

# 感染症流行の予測：感染症数理モデルにおける 定量的課題

西浦 博<sup>1</sup>・稲葉 寿<sup>2</sup>

(受付 2006年1月4日;改訂 2006年2月6日)

## 要 旨

直接伝播する感染症の数理疫学的研究に関して、特に流行の統計データを用いた予測と関連する推定について定量的な評価手法と考え方を総説する。予測を実施するためには伝播能力と感染待ち時間や感染性期間などに関する情報を詳細に得ておく必要がある。感染症数理モデルの多くは個体群動態に基づいて構築され、特に流行予測やモデルに基づくパラメータ推定を実施する場合は推定値の精度やデータの変動を考慮した計数過程を想定した推定量が用いられる。観察されたデータに基づく基本再生産数・効果的再生産数の推定手法および内在性の各想定の意味について詳述する。また、個体内の病理が緩徐に進行する HIV/AIDS や BSE のような感染症に関して潜伏期間の統計分布が得られれば、発症した症例数を基に逆算法を使用して感染者数を推定することが可能であり、それによって得られる短期予測の基本的構造を紹介する。実際の流行データに見られる伝播は感受性や接触頻度の異質性などを主な理由として均一性の想定が困難なため、マルチタイプの流行など、対象とする感染症の生物学的および疫学的特徴とデータ特性に応じたモデル構築および推定手法選択が必要である。どの手法を用いた場合でも「感染症がどのような振る舞いで流行するのか」を考慮した内在性の想定とその妥当性が定量的研究による予測値・推定値の鍵となる。

キーワード：感染症，数理モデル，基本再生産数，最尤推定法，マルチンゲール法。

## 1. はじめに

昨今のように感染症に関する制圧・共存が社会的問題として大きく取り上げられる以前から、感染症流行を予測することは公衆衛生上で1つの大きな課題であり、疫学・数理統計学領域の夢であり続けている。伝染病(ヒトからヒトへ伝播する感染症)の流行に関する数理的諸研究は最も古くから実施されており、その他の感染症も含めて集団レベルの現象における観察と説明に関して最も妥当な理論化に成功した疫学研究領域である(Bailey, 1975)。また、数理生物学の研究に関しても最も豊富な研究成果が得られた分野であり、流行に関する定性的分析は生態学的に大きな意義を有する説明を可能にしてきた(巖佐, 1998)。この要因の1つとして個体群生態学領域における数理モデルで利用されたような安定性分析が感染症疫学モデルの研究手法にも用いることが出来たことが挙げられる(Anderson and May, 1991; 稲葉, 2002)。最近の例で

<sup>1</sup> チュービンゲン大学 医学部医系計量生物学研究所：Westbahnhofstr. 55, Tübingen D-72070, Germany; nishiura.hiroshi@uni-tuebingen.de (現 長崎大学 熱帯医学研究所：〒852-8523 長崎市坂本 1-12-4)

<sup>2</sup> 東京大学大学院 数理科学研究科：〒153-8914 東京都目黒区駒場 3-8-1

は、重症急性呼吸器症候群(SARS; Riley et al., 2003; Lipsitch et al., 2003)や新型インフルエンザ(Longini et al., 2005; Ferguson et al., 2005),天然痘ウイルスを利用したバイオテロ(Kaplan et al., 2002; Eichner, 2003)などを始め,社会的要求度の高い感染症の予測・推定・流行抑止対策の評価が実施されており,徐々に現実性の高い議論が可能になってきた。

感染症の理論が進展した理由は古くに伝染病が主要死因であったこと以外にも幾つかの学問的特性があるためであり,あらゆる疾病の中でも感染症だけが宿主から宿主へと伝播する特異的な疾病であることが第一に挙げられる(Giesecke, 2002)。その為に,癌や循環器疾患など他の疾病とは異なり,感染者そのものが病気の危険因子となる。また,1度だけ感染するか予防接種による抗原刺激を受けるだけで一定期間の免疫を得ることがある点や,感染者は感染症例と認識されないままに感染源となる場合も多い点など,理論的に特異性が高いことも特徴である。これらの理由から,感染症の流行に関わるデータを分析する場合には内在性(Intrinsic; 感染症の伝播するメカニズムそのもの)及び外因性(Extrinsic; ワクチンをはじめとする流行抑止対策など)の妥当な想定に基づいた数理モデルを構築することが求められる。即ち,感染過程(Infection Process)と発症に至るまでの各疾病の特性を考慮した上でデータを取り扱う必要があり,それに従って諸々の統計学的推定手法をモデルに応じて使い分けなければならないために方法論は複雑になる傾向がある。本稿では,それらの中でも直接伝播する感染症の流行(Epidemic)あるいは集団発生(Outbreak)の観察データに基づいて実施する定量的諸問題を総説する。特に,数理的諸研究と公衆衛生活動のギャップを埋める部分に該当する確率的な(Stochastic)モデルと関連する推定を中心に紹介する。定性的研究についてはBailey(1975),Diekmann and Heesterbeek(2000)及び稲葉(2002)が詳しく,より数理生物学的な内容はMurray(2002, 2003)や巖佐(1997, 1998)が役立つ。また,風土病的流行(Endemic)状態にある感染症の分析に関して決定論的(Deterministic)モデルを中心にした基礎理論としてはAnderson and May(1991)が詳しい。

## 2. 数理モデルの構造と定義

### 2.1 直接伝播する感染症の数理モデル

感染症の数理モデルはその個体群動態(Population dynamics)を想定して,個体群レベルの流行メカニズムを考慮したボトムアップ式の過程で構築され,その中でも最も簡単な直接伝播する感染症の決定論的なモデルを最初に紹介する。この数理モデルは対象とする閉鎖人口(Closed population)について興味の対象である感染症に依存した3つの状態(S; 感受性宿主で感染する可能性のある人口, I; 感染して感染性を有する状態, R; 感染後に回復して免疫を獲得した者または死亡者)に区画(Compartment)で分けて検討することから開始する(図 1a)。それぞれの状態の頭文字をとって SIR モデル(Susceptible-Infectious-Recovered)と呼ばれることが多い(Anderson and May, 1991)。各区画間の時間当たりの変化は常微分方程式系によって以下のように与えられる:

$$(2.1) \quad \begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t)I(t), \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t). \end{aligned}$$

ここで $\beta$ は感染率, $\gamma$ は回復率や隔離率であり, $\beta I(t)$ は時間 $t$ における感染力(force of infection)と定義される。また, $\gamma^{-1}$ は感染期間が指数分布に従うと仮定した場合の平均値である。

これは20世紀前半の提唱者の名前をとってケルマック-マッケンドリック型モデルとも呼

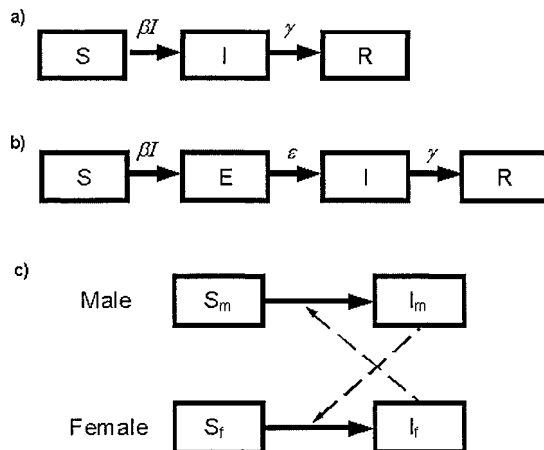


図 1. 基本的な感染症の個体群動態に関する区画 (Compartment). a) SIR 型モデル, b) SEIR 型モデル, c) 性感染症の SI 並列型モデル. S(感受性宿主), E(感染待ち時間), I(感染性期間), R(回復, 死亡, 免疫)など, 想定によって柔軟に変化を加えることが可能である. 矢印上のパラメータは各状態を推移する時間当たりの変化率を意味する.

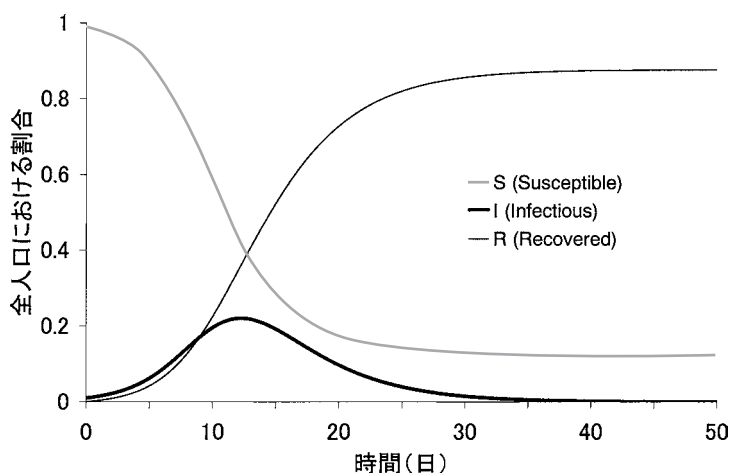


図 2. 感染症流行のシミュレーション. 式(2.1)の SIR 型モデルに基づく. パラメータはインフルエンザを参考にしてある( $\beta = 0.48$ ,  $\gamma^{-1} = 3.5$  [days],  $R_0 = 1.7$ ). I(感染性期間)は疫学調査における有病割合(Prevalence)で与えられる流行曲線とほぼ同義である.

ばれる(Kermack and McKendrick, 1927). 図 2 は典型的なシミュレーション結果であり, 流行規模やその速度は  $\beta$  や  $\gamma$  によって特徴付けられる. 短期的な流行を考慮しているモデルであるために背景にある出生・死亡などの人口動態はほぼ無視できると想定している. この系では, ある時刻  $t$  において感染者がいない状態, 即ち  $(S(t), 0, R(t))$  は全て定常状態である. ここで総人口サイズを 1 に固定した場合, 感染者が対象人口に(過去を含めて)存在しない平衡状態  $(S(t), I(t), R(t)) = (1, 0, 0)$  で(2.1)を線形化すると以下が得られる:

$$(2.2) \quad \frac{dI(t)}{dt} = (\beta - \gamma)I(t).$$

即ち、少数の感染者が侵入した初期に感染者は  $I(t) \approx I(0) \exp\{(\beta - \gamma)t\}$  という指数関数的増加(マルサス法則)に従う。このことから病気が集団に侵入可能となる条件は、マルサス係数が  $\beta - \gamma > 0$  となることであり、書き方を変えると

$$(2.3) \quad R_0 = \frac{\beta}{\gamma} > 1$$

と表すことが出来る(稲葉, 2002)。感染症が集団に侵入可能となる条件は、この基本再生産数(basic reproduction number;  $R_0$ )によって決定し、これは「(全ての個体が初期に感受性を有する状態で)1人の感染者当たりが生産する2次感染者数」を意味する( $R_0$  と発音する; Dietz, 1993)。  $R_0 < 1$  であれば流行が発生しても次第に感染者人口は減衰すると考えられ、  $R_0 > 1$  であれば感染者は初期に指数関数的に増大する。これはパラメータ  $\beta$  と  $\gamma$  に依存しており、この解の定性的挙動に関する現象は閾値現象(threshold phenomena)と呼ばれる(Anderson, 1991)。

## 2.2 計数過程を想定した場合

以上の決定論的モデルは解析的研究や想定された流行モデルの定性的な解の挙動に関する理解に役立つが、パラメータ推定を伴う定量的な検討をする際は推定値の精度や元々の対象データの変動を定量化する目的で確率論的なモデルに基づく推定を実施することが望ましい。そこで、時刻  $t$  までに感染を経験した者を  $N(t) = S(0) - S(t)$  とした場合(2.1)と同様の想定における流行をマルコフ型の連続時間モデルで表すと以下が得られる:

$$(2.4) \quad \begin{aligned} \Pr(dN(t) = 1, dR(t) = 0 | \mathcal{S}_{t-}) &\simeq \beta \bar{S}(t) I(t) dt, \\ \Pr(dN(t) = 0, dR(t) = 1 | \mathcal{S}_{t-}) &\simeq \gamma I(t) dt, \\ \Pr(dN(t) = 0, dR(t) = 0 | \mathcal{S}_{t-}) &\simeq 1 - \beta \bar{S}(t) I(t) dt - \gamma I(t) dt. \end{aligned}$$

つまり、 $N(t)$  と  $R(t)$  は単調に増加する計数過程(Counting Process)である(Andersson and Britton, 2000)。ここで  $\mathcal{S}_{t-}$  はシグマ代数で時刻  $t$  までの流行過程のヒストリーを表す ( $\{S(u), I(u); 0 \leq u \leq t\}$ )。総人口を 1 に固定すれば、 $R_0$  は決定論的モデルと同様に  $\beta/\gamma$  で得られる(Ball, 1983; Becker and Britton, 1999)。感染症の流行規模を予測するためには  $R_0$  を推定することを通じて興味の対象である感染症の伝播能力(transmission potential)を予め評価しておくことが望ましい。 $R_0$  を与えるパラメータの推定手法については第 3 節で紹介する。

## 2.3 数理モデルの構造に関する拡張

第 2.1 節で紹介した区画を利用する原理は伝播動態の様々な興味に従って拡張することが出来る。例えば、感染待ち時間(あるいは暴露後期間; latency period)を考慮する場合には、 $I$  の状態を Exposed(感染後、まだ感染性を持たない状態)と Infectious(感染性を有する状態)の 2 つに分けて扱うことによって 4 つの部分人口でモデルを構成することができる(SEIR モデル; 図 1b):

$$(2.5) \quad \begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t) I(t), \\ \frac{dE(t)}{dt} &= \beta S(t) I(t) - \varepsilon E(t), \\ \frac{dI(t)}{dt} &= \varepsilon E(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t). \end{aligned}$$

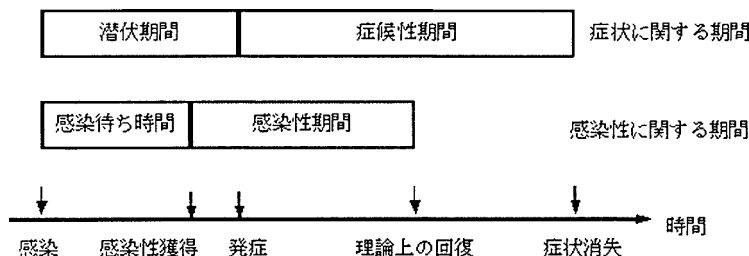


図 3. 症状と感染性を根拠にしたパラメータの定義．潜伏期間(incubation period)は感染から発症までの期間なのに対して感染待ち時間(latency period)は感染から感染性獲得までを意味する．症候性期間(symptomatic period)は症状を有する期間なのに対して感染性期間(infectious period)は感染性を有する期間を指す．通常、潜伏期間は感染待ち時間よりも若干長い傾向にあり、症候性期間は感染性期間よりも長い傾向にある．

ここで  $\varepsilon$  は暴露後に感染性を得る率であり、 $\varepsilon^{-1}$  は感染待ち時間が指数分布に従うと仮定した場合の平均値である．ここで注意すべき点として、感染待ち時間(latency period)と潜伏期間(incubation period)の違い、及び感染性期間(infectious period)と症候性期間(symptomatic period)の違いが挙げられる．感染待ち時間と感染性期間は感染性を根拠に分類されており、潜伏期間と症候性期間は症状を根拠に分類されている．これらは各々ほぼ同一と見なされる場合もあるが、図 3 にあるように感染性期間は潜伏期間の途中から開始する感染症も多く、また、症候性期間の間に感染性が消失する疾病が多くを占める(Anderson and May, 1991)．これらは感染者からの病原体検出などの情報に基づく微生物学的な根拠によって判断されることもあるが、一方で理論疫学的にも感染性期間を評価することが可能であり(Simpson, 1952)、定性的にも重要な役割を担うので第 4 節で詳述する．

同様に、感染後に免疫の獲得が終生で得られない場合は  $R$  から  $S$  へ推移することを考慮せねばならず SIRS モデルとなるし、性感染症の場合は男女を異なる人口と考えて SI が並列に並ぶようなモデルができる(図 1c)．また、マラリアのように他の生物を媒介する感染症の数理モデルも 2 つの異なる種(ヒトと媒介蚊)の個体群動態を考察することで構築でき、その定性的な解の挙動は現在でも基礎となるマラリアやデング熱の流行抑止対策評価ツールとして利用されてきた(Dietz et al., 1974; Dietz, 1974; Molineaux and Gramiccia, 1980)．

### 3. 感染症の伝播能力

#### 3.1 基本再生産数 ( $R_0$ ) の考え方

$R_0$  が持つ意義を最も簡単に理解するためには決定論的モデルで考えた予防接種率の達成目標の例が有名である．ある集団において効果(efficacy)が  $\varepsilon$  ( $0 < \varepsilon < 1$ ) のワクチン接種を接種率  $p$  ( $0 < p < 1$ ) で実施した場合、未接種者割合は  $1 - p$  であり、免疫を保持しない者の割合は  $1 - \varepsilon p$  である．流行は免疫を保持しない者の間で見られるので、均一(Homogeneous)な接触パターンにおける再生産数は  $(1 - \varepsilon p) \times R_0$  となる．これが 1 未満になれば感染症は終息するのだから、流行終息の条件は  $(1 - \varepsilon p) \times R_0 < 1$  で与えられる．これを  $p$  について解けば  $p > 1/\varepsilon \times (1 - 1/R_0)$  であり、ワクチン効果が明らかであれば  $R_0$  に従ってワクチン接種率目標を大まかに設定することに役立ってきた．関連する定性的検討は Anderson and May (1982) や Nokes and Anderson (1988) が詳しい．本来、ワクチンは部分的効果などがあるために伝播動態はさらに複雑になるが、簡単な感染防御そのものの例としてサージカルマスク着用などの個人保護具に代表される標準予防策が挙げられる．例えば SARS の  $R_0$  は平均の推定値が約 3 で与えられることが流行

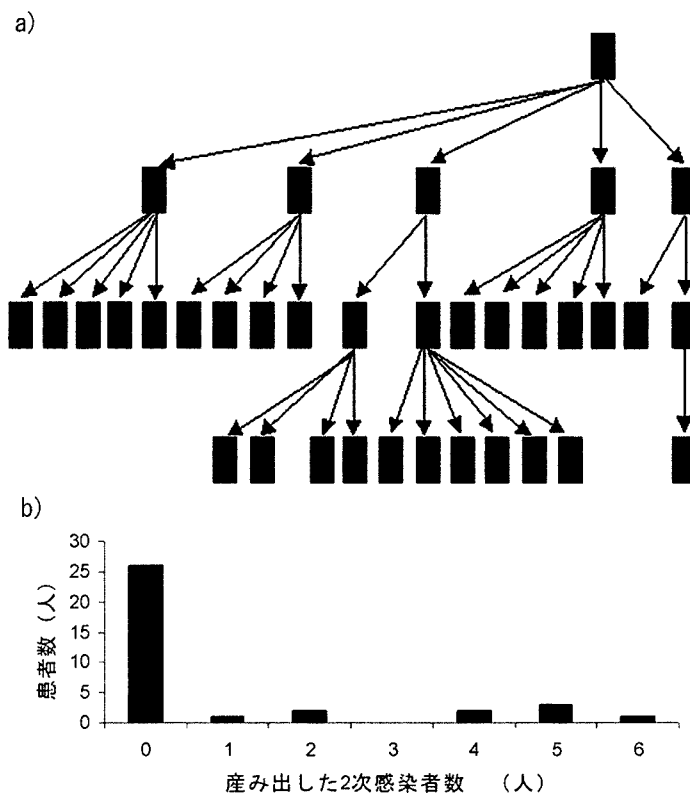


図 4. 天然痘の小規模流行における感染ネットワーク(Moscow, 1960年1月). a)感染ネットワーク(誰が誰に感染させたかを示す), b)感染者1人当たりの2次感染者数の頻度分布. a)では単純化して図示するため世代時間は無視して, 世代別に階層化している. モスクワにおける流行のうち院外での感染ネットワークのみを抜粋して再構築している(Baroyan and Serenko, 1961). b)では2次感染者数の頻度が歪度の高い分布を示している.

中の研究で明らかにされたが(Riley et al., 2003; Lipsitch et al., 2003), サージカルマスクの着用による感染予防効果を0.7であるすれば, 感受性宿主の全てがマスクを適切に着用することによって再生産数は $(1 - 0.7) \times 3.0 = 0.9$ となる. 病院内のように接触頻度が一定レベル以上は減弱できない場所であっても, 以上のマスク着用によってSARS流行中にヒトの防御行動が素早く変化することによって流行終焉(生態学的な絶滅の用語から Extinction と表現される)を迎えたことが検証されている(Nishiura et al., 2005).

$R_0$ の統計学的分布は「誰が誰に感染させたのが(who infected whom)」という感染ネットワークがわかれば, これを参考にすることができる. 図4aは1960年のモスクワ市における小規模な天然痘の流行で観察された感染ネットワークである(Baroyan and Serenko, 1961). 感染ネットワークは発症者毎に感染源を特定したものであって接触ネットワーク(=非感染者・非発症者も含む全ての個体間接触に関する情報)ほど豊富な情報を含まないが, 各感染者によって産み出された2次感染者数を直接に数えることができる. また, 発症から発症までの間隔を意味する世代時間(Generation intervalあるいはSerial interval)の情報が得られる(Fine, 2003). 図4bは図4aの情報から得られた2次感染者数の頻度に関する分布である. これは流行時刻や

背景にある検疫・隔離などの公衆衛生対策を考慮していないために厳密には  $R_0$  の分布ではないが定義上は類似の情報を意味する．多くの感染者は 2 次感染者を生産せず，一般的に(特にこういった小規模流行では)  $R_0$  の分布は右方向に裾を引く(Right skewed)分布を示す．

### 3.2 基本再生産数 ( $R_0$ ) の推定

多くの統計学的研究から  $R_0$  の推定値は得られており，例えば Endemic な感染症の  $R_0$  は年齢階級別の感染者数あるいは抗体陽性者数などの調査に基づいて年齢階級別の感染力を推定することで得られる(Farrington et al., 2001; Farrington and Whitaker, 2005)．例えば，麻疹の  $R_0$  は 9-16，水痘は 7-11，百日咳は 10-18，ジフテリアは 6.1 などと推定されている(Anderson, 1982)．最近注目されている感染症では，スペイン風邪などの流行データを根拠に新型インフルエンザでは 1.2-1.8 と推定されており，天然痘では過去の詳細な流行記録から 3-6 と推定された(Mills et al., 2004; Becker, 1977; Eichner and Dietz, 2003)．

流行(Epidemic)における  $R_0$  は最も単純には決定論的な(2.2)についてパラメータ  $\beta$  と  $\gamma$  を感染者有病割合のデータで回帰をすれば得られるが，上述の通りに精度や変動を考慮してデータを扱う必要性などの理由から確率論的な考え方に基づく推定量がこれまでに与えられてきた(Becker, 1989) (2.4)の計数過程を書き直すと対数尤度は以下のように得られる：

$$(3.1) \quad l(\beta, \gamma) = \int_0^\tau [\log \{\beta \bar{S}(u) I(u)\} dN(u) - \beta \bar{S}(u) I(u) du + \log \{\gamma I(u)\} dR(u) - \gamma I(u) du].$$

これは流行全体が時間及び空間軸上で最初から最後まで完全に観察された条件下における確率論的モデルを通じたパラメータ推定(具現化(Realization)と呼ぶ)として知られる．ここで「完全に観察された」とは，各個体の感染性獲得から消失までが確実に得られていることを意味する．最尤推定量はそれぞれ以下で与えられる：

$$(3.2) \quad \hat{\beta} = \frac{N(\tau)}{\int_0^\tau \bar{S}(u) I(u) du},$$

$$(3.3) \quad \widehat{\gamma^{-1}} = \frac{\int_0^\tau I(u) du}{R(\tau) - r_0}.$$

ここで  $r_0$  は  $R(0)$ ，つまり初期人口における非感受性宿主を意味する．また，推定量の標準誤差は以下で与えられる：

$$se(\hat{\beta}) = \frac{\hat{\beta}}{\sqrt{N(\tau)}},$$

および

$$se(\widehat{\gamma^{-1}}) = \frac{\widehat{\gamma^{-1}}}{\sqrt{R(\tau) - r_0}}.$$

「全てが感受性を有する宿主」で「流行の全てが観察された」などを中心とした想定が妥当である場合，上式により  $R_0$  推定値が得られる(Becker, 1989)．流行が部分的にのみ観察された場合は尤度方程式が閉鎖的な式で得ることが出来ないため，マルチンゲール法を駆使するなど他の方法が必要となる(Becker, 1993)．その他，実際のデータにおいては感染率  $\beta$  が流行時刻によって変動することを考慮せねばならないし(Becker and Yip, 1989)，感受性が宿主毎に異なる場合にはマルコフ連鎖モンテカルロを駆使する方法も提唱されている(O'Neill and Becker, 2001)．対象とする流行や感染症毎にデータの性質が異なるため，得られたデータの内容と特性によって異なる推定手法を柔軟に対応させなければならないのが本研究の特色である．

### 3.3 効果的再生産数 ( $R_t$ ) の意義

第 3.1 節の議論を別の角度から見ることによって発展させる．例えば，関数  $\lambda_u$  を「各個体

が感染後時刻  $u$  においてどれくらい感染性を有するか」を意味するものと定義する．これは感染待ち時間  $X$  内においては 0 であるはずで，その後に感染性期間  $Y$  を経るまでの  $X+Y$  までの間は平均で  $\beta$  だとする．つまり，関数  $\lambda_u$  は最も単純な表現では以下で与えられる：

$$(3.4) \quad \lambda_u = \begin{cases} \beta, & X \leq u \leq X+Y, \\ 0, & u < X \text{ and } X+Y < u. \end{cases}$$

実際には  $\beta$  は感染性期間  $Y$  において定数ではないが(第 4 節参照)，ここでは簡便に平均値として想定する．この  $\lambda_u$  を全て積分することによって平均の伝播能力を得ることが出来るが，感染源  $i$  の伝播能力を表すランダム量は以下のように表される：

$$(3.5) \quad \Theta_i = \int_0^\infty \lambda_u^{(i)} du = \beta Y_i.$$

つまり，伝播能力は

$$(3.6) \quad \theta = E(\Theta) = \beta E(Y),$$

によって与えられる．この  $\theta$  も  $R_0$  推定値の定義に近い概念である．これは図 4b に示したような分布から得られる平均値に等しい．

しかし，ここで注意すべきことは  $\theta$  が伝播能力の平均値であって時刻  $t$  に独立とする暗黙の想定である．例えば流行が終息に向かう場合，流行開始間もなくの頃と流行終焉前の時刻とでは伝播能力が異なることが通常である(感染者数の減少理由は感受性宿主数の感染による減少のみでない)．流行終焉前には直接伝播する感染症であれば，行動の変化(接触回数低下，防衛行動など)や公衆衛生対策(学級閉鎖，検疫，隔離など)が認められる．上の  $\theta$  だと，そういった背景要因の変化に推定値がバイアスされるのが通常である．そこで，効果的再生産数(Effective reproduction number,  $R_t$ )が定義される：これは「(ある時刻  $t$  における，一定の対策下での) 1 人の感染者による 2 次感染者数」を意味する概念である．重要な  $R_0$  との違いが(付きで書かれるのは，必ずしも時刻  $t$  の関数ではなく，また必ずしも明白に一定の対策下という外因的な要素を考慮したものでない，という曖昧な  $R_t$  の定義による(むしろ  $R_0$  でないことを意味する場合に  $R_t$  が用いられることが多い)．

例えば， $R_t$  は図 4a のような感染ネットワークから 2 次感染者数を直接数えることによって各時刻で得られる．英国で流行した家畜における口蹄疫に関する研究では，膨大な感染ネットワークから  $R_t$  の推定値を時間  $t$  について得ることによって  $R_t$  が流行中どこで 1 未満に至ったかを評価した(Haydon et al., 2003)．また，世代時間の分布が得られていれば感染ネットワークが全く得られない場合にも理論的にネットワークを構築することによって  $R_t$  を推定することができる．例えば，全ての感染者が独立に 2 次感染者を感染させていたと想定できる場合，世代時間の確率密度関数が  $f(t|\theta)$  で与えられるとすると(ここで  $\theta$  は母パラメータ)，ある感染者  $k$  が過去の感染者  $l$  によって感染させられた確率が  $f(t_k - t_l|\theta)$  で与えられる．ここで  $t_k$  および  $t_l$  はそれぞれ感染者  $k$  および  $l$  の発症時刻である．これを利用すれば， $m$  人の感染源候補(Candidate)がいる中で， $k$  が  $l$  によって感染させられた確率は以下で与えられる：

$$(3.7) \quad p_{(k,l)} = \frac{f(t_k - t_l|\theta)}{\sum_{m \neq k} f(t_k - t_m|\theta)}.$$

これによって，各感染者の 2 次感染者数が得られ，

$$(3.8) \quad R_t = \sum_k \text{Binomial}[p_{(k,l)}],$$



によって  $R_t$  が与えられる。Wallinga and Teunis (2004) は同方法を利用して SARS の流行において WHO が緊急事態のアラートを発した直後に  $R_t$  推定値が 1 未満に低下したことを示した。これは小規模流行にも応用できるほか (Nishiura and Kakehashi, 2005), 感染ネットワークが部分的のみに知られている場合は世代時間を部分的にのみ利用して再構築することも可能である (Nishiura et al., 2006)。

#### 4. 主要パラメータの役割と各想定に関する拡張

##### 4.1 潜伏期間

(2.1) における最も単純な数理モデルでは感染待ち時間について指数分布を想定しているが、現実には他の密度関数でより正確に説明できることの方が多し。より精密な感染待ち時間の分布によって解の定性的な挙動も変化することから (Lloyd, 2001a), 数理モデルによる現実的な流行予測にはフィールド調査に基づく各パラメータの分布を特徴付けた上で利用することが必要であると言われている (Wearing et al., 2005)。厳密には感染待ち時間と潜伏期間は図 3 の概略にあるように異なるが、その個体間の性質が類似しているために感染待ち時間の統計分布は頻りに潜伏期間と同等に扱われる (あくまで簡便性を追求したもの)。図 5 に直接伝播する感染症に関する潜伏期間の分布を例示した。潜伏期間の多くはフィールド調査でより明らかにしやすいことから、古くから多くのデータが蓄積されており、そのうちでも感染型 (生体内で病原体が増殖する) に分類される潜伏期間の多くは対数正規分布に従う傾向にあることが知られている (Sartwell, 1950)。この分布を特徴付ける統計学的諸研究は日本でも盛んに実施されてきた (阪本, 1985; 丹後, 1998)。

潜伏期間は共通感染源で一斉暴露 (Point source) の食中毒集団発生時に暴露時刻を推定することに役立つ (共通感染源だが一斉暴露でない場合を Common source と呼ぶ)。丹後 (1998) の研究では、以下のような 3 つのパラメータによる対数正規分布を潜伏期間に想定することに

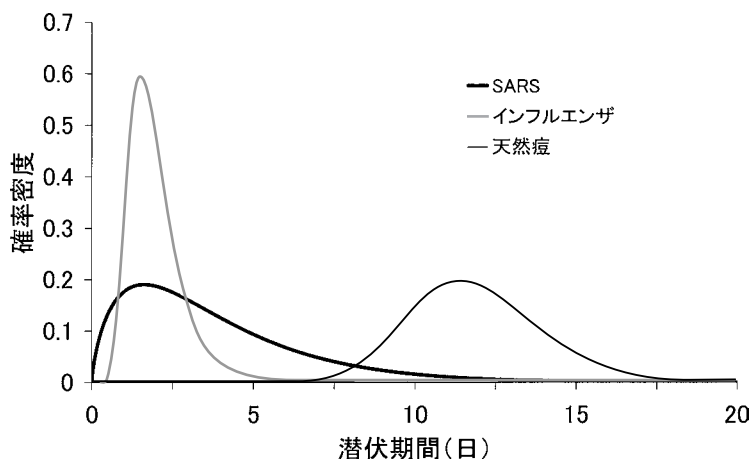


図 5. 潜伏期間の分布。SARS、インフルエンザ、天然痘の潜伏期間。インフルエンザおよび天然痘は対数正規分布で密度関数が与えられ、平均および標準偏差のパラメータはそれぞれ  $\mu = 0.6, \sigma = 0.4$  および  $\mu = 2.5, \sigma = 0.2$  である (単位は  $\log(\text{days})$ )。SARS に関しては引き続き統計学的研究の目的でガンマ分布が用いられた (Donnelly et al., 2003)。尺度及び位置パラメータは  $g = 2.0$  及び  $s = 2.1$ 。

よって、暴露時刻  $\tau$  を推定することを提唱している:

$$(4.1) \quad L_i(t_i; \mu, \sigma^2, \tau) = \frac{1}{\sqrt{2\pi\sigma(t_i - \tau)}} \exp \left\{ -\frac{[\ln(t_i - \tau) - \mu]^2}{2\sigma^2} \right\}.$$

ここで  $\mu$  および  $\sigma^2$  は対数正規分布の平均および分散であり、 $t_i$  は感染者  $i$  の発症時刻である。この感染者  $i$  に関する尤度について潜伏期間の代わりに世代時間を利用すれば、直接伝播する感染症で(3.7)にある  $R_t$  の推定式と原理が共通である点は興味深い。直接伝播する感染症における潜伏期間から直接的に得られる実践的意義としては、暴露を受けた可能性のある者の行動を制限する「検疫(Quarantine)」に必要な目安の時間がわかる点が挙げられる(西浦, 2004)。

また、AIDS(後天性免疫不全症候群)や BSE(ウシ海綿状脳症) vCJD(異型クロイツフェルトヤコブ病)に代表されるような緩徐に病理が進行して発症する感染症では、逆計算法(Backcalculation)によって感染者(頭)数推定及び短期将来予測が可能であり、その目的で潜伏期間が利用される(Brookmeyer and Gail, 1988)。HIV の感染から AIDS 発症までの潜伏期間をフィールド研究に基づく統計分布(その確率密度関数を時刻  $\tau$  において  $\omega(\tau)$  とする)と仮定すれば、ある時刻  $t$  での HIV 感染者数  $h(t)$  と潜伏期間を利用することによって、時刻  $t$  での AIDS 発症者数  $a(t)$  が以下のように表される:

$$(4.2) \quad a(t) = \int_0^t h(u)\omega(t-u)du.$$

HIV 感染症(あるいは BSE 感染)の特徴として、感染してから発症する迄の数年間にはほぼ無症候であり感染に気付かれないままにいる点が挙げられる。一方で、AIDS あるいは BSE のように顕著な発症があった場合、先進国で疾病報告制度・サーベイランスが十分に優れているとすれば一定の信頼性を持って AIDS と BSE の発症者(頭)数が時間毎に得られると想定できる。逆計算法は(4.2)を逆に解くこと(deconvolution)によって HIV 感染者数の時間的推移を推定できる点に利点がある。例えば、AIDS サーベイランスの報告データは 4 半期など一定の間隔で報告されることを考慮して(4.2)を離散的に書くと以下ようになる:

$$(4.3) \quad a_t = \sum_{u=1}^t h_u \omega_{t-u}.$$

HIV 感染者数が時間毎に一定でないポアソン過程に従って増加していると仮定した場合、各時刻の AIDS 患者数は独立なポアソン変量と考えられ、流行開始後時刻  $T$  までを観察したデータの尤度方程式は以下に比例すると考えられる:

$$(4.4) \quad \prod_{t=1}^T \left( \sum_{u=1}^t h_u \omega_{t-u} \right)^{r_t} \exp \left( - \sum_{u=1}^t h_u \omega_{t-u} \right).$$

ここで  $r_t$  は時刻  $t$  で観察された AIDS 患者数である。これによって HIV 感染者数が推定され、その潜伏期間にプールされた感染者数から短期的な将来予測が実施できる。実際には、さらに複雑な感染過程の想定を加味した上で妥当性を検討しなければならない。HIV/AIDS に関しては Brookmeyer and Gail (1994) 及び Dietz et al. (1994), BSE に関しては Donnelly et al. (2003) 及び Nishiura and Kakehashi (2004) が詳しい。

#### 4.2 感染性期間

潜伏期間と同様に感染性期間も適切な分布が存在するはずだが、潜伏期間と感染待ち時間の関係ほど感染性期間は症候性期間に類似ではない。また、第 2.3 節で説明した通り、病期によって感染性が異なることは生物学的に知られている。この複雑な期間の典型例として HIV 感染症が挙げられる(Shiboski and Jewell, 1992)。HIV 感染症は、感染してしばらく経過した後の

数か月間は大量のウイルス粒子が血中を循環する非常に感染可能性の高い時期であると知られている。しかし、免疫反応とウイルス潜伏のメカニズムによって、その後ウイルス負荷は低下して数年間そのレベルが保たれる。最終的に AIDS を発症する迄に再びウイルス負荷が増加する。即ち、効果的な治療が介在しない状況下における HIV 感染症の感染性は最も単純には 2 峰性の分布を示すと考えられる。定性的研究においても、この感染年齢 (infection age;  $d$ -state と呼ばれる) の影響が大きいことは知られている (Diekmann and Heesterbeek, 2000)。そのため、感染性を大きく左右するウイルス量は  $R_0$  の推定値にも大きく影響することが知られており (Lloyd, 2001b)、特に緩徐に進行する感染症では個体内のウイルス動態を無視した  $R_0$  推定はほぼ不可能に等しいことが示されている。また、感染しても無症候状態で感染性を有することは感染症の制圧に大きな影響を与えることが知られており、それは HIV/AIDS が理論的にコントロールが難解であると考えられる根拠である (Fraser et al., 2004)。飛沫感染・空気感染する感染症における感染性期間の第 1 の直接的意義は、感染して発症した患者の行動を制限する「隔離 (Isolation)」に必要な時間の目安を見積もることが出来る点にある。これまでにウイルス学的にウイルス分離あるいは抗原検出によって感染性を証明する試みがなされているが (例えば SARS では Chu et al., 2005)、そのウイルス分離あるいはウイルス量と疫学的な感染性の評価指標の相関は未だ確実には得られておらず今後の課題である。

#### 4.3 様々な想定 of 拡張

第 2 節で紹介した数理モデルは全ての者が均一に接触する点をはじめとして最もシンプルな想定に基づいているが、ほとんどの流行で接触の均一性を検定すると棄却されると考えられている (Britton, 1997)。個体毎に伝播を特徴付ける要素に関して異質性 (Heterogeneity) が認められる場合はそれを考慮しなければならない。例えば、接触頻度の違いを原因とする感染率の異質性を処理する最も単純な方法は、均一な接触パターンを有する複数の亜集団 (Subpopulation) に分けて考えることである (この方法をマルチタイプ型流行 (Multitype epidemic) モデルと呼ぶ)。例えば、感染率  $\beta_i$  をタイプ  $i$  の個体における感染率とすれば、タイプ  $i$  について (2.4) と同様の計数過程を想定できる：

$$(4.5) \quad \begin{aligned} N_i(t) = S_i(0) - S_i(t) \text{ の intensity は } & \beta_i \bar{S}_i(t-) \alpha^T I(t-), \\ R_i(t) \text{ の intensity は } & \gamma I_i(t-). \end{aligned}$$

ここで、 $\alpha$  および  $I(t)$  は感染源の相対的感染性および各タイプの感染性宿主を意味するベクトルである (各要素にタイプ  $j$  の感染性・感染性宿主、 $\{\alpha_j\}, \{I_j(t)\}$  を含む)。この時、以下のマルチンゲールが定義される：

$$(4.6) \quad \begin{aligned} M_{i,1}(t) &= N_i(t) - \int_0^t \beta_i \bar{S}_i(u) \alpha^T I(u) du, \\ M_{i,2}(t) &= R_i(t) - \gamma \int_0^t I_i(u) du. \end{aligned} \quad i = 1, \dots, k.$$

マルチンゲールの重要な特性として (1) 一様可積分性、及び (2) マルチンゲールの線形結合 (linear combination) もマルチンゲールであること、が挙げられる (Andersen et al., 1993)。この特性を利用すると以下が得られる (Britton, 1998a)：

$$(4.7) \quad U_i(t; \alpha, \beta) = \frac{\beta_i}{\gamma} \sum_j \frac{\alpha_j}{n} M_{j,2}(t) - \int_0^t \frac{1}{S_i(u-)} dM_{i,1}(u), \quad i = 1, \dots, k.$$

ここで  $n$  は初期感受性人口で  $S(0)$  に等しい。また、 $dN_i(u) = -dS_i(u)$  であるのだから (4.7) は

以下のようにも与えられる:

$$(4.8) \quad \begin{aligned} U_i(t; \alpha, \beta) &= \frac{\beta_i}{\gamma} \sum_j \frac{\alpha_j}{n} R_j(t) + \int_0^t \frac{dS_i(u)}{S_i(u-)} \\ &= \frac{\beta_i}{\gamma} \sum_j \frac{\alpha_j}{n} R_j(t) - A_i(t). \end{aligned}$$

ここで  $A_i(t)$  は以下で与えられる:

$$(4.9) \quad A_i(t) = \frac{1}{S_i(0)} + \frac{1}{S_i(0) - 1} + \cdots + \frac{1}{S_i(t) + 1}.$$

十分に大きなサイズの人口において初期にほとんどの者が感受性を有する場合,これが  $-\log(1 - \tilde{p}_i)$  で近似される( $\tilde{p}_i$  は観察されたタイプ  $i$  の個体群における流行終焉時の感染者割合である). また, 線形制約条件から  $\sum_j \alpha_j R_j(t)/n = \sum_j \alpha_j \pi_j \tilde{p}_j = \gamma$  であり( $\pi_i$  は全人口におけるタイプ  $i$  人口の割合; Becker and Marschner, 1990), 以上からタイプ  $i$  の感染率  $\beta_i$  の推定量は以下で与えられる:

$$(4.10) \quad \hat{\beta}_i \approx -\log(1 - \tilde{p}_i).$$

これは, ある流行が最後まで完全に観察された条件下で最終状態 (final state) がわかっている場合に適用できるので応用範囲が広い (Britton, 2001), 一方で最終状態のみに依存した推定値であることに注意しなければならない.

異質性の重要性が容易に理解できる例として, ワクチン接種政策の事例が挙げられる. 例えば, ワクチンの効果をフィールド疫学に基づいて推定する場合でも, ワクチン接種者の中で複数の異なる感染力に従って感染していることが明らかな場合は 2 つ以上の亜集団に分類することによって効果を推定する必要がある (Halloran et al., 1992). また, 接触パターンなど異質性の原因が明らかで直接観察できる場合は, 特定の集団を接種対象にすることによって最も合理的なワクチン接種戦略を構築することが可能である (Britton, 1998b; Ball et al., 2004). 最近では定性的な研究として接触ネットワークに着目した研究が急速に発展している. 例えば, 性感染症の現実的な伝播を考えた場合, 特定の個体群だけが他の個体群と比較して明らかに高頻度の接触回数を有する場合がある (Gupta et al., 1989). これは過去にコア (Core) 集団と定義され, コアとそれ以外で構成される 2 つの亜集団におけるマルチタイプ型モデルとして検討されてきた (Hethcote and Yorke, 1984; Stigum et al., 1994). 最近の学問的興味の多くは, 接触回数に関する頻度分布が得られた場合に, その頻度が Pareto の法則に基づくようなべき乗分布 ( $p = Cx^{-r}$ ;  $C$  は定数,  $x$  は接触回数,  $r$  はべき指数) に従う (Power law) ものとして一般化された場合にある. ある集団における接触パターンがべき乗則に従って特徴的なスケールに依存しない (べき乗分布は分散が無限大であり, ある感染源のみが異常な数の接触回数を有する) 性質にある場合をスケールフリー (Scale free) であるという (Pastor-Satorras and Vespignani, 2001). この性質が成り立った条件下で, 現実世界では接触パターンがわからない感染症の場合は, その制圧は困難を極めると考えられる (May and Lloyd, 2001). 類似の問題で接触頻度の根拠となる接触ネットワークとは異なり, 発症した者の感染ネットワークから得られる 2 次感染者数の頻度に関する分布について「異常に多くの 2 次感染者数を産み出した感染源」がスーパースプレッダー (Superspreader) と定義される. 感染ネットワークの分布は図 4b に示した通り右方向に裾を引く歪度の大きな分布であることが知られており, 同様の情報を検討した結果, 多くの流行においてスーパースプレッダーは頻りに存在するものであることが知られている (Lloyd-Smith et al., 2005). 接触頻度に関する定量的な説明は接触回数の計数が直接的に観察可能である性感染症から明らかになるであろうと期待されている (Schneeberger et al., 2004).

また、今までの議論は閉鎖人口について行なわれてきたが、現実には SARS 流行の地理的拡大に見られるように人口移動を伴うことが多い (Watts et al., 2005)。同様のことは HIV 感染者数が諸先進国と比較して少ないわが国に見られる来日外国人中の HIV 感染者数の報告からも窺い知ることが出来る (Nishiura et al., 2004)。さらに、流行は各地の人口密度を始めとする諸要因に影響を受けるために地理的な異質性が認められる。これら人口移動と地理的異質性を考慮したメタ個体群 (Metapopulation) モデルは感染症流行に応用されて久しいが (Sattenspiel and Dietz, 1995)、一方で定量的な妥当性の検討は非常に難しい。最近の Nakaya et al. (2005) にあるように、推定を伴うモデルにおいて接触パターン・発病・人口移動に関する一定の想定を如何に妥当に扱えるかは今後の課題である。

以上のように現実性は 1 つひとつの現実的要素を考慮する毎に複雑性を増すため、数式やプログラムが複雑化するだけでなく最適化 (Optimization) も困難になることがある。その為、例えば伝播能力を推定する場合は、個々の流行データを用いるよりは何らかの特異的な調査環境・条件の設定によって介入研究を実施し、実験的にパラメータ推定を実施することも多い。その典型例が家屋内伝播 (Household transmission) である。その利点の 1 つは、ある家屋内で伝播が起こっている場合は集団内の他の状況に独立に伝播が起こっているものと想定できる点にある (Becker, 1989)。つまり、特異的な家屋内の条件設定において (4.5) にあるような最終状態における感染者割合から伝播能力が推定できると想定する。しかし、ヒトの場合は厳格な移動制限などの環境が得られないために屋外感染を調整する必要があり、屋内と屋外で異なる感染率を想定した確率論的モデルを構築することによって推定する試みも報告されている (Addy et al., 1991)。

## 5. まとめ

感染症疫学に関する数理モデルは、力学系・確率論と確率過程・位相幾何学・統計学的推定などの分野を問わずに研究が同じ土俵で進められており、非常に応用的側面に富んだ領域である。感染症は統計データに富む一方で、将来予測や抑止対策評価という大きな目標と動機があるために現実的想定に関する批判的見方が厳しい。紹介した数理モデルは必ずしも現実を完全に説明できるものではなく、推定を目的とした統計モデルでさえも何らかの非現実的想定に基づかなければならないことも多い。これは感染症だけでなく一般的な多くの数理モデルに共通したことも知れないが、内在性の想定をどこまで論理的・現実的に描写できるかによって定性的・定量的に妥当性の高いモデル構築をする点に学問的興味は集中している。わが国の応用的な感染症数理モデルは潜伏期間で紹介したような統計モデルを例外として、これまで内在性想定に関する議論や定性的な解の挙動に関する理解が十分に深いとは言えなかった。過去に HIV/AIDS や結核を中心として流行予測に線形回帰分析がいくつか利用されてきた例に見られるように、伝播動態に関する認識は低く、それにも関わらず定量的側面に関する大きな拘りが社会的ニーズとして発展したためか、本領域は諸先進国研究と比較して明らかに遅れをとっている。最近になって本分野に関する高い関心が寄せられつつあるが、内在性および外因性の想定に基づいて「感染症がどのような振る舞いで流行するのか」に関する国際標準の考え方をまず私たちは学びとる必要がある。

さて、直接伝播する感染症の流行や集団発生を対象として、数理疫学的研究で用いられる想定や手法が個々の流行特性によって異なる点を明らかにしながら基礎的理論を紹介した。実際の流行データを扱う場合の具現化 (Realization) および予測は個別に異なり、確率論的モデルを利用した場合でも解析的なアプローチや個々の内在する一因子に関する定性的役割を探索する動機に基づくものであることも少なくない。また、最近のモデルベースの数値的なシ

ミュレーション研究では、個体の異質性をそのままモデル想定内で考慮した個人ベースモデル(individual-based model あるいは agent-based simulation)が盛んに用いられるようになった(Grimm and Railsback, 2005)。個人ベースモデルから個々のパラメータ推定を実施する最適化手法も急速に発展している。また、離散的なマルコフ連鎖や紹介した逆計算法に代表されるように、許容できる内在性の想定があれば系を想定せずに短期的な将来予測も可能である(Jackson et al., 2003)。また最近では、個々のパラメータ推定や家屋内伝播から得られる予測ではマルコフ連鎖モンテカルロが頻繁に推定に利用される傾向がある(Smith et al., 2004; Cauchemez et al., 2004)。感染症の数理的諸研究は疫学モデルに留まらず、類似の系を想定した研究は生体内の病原体の動態を説明することに応用範囲が広げられ、システムバイオロジーとの境界も不明瞭にさなりつつある(Nowak and May, 2000; Velasco-Hernandez et al., 2001)。最後に、本稿では典型的な直接伝播する感染症(主に飛沫感染)の数理モデルに関する議論が中心であったが、空気感染(Wells, 1934; Riley et al., 1978)、水系感染(Eisenberg et al., 2002; Eisenberg et al., 2005)や媒介動物による感染(Smith and McKenzie, 2004; McKenzie and Bossert, 2005)の数理的諸研究も現実と対比した内容が議論できつつある。

#### 謝 辞

チュービンゲン大学の Klaus Dietz 教授及び Martin Eichner 教授から構成に関するアドバイスをいただいたことに深謝したい。本研究の一部は万有生命科学振興国際交流財団の研究補助を受けて実施された。

#### 参 考 文 献

- Addy, C. L., Longini, I. M. and Haber, M. (1991). A generalized stochastic model for the analysis of infectious disease final size data, *Biometrics*, **47**, 961–974.
- Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer, Berlin.
- Anderson, R. M. (1982). Transmission dynamics and control of infectious disease agents, *Population Biology of Infectious Diseases* (ed. R. M. Anderson), 149–176, Springer, Berlin.
- Anderson, R. M. (1991). Discussion: The Kermack-McKendrick epidemic threshold theorem, *Bulletin of Mathematical Biology*, **53**, 3–32.
- Anderson, R. M. and May, R. M. (1982). Directly transmitted infectious diseases: Control by vaccination, *Science*, **215**, 1053–1060.
- Anderson, R. M. and May, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, New York.
- Andersson, H. and Britton, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*, Lecture Notes in Statistics, No. 151, Springer, New York.
- Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd ed., Griffin, London.
- Ball, F. G. (1983). The threshold behaviour of epidemic models, *Journal of Applied Probability*, **20**, 227–241.
- Ball, F., Britton, T. and Lyne, O. (2004). Stochastic multitype epidemics in a community of households: Estimation and form of optimal vaccination schemes, *Mathematical Biosciences*, **191**, 19–40.
- Baroyan, O. V. and Serenko, A. F. (1961). An outbreak of smallpox in Moscow during 1959–1960,

- Zhurnal Mikrobiologii, Epidemiologii, i Immunobiologii*, **32**, 72–79 (in Russian).
- Becker, N. G. (1977). Estimation for discrete time branching processes with application to epidemics, *Biometrics*, **33**, 515–522.
- Becker, N. G. (1989). *Analysis of Infectious Disease Data*, Chapman and Hall, New York.
- Becker, N. G. (1993). Martingale methods for the analysis of epidemic data, *Statistical Methods in Medical Research*, **2**, 93–112.
- Becker, N. G. and Britton, T. (1999). Statistical studies of infectious disease incidence, *Journal of the Royal Statistical Society Series B*, **61**, 287–301.
- Becker, N. G. and Marschner, I. C. (1990). The effect of heterogeneity on the spread of disease, *Lecture Notes in Biomathematics*, **86**, 90–103.
- Becker, N. G. and Yip, P. (1989). Analysis of variation in an infection rate, *Australian Journal of Statistics*, **31**, 42–52.
- Britton, T. (1997). A test of homogeneity versus a specified heterogeneity in an epidemic model, *Mathematical Biosciences*, **141**, 79–99.
- Britton, T. (1998a). Estimation in multitype epidemics, *Journal of the Royal Statistical Society Series B*, **60**, 663–679.
- Britton, T. (1998b). On critical vaccination coverage in multitype epidemics, *Journal of Applied Probability*, **35**, 1003–1006.
- Britton, T. (2001). Epidemics in heterogeneous communities: Estimation of  $R_0$  and secure vaccination coverage, *Journal of the Royal Statistical Society Series B*, **63**, 705–715.
- Brookmeyer, R. and Gail, M. H. (1988). A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic, *Journal of the American Statistical Association*, **83**, 301–308.
- Brookmeyer, R. and Gail, M. H. (1994). *AIDS Epidemiology: A Quantitative Approach*, Oxford University Press, New York.
- Cauchemez, S., Carrat, F., Viboud, C., Valleron, A. J. and Boelle, P. Y. (2004). A Bayesian MCMC approach to study transmission of influenza: Application to household longitudinal data, *Statistics in Medicine*, **23**, 3469–3487.
- Chu, C. M., Cheng, V. C. C., Hung, I. F. N., Chan, K. S., Tang, B. S. F., Tsang, T. H. F., Chan, K. H. and Yuen, K. Y. (2005). Viral load distribution in SARS outbreak, *Emerging Infectious Diseases*, **11**, 1882–1886.
- Diekmann, O. and Heesterbeek, J. A. P. (2000). *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, Wiley Series in Mathematical and Computational Biology, Wiley, New York.
- Dietz, K. (1974). Transmission and control of arbovirus diseases, *Proceedings of the Society for Industrial and Applied Mathematics, Epidemiology* (ed. D. Ludwig), 104–106, Society for Industrial and Applied Mathematics, Philadelphia.
- Dietz, K. (1993). The estimation of the basic reproduction number for infectious diseases, *Statistical Methods in Medical Research*, **2**, 23–41.
- Dietz, K., Molineaux, L. and Thomas, A. (1974). A malaria model tested in the African savannah, *Bulletin of the World Health Organization*, **50**, 347–357.
- Dietz, K., Seydel, J. and Schwartlander, B. (1994). Back-projection of German AIDS data using information on dates of tests, *Statistics in Medicine*, **13**, 1991–2008.
- Donnelly, C. A. (2003). Extending backcalculation to analyse BSE data, *Statistical Methods in Medical Research*, **12**, 177–190.
- Donnelly, C. A., Ghani, A. C., Leung, G. M., Hedley, A. J., Fraser, C., Riley, S., Abu-Raddad, L. J., Ho, L. M., Thach, T. Q., Chau, P., Chan, K. P., Lam, T. H., Tse, L. Y., Tsang, T., Liu, S. H.,

- Kong, J. H., Lau, E. M., Ferguson, N. M. and Anderson, R. M. (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong, *Lancet*, **361**, 1761–1766.
- Eichner, M. (2003). Case isolation and contact tracing can prevent the spread of smallpox, *American Journal of Epidemiology*, **158**, 118–128.
- Eichner, M. and Dietz, K. (2003). Transmission potential of smallpox: Estimates based on detailed data from an outbreak, *American Journal of Epidemiology*, **158**, 110–117.
- Eisenberg, J. N., Brookhart, M. A., Rice, G., Brown, M. and Colford, J. M. (2002). Disease transmission models for public health decision making: Analysis of epidemic and endemic conditions caused by waterborne pathogens, *Environmental Health Perspectives*, **110**, 783–790.
- Eisenberg, J. N., Lei, X., Hubbard, A. H., Brookhart, M. A. and Colford, J. M. (2005). The role of disease transmission and conferred immunity in outbreaks: Analysis of the 1993 Cryptosporidium outbreak in Milwaukee, Wisconsin, *American Journal of Epidemiology*, **161**, 62–72.
- Farrington, C. P. and Whitaker, H. J. (2005). Contact surface models for infectious diseases: Estimation from serologic survey data, *Journal of the American Statistical Association*, **100**, 370–379.
- Farrington, C. P., Kanaan, M. N. and Gay, N. J. (2001). Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data, *Journal of the Royal Statistical Society Series C*, **50**, 251–292.
- Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S. and Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia, *Nature*, **437**, 209–214.
- Fine, P. E. (2003). The interval between successive cases of an infectious disease, *American Journal of Epidemiology*, **158**, 1039–1047.
- Fraser, C., Riley, S., Anderson, R. M. and Ferguson, N. M. (2004). Factors that make an infectious disease outbreak controllable, *Proceedings of the National Academic of Science USA*, **101**, 6146–6151.
- Giesecke, J. (2002). *Modern Infectious Disease Epidemiology*, Edward Arnold, London.
- Grimm, V. and Railsback, S. F. (2005). *Individual-based Modeling and Ecology*, Princeton University Press, Princeton and Oxford.
- Gupta, S., Anderson, R. M. and May, R. M. (1989). Networks of sexual contacts: Implications for the pattern of spread of HIV, *AIDS*, **3**, 807–817.
- Halloran, M. E., Haber, M. and Longini, I. M. (1992). Interpretation and estimation of vaccine efficacy under heterogeneity, *American Journal of Epidemiology*, **136**, 328–343.
- Haydon, D. T., Chase-Topping, M., Shaw, D. J., Matthews, L., Friar, J. K., Wilesmith, J. and Woolhouse, M. E. (2003). The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak, *Proceedings of the Royal Society of London, Biological Science*, **270**, 121–127.
- Hethcote, H. and Yorke, J. A. (1984). *Gonorrhea Transmission Dynamics and Control*, Springer, New York.
- 稲葉 寿(2002). 『数理人口学』, 東京大学出版会, 東京.
- 巖佐 庸(1997). 『数理生態学シリーズ・ニューバイオフィジックス』, 共立出版, 東京.
- 巖佐 庸(1998). 『数理生物学入門—生物社会のダイナミクスを探る』, 共立出版, 東京.
- Jackson, C. H., Sharples, L. D., Thompson, S. G., Duffy, S. W. and Couto, E. (2003). Multistate Markov models for disease progression with classification error, *Journal of the Royal Statistical Society Series D (The Statistician)*, **52**, 193–209.
- Kaplan, E. H., Craft, D. L. and Wein, L. M. (2002). Emergency response to a smallpox attack:



- The case for mass vaccination, *Proceedings of the National Academy of Science USA*, **99**, 10935–10940.
- Kermack, W. O. and McKendrick, A. G. (1927). Contributions to the mathematical theory of epidemics – I, *Proceedings of the Royal Society Series A*, **115**, 700–721.
- Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M., Ma, S., James, L., Gopalakrishna, G., Chew, S. K., Tan, C. C., Samore, M. H., Fisman, D. and Murray, M. (2003). Transmission dynamics and control of severe acute respiratory syndrome, *Science*, **300**, 1966–1970.
- Lloyd, A. L. (2001a). Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics, *Theoretical Population Biology*, **60**, 59–71.
- Lloyd, A. L. (2001b). The dependence of viral parameter estimates on the assumed viral life cycle: Limitations of studies of viral load data, *Proceedings of the Royal Society of London Series B*, **268**, 847–854.
- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. and Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence, *Nature*, **438**, 355–359.
- Longini, I. M., Nizam, A., Xu, S., Ungchusak, K., Hanshaworakul, W., Cummings, D. A. and Halloran, M. E. (2005). Containing pandemic influenza at the source, *Science*, **309**, 1083–1087.
- May, R. M. and Lloyd, A. L. (2001). Infection dynamics on scale-free networks, *Physical Review E*, **64**, 066112.
- McKenzie, F. E. and Bossert, W. H. (2005). An integrated model of Plasmodium falciparum dynamics, *Journal of Theoretical Biology*, **232**, 411–426.
- Mills, C. E., Robins, J. M. and Lipsitch, M. (2004). Transmissibility of 1918 pandemic influenza, *Nature*, **432**, 904–906.
- Molineaux, L. and Gramiccia, G. (1980). *The Garki Project. Research on the Epidemiology and Control of Malaria in the Sudan Savanna of West Africa*, World Health Organization, Geneva.
- Murray, J. D. (2002). *Mathematical Biology I: An Introduction (Interdisciplinary Applied Mathematics)*, Springer-Verlag, New York.
- Murray, J. D. (2003). *Mathematical Biology II: Spatial Models and Biomedical Applications (Interdisciplinary Applied Mathematics)*, Springer-Verlag, New York.
- Nakaya, T., Nakase, K. and Osaka, K. (2005). Spatio-temporal modeling of the HIV epidemic in Japan based on the national HIV/AIDS surveillance, *Journal of Geographical Systems*, **7**, 313–336.
- 西浦 博 (2004). SARS (重症急性呼吸器症候群) の汎世界的流行：科学の貢献と社会的問題, *科学*, **74**, 965–969.
- Nishiura, H. and Kakehashi, M. (2004). Interpreting the trends of bovine spongiform encephalopathy in Japan: Application of the backcalculation method to analyze case records, *Journal of Veterinary Epidemiology*, **8**, 65–76.
- Nishiura, H. and Kakehashi, M. (2005). Real time estimation of reproduction numbers based on case notifications—effective reproduction number of primary pneumonic plague, *Tropical Medicine and Health*, **33**, 127–132.
- Nishiura, H., Imadzu, L., Yoshiyama, T., Yanai, H., Sawazaki, Y., Kakehashi, M. and Ishikawa, N. (2004). HIV notification and migration: Epidemiological analysis of migrants from Southeast Asian countries during 1986–2001, *The Journal of Population Studies*, **35**, 1–10.
- Nishiura, H., Kuratsuji, T., Quy, T., Phi, N. C., Van, B. V., Ha, L. E., Long, H. T., Yanai, H., Keicho, N., Kirikae, T., Sasazuki, T. and Anderson, R. M. (2005). Rapid awareness and transmission of severe acute respiratory syndrome in Hanoi French Hospital, Vietnam, *American Journal of Tropical Medicine and Hygiene*, **73**, 17–25.
- Nishiura, H., Schwehm, M., Kakehashi, M. and Eichner, M. (2006). Transmission potential of pri-

- mary pneumonic plague: Time-inhomogeneous evaluation based on historical documents of the transmission network, *Journal of Epidemiology and Community Health*, **60**, 640–645.
- Nokes, D. J. and Anderson, R. M. (1988). The use of mathematical models in the epidemiological study of infectious diseases and in the design of mass immunization programmes, *Epidemiology and Infection*, **101**, 1–20.
- Nowak, M. A. and May, R. M. (2000). *Virus dynamics: Mathematical Principles of Immunology and Virology*, Oxford University Press, Oxford.
- O'Neill, P. D. and Becker, N. G. (2001). Inference for an epidemic when susceptibility varies, *Bio-statistics*, **2**, 99–108.
- Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks, *Physical Review Letters*, **86**, 3200–3203.
- Riley, E. C., Murphy, G. and Riley, R. L. (1978). Airborne spread of measles in a suburban elementary school, *American Journal of Epidemiology*, **107**, 421–432.
- Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., Hedley, A. J., Leung, G. M., Ho, L. M., Lam, T. H., Thach, T. Q., Chau, P., Chan, K. P., Lo, S. V., Leung, P. Y., Tsang, T., Ho, W., Lee, K. H., Lau, E. M. C., Ferguson, N. M. and Anderson, R. M. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions, *Science*, **300**, 1961–1966.
- 阪本州弘 (1985). 『疫学と疫学モデル』, 金芳堂, 京都.
- Sartwell, P. E. (1950). The distribution of incubation periods of infectious diseases, *American Journal of Hygiene*, **51**, 310–318.
- Sattenspiel, L. and Dietz, K. (1995). A structured epidemic model incorporating geographic mobility among regions, *Mathematical Biosciences*, **128**, 71–91.
- Schneeberger, A., Mercer, C. H., Gregson, S. A., Ferguson, N. M., Nyamukapa, C. A., Anderson, R. M., Johnson, A. M. and Garnett, G. P. (2004). Scale-free networks and sexually transmitted disease: A description of observed patterns of sexual contacts in Britain and Zimbabwe, *Sexually Transmitted Diseases*, **31**, 380–387.
- Shiboski, S. C. and Jewell, N. P. (1992). Statistical analysis of the time dependence of HIV infectivity based on partner study data, *Journal of the American Statistical Association*, **87**, 360–372.
- Simpson, R. E. (1952). Infectiousness of communicable diseases in the household (measles, chicken-pox, and mumps), *Lancet*, **ii**, 549–554.
- Smith, D. L. and McKenzie, F. E. (2004). Statics and dynamics of malaria infection in Anopheles mosquitoes, *Malaria Journal*, **3**, 13.
- Smith, T., Dietz, K., Vounatsou, P., Mueller, I., English, M. and Marsh, K. (2004). Bayesian age-stage modeling of Plasmodium falciparum sequestered parasite loads in severe malaria patients, *Parasitology*, **129**, 289–299.
- Stigum, H., Falck, W. and Magnus, P. (1994). The core group revisited: The effect of partner mixing and migration on the spread of gonorrhoea, Chlamydia, and HIV, *Mathematical Biosciences*, **120**, 1–23.
- 丹後俊郎 (1998). 潜伏期間に対数正規分布を仮定した集団食中毒の暴露時点の最尤推定法, 日本公衆衛生雑誌, **45**, 129–141.
- Velasco-Hernandez, J. X., Garcia, J. A. and Kirschner, D. E. (2001). Remarks on modeling host-viral dynamics and treatment, *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction* (eds. C. Castillo-Chavez, S. Blower, P. van den Driessche, D. Kirschner and A. A. Yakubu), Volume 125, 287–308, Springer, New York.
- Wallinga, J. and Teunis, P. (2004). Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures, *American Journal of Epidemiology*, **160**, 509–516.

- Watts, D. J., Muhamad, R., Medina, D. C. and Dodds, P. S. (2005). Multiscale, resurgent epidemics in a hierarchical metapopulation model, *Proceedings of the National Academy of Science USA*, **102**, 11157–11162.
- Wearing, H. J., Rohani, P. and Keeling, M. J. (2005). Appropriate models for the management of infectious diseases, *PLoS Medicine*, **2**, e174.
- Wells, W. F. (1934). On airborne infection. Study II. Droplet nuclei, *American Journal of Hygiene*, **20**, 611–618.

## Prediction of Infectious Disease Outbreak with Particular Emphasis on the Statistical Issues Using Transmission Model

Hiroshi Nishiura<sup>1</sup> and Hisashi Inaba<sup>2</sup>

<sup>1</sup>Department of Medical Biometry, University of Tübingen

<sup>2</sup>Department of Mathematical Sciences, University of Tokyo

This article reviews statistical issues that arise in evaluation methods as well as model building of directly transmitted infectious diseases to predict and manage actual outbreaks. To perform model-based predictions, it is essential to detail a priori information of transmission potential, latency and infectious periods. Population dynamics is frequently applied to the spread of disease and, in particular, the counting process is assumed to derive an estimator for the key parameter to quantify the precision of estimates and the variation in data. Thus, here we describe methods for estimating basic (and effective) reproduction number(s) based on observational epidemic records and how to interpret the intrinsic assumptions. Moreover, we introduce a backcalculation method for slowly progressive diseases (i.e., HIV/AIDS and BSE), which enables us to estimate the total number of infected individuals and obtain short-term projections. As the majority of epidemic records violate the homogeneously mixing assumption, it is fruitful to select model structure and statistical method flexibly, depending on the infectious disease of interest and the characteristics of available data. In all methods, intrinsic assumptions and validity play the most important role in prediction and estimation.