

# 研究報告書

## 「非疫学データによる感染症流行動態解析の新展開」

研究タイプ: 通常型

研究期間: 平成27年10月～平成31年3月

研究者: 大森 亮介

### 1. 研究のねらい

感染症の制圧の為に、流行状態の把握と予測、介入の有効性を理解する必要があるが、それらは全て感染症の流行動態の把握から始まる。これまでに感染症流行動態は、感染者数の時系列変化を記述する SIR モデルと呼ばれる数理モデルを、感染者数時系列データに当てはめ解析されることが主流であった。しかしながら、感染者数の時系列データは様々な問題が付随する。感染個体である宿主がヒトである場合は感染者数データを入手する事自体が困難である。病院の来院者数データ等の既存のデータを使用する場合は、母集団の選定にバイアスがある為に正確な感染者数が把握できず、新たな調査を開始する場合は莫大なコストがかかる。さらには、ヒト以外の宿主では、人間への影響が直接的でないために、感染個体数の情報は非常に少ない。昨今の新興感染症の多くはヒト以外の宿主の感染症がヒト集団に侵入することで問題となっているため、ヒト以外の宿主での流行状態の把握が必須となる。しかしながら感染個体数のデータが存在しない為に、旧来の疫学解析ができず、代わりとなるデータとその解析手法の開発が急務となっている。例え過去の感染者数の時系列情報が入手可能であったとしても、感染者数以外のデータを使用しなければ解析が困難な事が多い。感染症の伝搬過程は未感染宿主と感染宿主の接触によって起こるが、宿主同士の接触は宿主の行動によるものであり、この情報は感染者数の時系列情報から抽出することは不可能である。近年において、現実的な個体による接触率の差を考慮した解析がされはじめ、これまでの単純化した宿主の行動を仮定した流行動態の解析結果と大きく異なる事が明らかになっている。特に HIV をはじめとする性感染症は、性的接触ネットワークを介して流行するが、このネットワークは個体同士の繋がりに大きな差が見られる為に、旧来のモデルによる解析結果と現実の流行動態は大きく異なる。性的接触ネットワーク構造の理解なしでは性感染症の流行動態は把握できず、宿主行動の理解が求められている。近年は感染症の流行制御が重要視され、感染個体数以外の流行に関するデータを取得され、利用可能になりつつある。しかし、これらのデータを解析する手法の開発は未だ途上の段階にある。本研究はこれらの手法の開発及び活用を通じ、数学と感染症分野との協働による感染症流行のコントロールの現状における限界を打破することを目的とする。

### 2. 研究成果

#### (1) 概要

感染個体数のデータが存在しない場合での疫学解析の為に、遺伝子配列情報から疫学情報を推定する手法の確立を試みた。進化学で淘汰の度合いの指標として用いられる Tajima's D を用いて、病原体遺伝子情報から疫学情報、特に基本増殖率  $R_0$  の推定を行う方

法を提案した。Tajima's D は選択圧だけでなく、集団サイズの時間変化にも影響を受けることが知られており、この性質により病原体遺伝子の時系列変化が考察されてきた。しかしながらこの性質は現実のデータの解析には適していない仮定の下で導出されたものである。そこで、感染症流行ダイナミクスの数理モデルである SIR モデルに病原体の塩基配列情報の時系列変化を組み込み、Tajima's D の集団サイズの時間変化に対する依存性を解析した。Tajima's D の集団サイズの時間変化に対する依存性は集団サイズの有限性と集団ダイナミクスの確率性によるものと推察され、Tajima's D からの  $R_0$  の推定は可能であることが示唆された[1]。また、Tajima's D からの  $R_0$  の推定法を実際の感染症の塩基配列情報に適用し、推定された  $R_0$  を従来の方法である感染者数の時系列データから推定された  $R_0$  の値と比較し、この方法の妥当性を確認した[2]。

詳細な性的接触ネットワークを考慮した性感染症疫学解析の為に、性行動データから推定された国別、性別、婚姻状態別の性交渉パートナー数と国別の HIV 有病率の関連性を解析した。理論的には HIV 有病率は性交渉パートナー数の平均と分散によって説明されるはずであるが、現実の性行動データからはそのような関連性は観察できず、性行動データは HIV 有病率を説明出来るほどの正確性を有していないことが判明した。これは、性行動データ以外のデータからの性的接触ネットワークの推定および性感染症流行動態の推定が必要となることを意味する。その推定法の一つとして、HIV 以外の性感染症の流行からの HIV の流行の推定がある。この推定法の妥当性の検証として、性的接触ネットワークモデル上での HIV と性器ヘルペス2型(HSV-2)の流行ダイナミクスを記述した数理モデルを構築し、HIV と HSV-2 の流行の関連性を解析した。これにより、多様な性的接触ネットワークの多様性を考慮したとしても、性器ヘルペスの流行レベルは4割程度の性的接触ネットワークにおける HIV の流行レベルを説明する事が判明した[4]。さらに、HSV-2 の流行レベルとネットワーク統計量と組み合わせることで9割程度の性的接触ネットワークにおける HIV の流行レベルを説明することが出来ることが判明した[4]。

## (2) 詳細

研究テーマ A「感染個体数のデータが存在しない場合での疫学解析の為に遺伝子配列情報から疫学情報を推定する手法の確立」

Tajima's D は集団遺伝学、進化学において選択圧の有意性を検定するために広く用いられている統計量である。Tajima's D は選択圧だけでなく、集団サイズの時間変化にも影響を受けることが知られており、その影響は生物集団サイズが拡大すると Tajima's D は負になり、縮小すると正になると言われており、この性質により病原体遺伝子の時系列変化が考察されてきた。しかしながらこの性質は集団サイズの変化率が著しく高い事を仮定しており、現実のデータの解析には適していない。そこで、感染症疫学解析で最も多用される流行ダイナミクスの数理モデルである SIR モデルに病原体の塩基配列情報の時系列変化を組み込み解析した。そのモデルを解析したところ、もし病原体の集団サイズが連続であり感染症の流行ダイナミクス及び進化ダイナミクスが決定論的に起きていると仮定すると、Tajima's D は集団サイズの時間変化と独立である事を証明した[1]。次に、病原体の集団サイズが有限であり、感染症の流行ダイナミクス及び進化ダイナミクスが確率論的に起きていると仮定すると、Tajima's D

の集団サイズの時間変化に対する依存性が観察された。これらの結果から、Tajima's D の集団サイズの時間変化に対する依存性は集団サイズの有限性と集団ダイナミクスの確率性によるものと推察された[1]。これは病原体遺伝子配列の時系列情報から感染宿主数の時系列を推定し得る事を意味する。また、もし病原体の集団サイズが有限であり、感染症の流行ダイナミクス及び進化ダイナミクスが確率論的に起きているとしても、”生物集団サイズが拡大すると Tajima's D は負になり、縮小すると正になる”というこれまでの Tajima's D の解釈の通説は現実に観察される病原体遺伝子データにおいては全く当てはまらない事が判明した。この方法の妥当性をアルゼンチンで採取されたヒトインフルエンザ AH1N12009pdm の塩基配列情報に適用し推定された  $R_0$  を従来の方法である感染者数の時系列データから推定された  $R_0$  の値と比較し、この方法の妥当性を確認した[2]。

遺伝子配列情報による疫学解析は、比較的単純な流行ダイナミクスが観察される病原体では確立した手法が適用できる。しかし、複雑な流行ダイナミクスを示す病原体においては、流行予測や制御はその流行ダイナミクスの理解が前提条件となる。複雑な流行ダイナミクスを引き起こす因子の一つに、再感染による宿主の感受性の変化が多様である事が挙げられるが、そのダイナミクスの理解は十分ではない。よって、再感染による宿主の病原体に対する感受性の変化を考慮した感染症流行の数理モデルを構築し、ダイナミクスの分類を行なった。これにより、今まで理論的にも発見されていなかった、感染者数は一度減少した後アウトブレイクが起きるといった新たなダイナミクスを発見し、”Delayed outbreak”と名付けた。

研究テーマ B「病原体の流行干渉が引き起こす複雑な流行ダイナミクスの予測手法の確立」  
複数の系統が存在する病原体では、系統間で互いに流行が干渉し流行ダイナミクスが複雑となり、予測が困難である事が知られている。インフルエンザウイルス B 型は現在、Yamagata 系統と Victoria 系統の二種類が存在し、どちらの系統が主な流行系統となるかを予測することはワクチン株選定の為に重要である。インフルエンザウイルスの流行予測の困難さは、流行干渉による複雑な流行ダイナミクスに加え、感染確率が時間変動する外的要因により決定されることと、各シーズンの出現時期が流行動態に多大な影響を与えるがその出現時期は予測困難であることに起因する。そこで、系統間の流行干渉を考慮した系統毎の感染者数の時系列を記述する確率過程の数理モデルを構築し、過去5年分の系統ごとの感染者数の時系列にあてはめることにより、感染確率と複数の因子との関連性を系統毎に推定した[3]。推定したパラメータにより 2015-2016 シーズンにおける Yamagata 系統と Victoria 系統の流行規模の比の予測を行ったところ、本手法はある程度の精度での予測が可能であった。また、各系統の出現時期の情報を使用できない場合でも、本手法で確立した数理モデルによる出現時期の予測値を用いた予測も、出現時期の実データを用いた予測より精度は劣るものの、ある程度の予測精度をもつことが判明した。また、推定された感染確率と複数の因子との関連性のパラメータ値を系統間で比較したところ、Yamagata 系統の方が Victoria 系統に比べ免疫を獲得しにくいことが示唆された。

研究テーマ C「複雑な性行動を考慮した性感染症疫学解析手法の確立」

性行動データから推定された国別、性別、婚姻状態別の性交渉パートナー数と国別の HIV 有病率の関連性を解析したところ、性行動データは HIV 有病率を推定できる精度を保持できていないことが判明し、性行動データ以外のデータからの性的接触ネットワークの推定および性

感染症流行動態の推定が必要であることが判明した。代替推定法の一つとして、HIV 以外の性感染症の流行からの HIV の流行の推定に着目し、その推定法確率のために複雑な性的接触ネットワーク上での性感染症流行の関連性を解析した。性的接触ネットワーク生成過程の数理モデルに HIV と HSV-2 の流行ダイナミクスを記述した数理モデルを組み込むことで、HIV と HSV-2 の流行の関連性を解析したところ、HIV と HSV-2 ではネットワークに対する流行度合いの依存性が違う事が判明した(HIV の流行レベルの大半は未婚者のパートナー数の平均と分散によって決まることに対し、性器ヘルペスの流行レベルの大半は未婚者、既婚者のパートナー数の平均と分散、クラスター係数によって決まる)[4]。また、上記ネットワークモデルの広域なパラメータ領域を探索したところ、多様な性的接触ネットワークの多様性を考慮したとしても、性器ヘルペスの流行レベルは4割程度の性的接触ネットワークにおける HIV の流行レベルを説明する事が判明した。さらに、性器ヘルペスの流行レベルにネットワーク統計量(次数相関、クラスター係数、コンカレンシー)と組み合わせることにより9割程度の性的接触ネットワークにおける HIV の流行レベルを説明することが出来ることが判明した[4]。

上記のモデルでは HIV 感染と HSV-2 感染は完全に独立であることを仮定しているが、“HSV-2 に感染すると HIV に感染しやすくなる”という仮説が疫学データから提示され、議論が盛んに行なわれている。この疫学データとは、HIV と HSV-2 の個人レベルの感染状況を示すデータであるが、HIV と HSV-2 の感染経路は性行為という同一のものであるために、HIV に感染した人はある程度の高リスクな性行動をしていた可能性があるために HSV-2 の感染リスクも高い可能性がある。そこで、これまでに構築した HIV と HSV-2 のモデルを用い HIV と HSV-2 の個人レベルの感染状況を調べ、もし HIV と HSV-2 の感染が独立であったとしても“HSV-2 に感染すると HIV に感染しやすくなる”様に解釈されてしまう様な状況が起こりうるかを調べたところ、約 6.5 倍感染しやすくなる様に解釈されてしまうことが判明した[5]。また、もし“HSV-2 に感染すると 2.5 倍 HIV に感染しやすくなる”と仮定したとしても、約 8 倍感染しやすくなると解釈されてしまい、疫学データの解釈には開発した詳細な性的接触ネットワーク形成と性感染症流行を同時に記述したモデルを用いて解析する必要があることを示した。

### 3. 今後の展開

開発した感染個体数のデータが存在しない場合での疫学解析の為に遺伝子配列情報から疫学情報を推定する手法を実際の様々な病原体遺伝子配列情報に当てはめ、疫学解析を行う。また、この推定法を拡張し、流行を引き起こす突然変異箇所の推定法を開発する。

複雑な性行動を考慮した性感染症疫学解析手法の精度向上の為に、HSV-2 以外の性感染症同士の流行の関連性の理論解析を行う。

### 4. 自己評価

感染個体数のデータが存在しない場合での疫学解析の為に遺伝子配列情報から疫学情報を推定する手法の開発は、開発とその手法の妥当性の検証が行えた、目標を達成できた。複雑な性行動を考慮した性感染症疫学解析手法に関しては、これまでの手法の評価及び代替案の提案と精度の検証が行えた点ではおおよその目標は達成できた。今後は精度の向上を続ける。研究費は当初の予定通りに執行された。本課題で開発された疫学解析手法を元に、実際に疫学解析に適用し、よりの確な感染症流行に対する介入、制御の方法を提案できると

予想する。

## 5. 主な研究成果リスト

### (1) 論文(原著論文)発表

- |  |
|--|
| 1. Ryosuke Omori, Jianhong Wu. Tajima's D and site-specific nucleotide frequency in a population during an infectious disease outbreak. <i>SIAM Journal on Applied Mathematics</i> . 2017, 77(6), 2156-2171    |
| 2. Kiyeon Kim, Ryosuke Omori, Kimihito Ito. Inferring epidemiological dynamics of infectious diseases using Tajima's D statistic on nucleotide sequences of pathogens. <i>Epidemics</i> . 2017, 21, 21-29      |
| 3. Mayumbo Nyirenda, Ryosuke Omori, Heidi L. Tessmer, Hiroki Arimura, Kimihito Ito. Estimating the Lineage Dynamics of Human Influenza B Viruses. <i>PLOS ONE</i> . 2016, 11(11) e0166107                      |
| 4. Ryosuke Omori, Laith J. Abu-Raddad. Sexual network drivers of HIV and herpes simplex virus type 2 (HSV-2) transmission. <i>AIDS</i> . 2017, 31(12), 1721-1732   |
| 5. Ryosuke Omori, Nico Nagelkerke, Laith J. Abu-Raddad. HIV and Herpes Simplex Virus Type 2 Epidemiologic Synergy: Misguided Observational Evidence? <i>Sexually Transmitted Infections</i> . 2017, 94 372-376 |

### (2) 特許出願

研究期間累積件数:0件(公開前の出願件名については件数のみ記載)

### (3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. 学会発表 Omori R, Nakata Y. "Heterogeneity in susceptibility induces unpredictable outbreak". 11<sup>th</sup> European Conference on Mathematical and Theoretical Biology; July 2018; Lisbon, Portugal.
2. 学会発表 Omori R, Abu-Raddad LJ. "Sexual network drivers of HIV and herpes simplex virus type 2 (HSV-2) transmission: a comparative mathematical modeling analysis". The STI & HIV World Congress 2017; June 2017; Rio de Janeiro, Brazil.
3. 学会発表 Omori R, Wu J. "Tajima's D & site-specific nucleotide frequency of pathogen during its outbreak". The 2017 Society for Mathematical Biology Annual meeting; July 2017; Salt Lake City, U.S.A.
4. 学会発表 Omori R. "Analysis of lineage dynamics of human influenza B viruses towards the prediction of epidemic dynamics". One Health EcoHealth Congress 2016; December 2016; Melbourne, Australia.

5. プレスリリース「性器ヘルペスと HIV の流行の関連性は性的接触ネットワークの構造で大きく異なることを解明」