

がん登録実務者支援としての データ分析ワークショップ

佐賀大学¹⁾、佐賀県²⁾、国立がん研究センター³⁾
 滝崎光浩¹⁾、佐々木和典¹⁾、中田聡子¹⁾、
 原 晴太¹⁾、長本敬²⁾、松田健大³⁾



背景①

- 「がん登録」は、がんの罹患や転移の状況を登録・把握し、分析する仕組みであり、がんの患者数や罹患率、生存率、治療効果の把握など、がん対策の基礎となるデータを把握するために必要なものである。
- がん対策を推進するためには、正確ながんの実態把握が必要であり、「がん登録」がその中心的な役割を果たす。
- 「がん登録等の推進に関する法律」施行（2016年1月）により、届出漏れがなくなり、がん罹患の状況を把握がより正確化する期待。

地域がん登録～全国がん登録

- 「がん登録等の推進に関する法律」施行前：地域がん登録
 - ・都道府県毎にデータを収集
 - ・医療機関からの任意の届出に基づく
- 全国がん登録（2016年1月以降）：「がん登録等の推進に関する法律」
 - ・すべての病院と指定診療所：届出義務
 - ・届出漏れがなくなり、がん罹患の状況を把握がより正確化

背景②

- がん登録の主要な目的の一つに、がん対策の企画・立案・実施のための調査研究や、がん医療の質の向上などを目的として行われる調査研究でのがん登録情報の利用がある。登録実務者の役割もデータの精度維持を含むデータ提供だけではなく、病院執行部や研究者と連携したデータ分析業務への期待も大きい。
- データサイエンスの重要性の認識が浸透してきている一方で、簡単に評価されたグラフ等を聞いた誤解させることを留意した情報も記載している。

目的

- がん登録におけるデータ分析は、根拠に基づいたがん医療のエビデンスの信頼性に直結するものであるため、それに携わる実務者に対するデータ分析に関する基本的知識の解説と演習を試行した。

データサイエンスとは

- データサイエンスは、数々が日常的に行っているデータの収集や分析を通して、新たな知見や発見を得るための学問である。また、問題解決や意思決定をサポートする手法でもある。私たちが健康やビジネス、社会のさまざまな分野で活用されている。データから得られる情報を活かすことでより効果的な意思決定を行うことができる。
- データサイエンスは、専門的な知識やスキルが必要であるが、基本的なデータの収集と分析は誰にでも取り扱えるものである。
- 自身の業務にどうデータを活用し、分析することが出来るか、より大きな課題に対しても実践的なスキルを身につけることができる。

データサイエンスのアプローチ

- データサイエンスでは、まずデータを集める。食事内容や体重を記録するために、スマートフォンのアプリや書き置きノートを使用する。次に、集めたデータを分析する。データ分析の手法やツールを用いて、食事のパターンと体重の変化に関連性があるかどうかを調べる。
- 分析の結果、特定の食事パターンと体重の増加が関連していることがわかった場合、自身の食事習慣を見直すことができる。また、他の人々のデータと比較することで、一般的な傾向やベストプラクティスを見つけることも可能。

がん登録データとデータサイエンス

- データサイエンスは、がん登録実務者にとっても重要な役割を果たす。
- データを分析し、トレンドやパターンを把握することで、診療の改善や治療戦略の最適化につながる分析結果を担当医に示すことができる。
- また、他の医療機関や地域との比較によって、自院の診療成績や特徴を客観的に評価することができる。

ワークショップ開催の目的

- データサイエンスの活用には、統計学やデータ分析の基礎知識が必要である。
- データサイエンスのスキルや知識を身につけるための学習リソースを提供する目的でワークショップを開催した。
- 講義を聴くだけでなく実務経験もデータサイエンス活用能力の獲得には重要である。
- 実際にがん登録データや患者データを取り扱い、分析の経験を通して、データサイエンスのスキルを深くすることができる。

ワークショップがもたらす副次的効果

- 他の実務者とデータ分析のテーマについて意見交換できる。
- データ分析に関する疑問点などを専門家に講義できる。
- コラボレーションや情報共有も大切であり、意見交換や経験の共有によって、より豊富な知識や洞察を得ることができる。

ワークショップの実際

- 2日間完結コースとしてR-Workshopを2022年8月に実施した。
 - ・抽出統計分析全般について入門的知識の解説を行う。Rによる統計解析をサンプルデータの分析を行いながら解説した。
 - ・高橋が、4-5分（政府統計の総合窓口）よりがん登録データを抽出し、それを用いた分析結果を提示した。また、Excelでの分析手法を併用し、次にその分析手法を解説した。
- 事後アンケートの結果を踏まえ、佐賀県がん登録研究会開催に合わせて、Workshopの補足的内容をmini-workshopとして、2022年12月、2023年1月、2月にオンライン（webexによるwebinar）で開催した。

ワークショップ実施概要とロードマップ

日付	内容
2022/08/05 13:00～17:00	R-Workshop(1) ①
2022/08/12 13:00～17:00	② 統計学入門～統計学基礎～データ分析の基礎 ③ 基礎～データサイエンスの活用④
2022/08/19 13:00～17:00	R-Workshop mini
2022/09/01 13:00～17:00	⑤ がん登録実務者向けデータ分析講座⑥
2022/09/08 13:00～17:00	⑦ がん登録実務者向けデータ分析講座⑧

がん登録研究会向けデータ分析講座 (R-Workshop) 開催 (オンライン)	2022年度開催	2023年度開催
2022年8月5日 13:00～17:00 R-Workshop(1) ①	2022年12月 13:00～17:00 R-Workshop mini ⑤	2023年1月 13:00～17:00 R-Workshop mini ⑥
2022年8月12日 13:00～17:00 ② 統計学入門～統計学基礎～データ分析の基礎 ③ 基礎～データサイエンスの活用④	2022年12月 13:00～17:00 R-Workshop mini ⑤	2023年1月 13:00～17:00 R-Workshop mini ⑥

Workshop実施方法：事前準備

- 各受講者のPCでR及びR-Studioが動作する環境を事前に整えておく必要があるため、インストール方法を解説した動画を事前に公開した。



Workshop実施方法：web開催特有の対応

- Covid-19により対面開催ができなかったため、webexのウェビナー形式で実施した。
- 資料共有をアプリケーション画面毎に行うと切り替えのための時間ロスが生じるため、画面全体を共有してアプリケーション切り替え時の待ち時間を減らすように工夫した。
- コード入力時は入力内容がよく見えるように、画面拡大等を適宜行うよう配慮した。

画面例

- Excelでの分析手法の例示



画面例

- R-Studioでの分析手法の例示



演習Workshopをオンラインで実施できるか？

- PC操作を行うWorkshopをオンラインで実施するにあたり不安があった。
- PC操作がうまくいかないとき、実際の画面が確認できないので対応が困難。
- 受講者の様子が確認しにくいので、進行のペース配分が難しい。
- 自力ではできないが少しサポートがあればできることも多いが、オンラインではその対応が難しい。

対面 vs. オンライン

対面のメリット	対面のデメリット
<ul style="list-style-type: none"> ・受講者の様子が現場で確認できる ・PCのトラブル時に対応できる 	<ul style="list-style-type: none"> ・スタッフ数不足で全体把握は困難 ・その場で対応できるスタッフの確保は困難

オンラインならではのメリットもある
<ul style="list-style-type: none"> ・入室するコードなどをチャット機能を使って一斉に伝達できる。 ・制限：数値と見えない、タイプミスなどが発生する。 ・共有画面を拡大して注目部分を見やすくできる。

Excelで行うデータ分析

- Excelは日常業務で使い慣れている。
- 分析に使う場合もハードルが低い。
- 結果が保存しやすい。
- ・分析→レポート作成まで自動化して、グラフ作成/印刷/保存
→データ抽出/集約/集約/集約/集約
→データ抽出/集約/集約/集約/集約
→データ抽出/集約/集約/集約/集約
- ・年次報告等同じ分析を行う場合でも同じ操作が必要（作業が重ならない）
→操作が同じなら認識/印刷/印刷/印刷/印刷
- Excelによる高度な分析（統計学的検定等）は学会等で受け入れられていない。

Rで行うデータ分析

- 統計学的分析で定評がある。
- ・分析しやすいが、データ入力の負担は初心者には負担がある。
- ・使い始めるハードルが高い。
- ・分析を行うにはプログラミングコードの記述が必要など
- ・すべてコードとして記述しているため、データ選択ミス等を見逃し確認できる。
- ・全ての分析過程を第三者が追跡できる。
- ・年次報告等同じ分析を行う場合は、2回目以降はデータを用意するだけで手間が大幅に削減、効率化できる。

受講後アンケートの結果

- 満足度について
 - ・楽しかった：第1回WS 71.4% 第2回WS 66.7%
- 自由記載コメント
 - ・疑問事項が多かった。
 - ・オンデマンド学習資料が多かった。

考察とまとめ

- ほとんどの受講者が初めて触れるプログラミング環境だったため難しいという印象を持つことは避けられなかったが、操作性の認識を高く、継続して学びたいという意見が多かったことよりデータ分析を学ぶことの重要性は伝わったと思われる。
- また、実務者と情報部門からの共同参加がもたらしたことも、どのように分析したいかのニーズを持っている実務者とそれを実施できる技術を持っている情報部門が協力する体制があるという点で良い結果である。
- Workshopの成否の要因は「準備万端か？」ということであり、対面かオンラインかではない。次年度以降の到達目標を定め、授業設計を適切に行いつつ、workshopを継続し理解度を高めていきたい。