

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 12 日現在

機関番号：84409

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24701042

研究課題名(和文)生存時間解析における樹木構造接近法によるがん患者の予後予測モデルの検討

研究課題名(英文)Prognostic model of cancer patients using the relative survival regression tree

研究代表者

伊藤 ゆり (Yuri, Ito)

地方独立行政法人大阪府立病院機構大阪府立成人病センター(研究所)・その他部局等・その他

研究者番号：60585305

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：がんの診断時の情報および予後が追跡されているがん登録資料を軸に、医療機関において収集された生活習慣アンケートや教授うち情報に基づく社会経済指標をリンクしたデータベースを用いて、がん患者の予後予測を行った。がんの予後を左右する重要な因子から順に分岐する樹木構造接近法を非小細胞肺癌患者のデータに適用し、予後を予測するモデルを推定した。予後に最も影響を与えるのは診断時の進行度であり、また、早期のがん患者であっても、過去または現在喫煙者においては非喫煙者と比べて予後が悪かった。この方法により、様々な予後因子の関連性を反映し、がん患者を予後の異なるグループに分類できることが示唆された。

研究成果の概要(英文)：We developed the regression tree based on the relative survival approach. Using this data mining approach, we can explore the prognostic factors and classify patients to significant subgroups from database with many candidate variables. We applied to the data of non-small-cell lung cancer patients from the hospital cancer registry data of our institute linking with the lifestyle survey for patients. Stage at diagnosis was the most important prognostic factor. Smoking status was a prognostic factor among the localised patients. We proposed that this approach can classified patients to subgroups with different prognosis.

研究分野：がん疫学

キーワード：生存解析 がん登録 予後予測モデル

1. 研究開始当初の背景

がん患者の予後因子については、部位ごとに各種臨床情報に基づき、ある程度特定はされている。しかし、臨床上重要とされる予後因子に加え、性別や年齢など基本的な患者の背景因子や喫煙や飲酒などの生活習慣を考慮すると実に多種多様な組み合わせとなり、どの要因ががん患者の予後に大きく影響を与えるか、またそれらの交互作用の詳細は明らかではない。

そのような状況下で、予後因子を探索し、予後予測を行う方法論が発展してきた。樹木構造接近法は多数の因子の中から探索的にアウトカムとの関連を見出し、樹木構造として視覚化する手法として用いられている。分類回帰樹木法 (CART: Classification And Regression Tree, Breiman et al. 1984) に始まり、多くの手法が開発されてきた。生存時間解析においても樹木構造接近法が取り入れられ、多数の因子の中から有意な関連性を有する因子を抽出し、それらの組み合わせにより異なる予後を有する患者群に分類することが可能となった (LeBlanc, 2006)。

樹木構造接近法は統計数理の分野において方法論の開発は進んでいるものの、実際の医療データに適用し、医療現場に還元しうる研究成果はまだ少なかった。一方、近年がん診療連携拠点病院を中心に整備され始めた院内がん登録資料のようにがん患者の医療を評価するためのデータベースや入院時アンケートから得られる生活習慣、治療内容、レセプト情報などの病院情報は十分に活用されてこなかった。

2. 研究の目的

本研究はがん登録資料を軸とし、診療科で袖手する臨床情報や入院時アンケートなどで把握される生活習慣に関する情報を用いて、実際のがん患者のデータに基づいた予後予測モデルを行い、医療現場に還元することを目的とする。

医療現場においては、臨床試験の結果では限られた特徴をもつ患者像への治療効果しかわからないため、あらゆる年齢層 (特に増加する高齢者層)、腫瘍のタイプ、生活習慣などとの組み合わせによる治療の効果についての情報が求められている。そのような情報還元を行うためにがん登録資料を臨床データベースと連結し既存の医療情報を最大限に活用した予後予測モデルを確立することを目的とする。

3. 研究の方法

がん患者の診断時の腫瘍情報、治療内容、生存時間及び転帰 (生存 / 死亡) を含む院内がん登録資料を基本のデータベースとし、これに各種病院情報 (入院時生活習慣アンケート、詳細治療内容など) をリンケージする。このデータセットを用いて、がん患者の生存時間に影響を与える因子を探索的に抽出し、

影響度の高い順に樹木構造を推定し、異なる予後を持つ患者集団に分類した。対象は肺非小細胞がん患者とした。

(1) 生存時間解析における樹木構造接近法

以下に樹木構造推定における概要を示す。分岐過程 (患者群を生存率の違いで分割する)

どの説明変数が生存率の最も異なる群に分割するかを探索する。最も説明度の高い変数で分割された患者集団ごとに上記を繰り返す。全ての分岐過程を経た大きな (複雑な) 樹木構造が出来上がる。

刈り込み過程

階層が多い複雑な樹木ほど適合度は高いが、複雑すぎるデメリットがある (トレードオフの関係)。そこで、複雑さと適合度のバランスが取れた樹木を選定する過程が必要となる。

最適樹木の決定 (交差確認法)

学習データ (全データの一部) により樹木を推定し、テストデータ (残りのデータ) で、樹木の正確さを検証し、最適樹木を決定する。

(2) 相対生存率の考え方を考慮した樹木構造接近法 (Relative survival regression tree)

研究協力者の杉本により、既存の Survival CART の方法を改良し、(1) における分岐の過程で、がんによる過剰死亡をアウトカムとした相対生存率モデル (Dickman et al. 2004 Stat Med.) を用いた方法が確立された。本研究はこの手法を用いて、分析を行った。

4. 研究成果

(1) Relative survival regression tree の分析ツール

本研究に用いた非小細胞肺がん患者のデータを使用し、研究協力者の杉本により、相対生存率モデルに基づく分岐過程を組み込んだ分析手法およびツールが開発された。従来用いられてきた方法では、がん以外の脂肪の影響を除去できなかった。通常、死因の特定が困難である疾病登録データベースを用いる場合には、相対生存率の考え方をを用いて、当該疾患により過剰に死亡した場合をアウトカムとして分析する。この考え方を Survival CART に組み込むことができたため、今後、他の疾患登録データベースを用いた予後予測モデルの推定にも適用可能となった。

(2) データベースのリンケージおよび候補変数の洗い出し

Dataset 1: 生活習慣アンケートと院内がん登録資料と社会経済指標のリンケージ

大阪府立成人病センター院内がん登録より 2004-2007 年診断患者の診断後 3~5 年の予後情報を得て、「健康と生活習慣」データベースとのリンケージを行った (胃がん 624 件、肺がん 499 件、大腸がん 204 件、食道が

ん 226 件)。上記データを用いて、相対生存率モデルによる単変量・多変量解析を行い、樹木構造に投入する変数の候補を検討した。

院内がん登録より性・年齢などの患者背景因子および診断時進行度および組織型、治療内容などの腫瘍情報を得た。生活習慣アンケートより、喫煙 (Pack year)、アルコール摂取、家族歴、食事や職業などに関する項目を得た。これらを相対生存率モデルの単変量解析により、 $p < 0.25$ の基準で候補変数として選ばれたのは、性、年齢、進行度、組織型、術後補助療法の有無 (放射線または化学療法の有無)、BMI、アルコール摂取量、喫煙 (Pack year)、合併症の有無、肺がん検診受診の有無、社会経済指標であった。

Dataset 2: 院内がん登録と社会経済指標のリンケージ

Dataset 1 は、診断からの経過年数が短く、イベント数が少なかったため、クリアーに患者の予後が分類できない可能性があったので、診断から 10 年以上経過している症例を用いて、分析を行った。候補となった変数は性、年齢、組織型、術後補助療法の有無、社会経済因子であった。

(3) Relative survival regression tree により推定された樹木構造

Dataset 1: 生活習慣アンケートと院内がん登録資料と社会経済指標のリンケージ

生活習慣の情報をリンケージしたデータベース (2004-2008 年診断患者 475 例) に基づく予後予測モデルにおいては、進行度が限局 (早期がん) の患者において喫煙の有無が予後を左右する因子として抽出された。また、喫煙者においても、主に喫煙が原因とされている扁平上皮がんの患者の 5 年相対生存率は 92.6% と、それ以外の患者の 5 年相対生存率 99.7% と比べ、早期であっても若干予後が悪かった。進行がん (領域・遠隔転移) においては、扁平上皮がんの 5 年相対生存率が 61.3% と最も低かった。しかし、診断からの経過年数が 3~5 年と短かったため、予後を分類するには不十分なイベント数であったと考える。

Dataset 2: 院内がん登録と社会経済指標のリンケージ

Dataset 1 では十分なイベント数ではなかったため、より長期間追跡したデータ (1991-2001 年診断患者 700 例) を用いて樹木構造接近法を適用した。推定された樹木では、臨床進行度、がんの組織型に加え、治療情報 (治癒的切除や補助療法の有無) も予後を分類する項目として選択された。この樹木モデルでは 10 年相対生存率が 0.7%~92.7% と予後が大きく異なる 7 群に分類された。

(4) 総括および今後の課題

本研究では、がん登録資料を軸とし、他の

データベースをリンケージした実際のがん患者のデータベースを用いた予後予測モデルの検討方法を提案した。今回使用したデータベースの性質上、臨床応用可能な予後予測モデルはまだ完成していないが、今後、生活習慣を含んだデータセットにおいてさらに追跡期間を延長し、再度分析を行う。また、当院呼吸器外科の診療科データベースと院内がん登録資料とのリンケージを行い、非小細胞肺がんについての予後予測モデルを分析中である。特に、進行度別や年齢階級別の予後予測モデルを検討する必要がある。

また、臨床医とともに本手法により推定された樹木構造を検討した折、樹木が複雑すぎるなどの意見も聞かれた。より臨床現場で役に立つ予後予測モデルのあり方について検討を続けたい。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 15 件)

1. Tabuchi T, Nakayama T, Fukushima W, Matsunaga I, Ohfuji S, Kondo K, Kawano E, Fukuhara H, Ito Y, Oshima A. Determinants of participation in prostate cancer screening: a simple analytical framework to account for healthy-user bias. *Cancer Sci.* 2015;106:108-14. doi:10.1111/cas.12561
2. Chihara D, Ito H, Izutsu K, Hattori M, Nishino Y, Ioka A, Matsuda T, Ito Y. Advance and stagnation in the treatment of patients with lymphoma and myeloma: Analysis using population-based cancer registry data in Japan from 1993 to 2006. *Int J Cancer.* 2015. doi:10.1002/ijc.29477
3. Tabuchi T, Hoshino T, Hama H, Nakata-Yamada K, Ito Y, Ioka A, Nakayama T, Miyashiro I, Tsukuma H. Complete Workplace Indoor Smoking Ban and Smoking Behavior among Male Workers and Female Nonsmoking Workers' Husbands: A Pseudo Cohort Study of Japanese Public Workers. *BioMed Research International.* 2014;2014:9. doi:10.1155/2014/303917
4. Otani K, Nishiyama K, Ito Y, Kawaguchi Y, Inaji H. Steroid treatment increases the recurrence of radiation-induced organizing pneumonia after breast-conserving therapy. *Cancer Med.* 2014;3:947-53. doi:10.1002/cam4.255
5. Marubashi S, Gotoh K, Akita H, Takahashi H, Sugimura K, Miyoshi N, Motoori M, Kishi K, Noura S, Fujiwara Y, Ohue M, Nakazawa T, Nakanishi K, Ito Y, Yano M, Ishikawa O, Sakon M. Analysis of Recurrence Patterns After Anatomical or

Non-anatomical Resection for Hepatocellular Carcinoma. *Ann Surg Oncol*. 2014. doi:10.1245/s10434-014-4214-4

6. Ito Y, Nakaya T, Nakayama T, Miyashiro I, Ioka A, Tsukuma H, Rachet B. Socioeconomic inequalities in cancer survival: A population-based study of adult patients diagnosed in Osaka, Japan, during the period 1993-2004. *Acta Oncol*. 2014;53:1423-33. doi:10.3109/0284186x.2014.912350

7. Ito Y, Miyashiro I, Ito H, Hosono S, Chihara D, Nakata-Yamada K, Nakayama M, Matsuzaka M, Hattori M, Sugiyama H, Oze I, Tanaka R, Nomura E, Nishino Y, Matsuda T, Ioka A, Tsukuma H, Nakayama T, the JCRG. Long-term survival and conditional survival of cancer patients in Japan using population-based cancer registry data. *Cancer Science*. 2014;105:1480-6. doi:10.1111/cas.12525

8. Tabuchi T, Ito Y, Ioka A, Nakayama T, Miyashiro I, Tsukuma H. Tobacco smoking and the risk of subsequent primary cancer among cancer survivors: a retrospective cohort study. *Ann Oncol*. 2013;24:2699-704. doi:10.1093/annonc/mdt279

9. Tabuchi T, Hoshino T, Nakayama T, Ito Y, Ioka A, Miyashiro I, Tsukuma H. Does removal of out-of-pocket costs for cervical and breast cancer screening work? A quasi-experimental study to evaluate the impact on attendance, attendance inequality and average cost per uptake of a Japanese government intervention. *Int J Cancer*. 2013;133:972-83. doi:10.1002/ijc.28095

10. Ito Y, Nakayama T, Miyashiro I, Ioka A, Tsukuma H. Conditional survival for longer-term survivors from 2000--2004 using population-based cancer registry data in Osaka, Japan. *BMC Cancer*. 2013;13:304. doi:10.1186/1471-2407-13-304

11. Ikeda A, Miyashiro I, Nakayama T, Ioka A, Tabuchi T, Ito Y, Tsukuma H. Descriptive epidemiology of bile duct carcinoma in Osaka. *Jpn J Clin Oncol*. 2013;43:1150-5. doi:10.1093/jjco/hyt126

12. Aono N, Ito Y, Nishino K, Uchida J, Kumagai T, Akazawa Y, Okuyama T, Yoshinami T, Imamura F. A Retrospective Study of the Novel Combination of Paclitaxel and S1 for Pretreated Advanced Non-Small Cell Lung Cancer. *Chemotherapy*. 2013;58:454-60. doi:10.1159/000345624

13. Utada M, Ohno Y, Shimizu S, Ito Y, Tsukuma H. Cancer incidence and mortality in Osaka, Japan: future trends estimation

with an age-period-cohort model. *Asian Pac J Cancer Prev*. 2012;13:3893-8.

14. Tabuchi T, Ito Y, Ioka A, Miyashiro I, Tsukuma H. Incidence of metachronous second primary cancers in Osaka, Japan: Update of analyses using population-based cancer registry data. *Cancer Sci*. 2012;103:1111-20. doi:10.1111/j.1349-7006.2012.02254.x

15. Hanaoka N, Ishihara R, Takeuchi Y, Uedo N, Higashino K, Ohta T, Kanzaki H, Hanafusa M, Nagai K, Matsui F, Iishi H, Tatsuta M, Ito Y. Intralesional steroid injection to prevent stricture after endoscopic submucosal dissection for esophageal cancer: a controlled prospective study. *Endoscopy*. 2012;44:1007-11. doi:10.1055/s-0032-1310107

〔学会発表〕(計 17 件)

1. Oze I, Ito H, Hosono S, Nishino Y, Hattori M, Ioka A, Nakayama T, Tanaka H, Ito Y. Trends in lung cancer survival in Japan 1993-2006 (J-CANSIS): small cell histologic subtype specific survival. 第 25 回日本疫学会学術総会. 2015. 名古屋

2. Ito Y, Miyashiro I, Ito H, Hosono S, Chihara D, Nakata-Yamada K, Nakayama M, Matsuzaka M, Hattori M, Sugiyama H, Oze I, Tanaka R, Nomura E, Nishino Y, Matsuda T, Ioka A, Tsukuma H, Nakayama T, the J-CANSIS research group. Trends in long-term survival for cancer patients in Japan using population-based cancer registry data in 1993-2006: the J-CANSIS study. 第 25 回日本疫学会学術総会. 23 Jan. 2015:AW-07. 名古屋

3. 伊藤ゆり, 中山富雄, 山崎秀男, 津熊秀明. 大阪府における大腸がん検診の精検受診勧奨方法と精検受診率の関係: 効果的な精検受診勧奨方法は何か?. 第 24 回日本疫学会学術総会. 2014.1.24-25. 2014. 仙台

4. 伊藤ゆり, 杉本知之, 宮代 勲, 中山富雄. 地域がん登録データに対する治療モデルの適用とその結果の考察. 2014 年度日本計量生物学会年会. 23-24, May (24) 2014;事例研究[P7]. 東京都立川市

5. 伊藤ゆり. 肺がん生存率の国際比較. 第 29 回肺がん集検セミナー. 16 Nov 2014. 京都

6. 中山雅志, 林裕次郎, 石津谷祐, 武田健, 山口唯一郎, 新井康之, 垣本健一, 伊藤ゆり, 西村和郎. 術後補助療法未施行筋層非浸潤性膀胱癌の膀胱内再発に対するリスク因子の検討. 第 51 回日本癌治療学会学術集會. 2013.10.24-26 2013;48:1616. 京都

7. 小向翔, 杉本知之, 服部聡, 伊藤

ゆり. 相対生存率に対するセミパラメトリック最尤推定. 日本計算機統計学会第 27 回大会. 2013. 弘前

8. 伊藤ゆり, 中谷友樹, 中山富雄, 田淵貴大, 宮代勲, 井岡亜希子, 池田章子, 津熊秀明. がん患者の生存率の社会経済因子による格差: 大阪府がん登録資料による検討. 第 23 回日本疫学会学術総会. 24-26 Jan 2013:76(0-06). 大阪

9. Sugimoto T, Ito Y. Relative survival regression trees with application to cancer registry data. 34th Annual Conference of the International Society for Clinical Biostatistics. 26 Aug (25-29 Aug) 2013:C.4.1. Munich, Germany

10. Miyashiro I, Ikeda A, Ioka A, Ito Y, Nakata K, Tabuchi T, Nakayama T, Yano M, Tsukuma H. Multimodality therapy for pancreatic cancer: Outcomes based on the hospital cancer registry of OMCC. The 72nd annual meeting of Japanese Cancer Association. 10.3-5. 2013. Yokohama, Japan

11. Komukai S, Sugimoto T, Hattori S, Ito Y. Semi-parametric Maximum Likelihood Estimation for Relative Survival rate. 34th Annual Conference of the International Society for Clinical Biostatistics. 26 Aug (25-29 Aug) 2013:C.4.3. Munich, Germany

12. Ito Y, Nakayama T, Miyashiro I, Tsukuma H. Prognostic model of cancer patients using hospital cancer registry: application of the relative survival regression tree. The 72nd annual meeting of Japanese Cancer Association. 10.3-5. 2013. Yokohama, Japan

13. Ito Y, Nakayama T, Ioka A, Tanaka H, Tsukuma H. Descriptive epidemiology of prostate cancer in Osaka, Japan: 1975-2010. The 35th annual meeting of International Association of Cancer Registries. 10.22-24. 2013. Buenos Aires, Argentina

14. Ito Y, Nakaya T, Nakayama T, Ioka A, Tsukuma H, Rachet B. Socioeconomic inequalities in cancer survival in Osaka, Japan: 1993-2004. The 35th annual meeting of International Association of Cancer Registries. 10.22-24. 2013. Buenos Aires, Argentina

15. Ito Y, Nakaya T, Ioka A, Nakayama T, Tsukuma H. Investigating spatial clusters of cancer incidence in Osaka Prefecture, Japan: An application of GIS for Cancer Control. IGU Kyoto Regional Conference. 4-9 Aug. 2013. Kyoto, Japan

16. Ito Y, Nakayama T, Miyashiro I, Tsukuma H, Rachet B. Trends in ' Cure from Cancer in Osaka, Japan in 1975-2004: 13 Major Sites of Solid Tumour. World Cancer

Congress, UICC. 26-30 August 2012:0C20, Track 2: Cancer care and Survivorship [Oral]. Montreal, Canada

17. Ito Y, Nakayama T, Miyashiro I, Tabuchi T, Ioka A, Tsukuma H. Conditional five-year relative survival for cancer survivors from 2000-2004 in Osaka, Japan, (がん X 年サバイバーにおけるその後の 5 年相対生存率: 2000-2004). 71st Annual Meeting of the Japanese Cancer Association. 19th Sep. 2012:P-1220 (Poster Session: P24-4 Descriptive Epidemiology) [Poster]. 札幌

〔図書〕(計 3 件)

1. Ishikawa O, Ohhigashi H, Takahashi H, Ito Y. Survival 3, 5, and 10 years after adjuvant regional and neoadjuvant chemotherapy in resectable pancreatic cancer patients: an institutional experience. In: Beger HG, Peng S, Sarr M, Nakao A, editors. Pancreatic Cancer, Cystic Neoplasm and Endocrine Tumors: Diagnosis and Management: Wiley-Blackwell; 2014 [in press].

2. 井岡亜希子, 伊藤ゆり, 津熊秀明. 第 3 章 がんの罹患と死亡動向の府県別分析. In: 祖父江友孝, 片野田耕太, 味木和喜子, 津熊秀明, 井岡亜希子, editors. がん統計白書 2012. 東京: 篠原出版新社; 2012. p. 43-61.

3. 伊藤ゆり. 第 9 章 がん生存率の国際比較. In: 祖父江友孝, 片野田耕太, 味木和喜子, 津熊秀明, 井岡亜希子, editors. がん統計白書 2012. 東京: 篠原出版新社; 2012. p. 205-15.

〔産業財産権〕
出願状況 (計 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況 (計 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
取得年月日:
国内外の別:

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

伊藤 ゆり (ITO Yuri)
地方独立行政法人大阪府立病院
機構大阪府立成人病センター(研究所)・がん
予防情報センター・研究員
研究者番号： 60585305

(2) 研究分担者

なし

(3) 連携研究者

なし