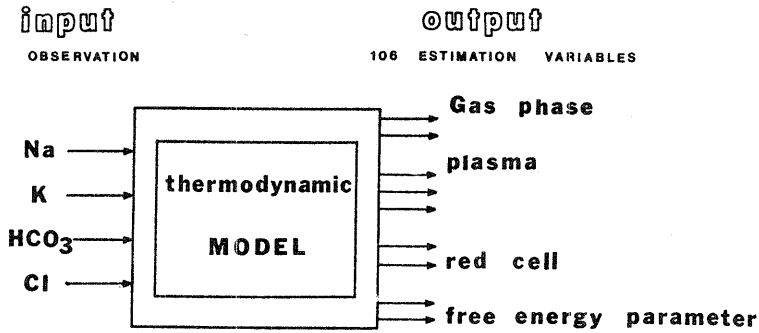


図4 モデルの基本構造

表1 とり入れた化学成分

GAS PHASE		RED CELLS		RED HEME GRP	
1	O2	35	O2	73	CARBOX
2	CO2	36	CO2	74	CARBO-
3	N2	37	N2	75	AMINO+
4	H2O	38	H2O	76	AMINO
		39	H+	77	CARBAM
		40	OH-		
		41	NA+		OXY HEME GRP
		42	K+		
5	O2	43	CA++	78	CARBOX
6	CO2	44	MG++	79	CARBO-
7	N2	45	CL-	80	AMINO+
8	H2O	46	LACTIC	81	AMINO
9	H+	47	ORGAN-	82	CARBAM
10	OH-	48	ORGAN-		
11	NA+	49	HCO3-		
12	K+	50	H2CO3		OXYSTABL GRP
13	CA++	51	CO3=	83	HMCOOH
14	MG++	52	HPO4=	84	HMCOO-
15	CL-	53	H2PO4-	85	BCOOH
16	LACTIC	54	PSTER=	86	BCOO-
17	ORGAN-	55	PSTER-	87	IMID+
18	HCO3-	56	SO4=	88	IMID
19	H2CO3	57	SESTR	89	PHENOL
20	CO3=	58	AMIN+	90	PHENO-
21	HPO4=	59	AMIN+-	91	EAMIN+
22	H2PO4-	60	AMIN-	92	EAMIN
23	SO4=	61	NH4+	93	GUAN+
24	SESTR	62	NH3	94	GUANID
25	AMIN+	63	UREA		
26	AMIN+-	64	GLUCOS		PL ION GRPS
27	AMIN-	65	CHOLES		
28	NH4+	66	ADENOS		
29	NH3	67	X-MISC	95	BCOOH
30	UREA	68	MG++ B	96	BCOO-
31	GLUCOS	69	HB4	97	CA++ B
32	CHOLES	70	HB4O2	98	IMID+
33	PLA PR	71	HB4O4	99	IMID
34	X-MISC	72	HB4O6	100	EAMIN+
			HB4O8	101	EAMIN
				102	PHENOL
				103	PHENO-
				104	GUAN+

情報量評価の基本的な考え方を図5に示した。左側のNa、K、HCO₃は一例として選んだ患者の観察値、中央はモデル、右側のClはモデルによつて推定した値である。実線の大きいboxは生体の全調節因子がとり入れられた場合のモデルで、破線で囲んだ部分は熱力学的因子のみを考慮したモデルである。

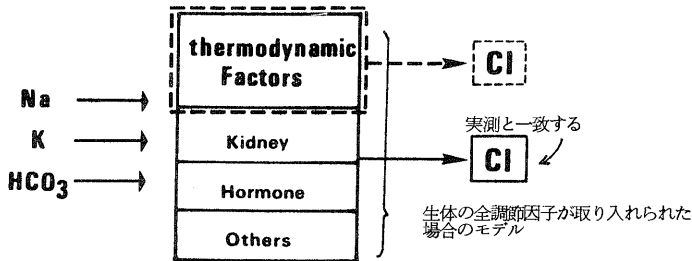


図5 情報量の評価法

さて、熱力学的な因子が全体に占める割合であるが、次のように考える。まずモデルが完全に生体の全情報を反映しているなら、このモデルをあてはめたClの推定値は実測値と一致するはずである。ところが熱力学的な知識のみによるモデルの推定値 \boxed{Cl} は実測値のもつ情報量よりも少ない。この2つの値を比較すれば熱力学モデルのもつ情報量を求めることができる。このような入力情報と出力情報、およびモデルの情報の関係を図6に示した。

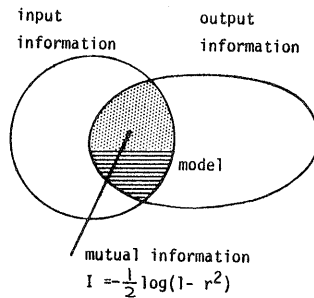


図6 情報の相互関係

熱力学知識の情報量は図6の横線部に当り、仮想的な完全なモデルの情報量は横線部分とメッシュ部分の和にあたる。左の円は入力情報で、いまNa、K、HCO₃を考えている。右のサークルは出力情報で、いまはClである。

具体的な数値計算は、300例の患者データを用いて行つた。この結果、熱力学モデルによる入出力情報の相関係数は $\gamma = 0.606$ となり、コントロールの完全モデルの場合は、実測値の重回帰から $\gamma = 0.658$ と見積られた。これらの相関係数は Shannon の理論に基づいて導かれた $I = -\frac{1}{2} \log(1 - \gamma^2)$ により相互情報量に変換した(図7)。

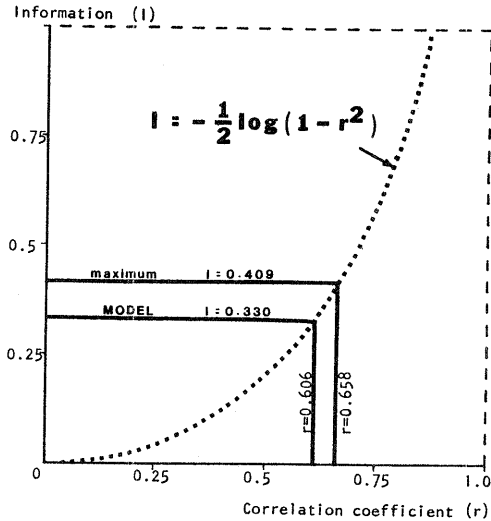


図7 相関係数と相互情報量の関係 (Clの場合)

2つの相関係数を情報量に変換し、割合になおした結果Clに関する熱力学知識の情報量は80.5%となつた。この場合の目的変数であるClは、これまで述べてきた「仮説」にあたつている。すなわち、与えられた知識と患者データを用いて血清Clの濃度に関する定量的な仮説を形成、この値と真の値を比較したことになる。目的変数には任意の変数を選ぶことができるが、選んだ変数によつて与えられた知識のもつ情報量は異なる。もちろんモデルに含まれない変数の情報はゼロである。なお、同熱力学モデルを他の電解質にもあてはめた結果、Na⁺に関して29.7%、Kは99.9%、HCO₃⁻は98.3%の情報量をもつという評価値が得られた。

V あとがき

診断においては、ふつう仮説の真の値は得られない。そのため対立仮説を対象として、相対的な評価を行わざるを得ないという事情がある。さらに、医師は情報量を含む複数の評価関数を有機的に用いている可能性があり、分析を要する。このほか知識をいかに記述するかという表現形式の問題も重要であるが、この研究に際しては医学知識の階層性や冗長性に対する充分の注意が必要と考えられる。

VI 文献

1. S.Takasugi, D.A.B.Lindberg et al: Linear model of the blood chemical components for estimating the steady-state distribution of in serum and red blood cell compartments based on thermodynamic principles, University of Missouri-Columbia Information Science Series Technical Report MOU-IS-TR-17, Nov. 1977.
2. S.Takasugi, D.Goldman, D.A.B.Lindberg & E.DeLand: Analysis of information content of blood chemical components by thermodynamic model, XII International Conference on Medical & Biological Engineering, Jerusalem, ISRAEL, Aug. 19-24, 1979.
3. F.Mizoguchi, T.Koyama, S.Takasugi & S.Kaihara: Research objectives, facilities and current works on the clinical diagnostic reasoning with application of Artificial Intelligence. Hospital Computer Center School of Medicine, JSPS-NSF Cooperative Science Program Technical Report Series, Aug. 1979.